Dell Acceleration Appliance for Databases 2.0

Configuration Guide

2015 - 08

Rev. A03

Contents

About this guide

This guide contains information about installing and configuring the Dell Acceleration Appliance for Databases (DAAD) software. This guide is intended for administrators responsible for server and storage systems. It is assumed the reader is familiar with basic server administration.

Typographical conventions

This document follows these conventions:

Convention	Usage	Examples
NOTE:	Important additional information or further explanation of a topic.	**NOTE:** A weekly backup is recommended.
CAUTION!	The task or operation might have serious consequences if conducted incorrectly or without appropriate safeguards. If you are not an expert in the use of this product, consult support for assistance.	**CAUTION!** Do not change configuration parameters.
Bold	A command or system input that you type, or text or a button you click on a graphical user interface (GUI).	Click **Help** for details about disaster recovery.
Italic	Italic font indicates any of the following: • A term with a specific meaning in the context of this document. • Emphasis on specific information. • Reference to another document. • Variables in a syntax statement for which values are substituted.	Detailed information about disaster recovery methods is available in the Administrator Guide. `network:ping` *`hostname`*
`Courier`	System output, file names or path names. Bold Courier for commands typed by user.	`> Recovery in progress` **`network:ping 10.1.100.14`**
< > Angle Brackets	A required entry or variable parameter	`installer-<version#>.run`
Square [ ] Brackets	An optional entry or variable parameter.	**`tar [zxvf] file.tgz`**
Curly { } Brackets	A list of options separated by a the pipe symbol "|" from which any one must be selected.	Click { **OK** | **Cancel** }.

1

About DAAD configurations

The steps for configuring DAAD depend on the networking hardware that is installed in the appliance. Generally, DAAD can be configured for one of the following data networking protocols:

- Fibre Channel
- iSCSI
- InfiniBand (IB)

About Fibre Channel configurations

Fibre channel DAAD configurations consist of the following components:

- Four Fusion ioMemory devices that are either 3.2 TB or 6.4 TB capacity. (All the Fusion ioMemory devices in the DAAD are the same capacity.)
- Two dual-port fibre channel cards.
- Optionally, a ConnectX-3, 40 GbE, dual-port adapter for a high availability link between two DAAD systems.

Figure 1-1. DAAD Fibre Channel Configuration

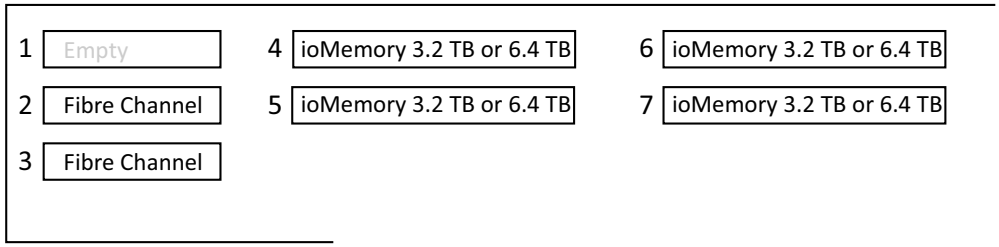
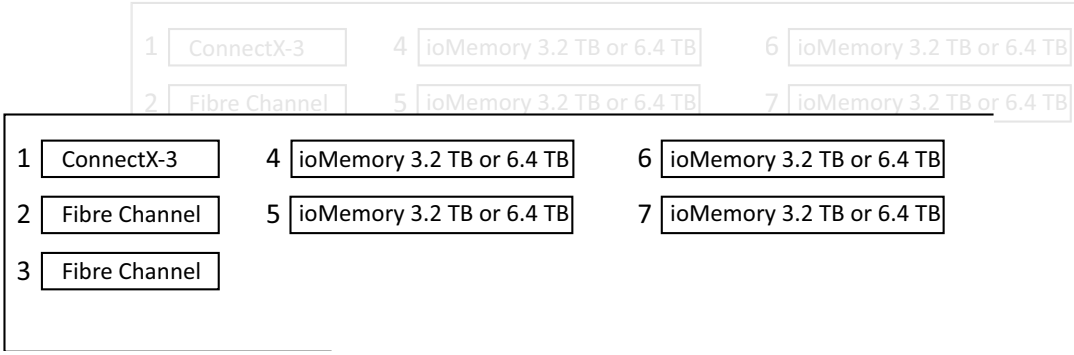


Figure 1-2. DAAD HA Fibre Channel Configuration



About iSCSI configurations

iSCSI DAAD configurations for consist of the following components:

- Four Fusion ioMemory devices that are either 3.2 TB or 6.4 TB capacity. (All the Fusion ioMemory devices in the DAAD are the same capacity.)
- Two ConnectX-3 adapters used for iSCSI protocol connections to the DAAD
- Optionally, one ConnectX-3 adapter used for a high availability link between two DAAD systems.

Figure 1-3. DAAD iSCSI Configuration

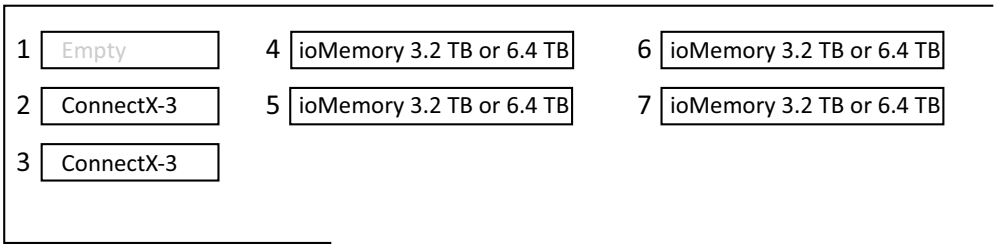


Figure 1-4. DAAD HA iSCSI Configuration

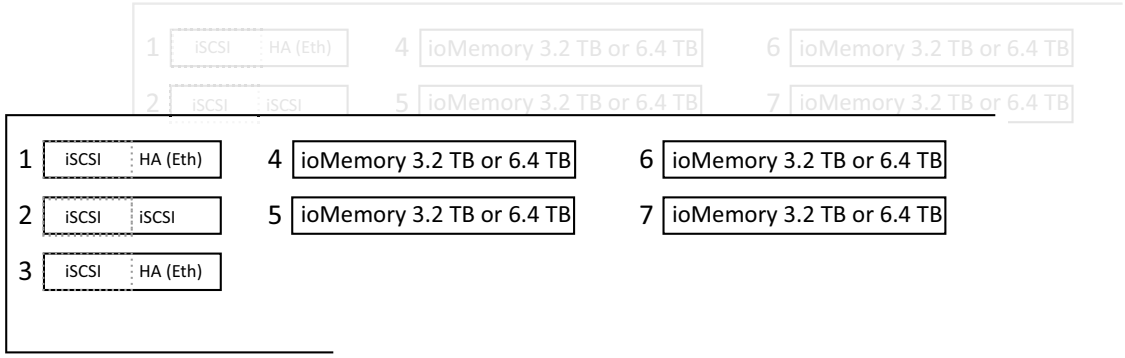


Figure 1-5.

About InfiniBand configurations

InfiBand DAAD configurations consist of the following components:

- Four Fusion ioMemory devices that are either 3.2 TB or 6.4 TB capacity. (All the Fusion ioMemory devices in the DAAD are the same capacity.)
- Two InfiniBand adapters used for IB protocol connections to the DAAD
- Optionally, for HA configurations, the system may be configured with three IB adapters with two ports on two of the adapters configured to be an Ethernet link between two DAAD systems in the cluster.

Figure 1-6. DAAD InfiniBand Configuration

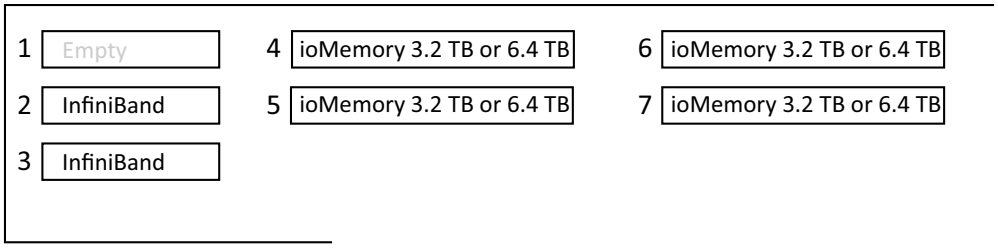
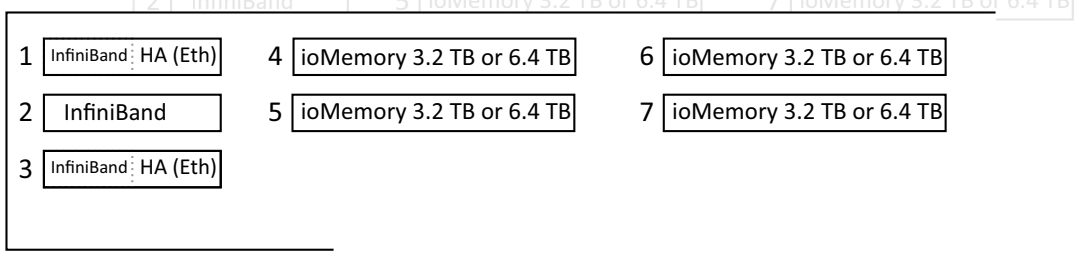


Figure 1-7. DAAD HA InfiniBand Configuration



2

About DAAD first boot

When booting DAAD for the first time,, perform some basic configuration tasks for the DAAD system. These tasks can be grouped into the following phases:

- Accepting the license agreement
- Configuring networking
- Setting clock and time zone
- Configuring HA clustering
- Setting the admin user password

The networking adapters installed in the your system will determine what is displayed on the Network Configuration screen. As directed in Table 2-1 on page 18, go to the section that matches the Network Configuration screen displayed on the console of your system.

Prerequisites

Before beginning, ensure that you have access to the following items:

- Keyboard and video monitor for First Boot of the Dell Acceleration Appliance for Databases
- IP address assignments, if static IP addresses will be used
- IP default gateway setting
- SSL certificate for remote access. A pre-configured SSL certificate is provided, but it will trigger security warnings; therefore, Dell recommends that you use your own SSL certificate.

To install the Dell Acceleration Appliance for Databases software on your target server, perform the following steps:

> **NOTE:** To navigate through the installation screens, press Tab until your selection is highlighted, and then press Enter. Or, you can press **Alt+<highlighted letter>**, such as **Alt+H** for Help, or **Alt+N** for Next. Press the down arrow key to scroll through a list.

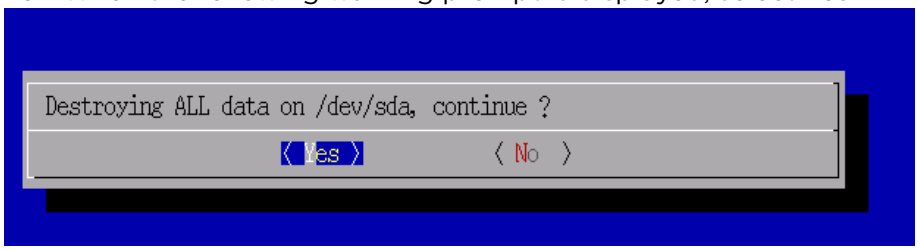
1. From the customer support site download the ISO image for the DAAD software.
2. Burn an installation DVD that contains the .ISO image. You can also softmount the ISO image through iDRAC or you can configure it on a bootable USB.

3 Power on the DAAD and boot the machine with the ISO image. The install screen is displayed:

4 Select **Install ION Accelerator**.

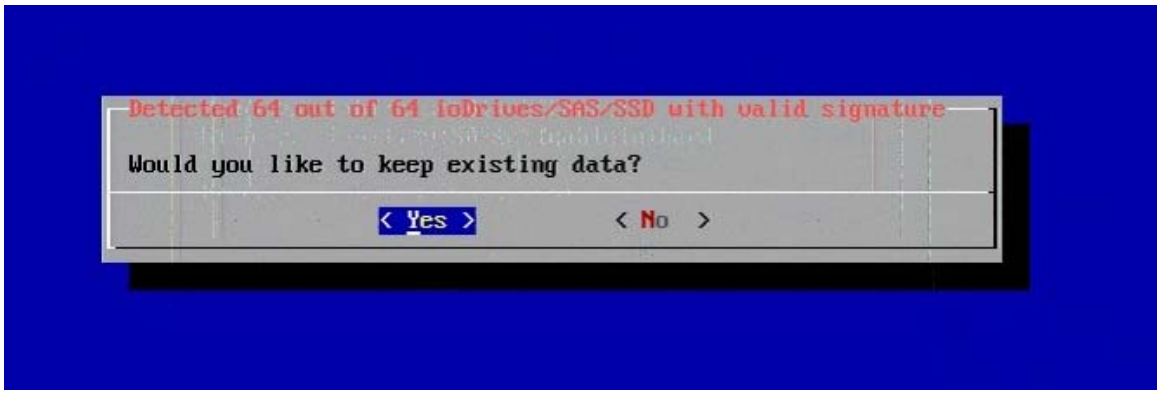
5 If you have multiple disks in the appliance, select the disk where the Dell Acceleration Appliance for Databases software will be installed, such as `/dev/sda`.

6 When the following warning prompt is displayed, select **Yes**:



A progress dialog box is displayed, first for loading the software and then for verifying the `/dev/sda` section on the boot disk.

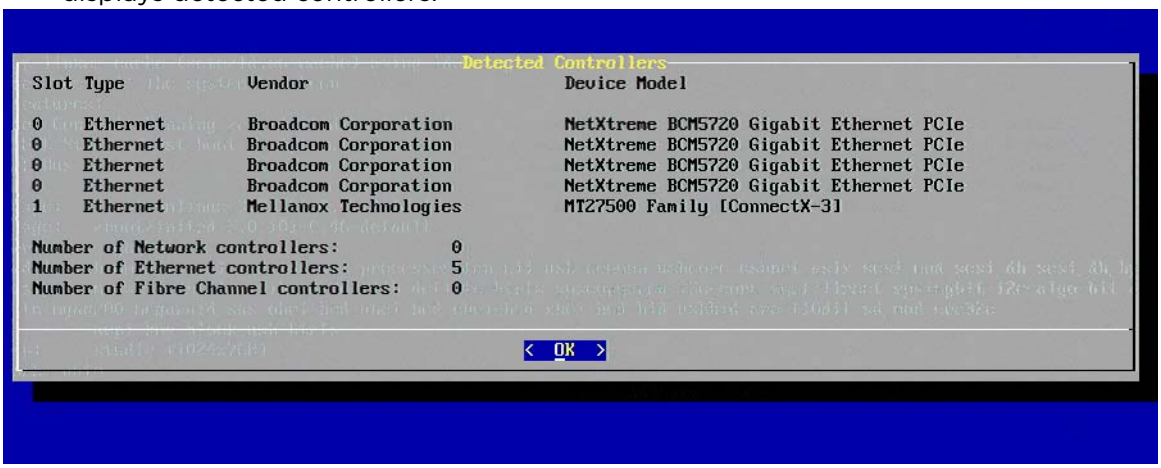
DAAD detects the drives available in the IF100 and displays a dialog box asking you to confirm deletion of data on the drives.



7 To remove the previous data from all the drives, select **No**.

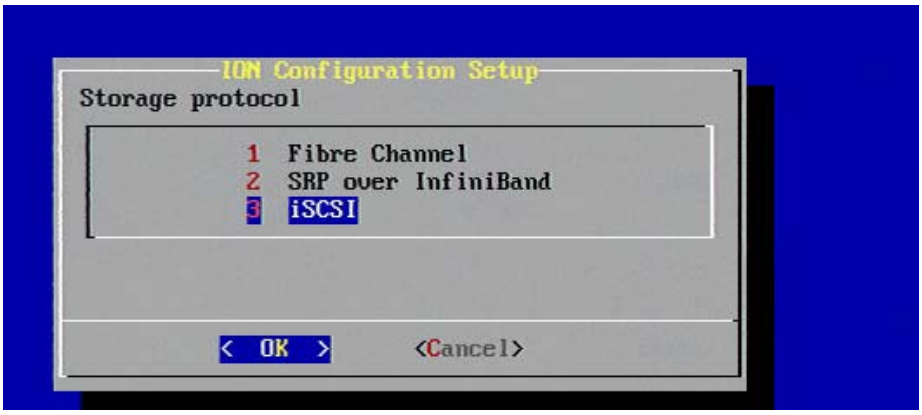
8 Confirm that you want to delete data by selecting **Yes**.

After DAAD software is installed, the software scans for existing network cards and displays detected controllers:



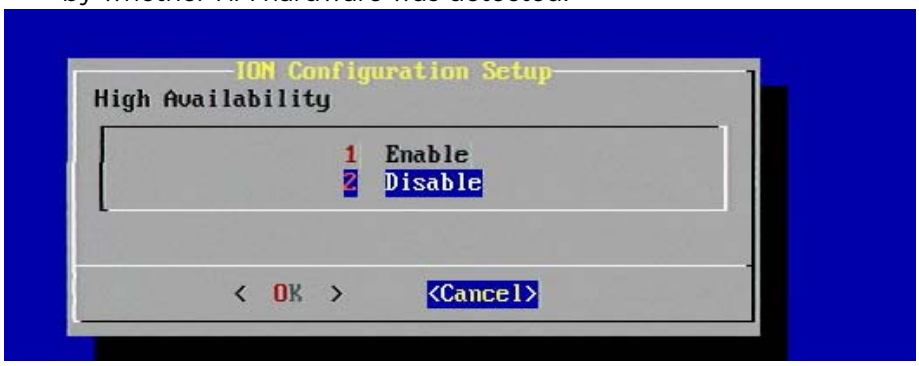
9 Select **OK** to proceed.

The Storage Protocol Selection dialog box is displayed showing the detected protocol as the default selection.



10 Select **iSCSI** or **Fibre Channel** and then select **OK**.

The High Availability dialog box is displayed, showing a default selection determined by whether HA hardware was detected:



11 Select **Disable** and then select **OK**.

The ION Configuration Summary dialog box is displayed, with a summary of the selected protocol and mode (HA or standalone)..

12 Select **Yes** to proceed with the Dell Acceleration Appliance for Databases configuration.

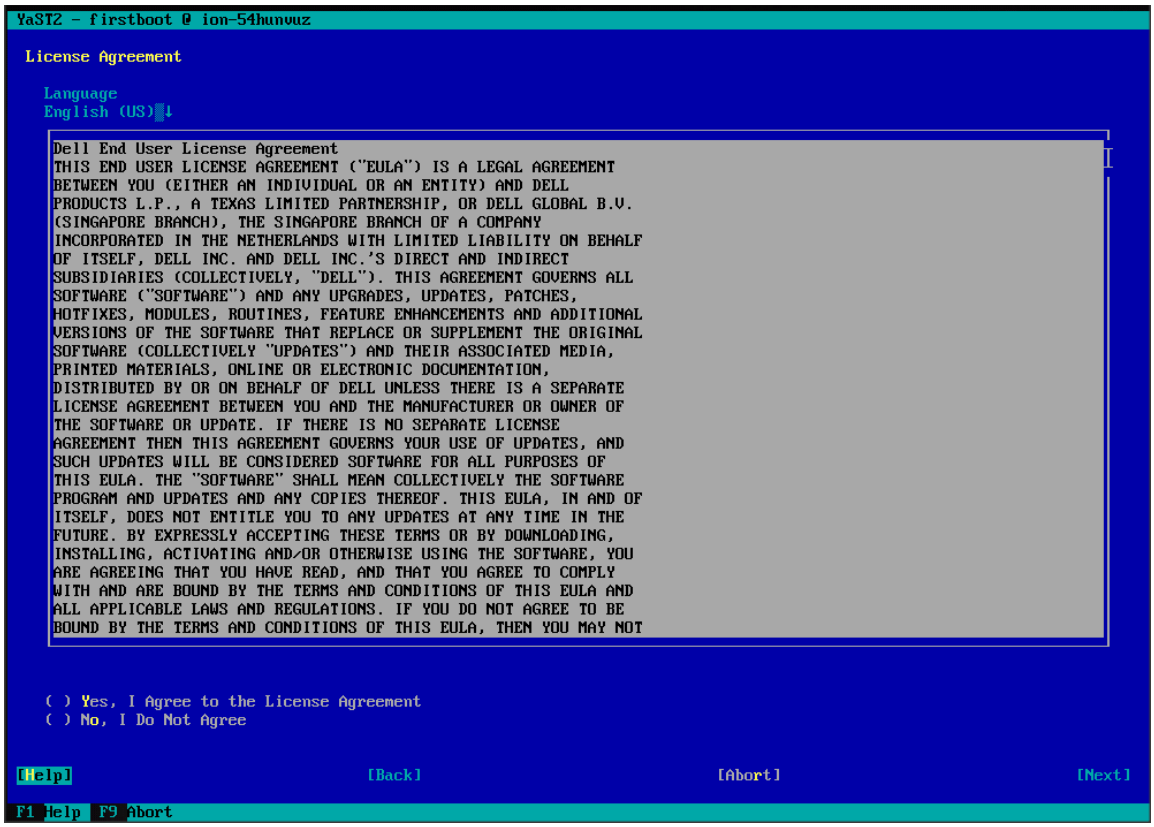
After DAAD completes configuration it will display the Network Configuration screen.

Accepting the license agreement

When you boot DAAD for the first time, a number of start-up messages are displayed. However, the first screen that requires user input is the End-User License Agreement (EULA) screen. When the EULA is displayed, read the agreement and accept it.

> **NOTE:** To navigate through the installation screens, press Tab until your selection is highlighted, and then press Enter. Or, you can press Alt+<highlighted letter>, such as **Alt+H** for help, or **Alt+N** for Next. Press the down arrow key to scroll through a list.

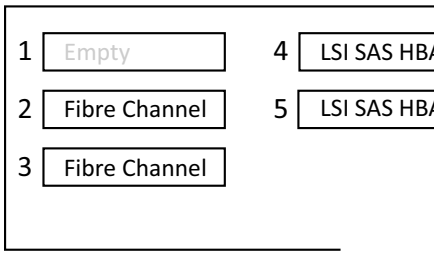
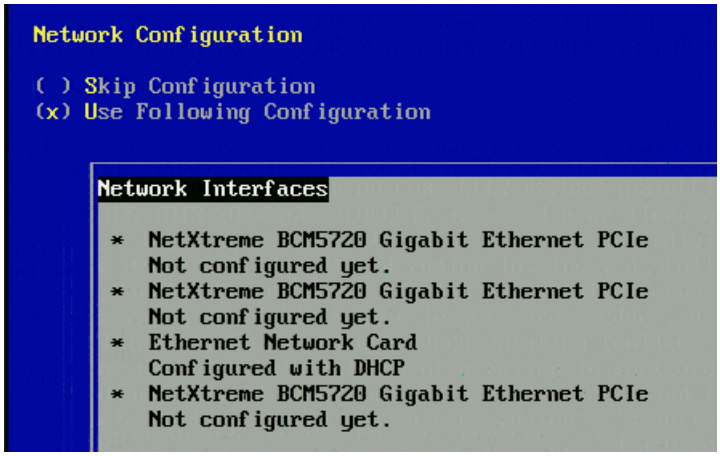
1. Check **Yes** at the bottom of the screen to accept the agreement.
2. Select **Next** to continue.



Configuring networking

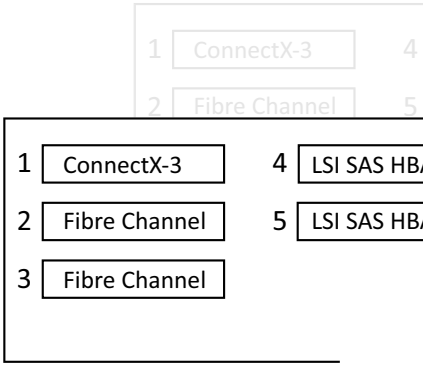
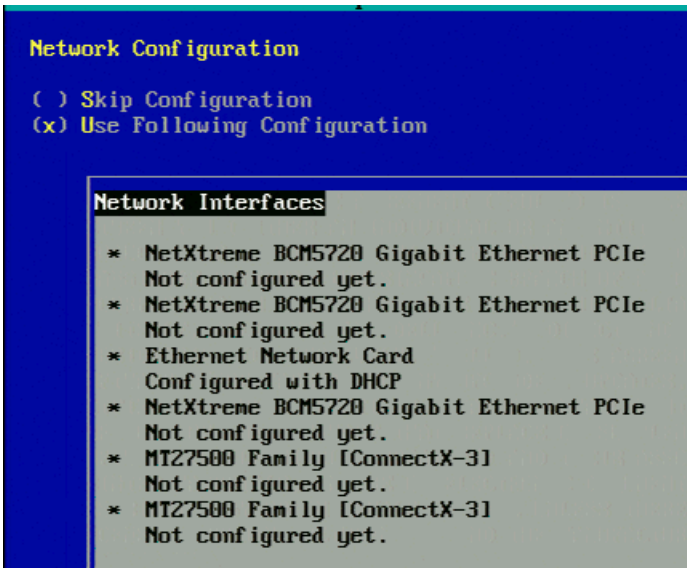
After accepting the EULA, the Network Configuration screen displays the Ethernet adapters that are installed in your system. By using Table 2-1, determine the section of this guide to use to configure DAAD networking.

Table 2-1. DAAD configuration sections

Slots 1-3	Network Configuration Screen
	

For a standalone Fibre Channel configuration, configure one of the built-in Ethernet ports for console access into the DAAD. Instructions about this task are available at Configuring standalone FC and IB on page 25.

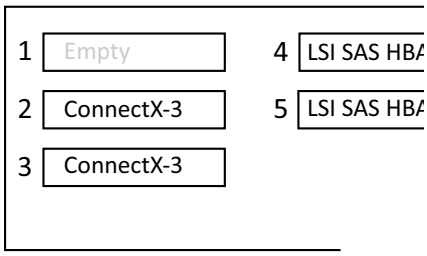
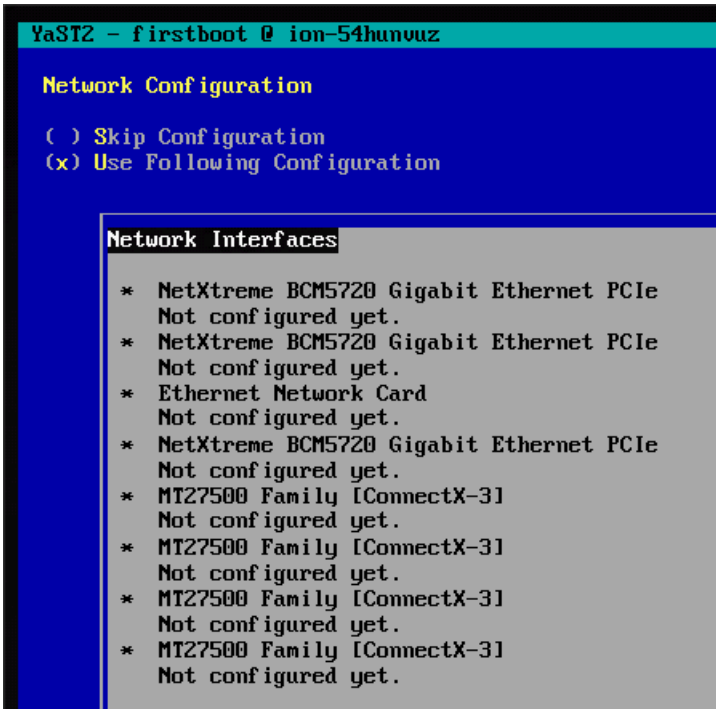
Table 2-1. DAAD configuration sections (Continued)

Slots 1-3	Network Configuration Screen
	

For an HA Fibre Channel configuration, configure one of the built-in Ethernet ports for console access into the DAAD and configure the cluster connection between your two DAAD systems. Instructions about this task are available at Configuring FC and IB for HA on page 73

NOTE: In this configuration, the ConnectX-3 entries represent two ports on the same card. It does not represent two cards.

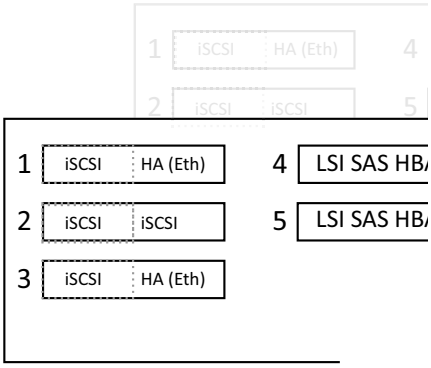
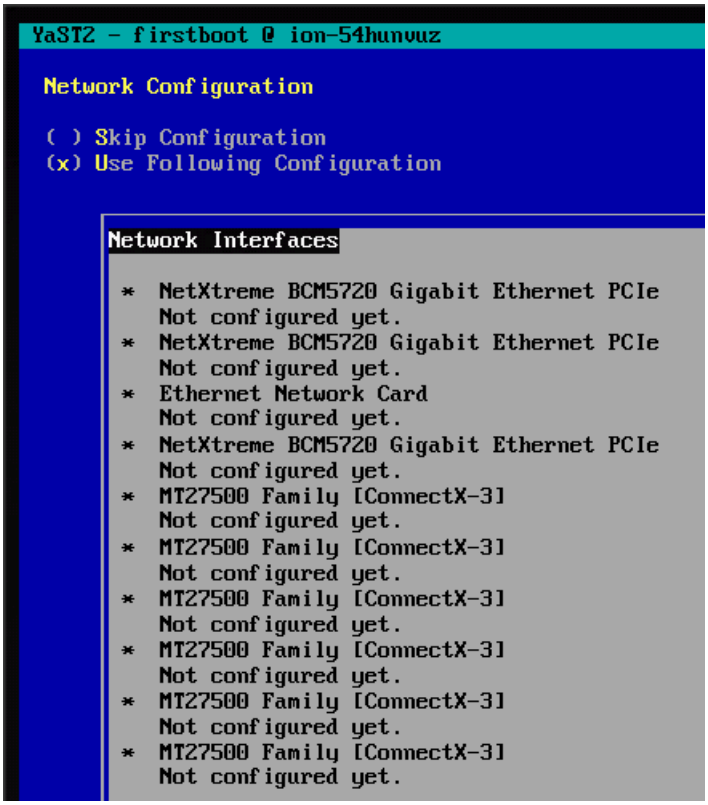
Table 2-1. DAAD configuration sections (Continued)

Slots 1-3	Network Configuration Screen
	

For a standalone iSCSI configuration, configure one of the built-in Ethernet ports for console access into the DAAD and then configure the ConnectX-3 iSCSI interfaces. Instructions about this task are available at Configuring standalone iSCSI on page 37

NOTE: In this configuration, the ConnectX-3 entries represent four ports on two different cards

Table 2-1. DAAD configuration sections (Continued)

Slots 1-3	Network Configuration Screen
	

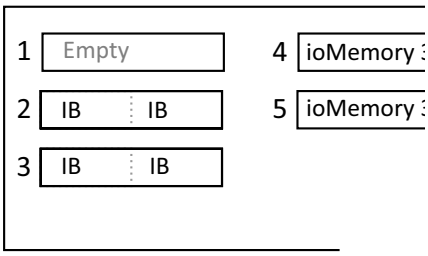
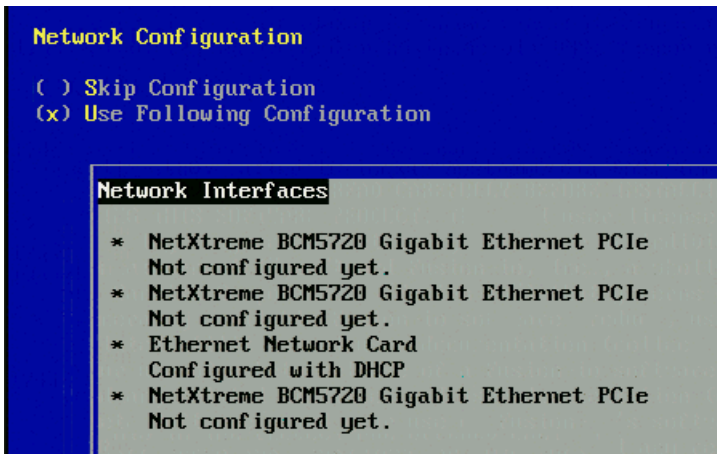
For an HA iSCSI configuration, configure the following:

- One of the built-in Ethernet ports for console access into the DAAD
- Four ConnectX-3 iSCSI interfaces
- Two ConnectX-3 ports for the HA cluster connection between your two DAAD systems.

Instructions about this task are available at Configuring iSCSI for HA on page 53

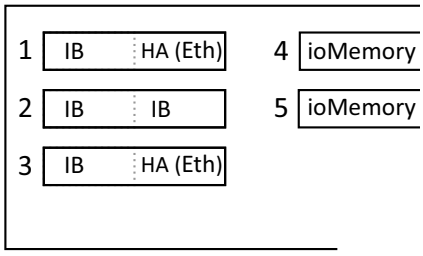
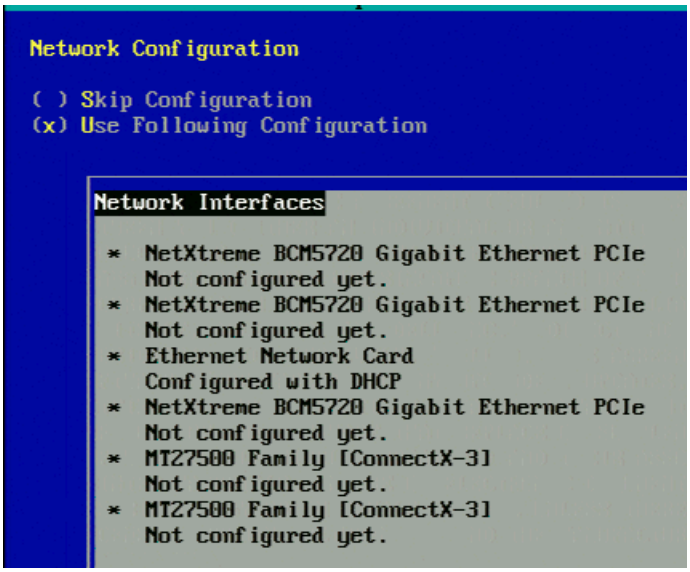
NOTE: In this configuration, the ConnectX-3 entries represent six ports on three different cards. Do not configure two ports on the same card for the HA cluster connection.

Table 2-1. DAAD configuration sections (Continued)

Slots 1-3	Network Configuration Screen
	

For a standalone InfiniBand configuration, configure one of the built-in Ethernet ports for console access into the DAAD. Instructions about this task are available at Configuring standalone FC and IB on page 25.

Table 2-1. DAAD configuration sections (Continued)

Slots 1-3	Network Configuration Screen
	
	For an HA InfiniBand configuration, configure one of the built-in Ethernet ports for console access into the DAAD and configure the cluster connection between your two DAAD systems. Instructions about this task are available at Configuring FC and IB for HA on page 73 **NOTE:** In this configuration, the ConnectX-3 entries represent a single port on two different cards.

3

Configuring standalone FC and IB

The steps for configuring a standalone fibre channel DAAD system and a standalone InfiniBand/SRP DAAD system are the same. To configure either your standalone fibre channel or standalone InfiniBand/SRP DAAD system, perform the following steps:

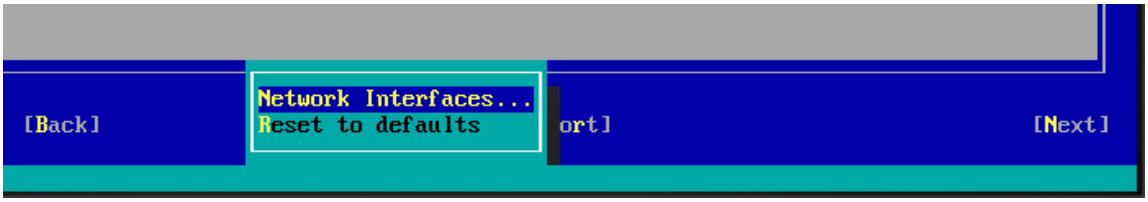
- Configure the management port
- Set the Hostname and Routing information for DAAD
- Set the Time Zone and NTP settings
- Set the admin user password
- Finish the configuration

Configuring the management port

The management port is one you have connected by an external cable to your network and is externally visible. This port must be configured so your network hardware communicates with the Dell Acceleration Appliance for Databases.

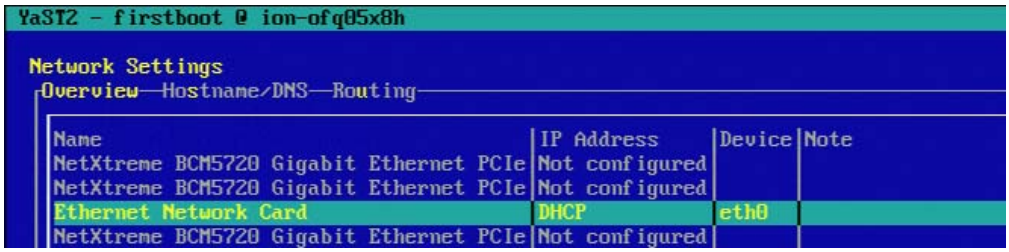
To configure the management port:

1 On the Network Configuration screen, select **Change** (bottom of the screen) and then **Network Interfaces**.



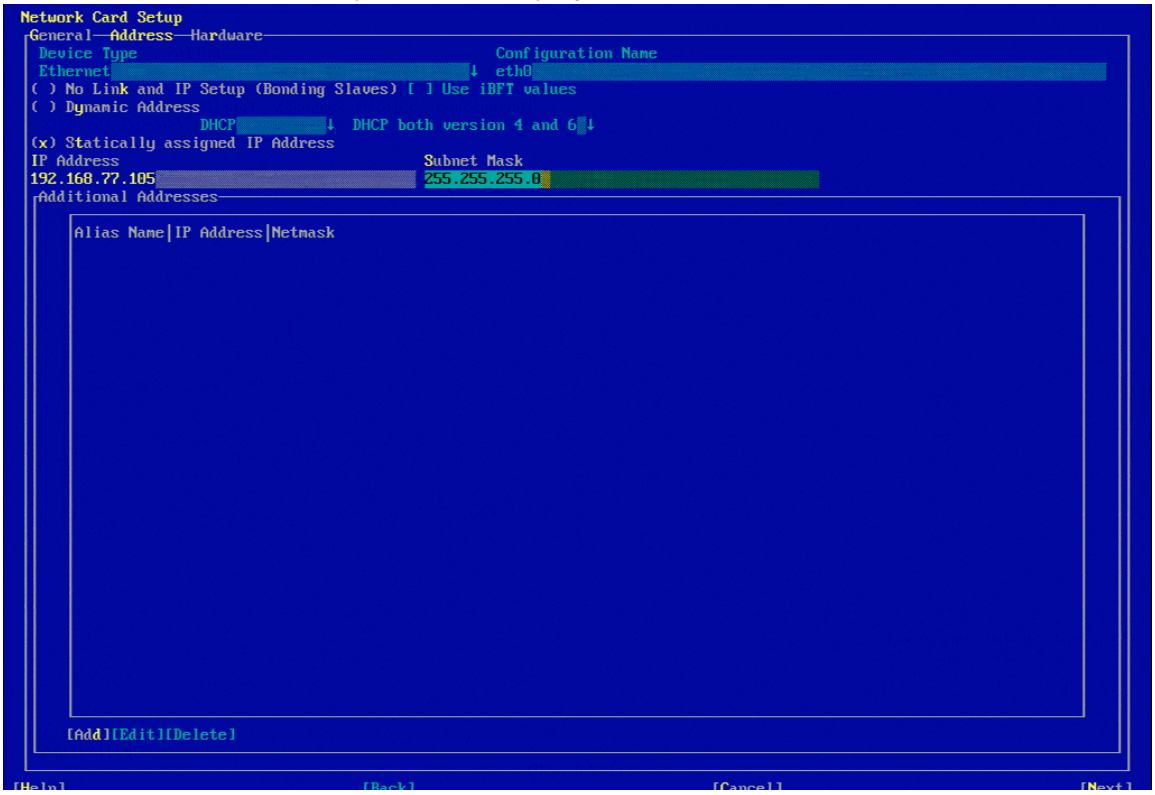
The Network Settings screen is displayed with the detected network adapters listed.

2 Select the management IP card from the list.



3 Select **Edit** (in the lower-left corner of the screen). [Add][Edit][Delete]

The Network Card Setup screen is displayed.



4 Select **Statically assigned IP Address**.

5 Enter the IP Address and Subnet Mask for your network.

6 Select **Next**.

Setting hostname and routing information

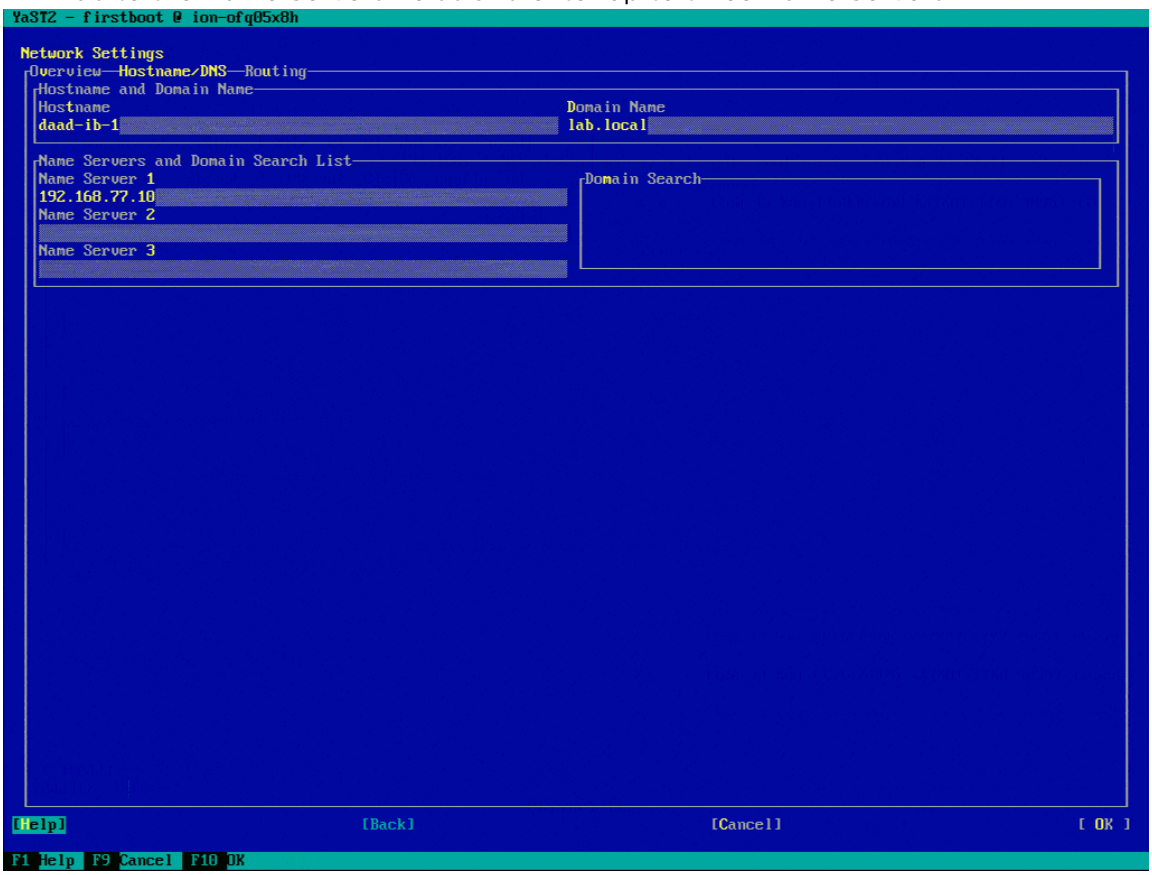
To configure Hostname and routing information:

1 On the Network Settings screen, press the right arrow key to select **Hostname**.

2 Tab to the Hostname field and enter the Hostname for the DAAD.

3 Tab to the Domain Name field and enter the Domain Name.

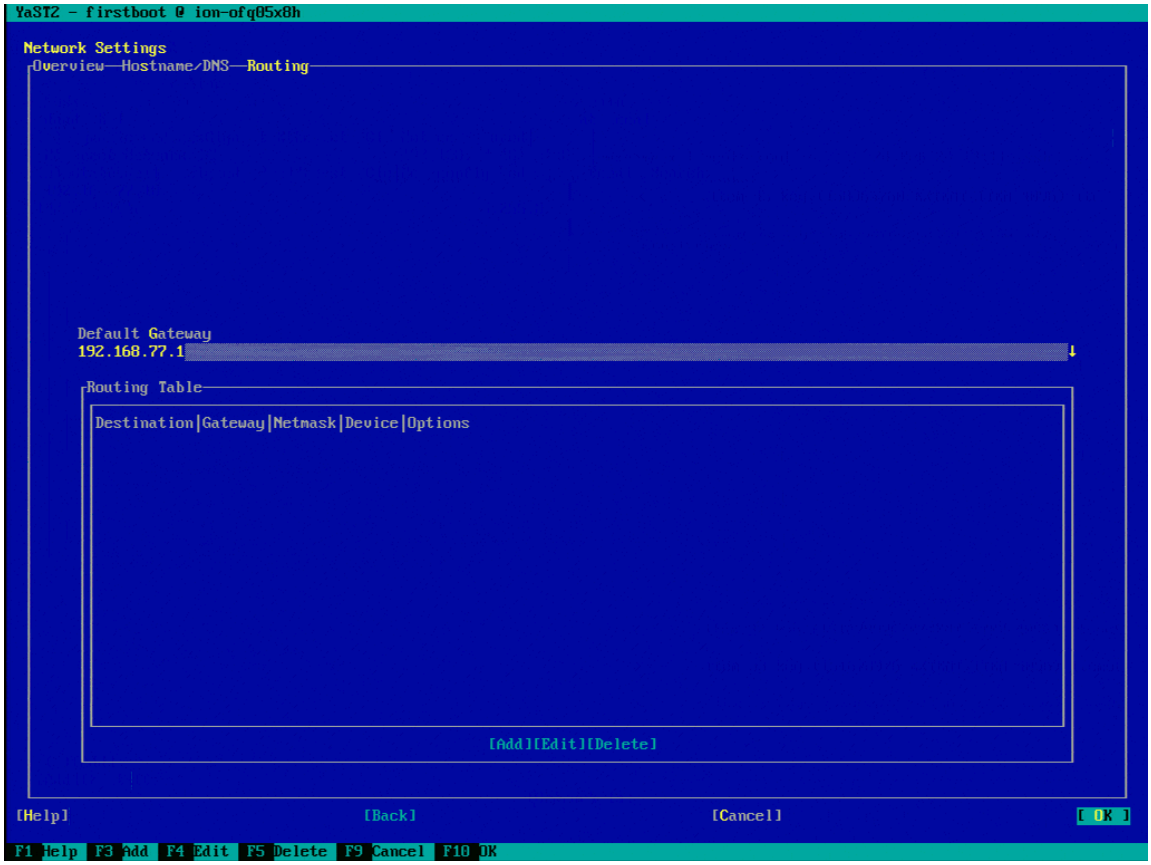
4 Tab to the Name Servers fields and enter up to three Name Servers.



5 Press Tab until the focus is back on the Overview line.

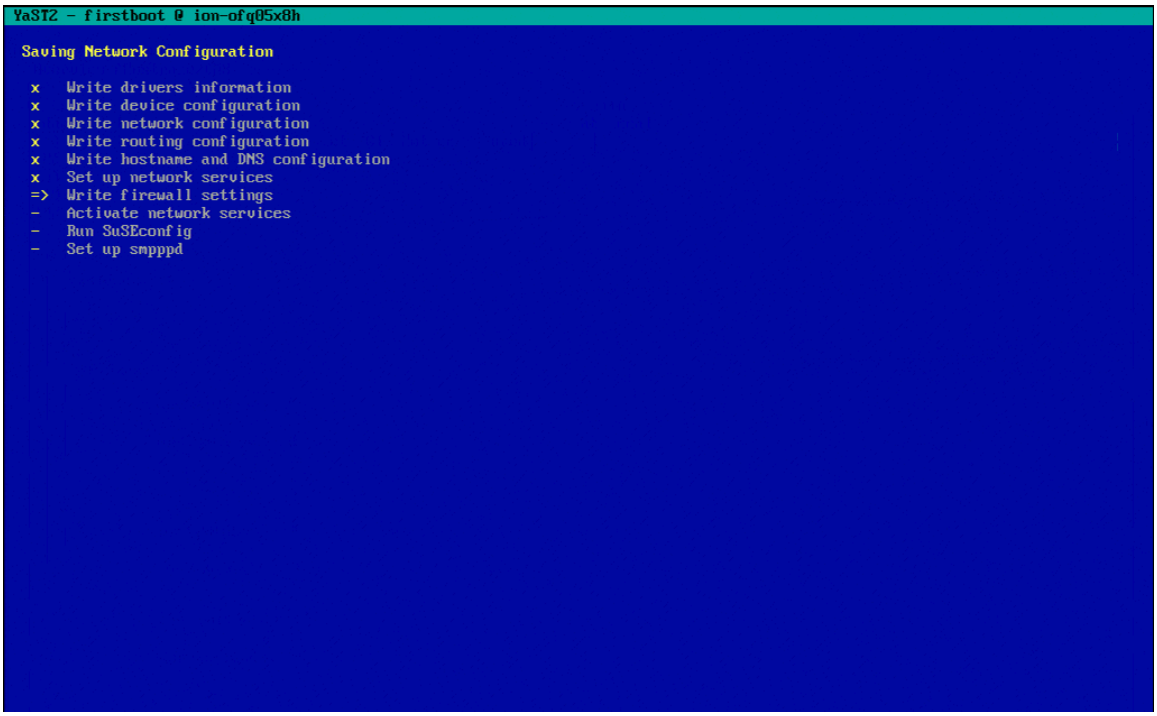
6 Press the right arrow key to select Routing.

7 Tab to the Default Gateway field and enter the Default Gateway.



8 Select **OK**.

The Saving Network Configuration Screen is displayed.

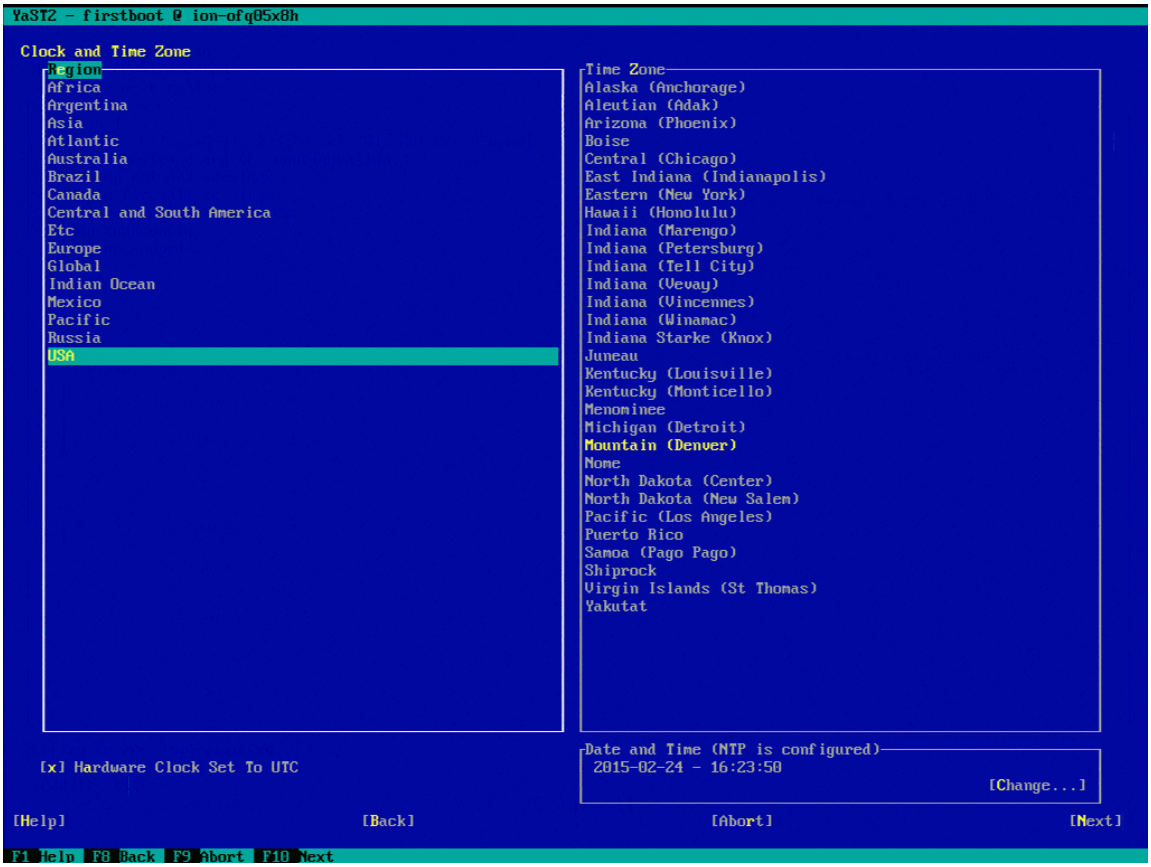


Setting time zone and NTP settings

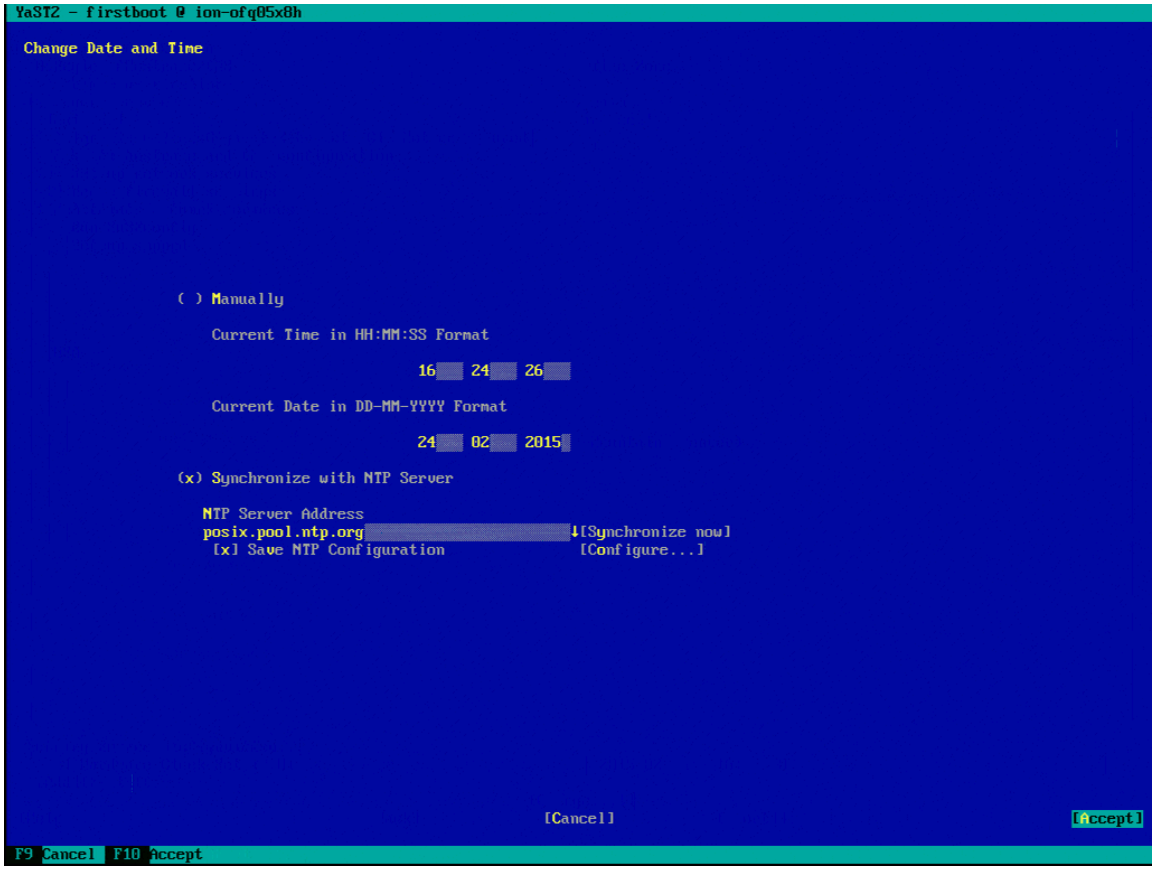
To set Time Zone and NTP:

1 Tab to the Region field and use the arrow keys to select the region where the DAAD will be located.

2 Tab to the Time Zone field and select the zone where the DAAD will be located.



3 Tab to the Date and Time field and select **Change**. The Change Date and Time screen is displayed.



CAUTION! If you are using an HA configuration you must configure DAAD to synchronize with NTP.

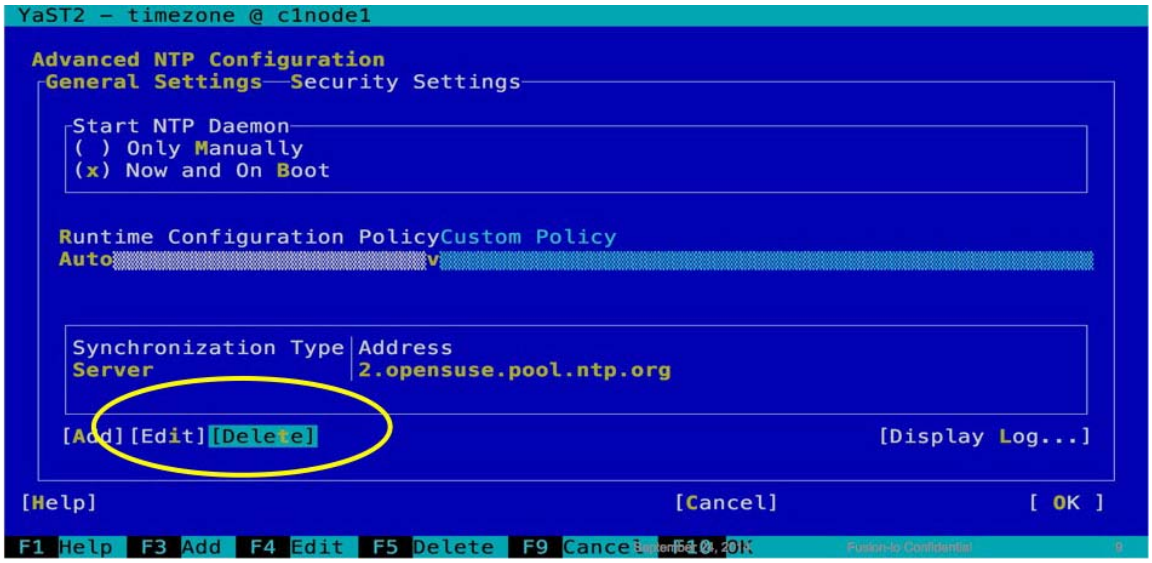
4 Tab to the Synchronize with NTP Server field and select it.

5 Tab to the NTP Server Address field and enter the address of the NTP server to synchronize with.

6 Tab to the Save NTP Configuration field and select it.

7 Tab to the Configure field and select it.

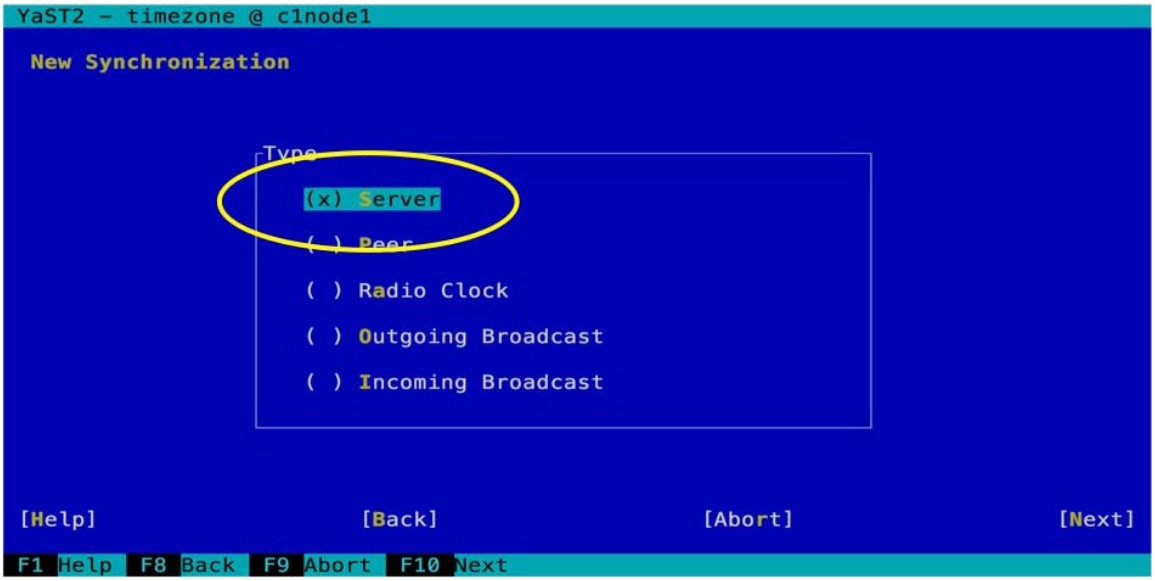
The Advanced NTP Configuration screen is displayed.



8 Select **Delete** to delete the default NTP configuration.

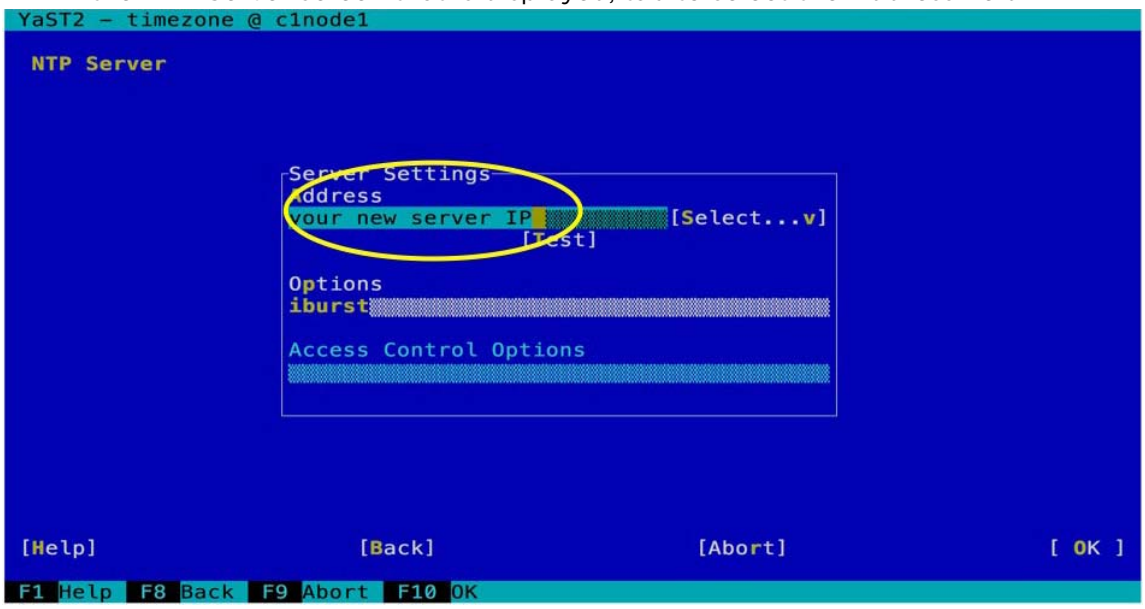
9 Select **Add** to add a new configuration.

10 In the New Synchronization screen that is displayed, select **Server**.



11 Select **Next**.

12 In the NTP Server screen that is displayed, tab to select the Address field.



13 Type the NTP server address.

14 Select **OK**.

15 If you have additional NTP servers, repeat steps 9-14 to configure them.

16 Select **Accept** to save the NTP changes and continue.

Setting the admin user password

To set the admin user password:

1 Type the Dell Acceleration Appliance for Databases password for the admin user.

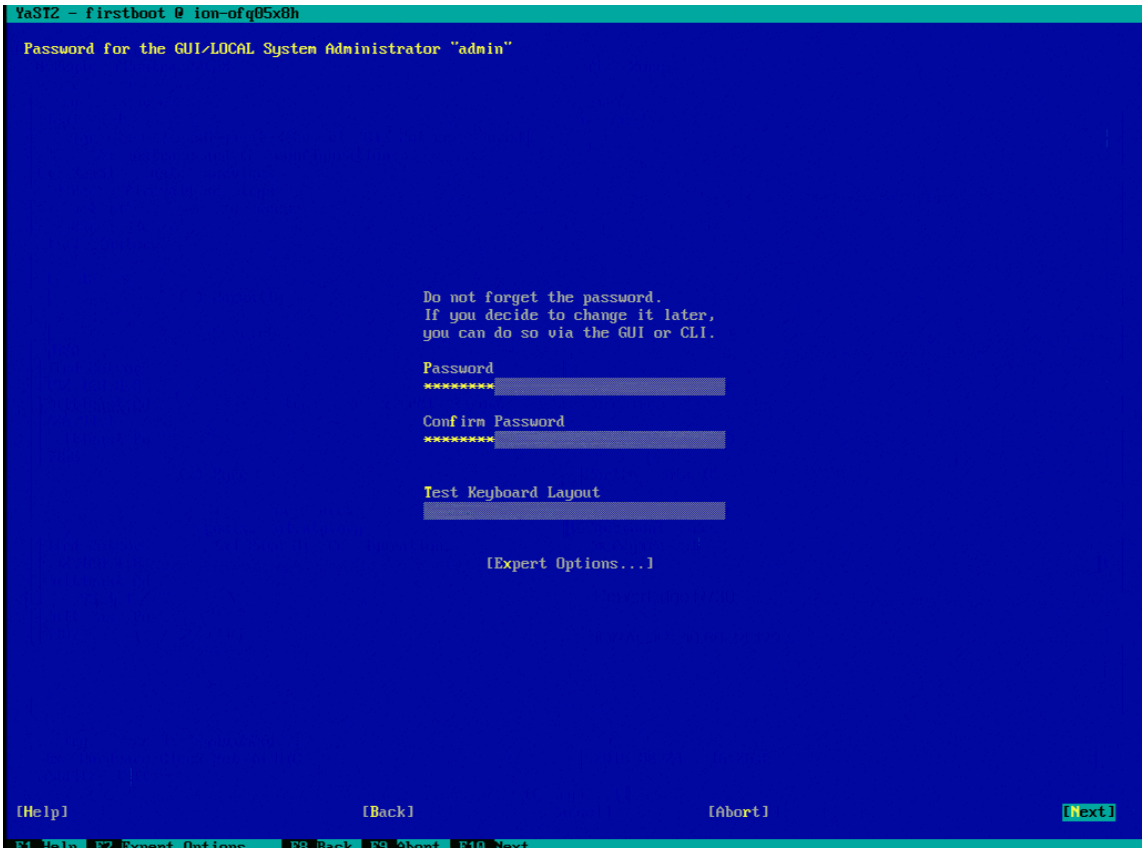
NOTE: If the password you selected is not sufficiently strong, a warning message is displayed so you can change the password, if necessary.

2 Retype the password you entered.

CAUTION! Be sure to record this password in a secure location in case it needs to be retrieved.

3 To test the Keyboard Layout or use the Expert Options, select those options on the screen.

4 Select **Next**.



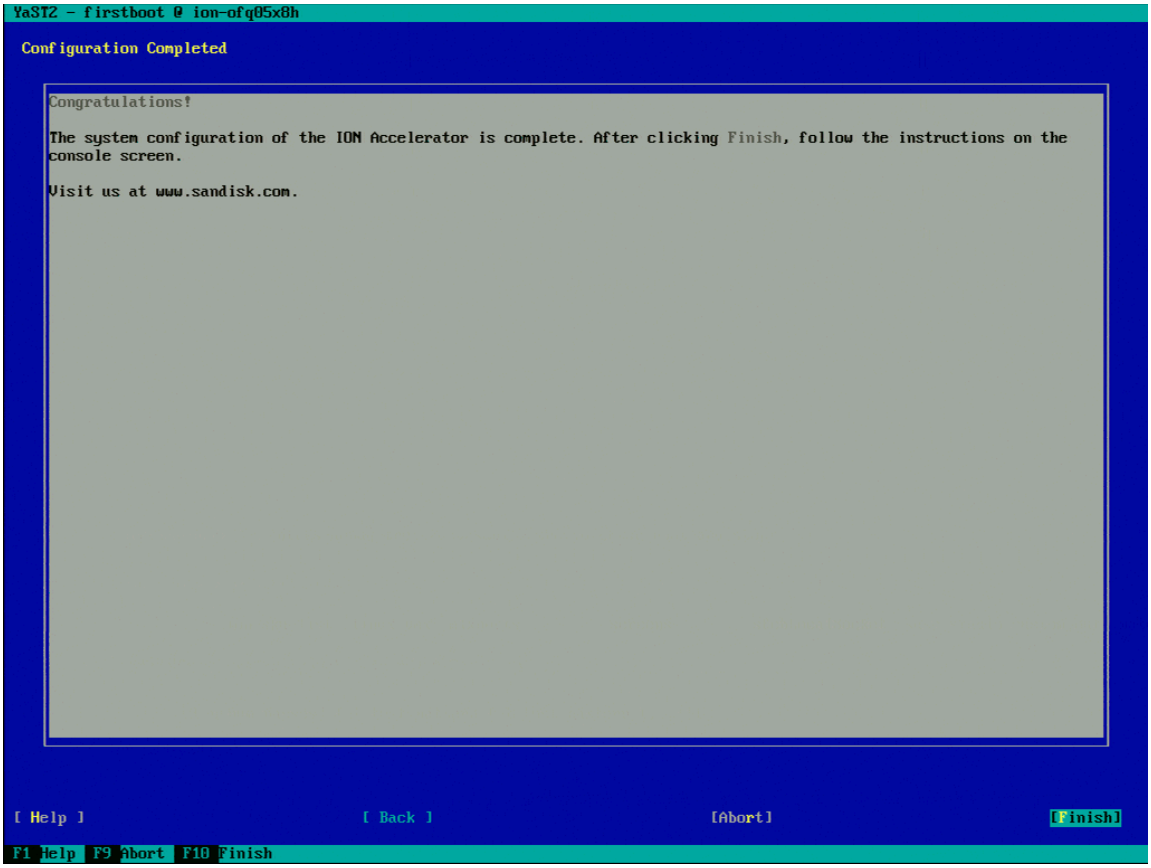
If you want to change this password later, change the admin user and GUI passwords separately. For more information, see Changing Passwords in the *Dell Acceleration Appliance for Databases GUI Guide*.

Completing the DAAD configuration

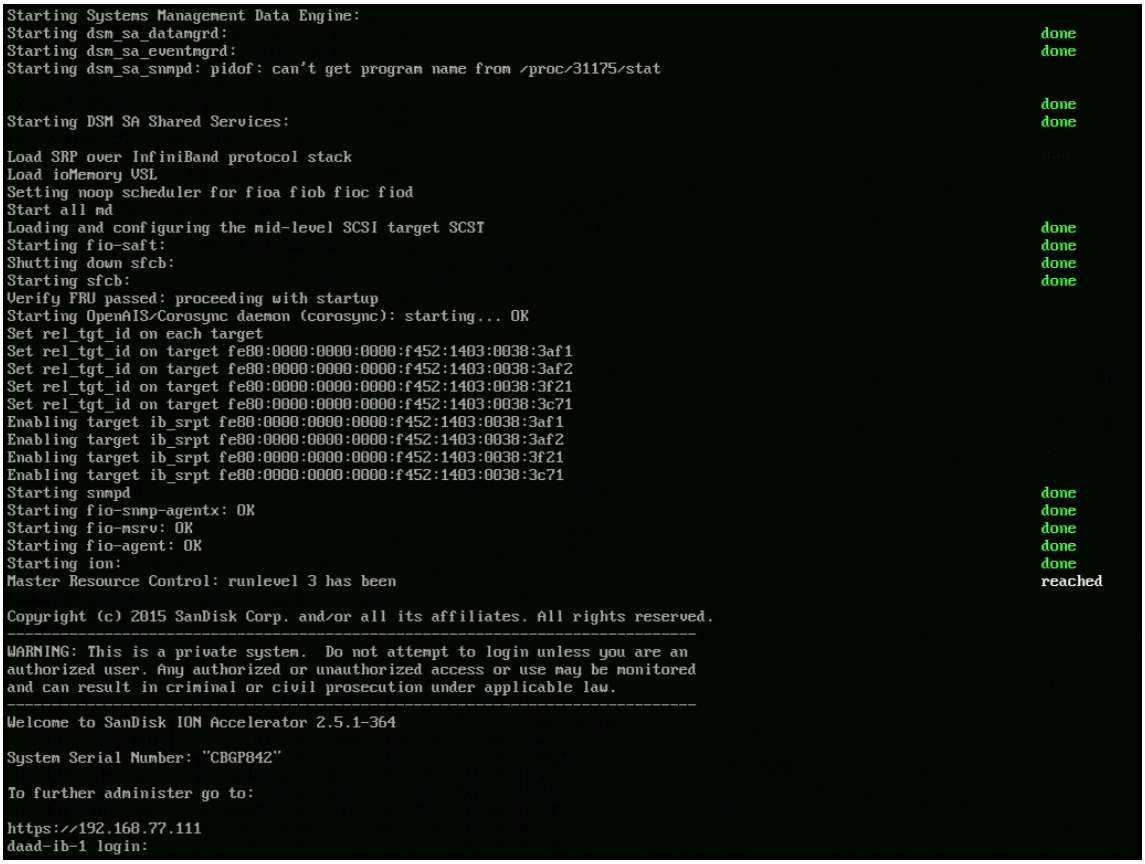
After the configuration phases are complete, the Configuration Completed screen is displayed.

To complete DAAD configuration:

1 Select **Finish**.



The DAAD system restarts and displays start-up messages. After a login prompt is displayed on the console, you can connect to DAAD for additional setup and configuration. An example screenshot is given here:



4

Configuring standalone iSCSI

To configure your standalone iSCSI DAAD system, perform the following steps:

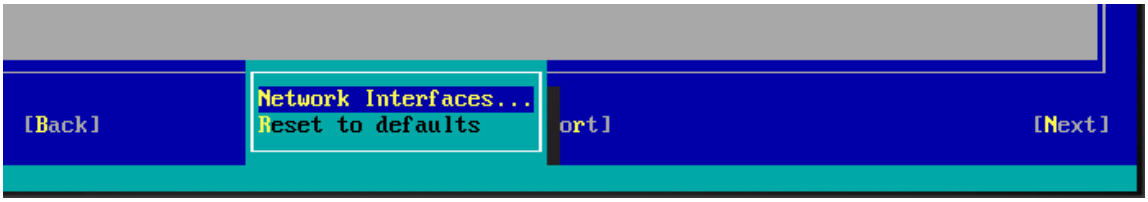
- Configure the management port
- Set the Hostname and Routing information for DAAD
- Set the Time Zone and NTP settings
- Set the admin user password
- Finish the configuration

Configuring the management port

The management port is one you have connected by an external cable to your network and is externally visible. This port must be configured so your network hardware communicates with the Dell Acceleration Appliance for Databases.

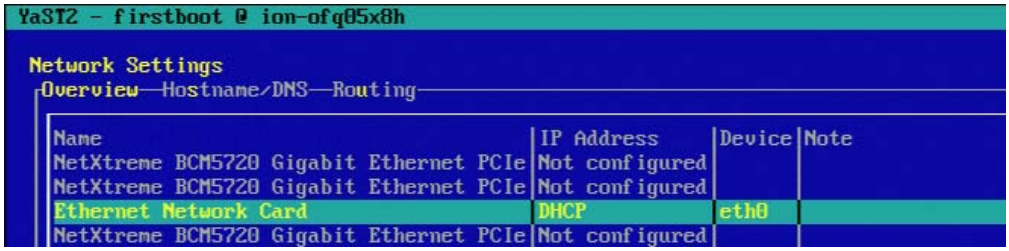
To configure the management port:

1 On the Network Configuration screen, select **Change** (bottom of the screen) and then **Network Interfaces**.



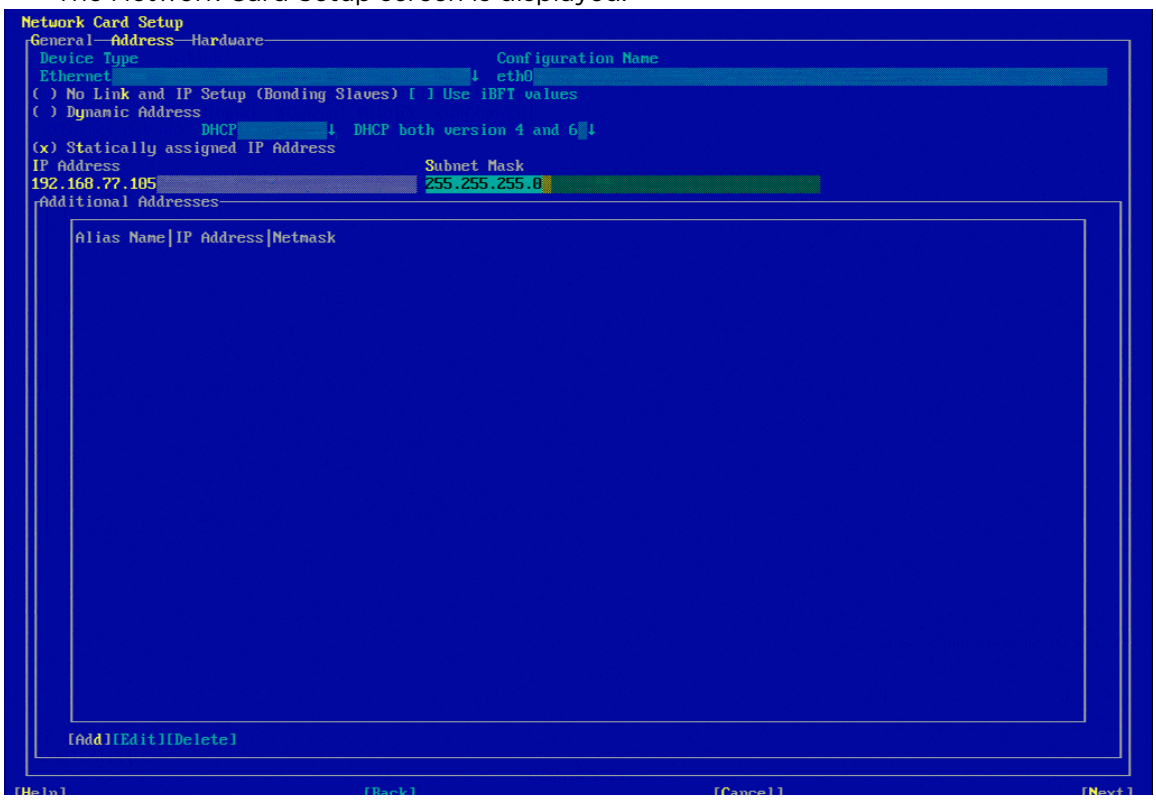
The Network Settings screen is displayed with the detected network adapters listed.

2 Select the management IP card from the list.



3 Select **Edit** (in the lower-left corner of the screen). [Add][Edit][Delete]

The Network Card Setup screen is displayed.



4 Select **Statically assigned IP Address**.

5 Enter the IP Address and Subnet Mask for your network.

6 Select **Next**.

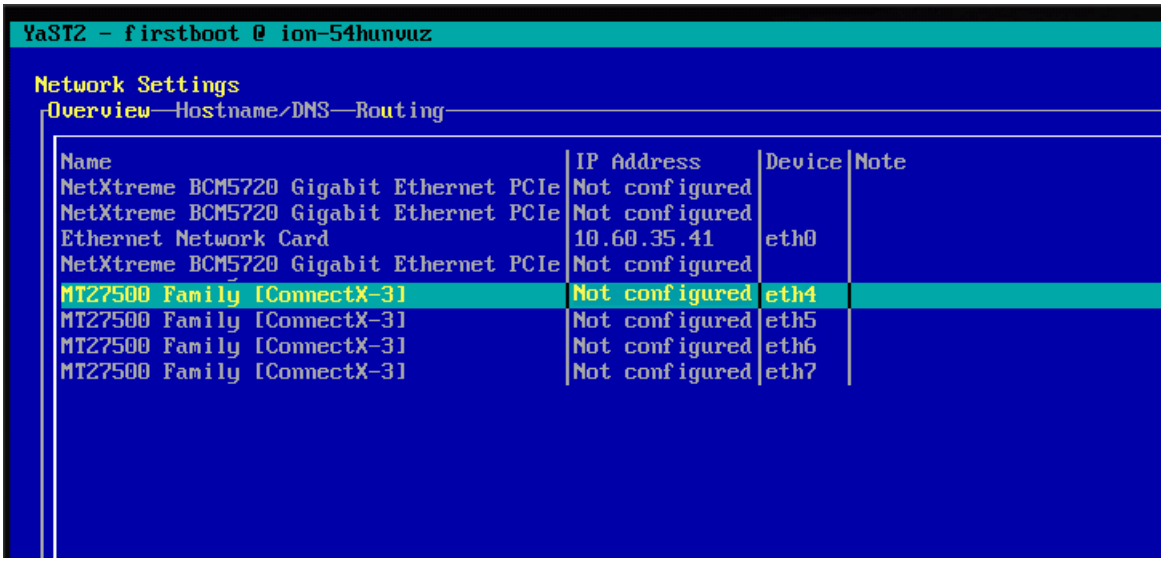
Configuring ConnectX-3 ports for iSCSI

To configure ConnectX-3 ports for iSCSI:

1 On the Network Configuration screen, select **Change** (bottom of the screen) and then **Network Interfaces**.

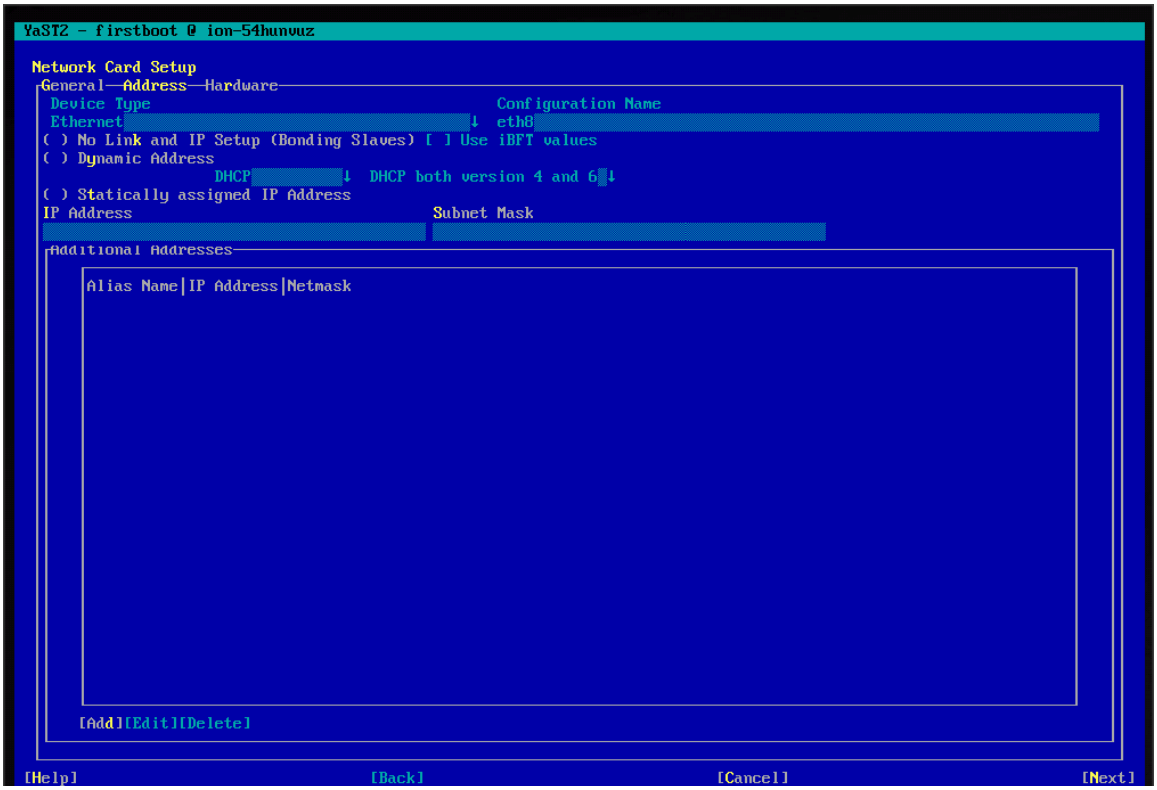
The Network Settings screen is displayed with the detected network adapters listed.

2 Select an unconfigured ConnectX-3 port from the list.



3 Select **Edit** (in the lower-left corner of the screen). [Add][Edit][Delete]

The Network Card Setup screen is displayed.



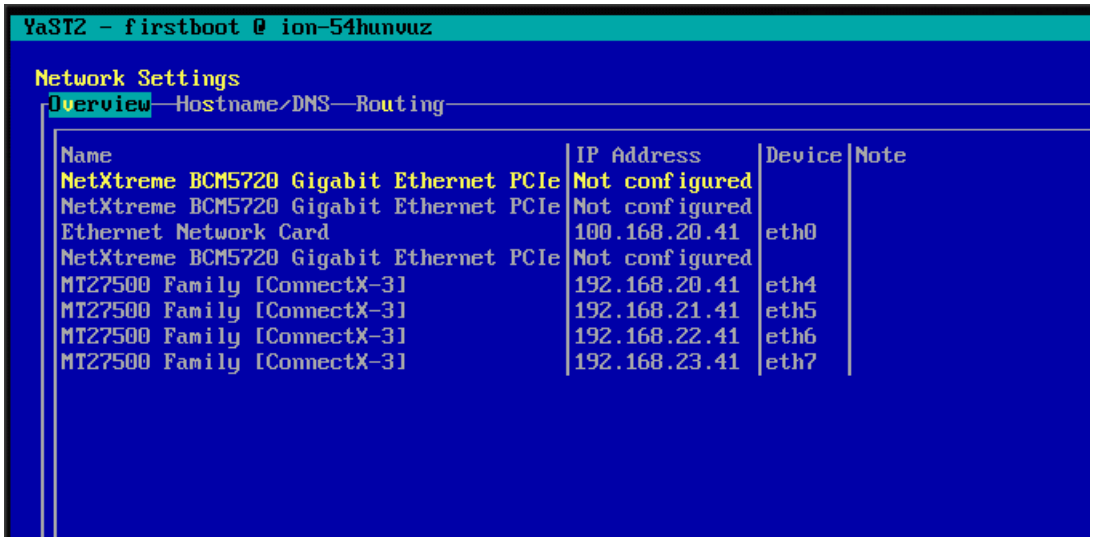
4 Select **Statically assigned IP Address**.

Each iSCSI port must be configured on its own subnet. Also, it is a best practice to have the last octet of the IP address match the last octet of the host.

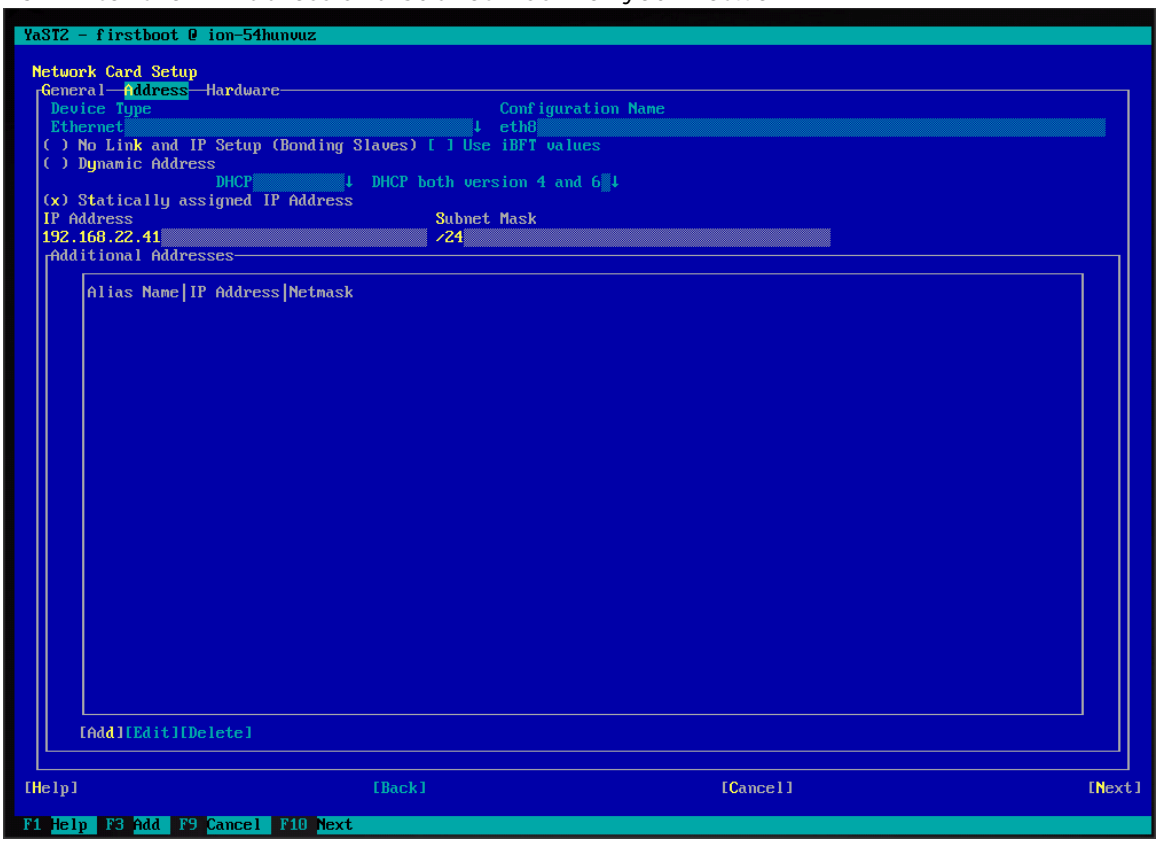
For example, with a host IP address of 100.128.10.41, you might configure the iSCSI ports for the following subnets:

- 192.168.**20**.41
- 192.168.**21**.41
- 192.168.**22**.41
- 192.168.**23**.41

Here is an example of unique subnet numbering for iSCSI ports in the user interface:



5 Enter the IP Address and Subnet Mask for your network.



6 Select **Next**.

7 On the Network Card Setup screen, press left-arrows to select **General**.

8 Tab to the Set MTU field.

9 Set MTU to **9000**.

> **NOTE:** Ensure that the switches and routers on your network are configured to correctly handle this MTU size.

10 Select **Next**.

11 Repeat steps 2–10 until all four iSCSI ports are configured.

Setting hostname and routing information

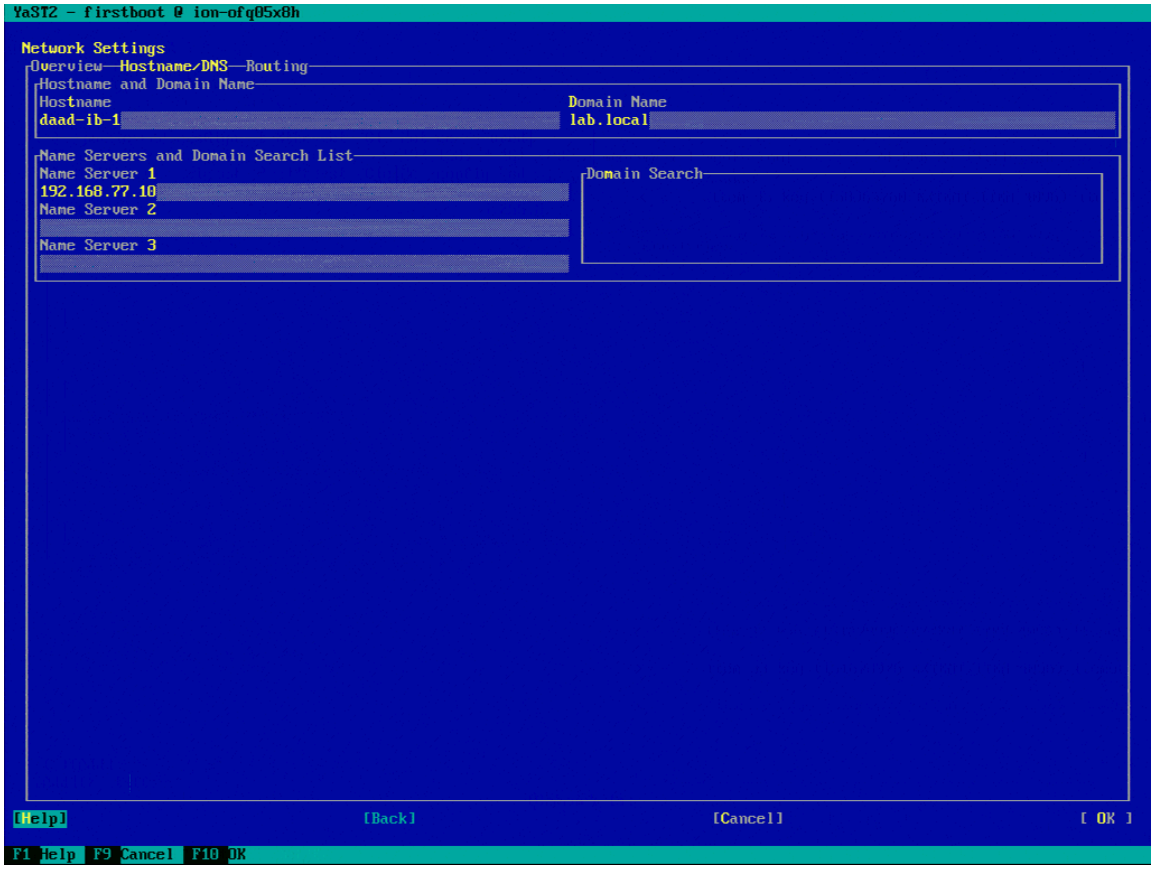
To configure Hostname and routing information:

1 On the Network Settings screen, press the right arrow key to select **Hostname**.

2 Tab to the Hostname field and enter the Hostname for the DAAD.

3 Tab to the Domain Name field and enter the Domain Name.

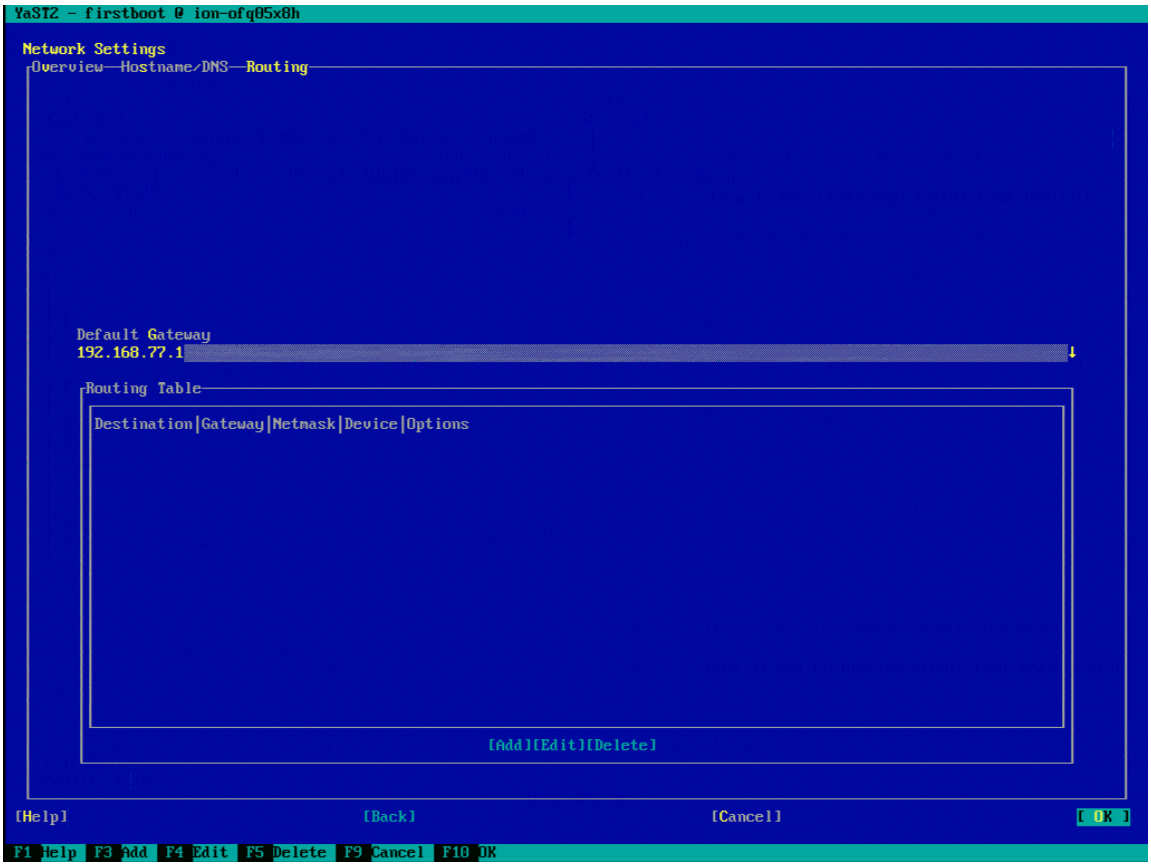
4 Tab to the Name Servers fields and enter up to three Name Servers.



5 Press Tab until the focus is back on the Overview line.

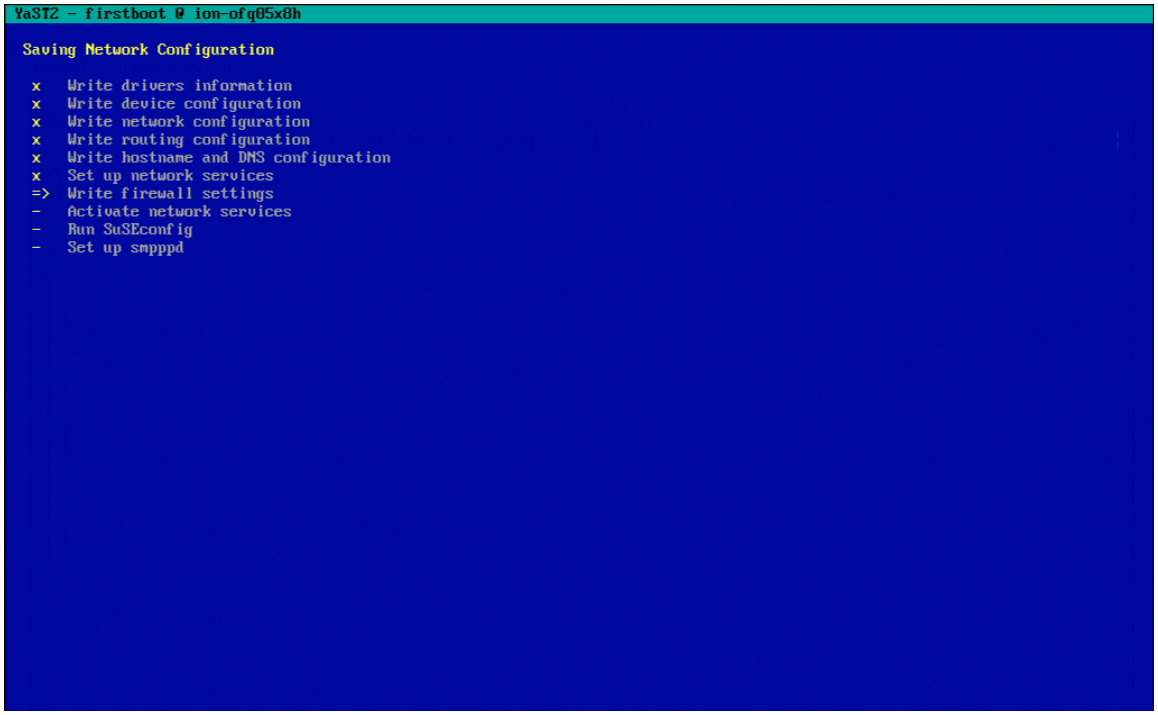
6 Press the right arrow key to select Routing.

7 Tab to the Default Gateway field and enter the Default Gateway.



8 Select **OK**.

The Saving Network Configuration Screen is displayed.

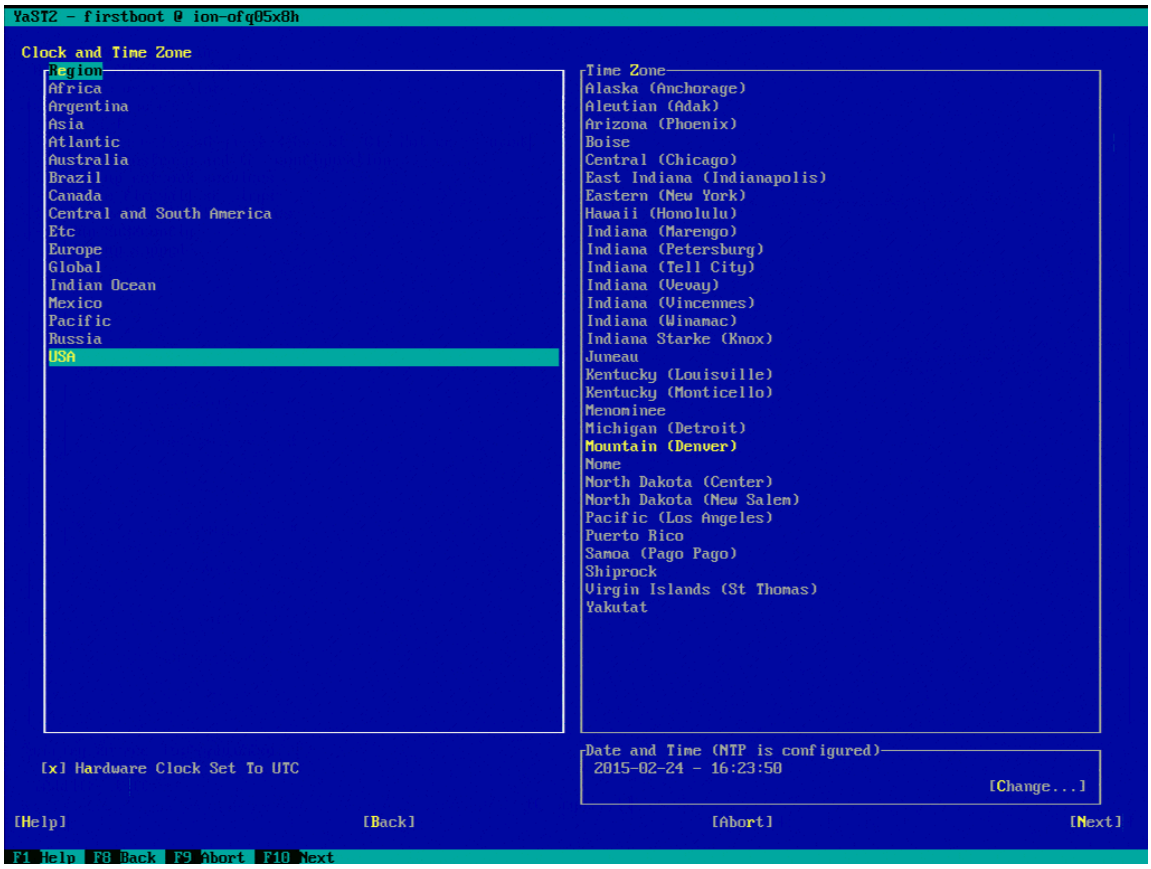


Setting time zone and NTP settings

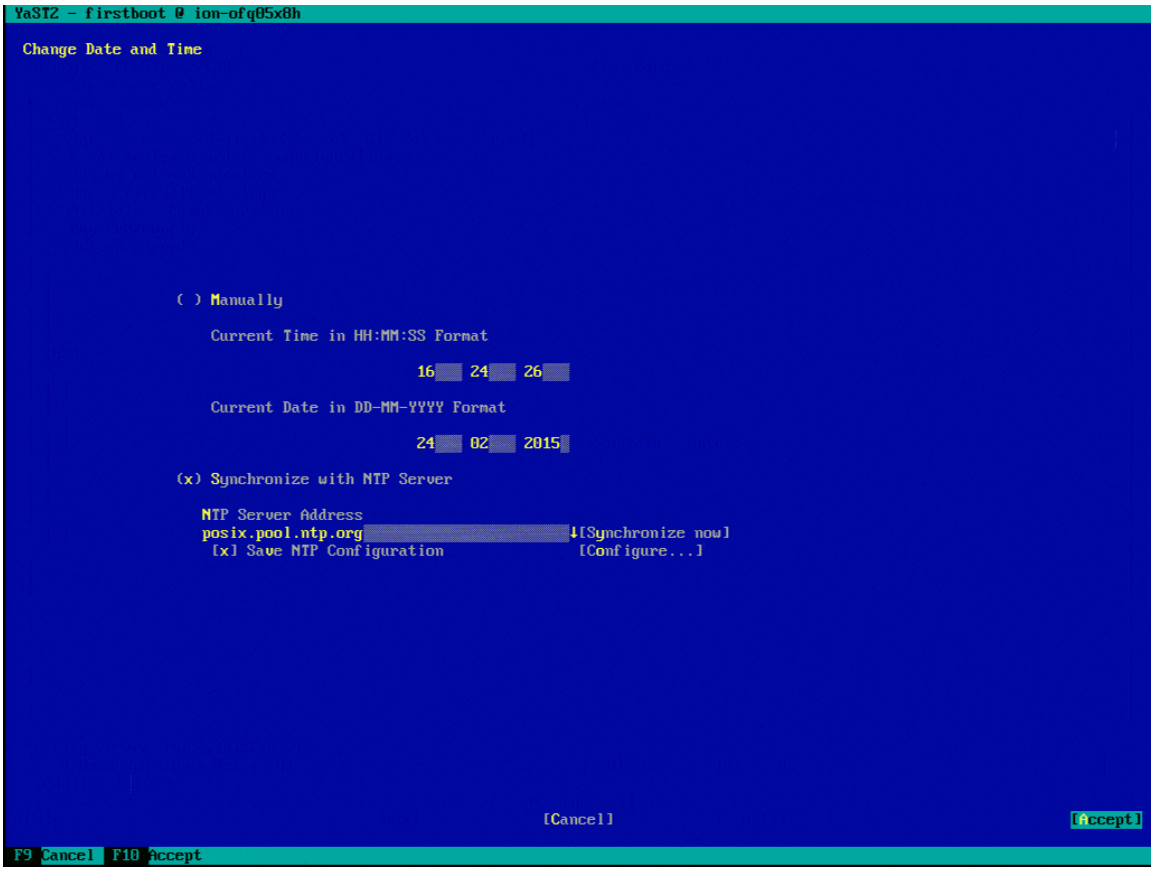
To set Time Zone and NTP:

1 Tab to the Region field and use the arrow keys to select the region where the DAAD will be located.

2 Tab to the Time Zone field and select the zone where the DAAD will be located.



3 Tab to the Date and Time field and select **Change**. The Change Date and Time screen is displayed.



CAUTION! If you are using an HA configuration you must configure DAAD to synchronize with NTP.

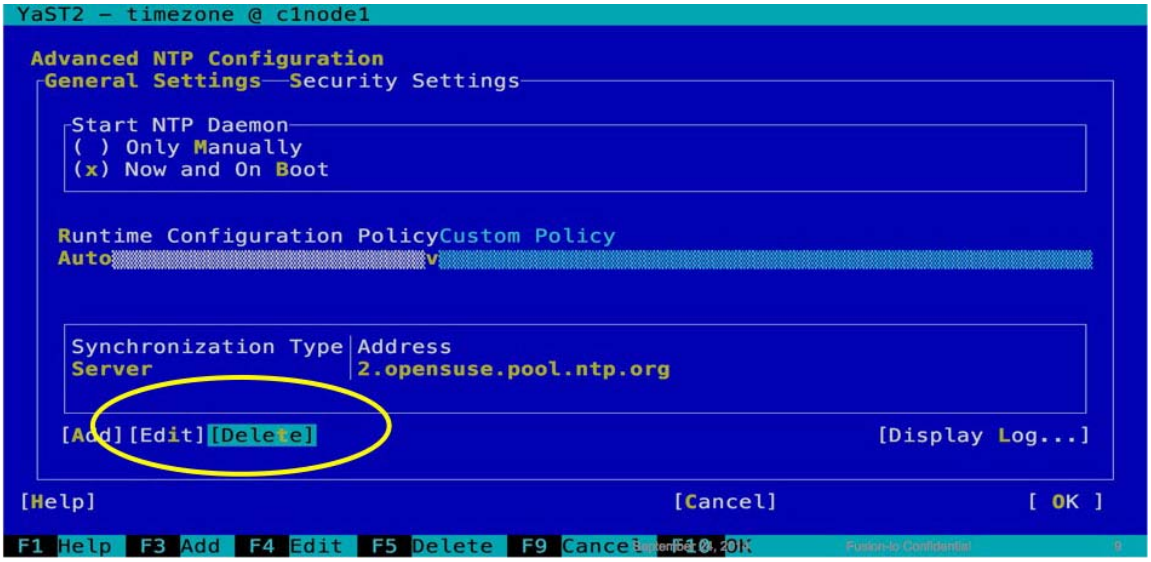
4 Tab to the Synchronize with NTP Server field and select it.

5 Tab to the NTP Server Address field and enter the address of the NTP server to synchronize with.

6 Tab to the Save NTP Configuration field and select it.

7 Tab to the Configure field and select it.

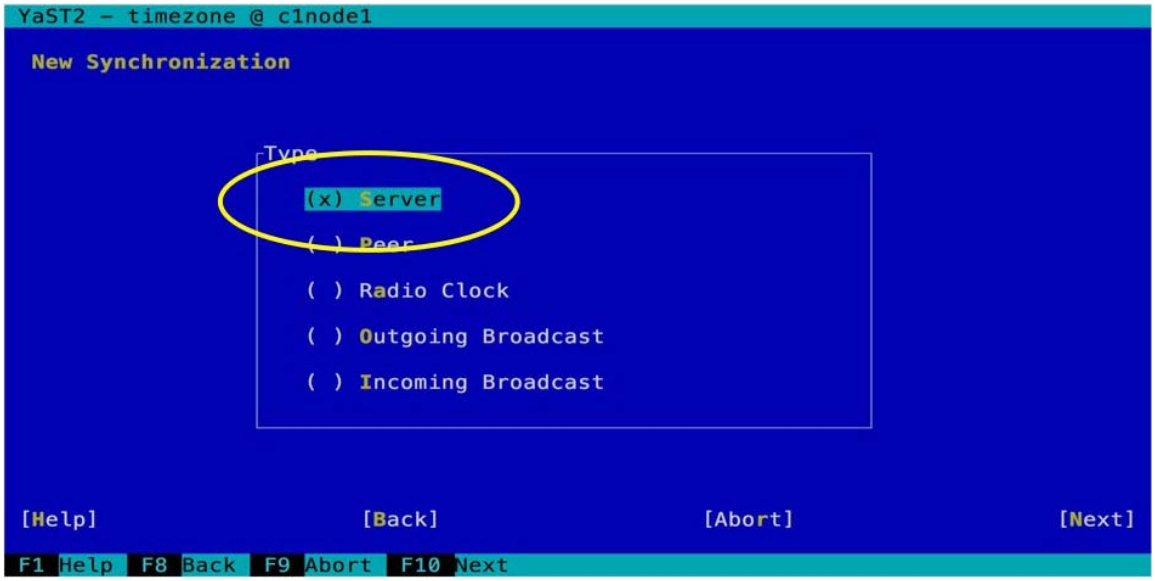
The Advanced NTP Configuration screen is displayed.



8 Select **Delete** to delete the default NTP configuration.

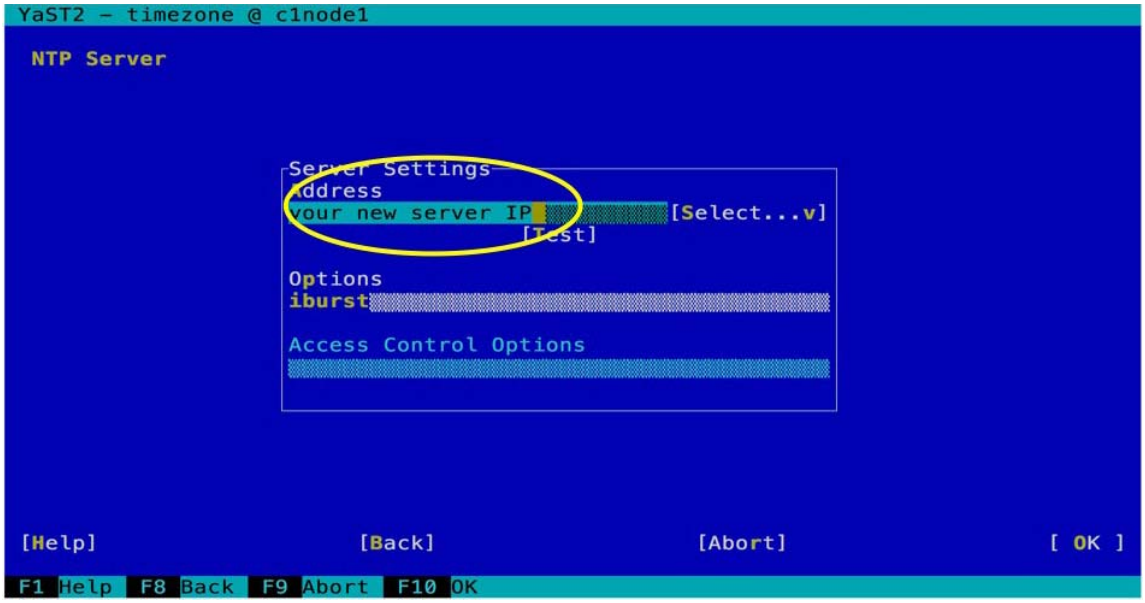
9 Select **Add** to add a new configuration.

10 In the New Synchronization screen that is displayed, select **Server**.



11 Select **Next**.

12 In the NTP Server screen that is displayed, tab to select the Address field.



13 Type the NTP server address.

14 Select **OK**.

15 If you have additional NTP servers, repeat steps 9-14 to configure them.

16 Select **Accept** to save the NTP changes and continue.

Setting the admin user password

To set the admin user password:

1 Type the Dell Acceleration Appliance for Databases password for the admin user.

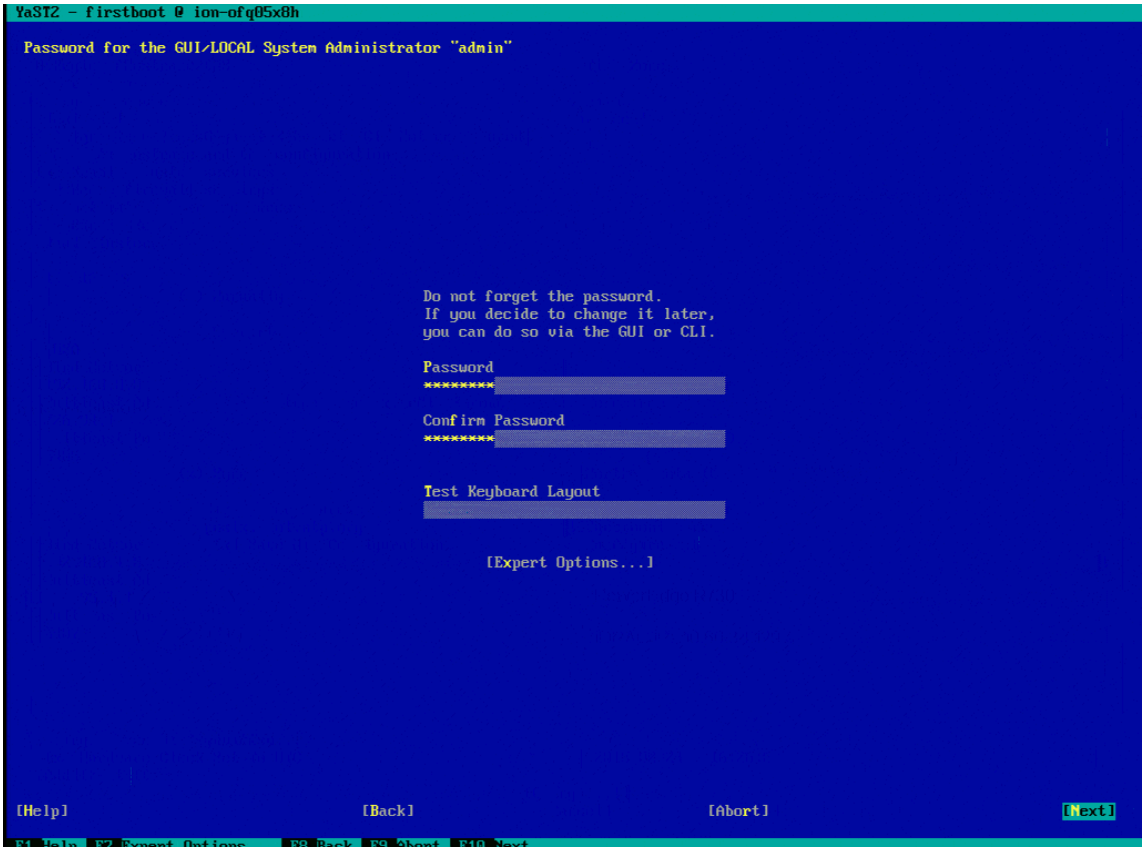
NOTE: If the password you selected is not sufficiently strong, a warning message is displayed so you can change the password, if necessary.

2 Retype the password you entered.

CAUTION! Be sure to record this password in a secure location in case it needs to be retrieved.

3 To test the Keyboard Layout or use the Expert Options, select those options on the screen.

4 Select **Next**.



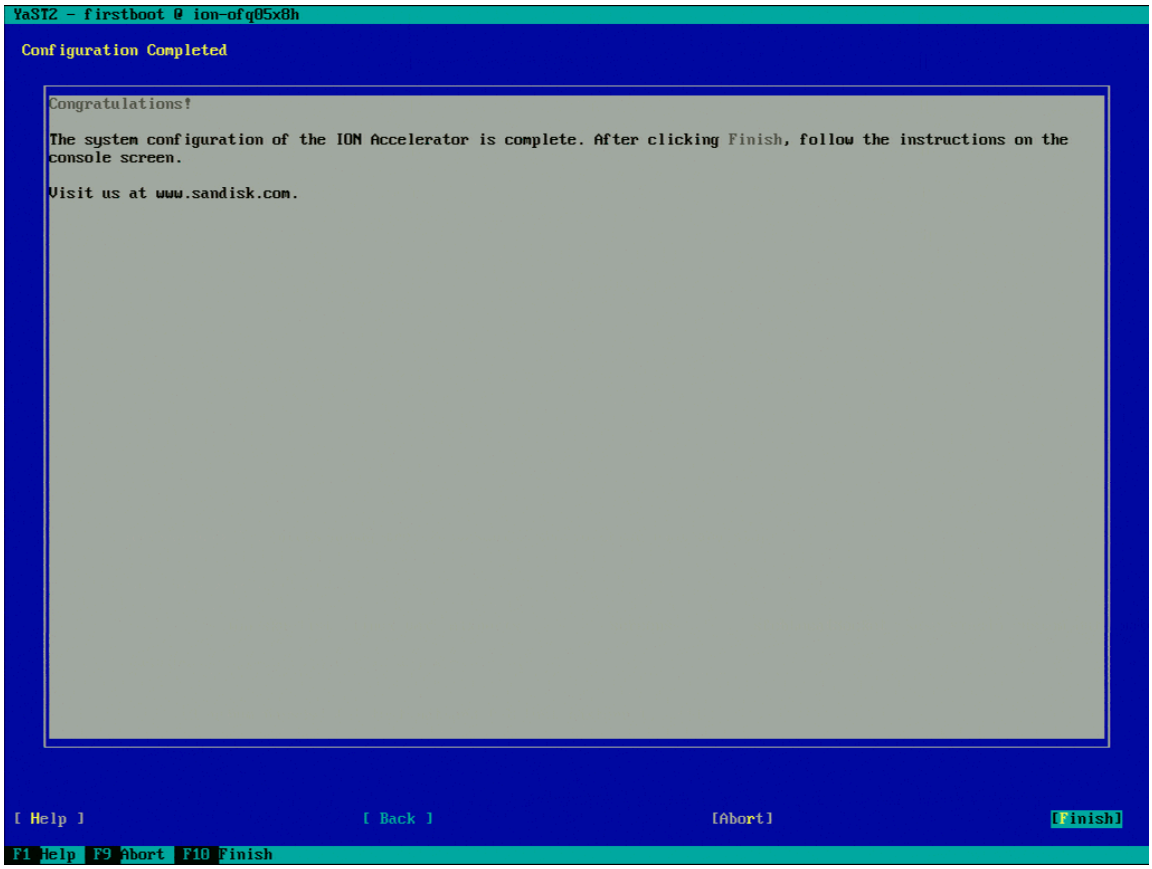
If you want to change this password later, change the admin user and GUI passwords separately. For more information, see Changing Passwords in the *Dell Acceleration Appliance for Databases GUI Guide*.

Completing the DAAD configuration

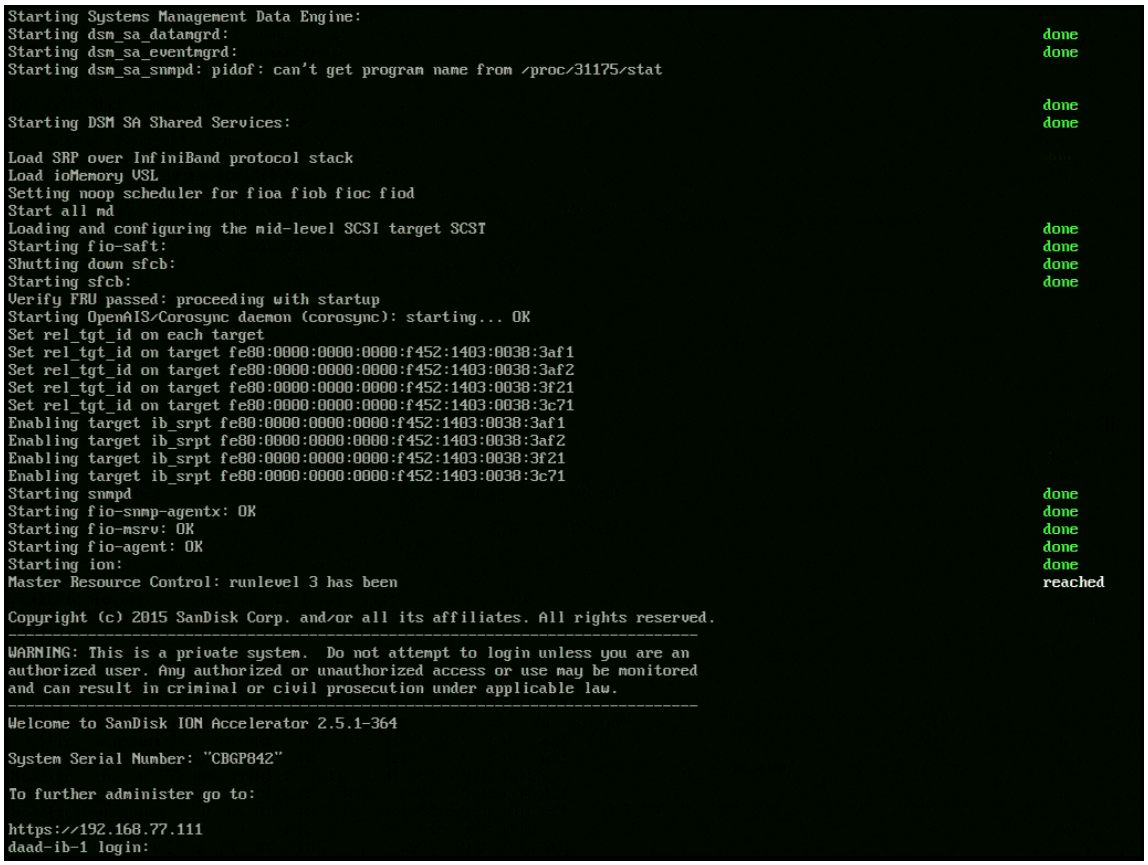
After the configuration phases are complete, the Configuration Completed screen is displayed.

To complete DAAD configuration:

1 Select **Finish**.



The DAAD system restarts and displays start-up messages. After a login prompt is displayed on the console, you can connect to DAAD for additional setup and configuration. An example screenshot is given here:



5

Configuring iSCSI for HA

To configure your iSCSI DAAD system for HA, perform the following steps:

- Configure the management port
- Configure the two ConnectX-3 ports for HA
- Configure the remaining ConnectX-3 ports for iSCSI
- Set the Hostname and Routing information for DAAD
- Set the Time Zone and NTP settings
- Enable the Cluster Setup
- Set the admin user password
- Finish the configuration

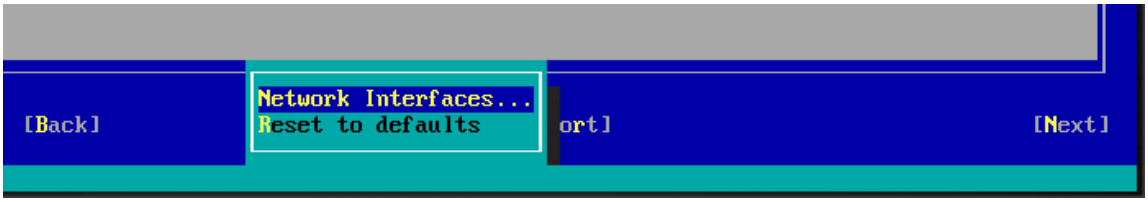
> **NOTE:** If you are planning to deploy DAAD in an HA configuration, ensure that both Infiniflash systems have the same number of Infiniflash drive cards in the same slots.

Configuring the management port

The management port is one you have connected by an external cable to your network and is externally visible. This port must be configured so your network hardware communicates with the Dell Acceleration Appliance for Databases.

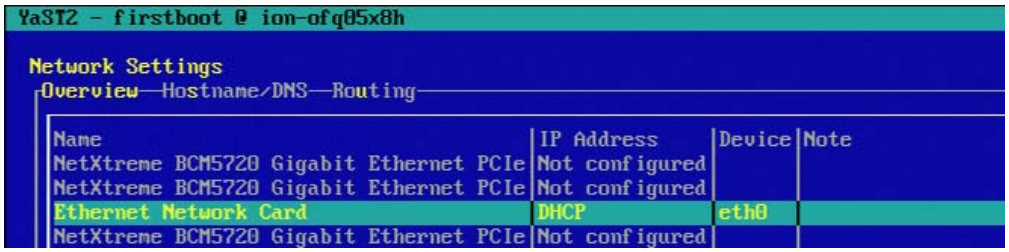
To configure the management port:

1 On the Network Configuration screen, select **Change** (bottom of the screen) and then **Network Interfaces**.



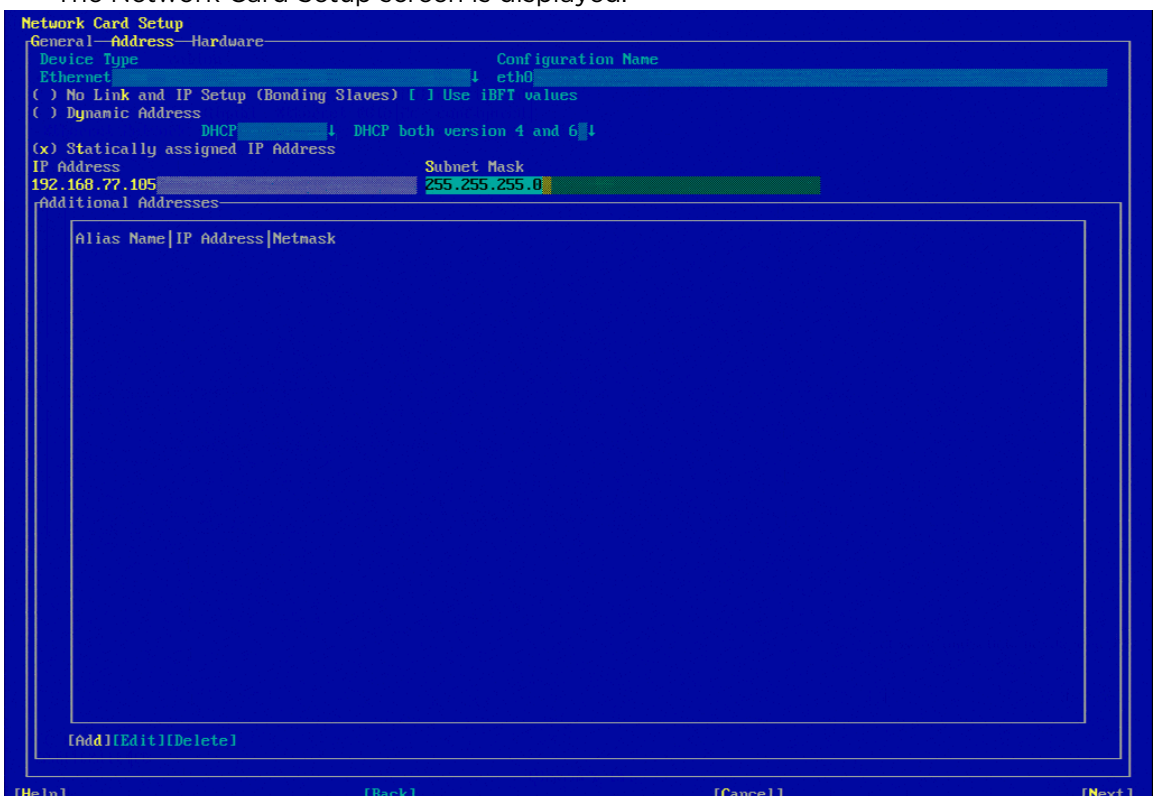
The Network Settings screen is displayed with the detected network adapters listed.

2 Select the management IP card from the list.



3 Select **Edit** (in the lower-left corner of the screen). [Add][Edit][Delete]

The Network Card Setup screen is displayed.



4 Select **Statically assigned IP Address**.

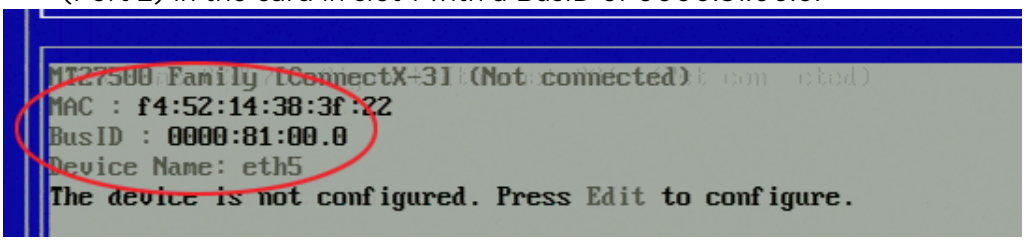
5 Enter the IP Address and Subnet Mask for your network.

6 Select **Next**.

Configuring ConnectX-3 ports for HA

To configure the ConnectX-3 ports for HA:

1 Tab to select the first ConnectX-3 Ethernet HA port. This will be the second port (Port 2) in the card in slot 1 with a BusID of 0000:81:00.0.



2 Select **Edit**.

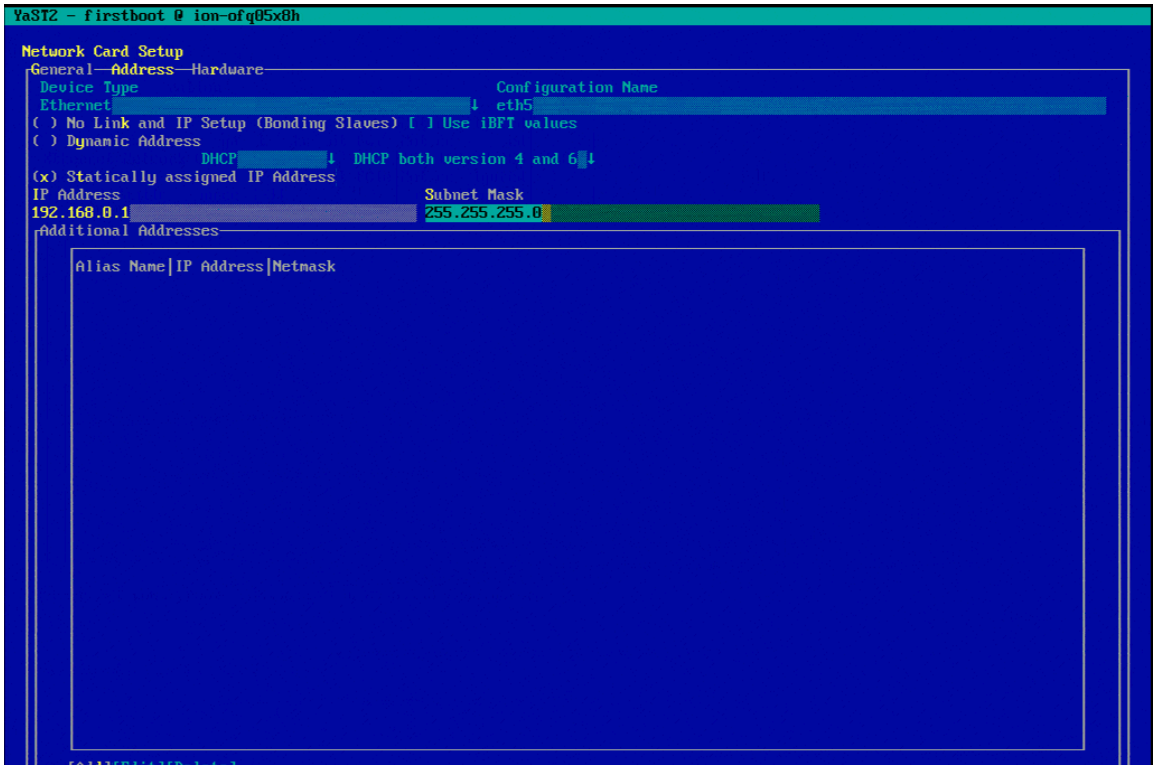
3 Tab until the IP Address field is selected**.**

4 Enter the IP Address for the port.

NOTE: Use the static IP addresses and subnet masks shown in the table here for the cluster interconnect ports. Using other IP addresses or subnet masks may lead to unexpected behavior.

	DAAD 1	DAAD 2
First HA Port	192.168.**1.1**	192.168.**1.2**
Second HA Port	192.168.**2.1**	192.168.**2.2**
Subnet Mask	255.255.255.0	255.255.255.0

5 Press Tab and enter **255.255.255.0** for the Subnet Mask.



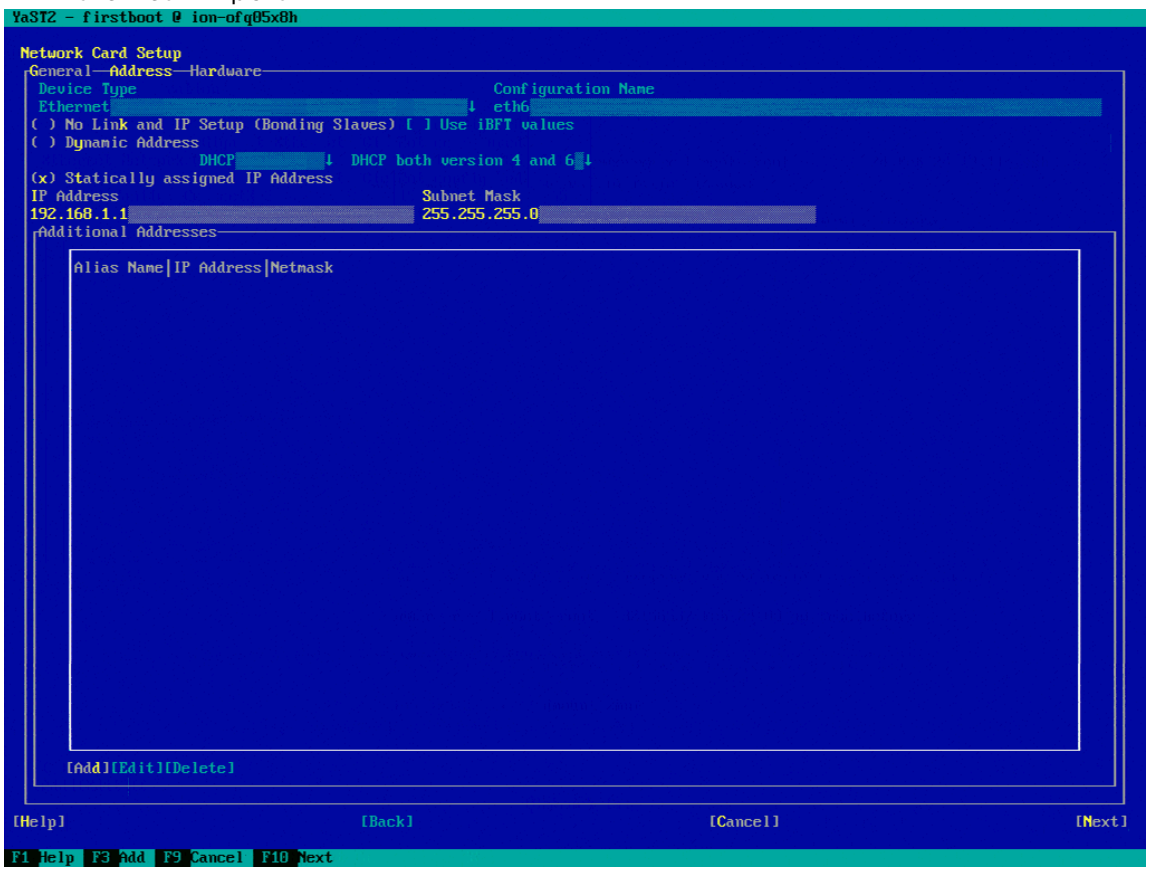
6 Select **Next**.

7 Tab to select the second ConnectX-3 Ethernet HA port.

In the case of HA Fibre Channel, this will be the first port (Port 1) in slot 1 whose MAC address is one higher than the one selected for the first Ethernet HA port.

In the case of HA iSCSI and HA InfiniBand/SRP this will be the second Ethernet port (Port 2) in the card in slot 3 with a BusID of 0000:83:00.0

8 Repeat steps 2-5, setting a different IP Address for the second ConnectX-3 Ethernet HA port.



9 Select **Next**.

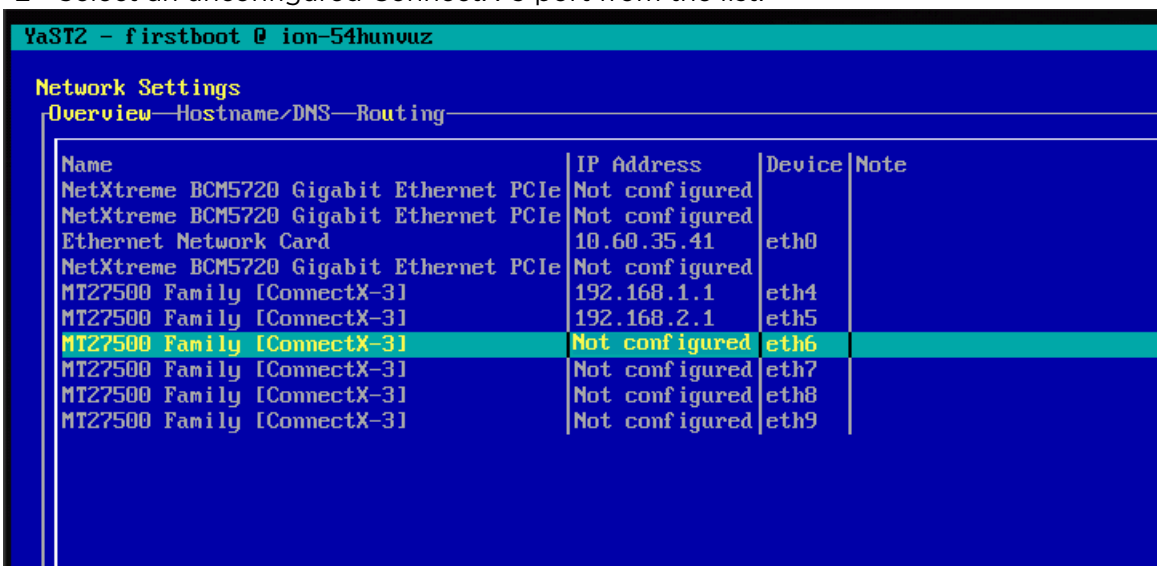
Configuring ConnectX-3 ports for iSCSI

To configure ConnectX-3 ports for iSCSI:

1 On the Network Configuration screen, select **Change** (bottom of the screen) and then **Network Interfaces**.

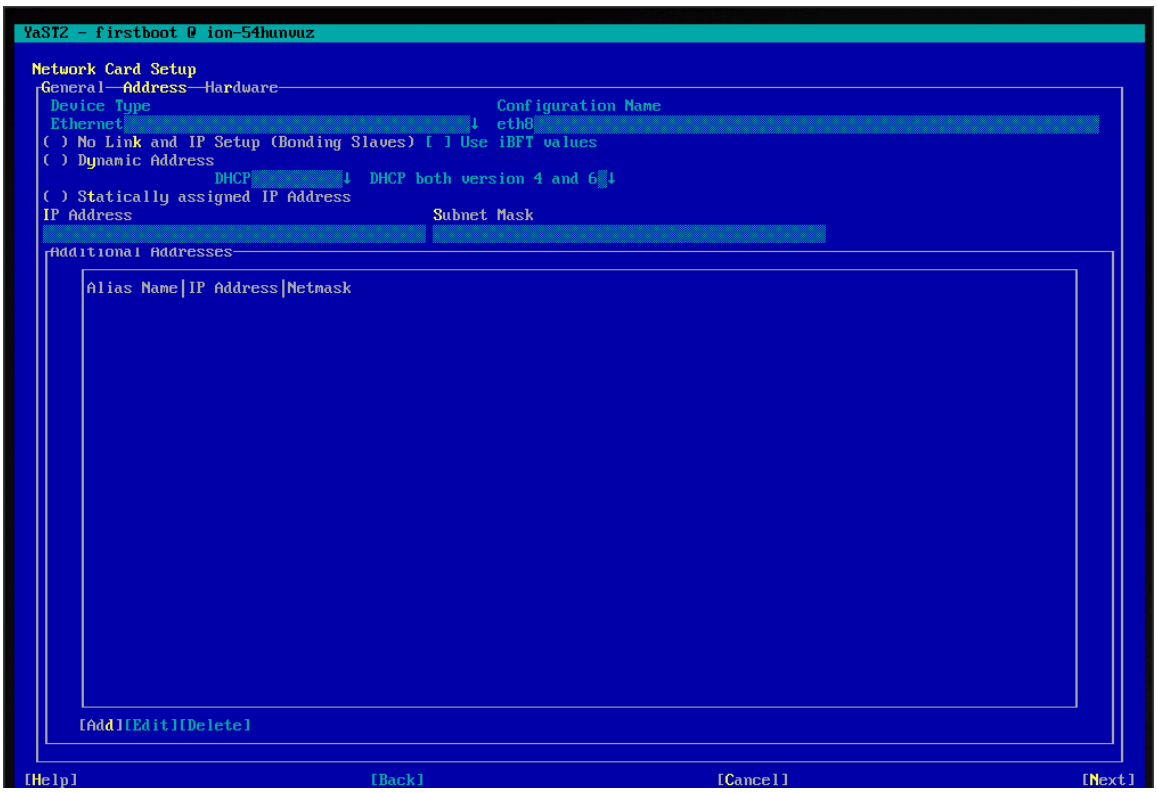
The Network Settings screen is displayed with the detected network adapters listed.

2 Select an unconfigured ConnectX-3 port from the list.



3 Select **Edit** (in the lower-left corner of the screen). [Add][Edit][Delete]

The Network Card Setup screen is displayed.



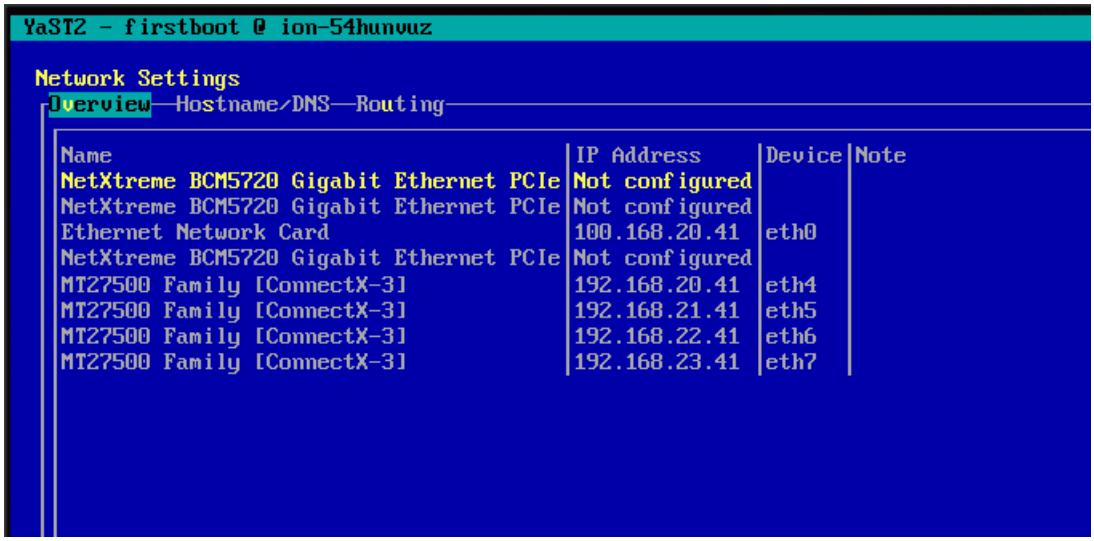
4 Select **Statically assigned IP Address**.

Each iSCSI port must be configured on its own subnet. Also, it is a best practice to have the last octet of the IP address match the last octet of the host.

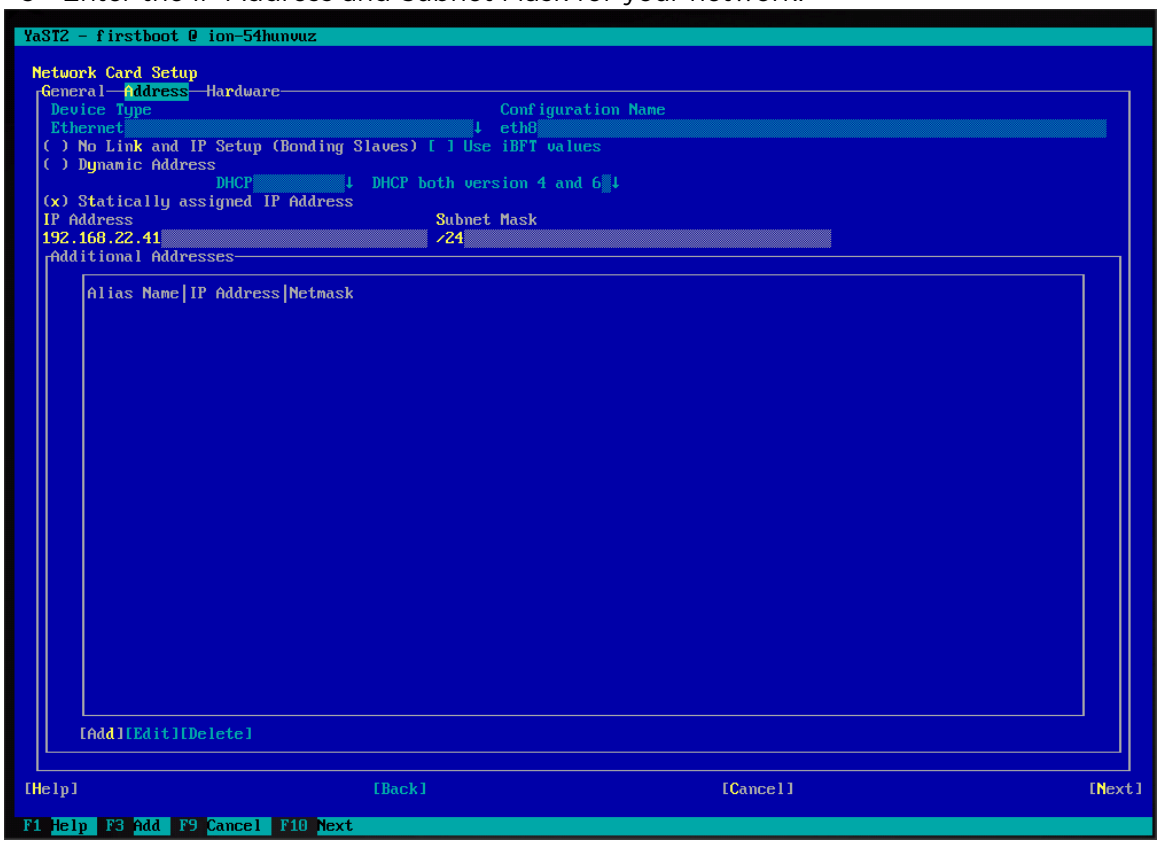
For example, with a host IP address of 100.128.10.41, you might configure the iSCSI ports for the following subnets:

- 192.168.**20**.41
- 192.168.**21**.41
- 192.168.**22**.41
- 192.168.**23**.41

Here is an example of unique subnet numbering for iSCSI ports in the user interface:



5 Enter the IP Address and Subnet Mask for your network.



6 Select **Next**.

7 On the Network Card Setup screen, press left-arrows to select **General**.

8 Tab to the Set MTU field.

9 Set MTU to **9000**.

> **NOTE:** Ensure that the switches and routers on your network are configured to correctly handle this MTU size.

10 Select **Next**.

11 Repeat steps 2–10 until all four iSCSI ports are configured.

Setting hostname and routing information

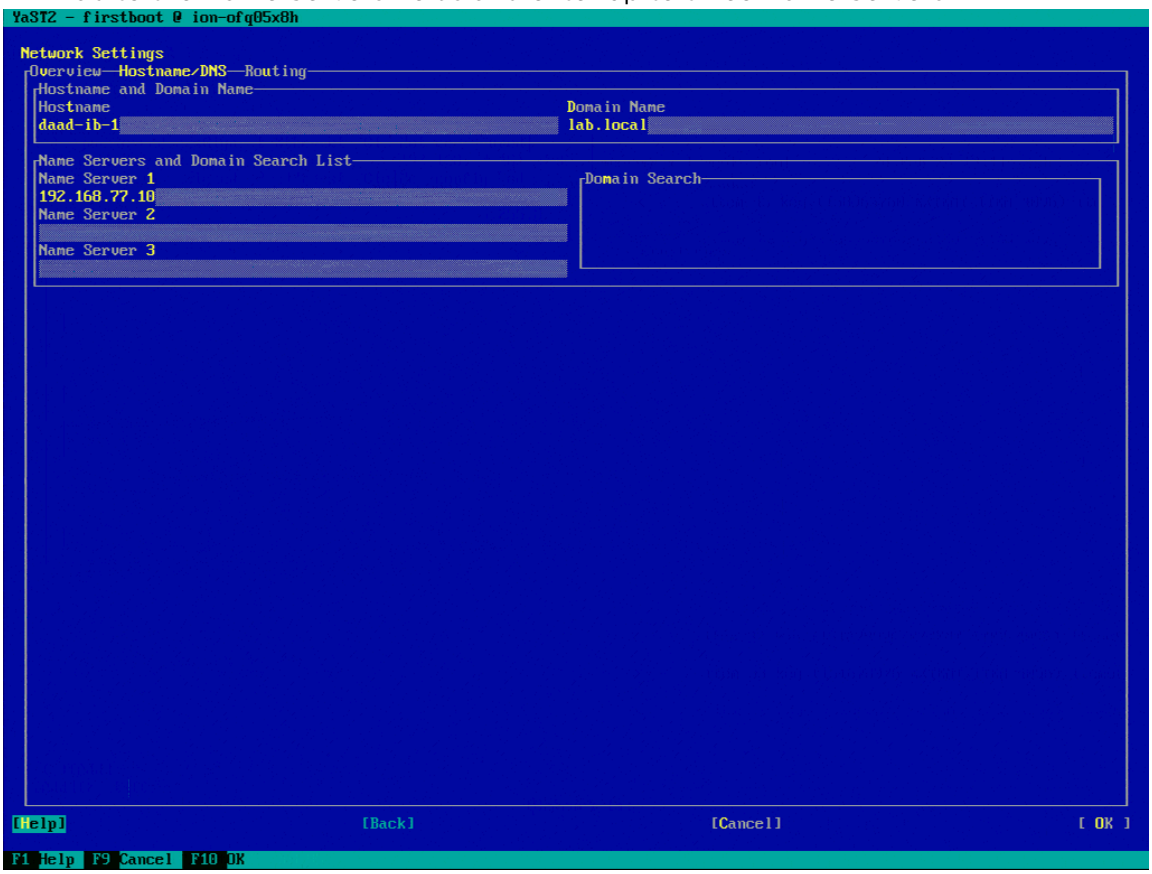
To configure Hostname and routing information:

1 On the Network Settings screen, press the right arrow key to select **Hostname**.

2 Tab to the Hostname field and enter the Hostname for the DAAD.

3 Tab to the Domain Name field and enter the Domain Name.

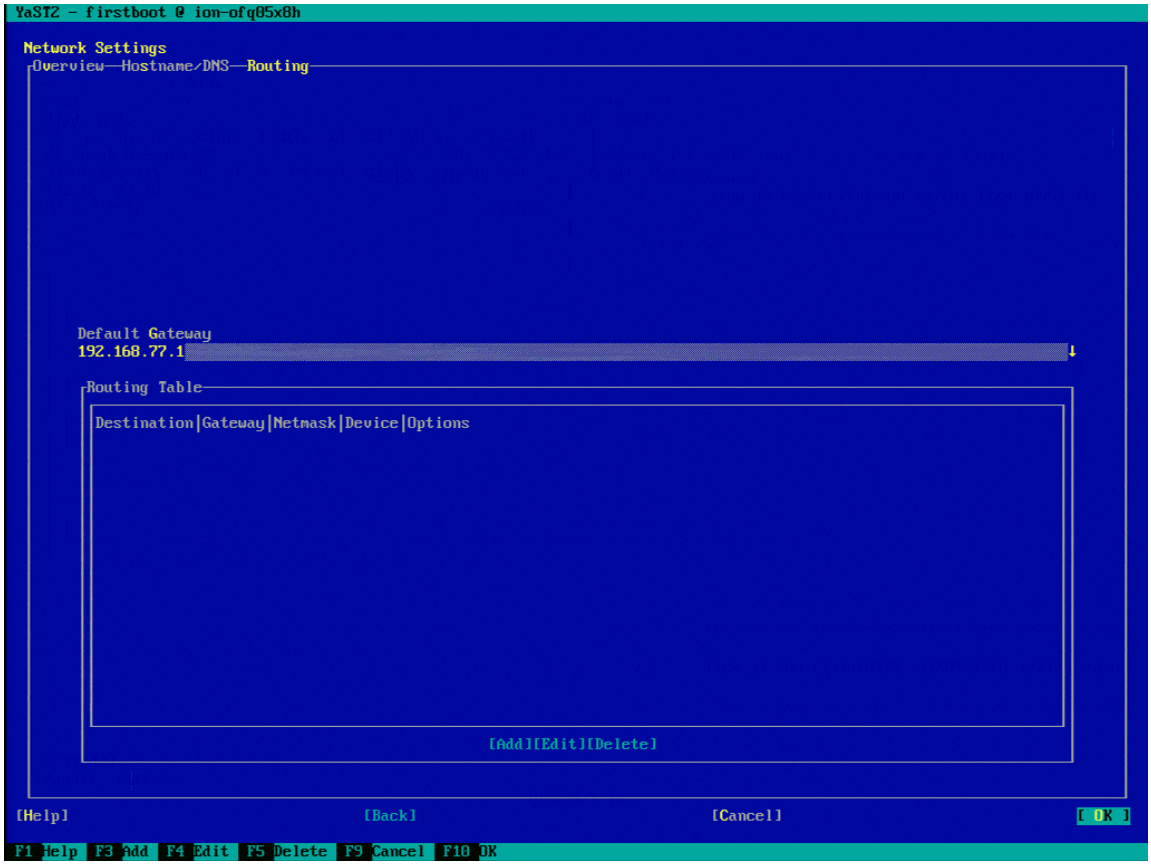
4 Tab to the Name Servers fields and enter up to three Name Servers.



5 Press Tab until the focus is back on the Overview line.

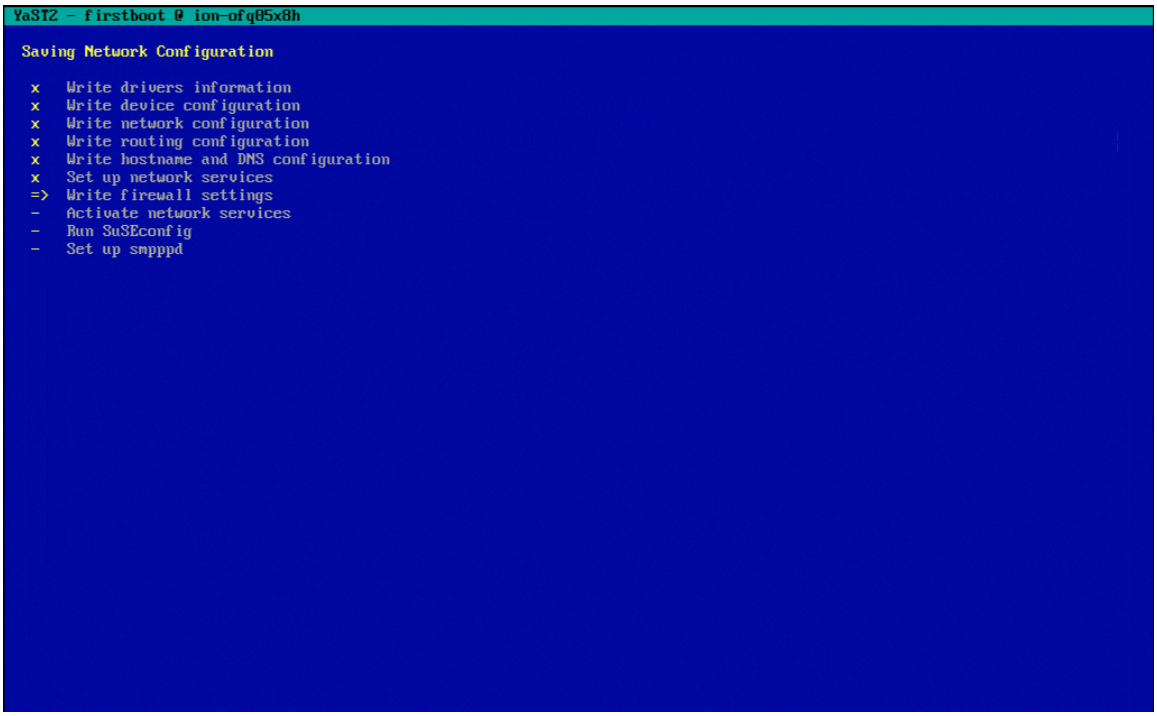
6 Press the right arrow key to select Routing.

7 Tab to the Default Gateway field and enter the Default Gateway.



8 Select **OK**.

The Saving Network Configuration Screen is displayed.

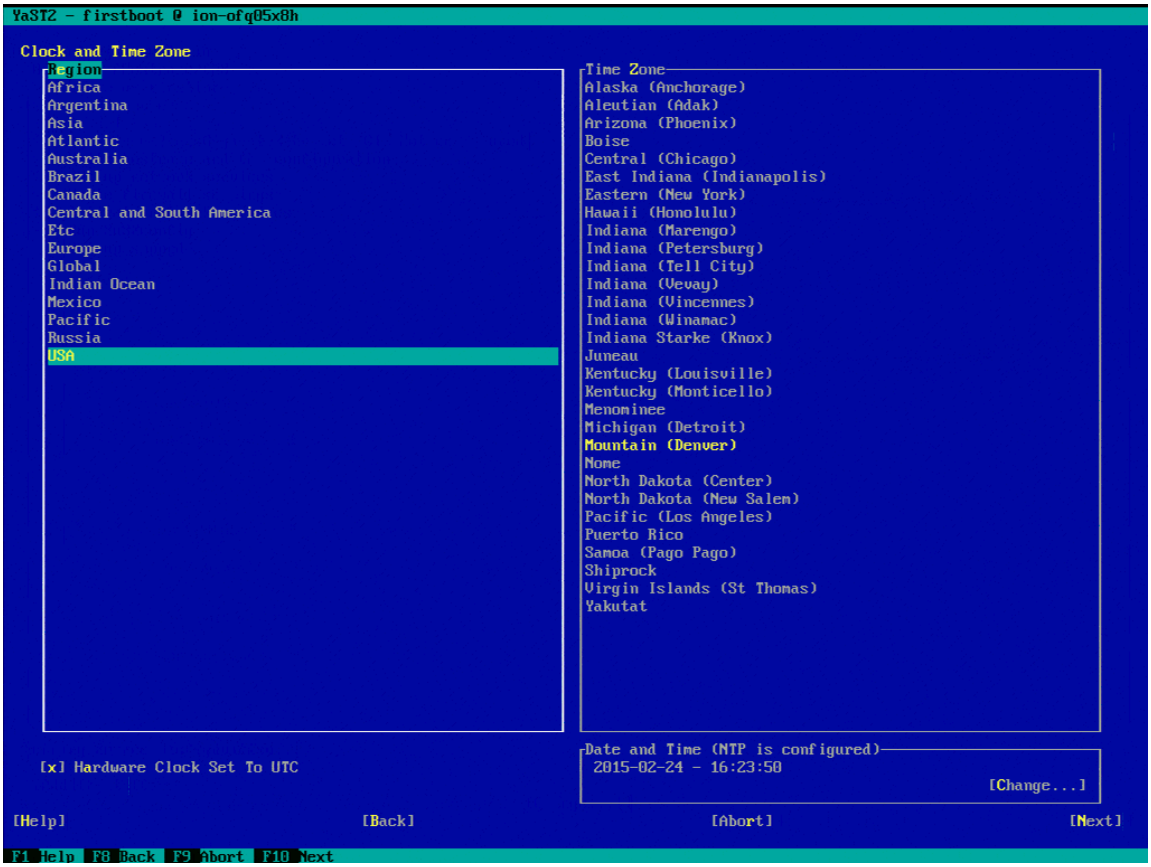


Setting time zone and NTP settings

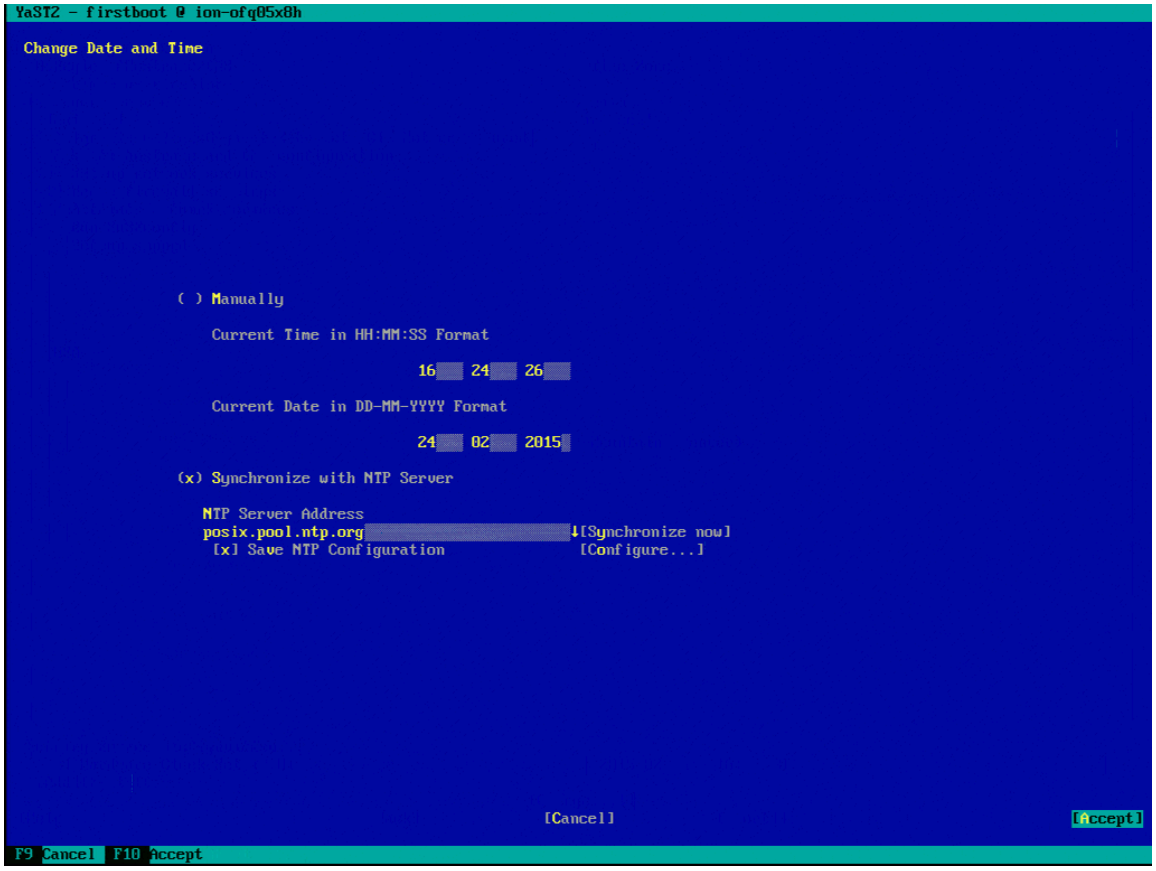
To set Time Zone and NTP:

1 Tab to the Region field and use the arrow keys to select the region where the DAAD will be located.

2 Tab to the Time Zone field and select the zone where the DAAD will be located.



3 Tab to the Date and Time field and select **Change**. The Change Date and Time screen is displayed.



CAUTION! If you are using an HA configuration you must configure DAAD to synchronize with NTP.

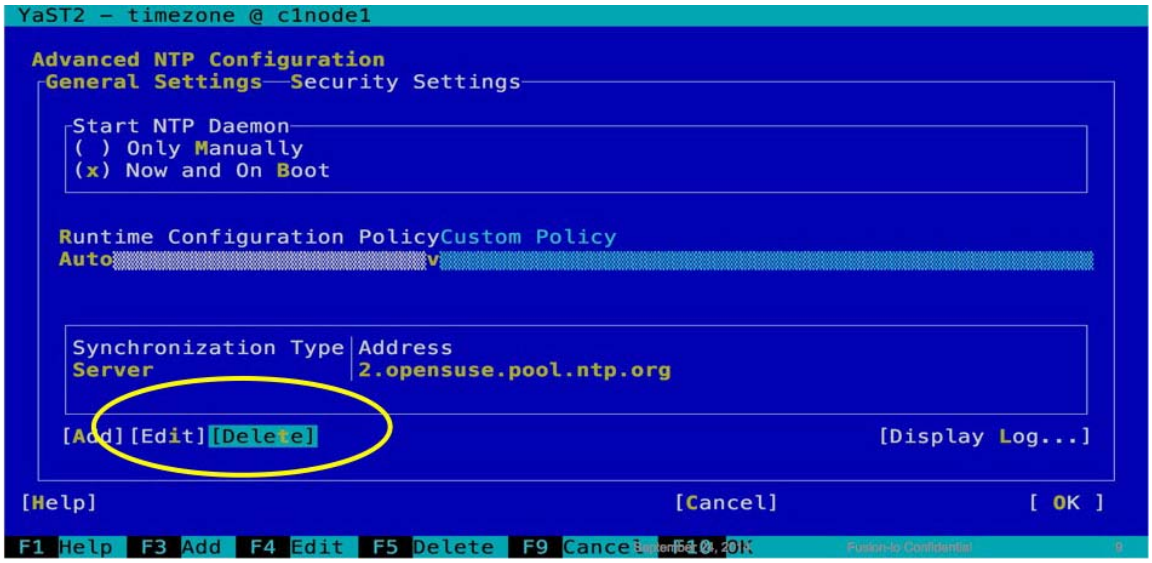
4 Tab to the Synchronize with NTP Server field and select it.

5 Tab to the NTP Server Address field and enter the address of the NTP server to synchronize with.

6 Tab to the Save NTP Configuration field and select it.

7 Tab to the Configure field and select it.

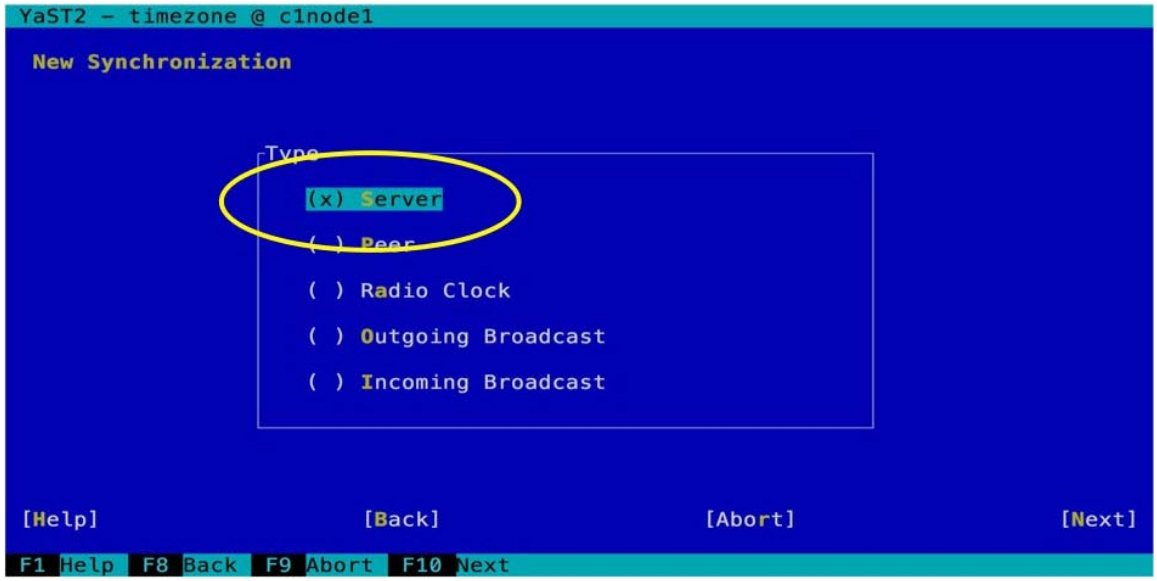
The Advanced NTP Configuration screen is displayed.



8 Select **Delete** to delete the default NTP configuration.

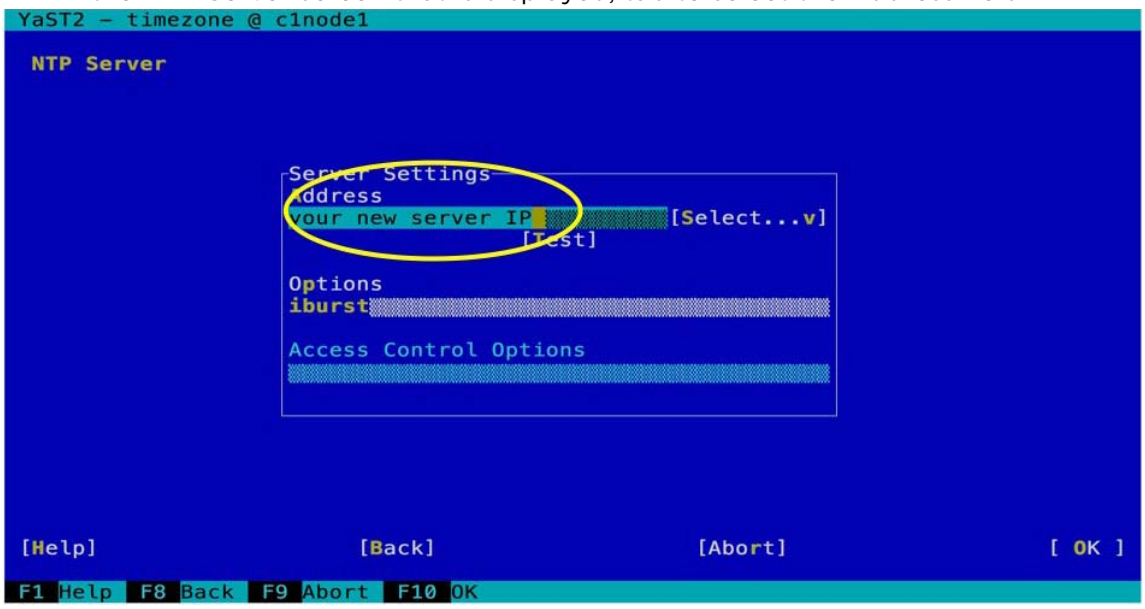
9 Select **Add** to add a new configuration.

10 In the New Synchronization screen that is displayed, select **Server**.



11 Select **Next**.

12 In the NTP Server screen that is displayed, tab to select the Address field.



13 Type the NTP server address.

14 Select **OK**.

15 If you have additional NTP servers, repeat steps 9-14 to configure them.

16 Select **Accept** to save the NTP changes and continue.

Enabling cluster setup

To enable the HA cluster:

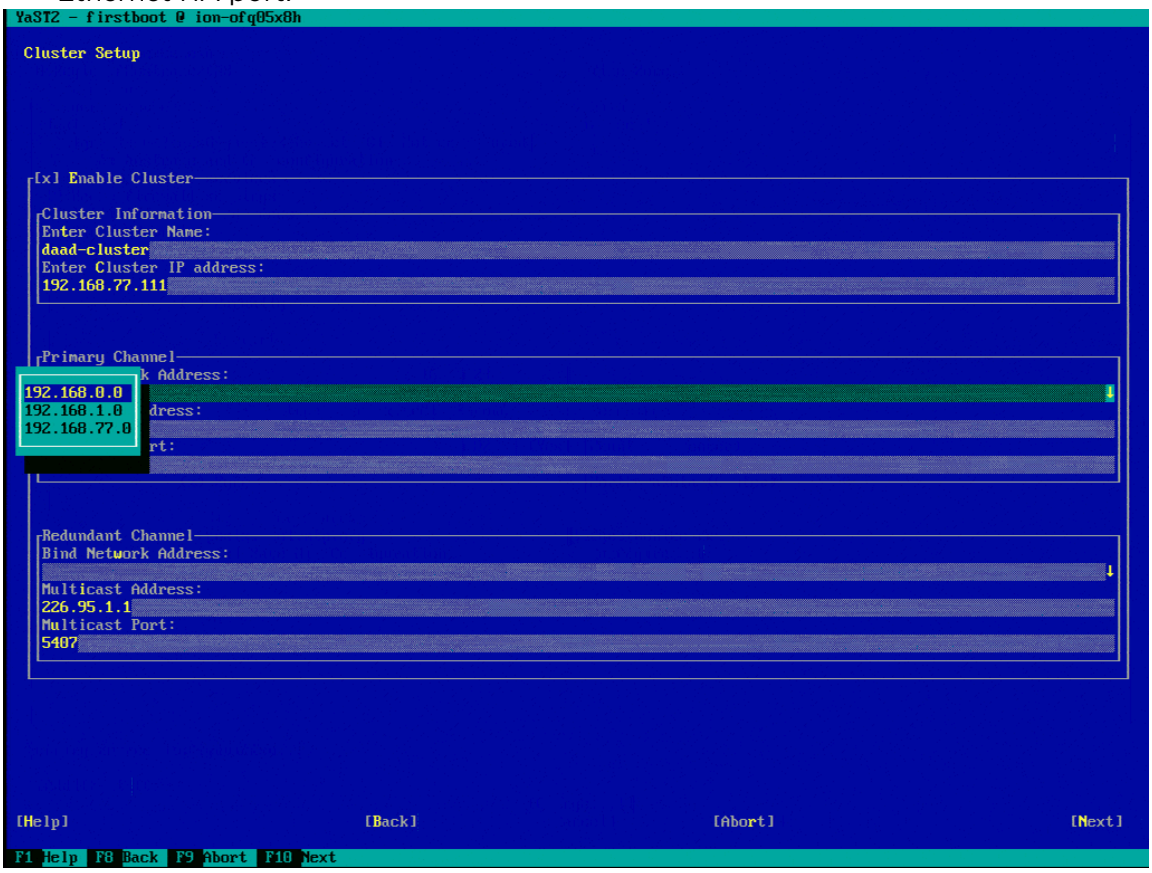
1 On the Cluster Setup screen, tab to the Enter Cluster Name field and enter a name for the Cluster.

> **CAUTION!** The cluster name must be a DNS entry, not DHCP, and must be on the same subnet as the management IP address for the Dell Acceleration Appliance for Databases nodes. Also, do not enter the name or IP address of one of the currently existing nodes.

2 Tab to the Enter Cluster IP Address field and enter the IP address for the cluster.

3 Tab to the Primary Channel field and press the down arrow key on Bind Network Address drop-down menu.

4 From the drop-down menu, select the address you set for the first ConnectX-3 Ethernet HA port.



5 Tab to the Redundant Channel field and press the down arrow key on Bind Network Address drop-down menu.

6 From the drop-down menu, select the address you set for the second ConnectX-3 Ethernet HA port.

7 Select **Next**.

Setting the admin user password

To set the admin user password:

1 Type the Dell Acceleration Appliance for Databases password for the admin user.

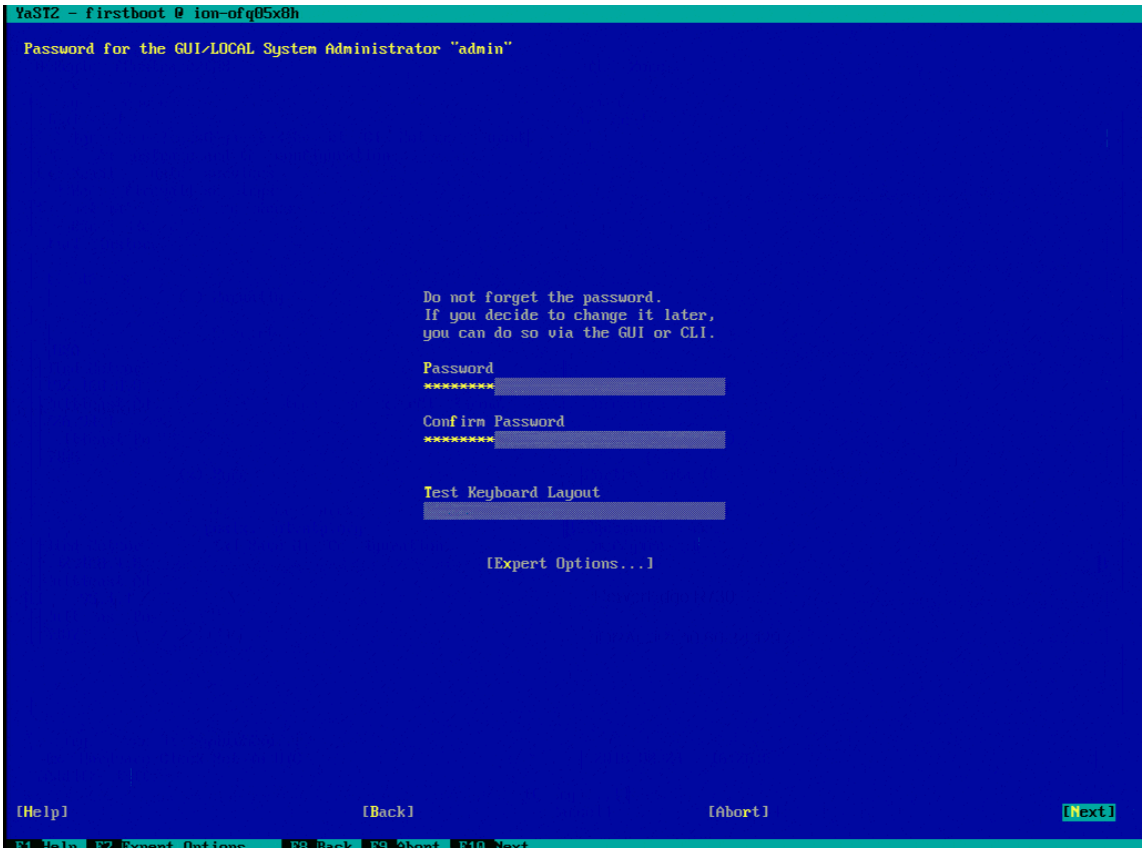
NOTE: If the password you selected is not sufficiently strong, a warning message is displayed so you can change the password, if necessary.

2 Retype the password you entered.

CAUTION! Be sure to record this password in a secure location in case it needs to be retrieved.

3 To test the Keyboard Layout or use the Expert Options, select those options on the screen.

4 Select **Next**.



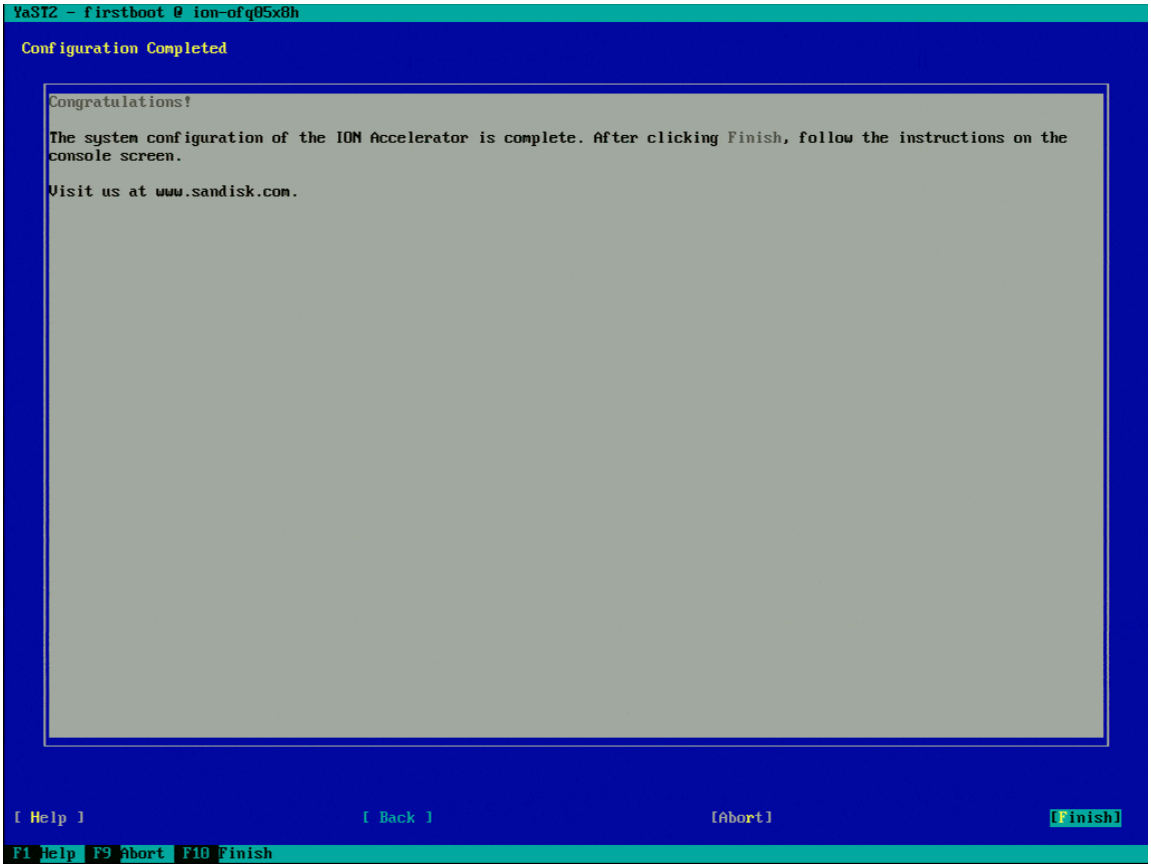
If you want to change this password later, change the admin user and GUI passwords separately. For more information, see Changing Passwords in the *Dell Acceleration Appliance for Databases GUI Guide*.

Completing the DAAD configuration

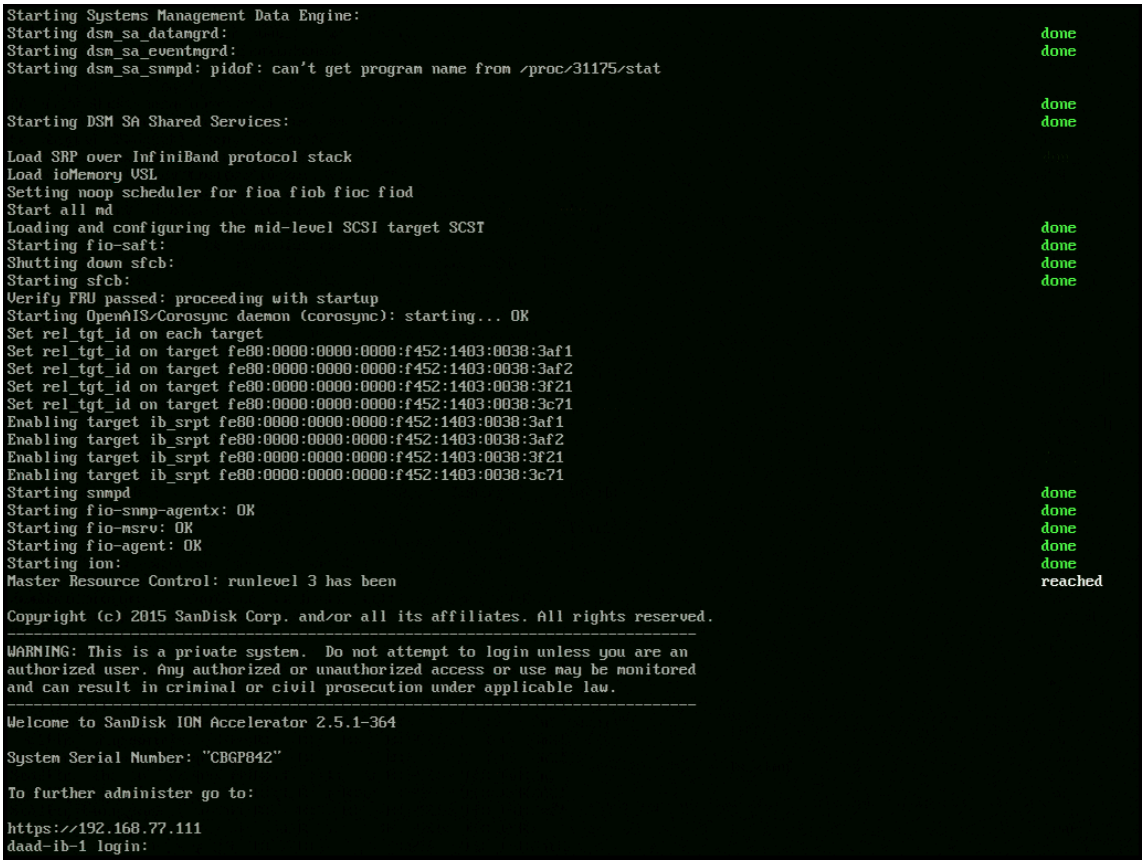
After the configuration phases are complete, the Configuration Completed screen is displayed.

To complete DAAD configuration:

1 Select **Finish**.



The DAAD system restarts and displays start-up messages. After a login prompt is displayed on the console, you can connect to DAAD for additional setup and configuration. An example screenshot is given here:



6

Configuring FC and IB for HA

The steps for configuring a fibre channel DAAD system for HA and an InfiniBand/SRP DAAD system for HA are very similar. They only difference is that with fibre channel the Ethernet HA ports are both on the Mellanox adpater in Slot 1, whereas with InfiniBand/SRP the one HA port is on the adapter in Slot 1 and the other port is on the adapter in Slot 3.

To configure either your fibre channel or your InfiniBand/SRP DAAD system for HA, perform the following steps:

- Configure the management port
- Configure the ports on the Mellanox card for HA
- Set the Hostname and Routing information for DAAD
- Set the Time Zone and NTP settings
- Enable the Cluster Setup
- Set the admin user password
- Finish the configuration

NOTE: If you are planning to deploy DAAD in an HA configuration, ensure that both Infiniflash systems have the same number of Infiniflash drive cards in the same slots.

Configuring the management port

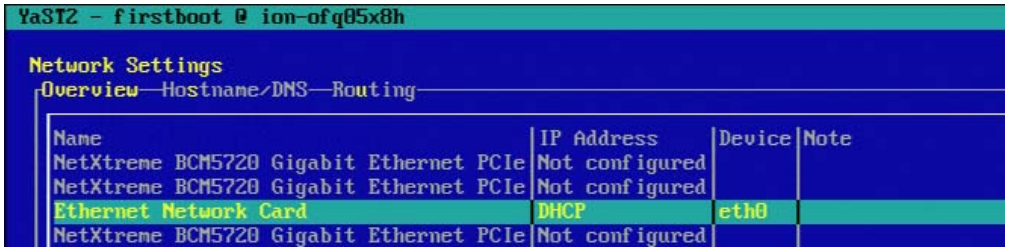
The management port is one you have connected by an external cable to your network and is externally visible. This port must be configured so your network hardware communicates with the Dell Acceleration Appliance for Databases.

To configure the management port:

1 On the Network Configuration screen, select **Change** (bottom of the screen) and then **Network Interfaces**.

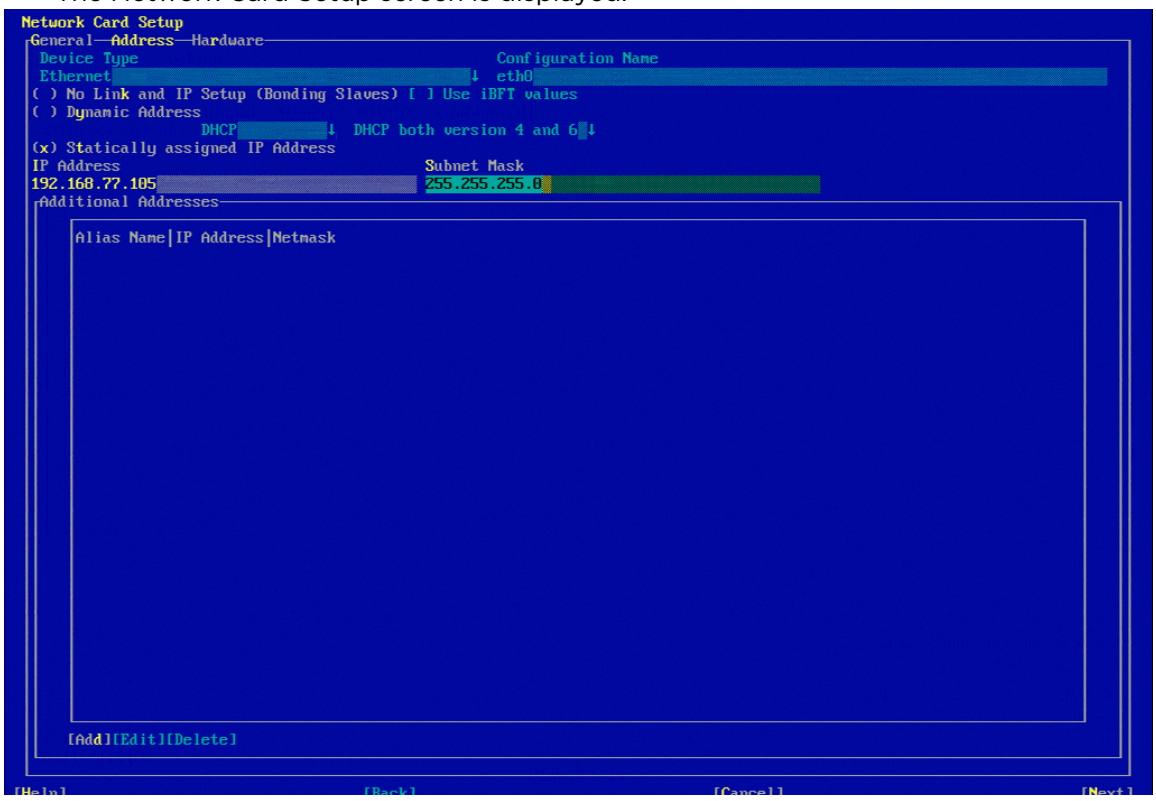
The Network Settings screen is displayed with the detected network adapters listed.

2 Select the management IP card from the list.



3 Select **Edit** (in the lower-left corner of the screen). [Add][Edit][Delete]

The Network Card Setup screen is displayed.



4 Select **Statically assigned IP Address**.

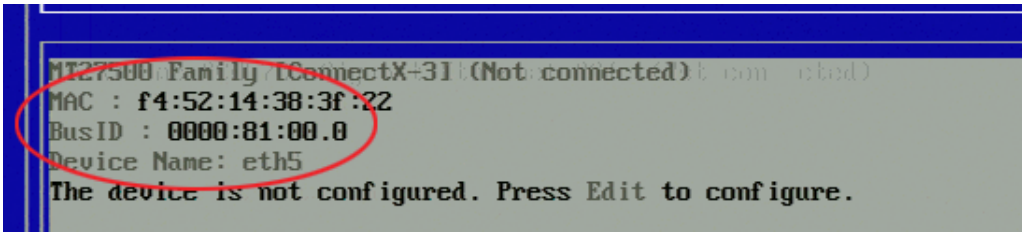
5 Enter the IP Address and Subnet Mask for your network.

6 Select **Next**.

Configuring ConnectX-3 ports for HA

To configure the ConnectX-3 ports for HA:

1 Tab to select the first ConnectX-3 Ethernet HA port. This will be the second port (Port 2) in the card in slot 1 with a BusID of 0000:81:00.0.



2 Select **Edit**.

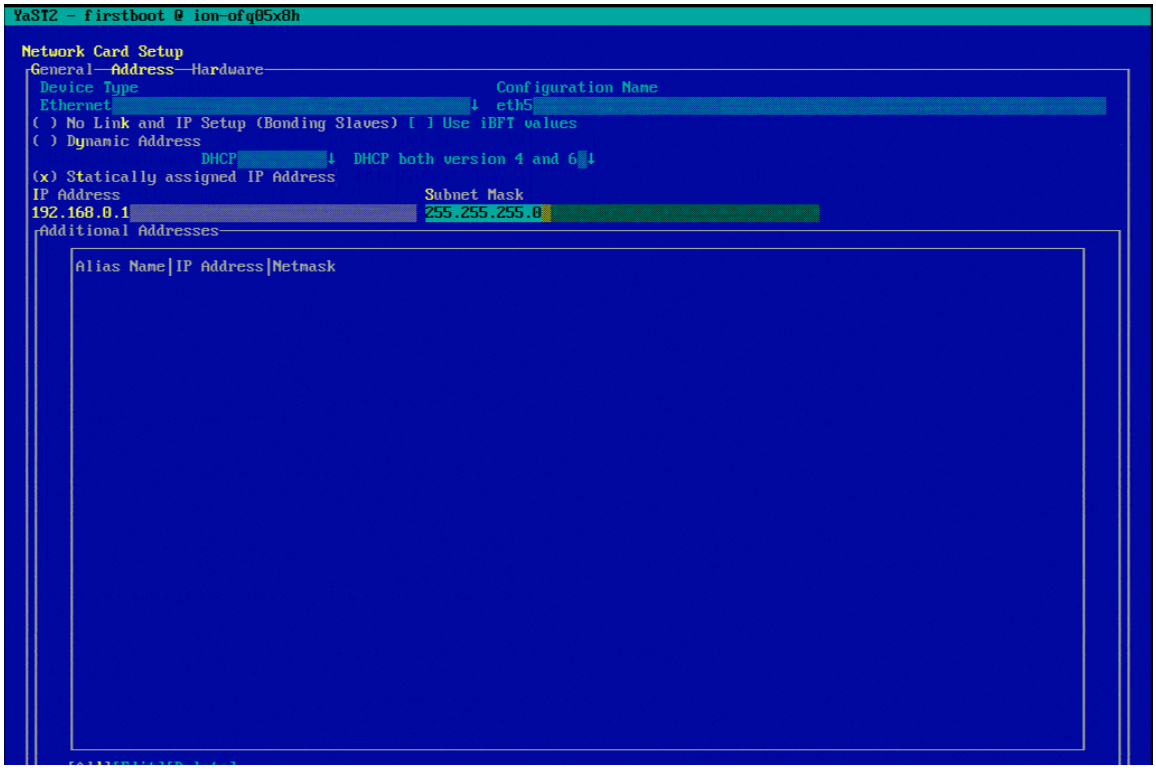
3 Tab until the IP Address field is selected**.**

4 Enter the IP Address for the port.

NOTE: Use the static IP addresses and subnet masks shown in the table here for the cluster interconnect ports. Using other IP addresses or subnet masks may lead to unexpected behavior.

	DAAD 1	DAAD 2
First HA Port	192.168.**1.1**	192.168.**1.2**
Second HA Port	192.168.**2.1**	192.168.**2.2**
Subnet Mask	255.255.255.0	255.255.255.0

5 Press Tab and enter **255.255.255.0** for the Subnet Mask.



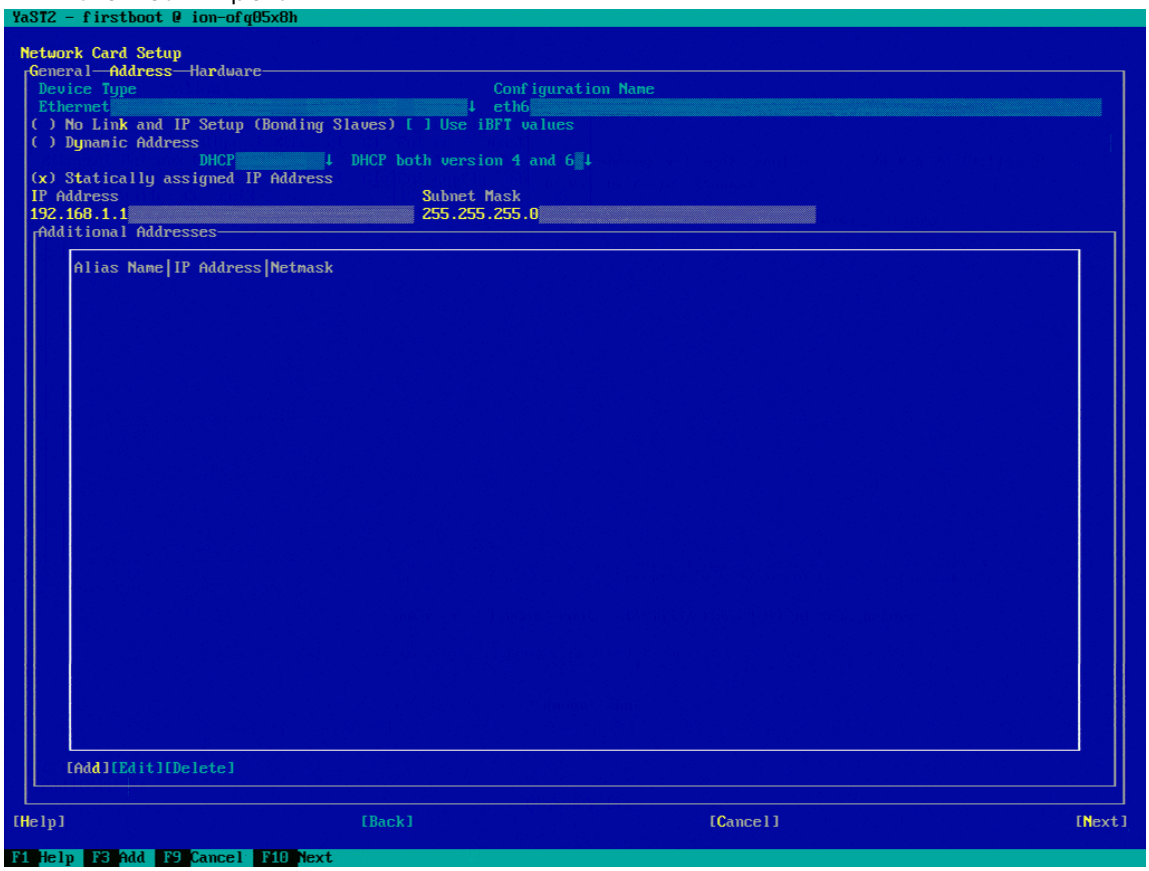
6 Select **Next**.

7 Tab to select the second ConnectX-3 Ethernet HA port.

In the case of HA Fibre Channel, this will be the first port (Port 1) in slot 1 whose MAC address is one higher than the one selected for the first Ethernet HA port.

In the case of HA iSCSI and HA InfiniBand/SRP this will be the second Ethernet port (Port 2) in the card in slot 3 with a BusID of 0000:83:00.0

8 Repeat steps 2-5, setting a different IP Address for the second ConnectX-3 Ethernet HA port.



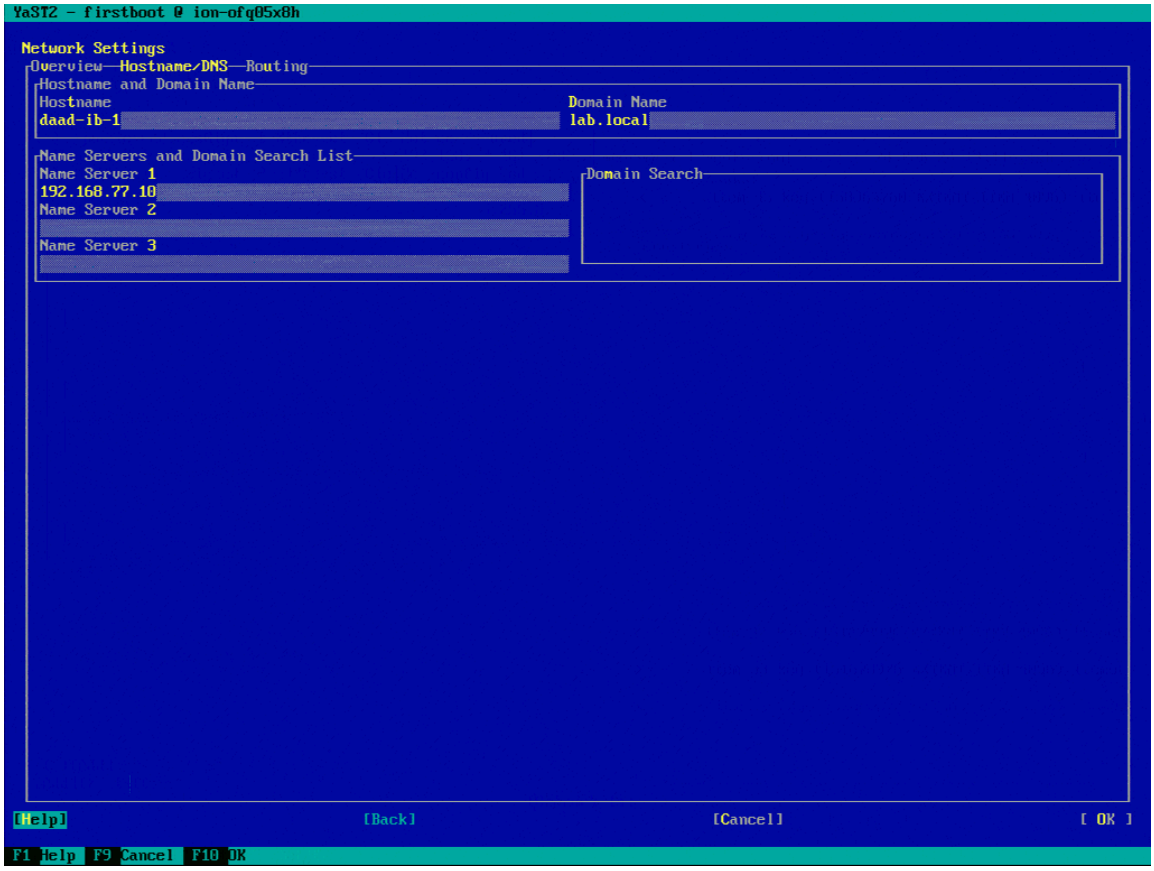
9 Select **Next**.

Setting hostname and routing information

To configure Hostname and routing information:

1 On the Network Settings screen, press the right arrow key to select **Hostname**.
2 Tab to the Hostname field and enter the Hostname for the DAAD.
3 Tab to the Domain Name field and enter the Domain Name.

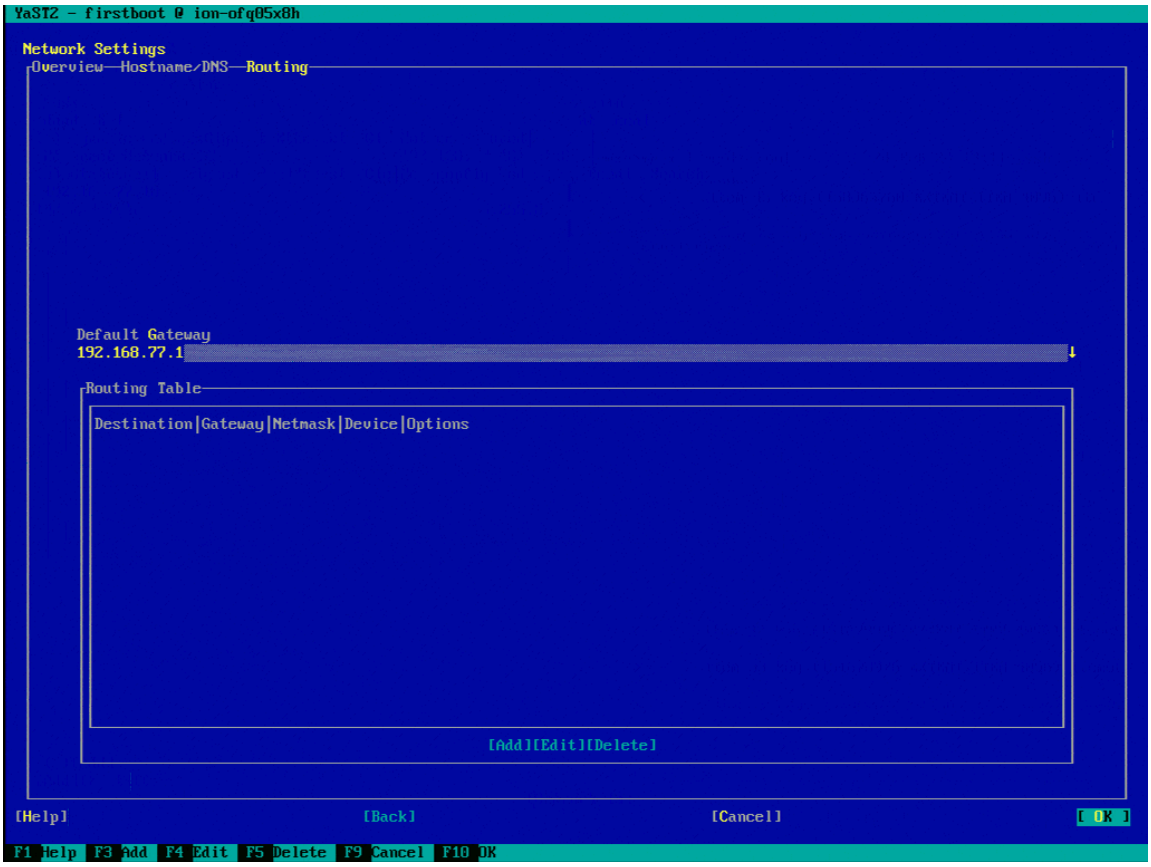
4 Tab to the Name Servers fields and enter up to three Name Servers.



5 Press Tab until the focus is back on the Overview line.

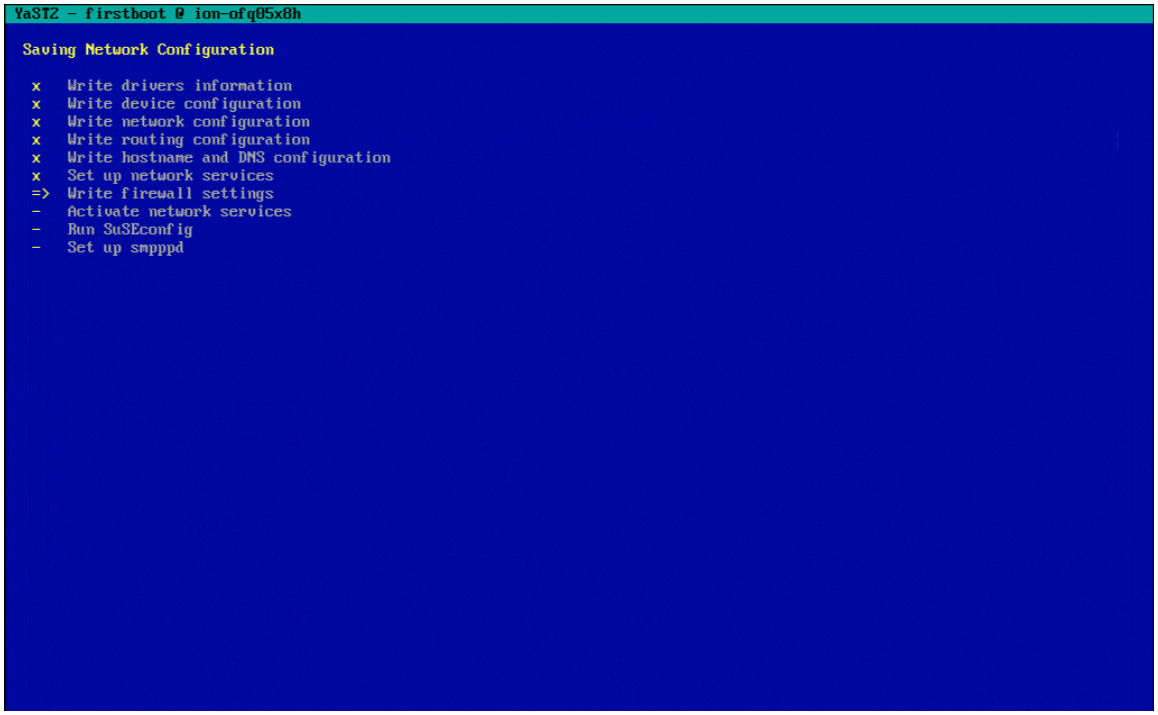
6 Press the right arrow key to select Routing.

7 Tab to the Default Gateway field and enter the Default Gateway.



8 Select **OK**.

The Saving Network Configuration Screen is displayed.

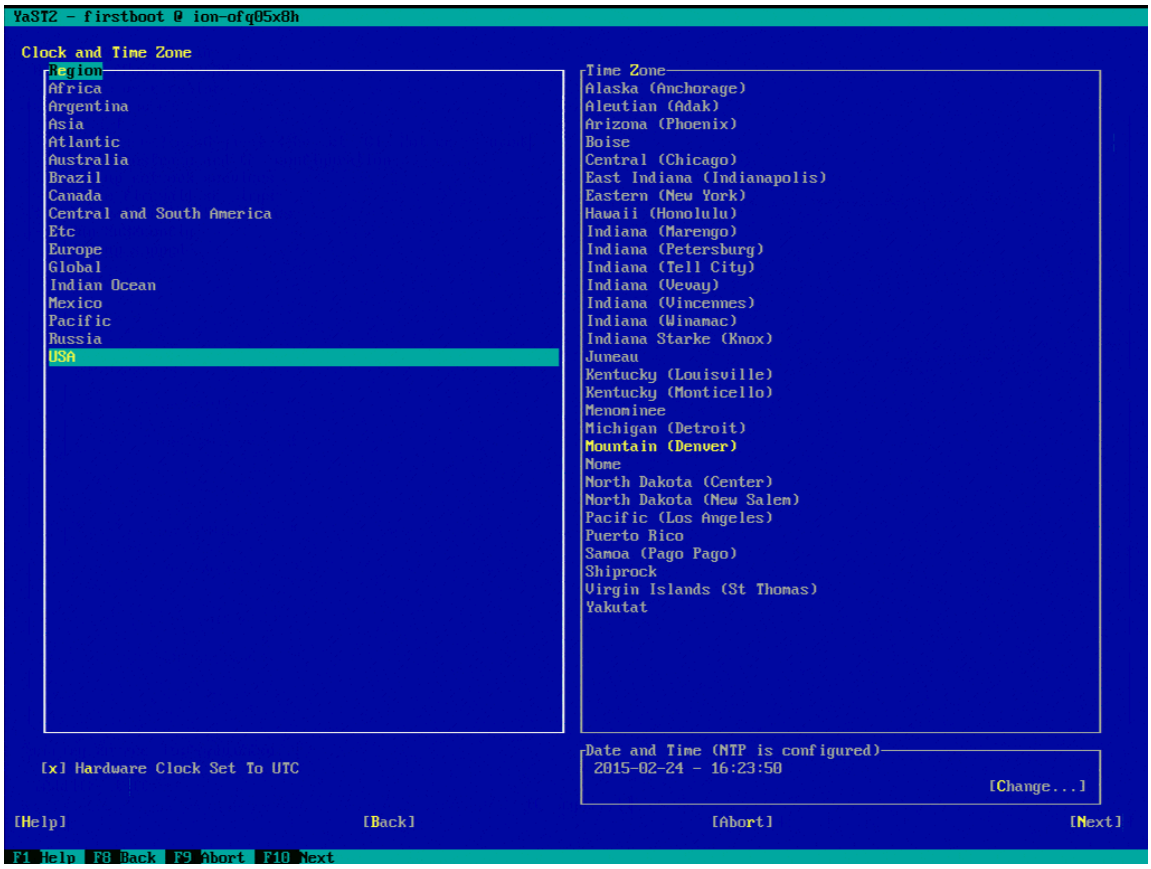


Setting time zone and NTP settings

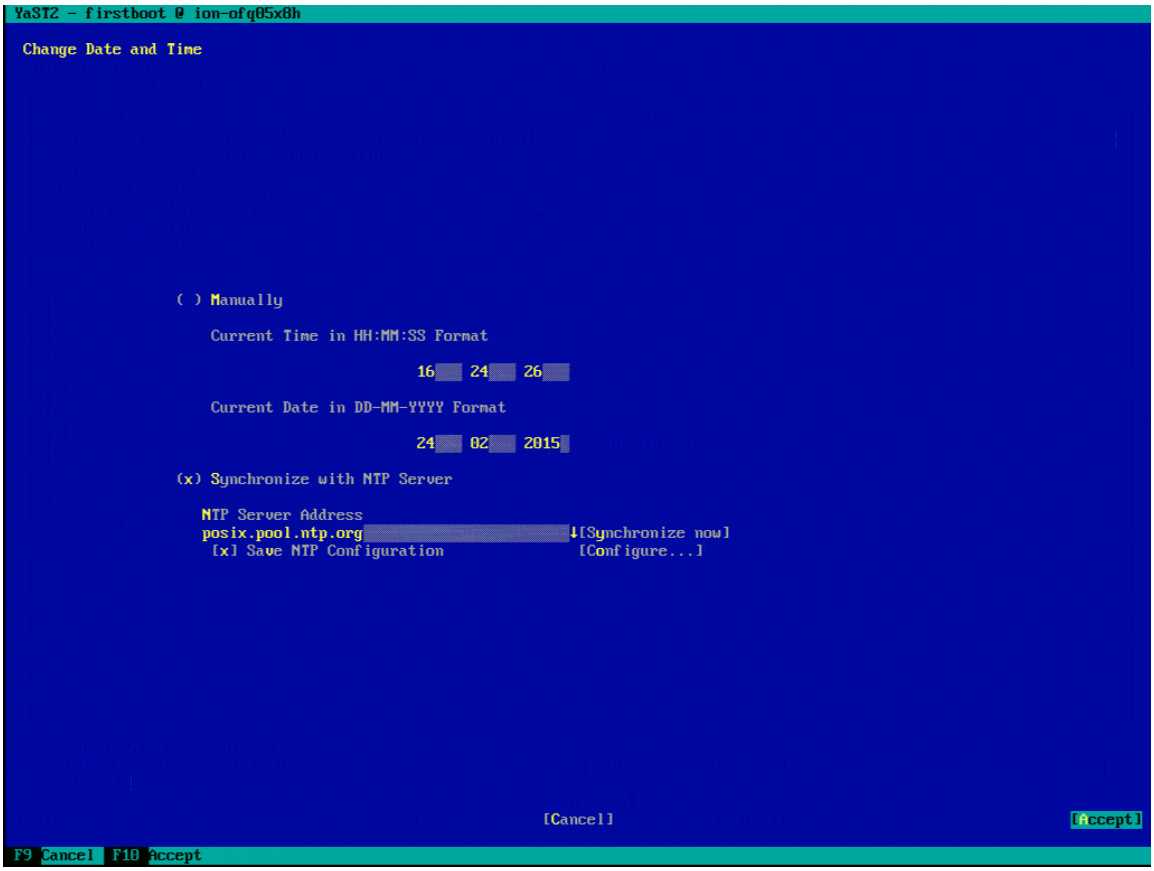
To set Time Zone and NTP:

1 Tab to the Region field and use the arrow keys to select the region where the DAAD will be located.

2 Tab to the Time Zone field and select the zone where the DAAD will be located.



3 Tab to the Date and Time field and select **Change**. The Change Date and Time screen is displayed.



CAUTION! If you are using an HA configuration you must configure DAAD to synchronize with NTP.

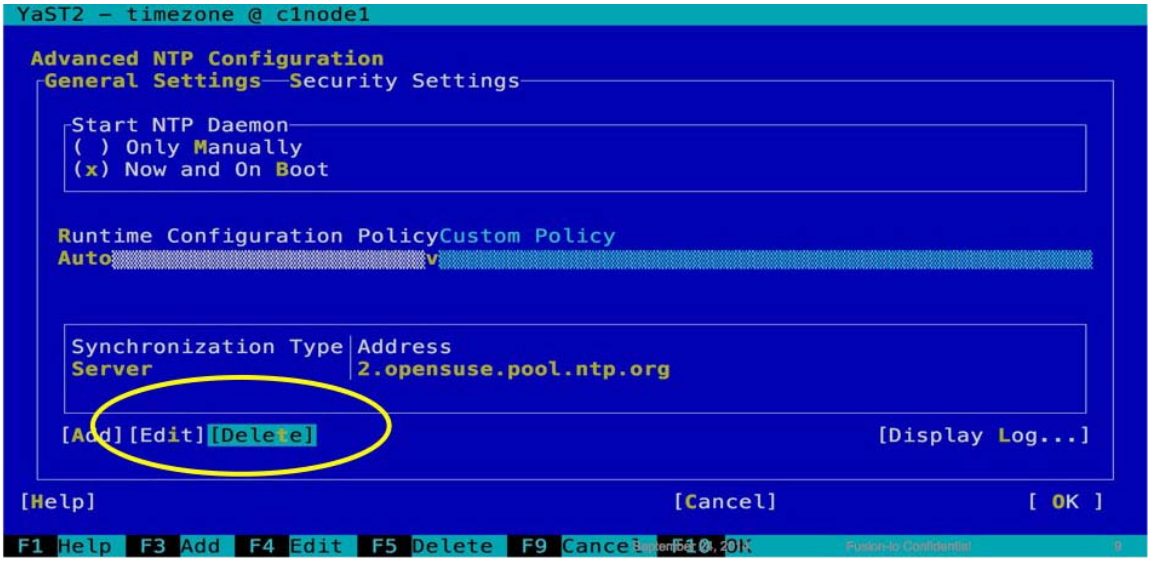
4 Tab to the Synchronize with NTP Server field and select it.

5 Tab to the NTP Server Address field and enter the address of the NTP server to synchronize with.

6 Tab to the Save NTP Configuration field and select it.

7 Tab to the Configure field and select it.

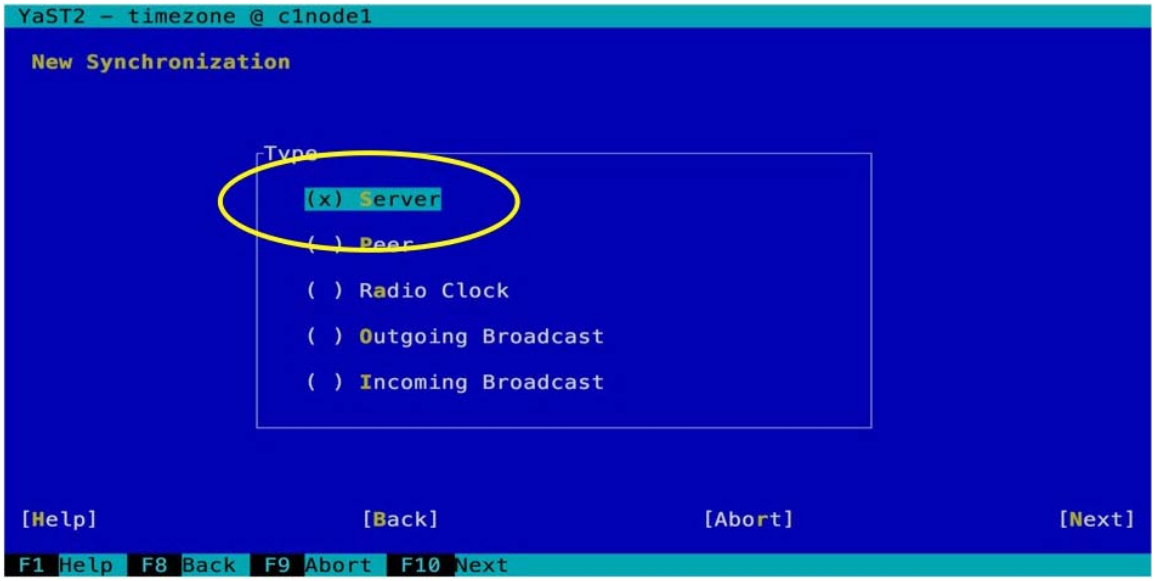
The Advanced NTP Configuration screen is displayed.



8 Select **Delete** to delete the default NTP configuration.

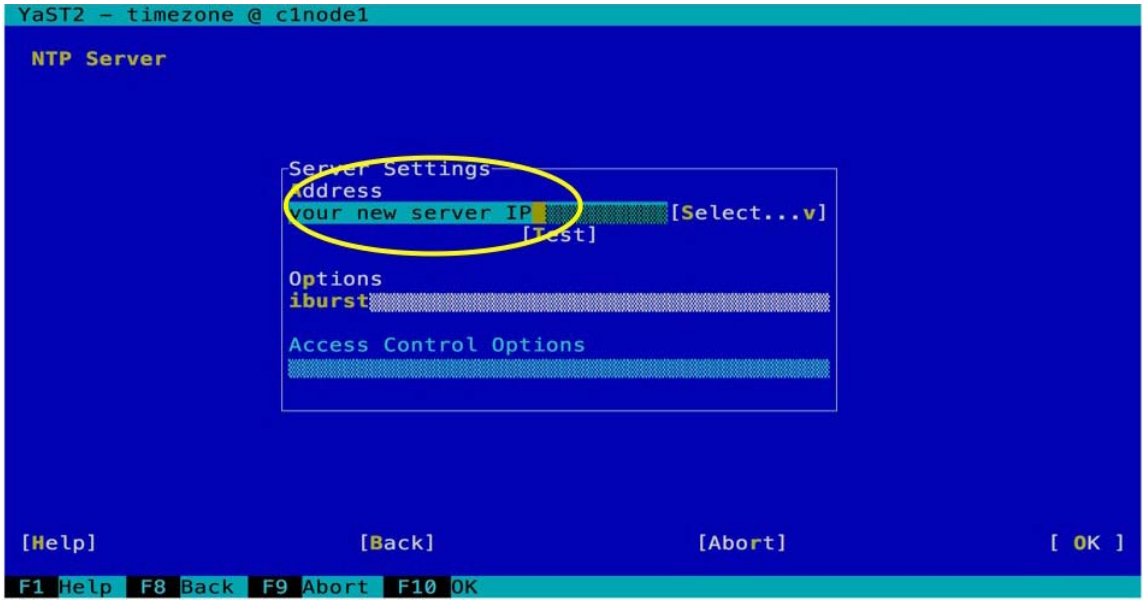
9 Select **Add** to add a new configuration.

10 In the New Synchronization screen that is displayed, select **Server**.



11 Select **Next**.

12 In the NTP Server screen that is displayed, tab to select the Address field.



13 Type the NTP server address.

14 Select **OK**.

15 If you have additional NTP servers, repeat steps 9-14 to configure them.

16 Select **Accept** to save the NTP changes and continue.

Enabling cluster setup

To enable the HA cluster:

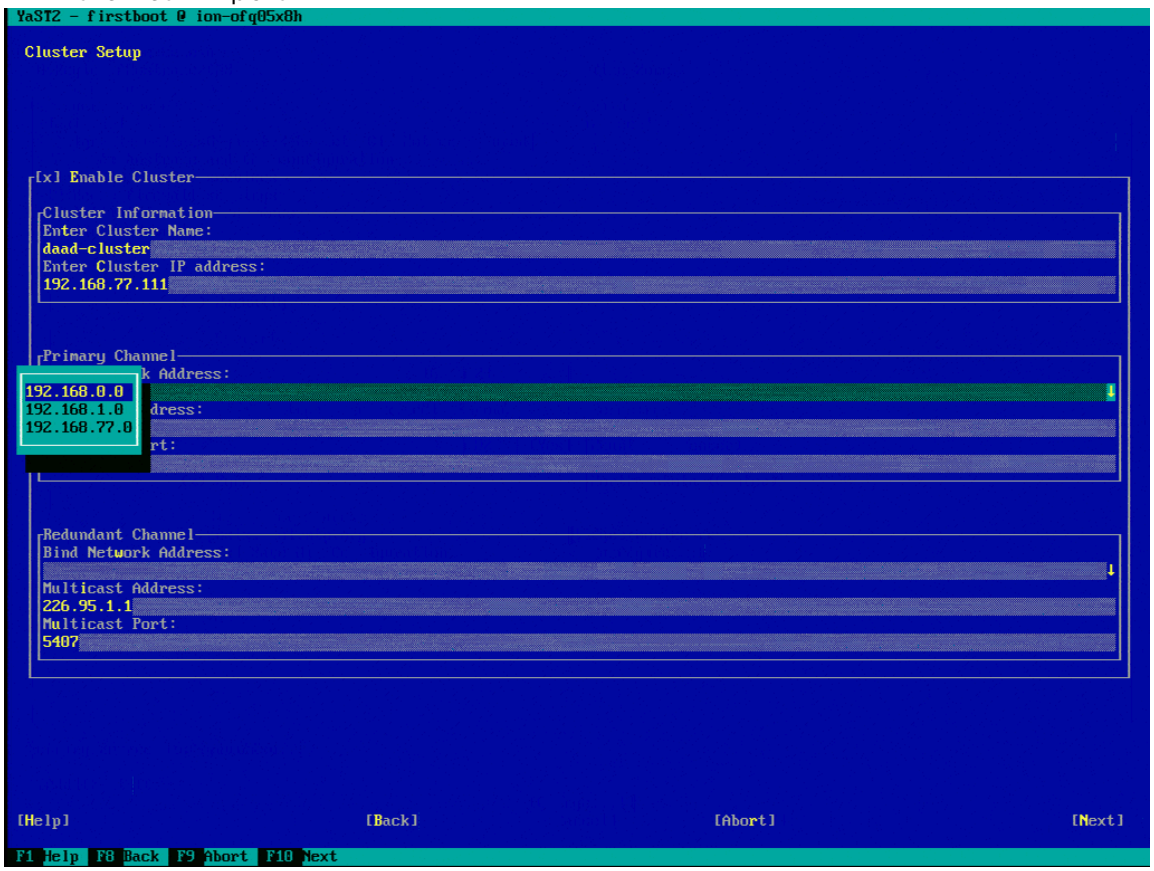
1 On the Cluster Setup screen, tab to the Enter Cluster Name field and enter a name for the Cluster.

> **CAUTION!** The cluster name must be a DNS entry, not DHCP, and must be on the same subnet as the management IP address for the Dell Acceleration Appliance for Databases nodes. Also, do not enter the name or IP address of one of the currently existing nodes.

2 Tab to the Enter Cluster IP Address field and enter the IP address for the cluster.

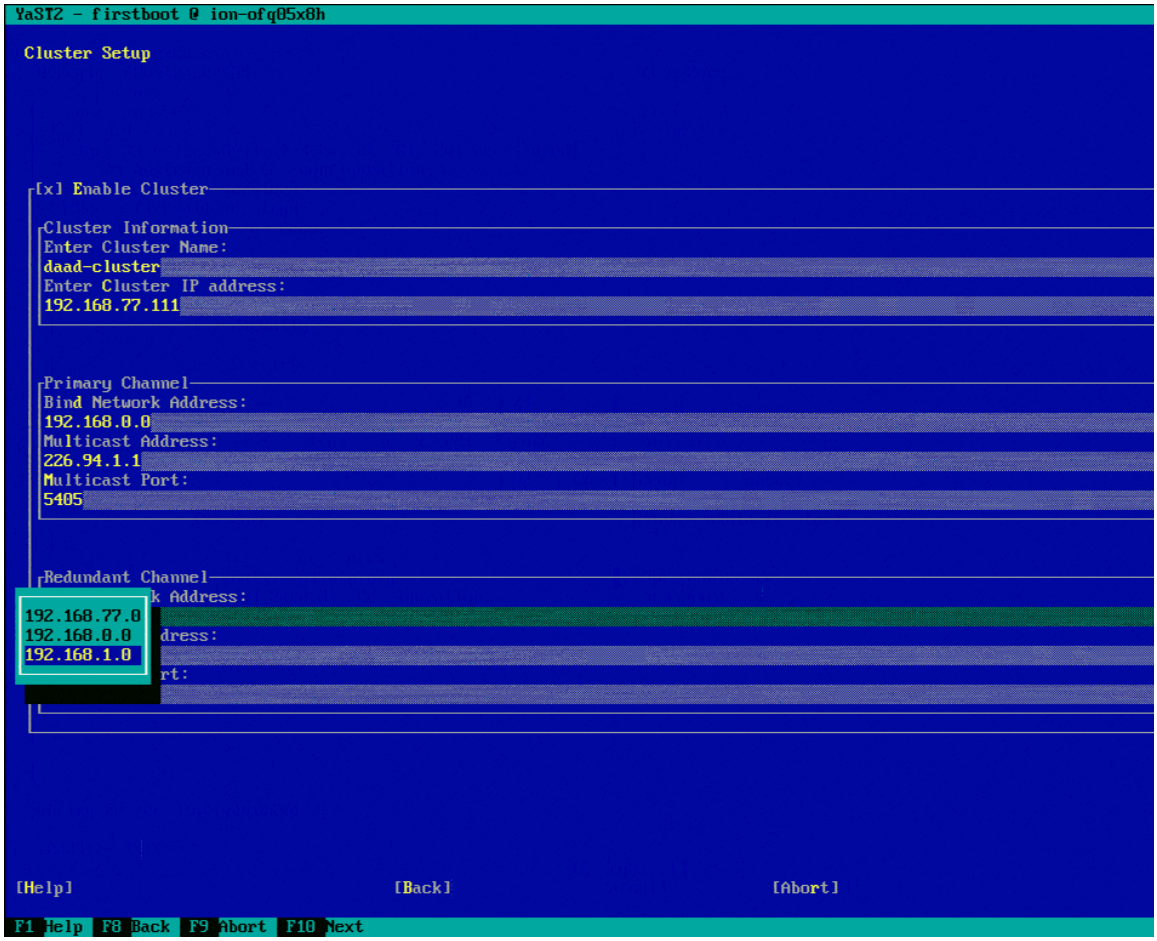
3 Tab to the Primary Channel field and press the down arrow key on Bind Network Address drop-down menu.

4 From the drop-down menu, select the address you set for the first ConnectX-3 Ethernet HA port.



5 Tab to the Redundant Channel field and press the down arrow key on Bind Network Address drop-down menu.

6 From the drop-down menu, select the address you set for the second ConnectX-3 Ethernet HA port.



7 Select **Next**.

Setting the admin user password

To set the admin user password:

1 Type the Dell Acceleration Appliance for Databases password for the admin user.

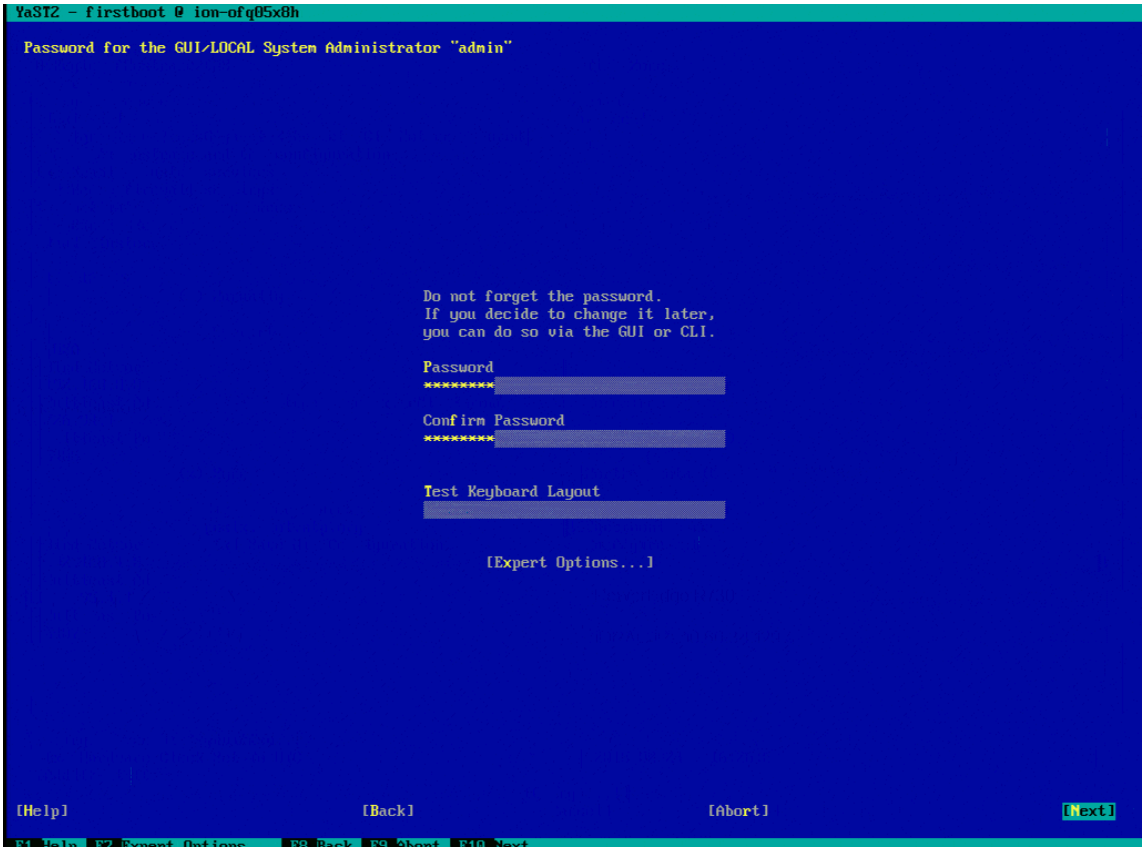
NOTE: If the password you selected is not sufficiently strong, a warning message is displayed so you can change the password, if necessary.

2 Retype the password you entered.

CAUTION! Be sure to record this password in a secure location in case it needs to be retrieved.

3 To test the Keyboard Layout or use the Expert Options, select those options on the screen.

4 Select **Next**.



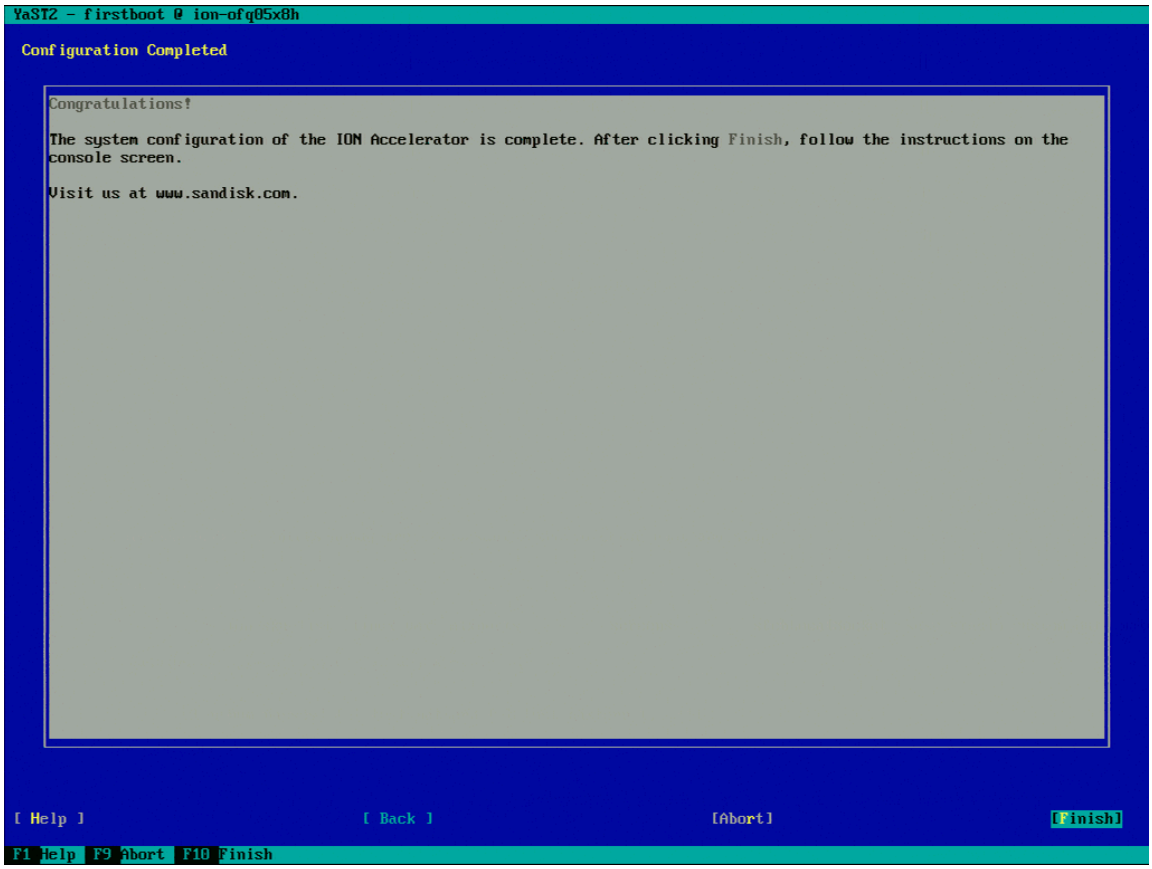
If you want to change this password later, change the admin user and GUI passwords separately. For more information, see Changing Passwords in the *Dell Acceleration Appliance for Databases GUI Guide*.

Completing the DAAD configuration

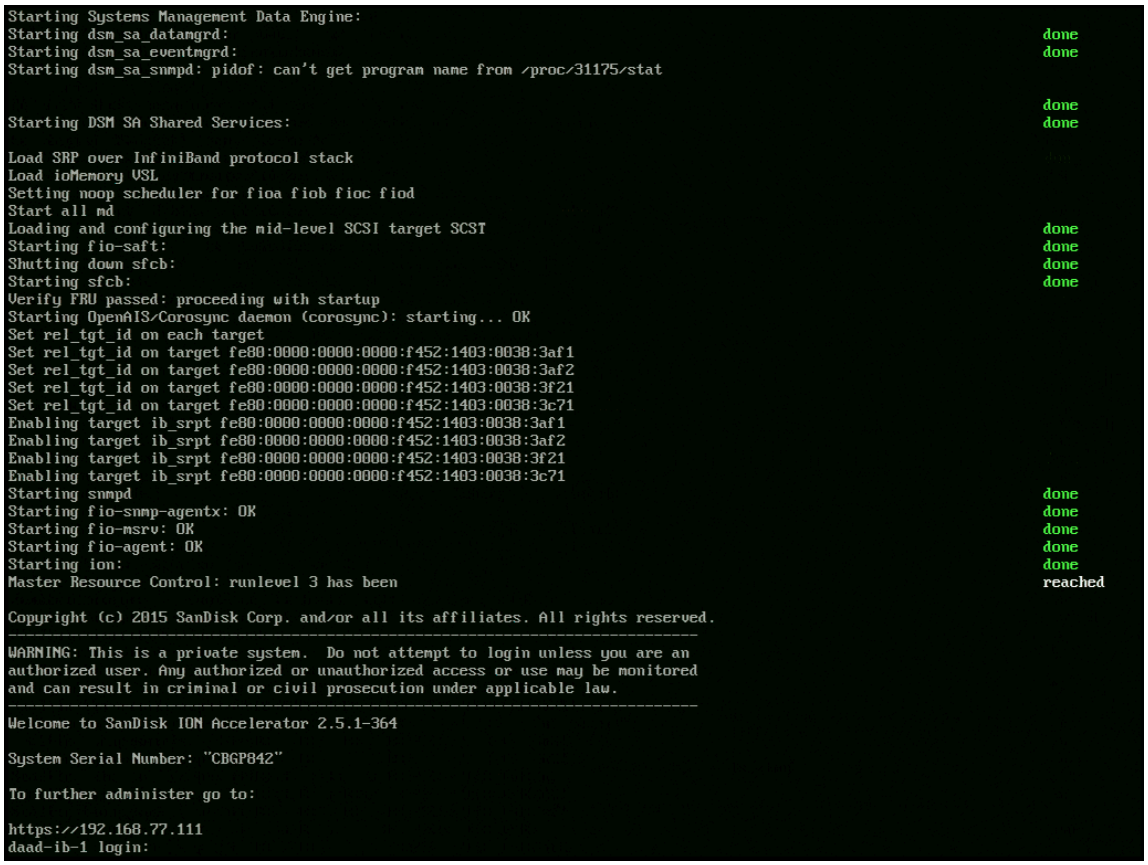
After the configuration phases are complete, the Configuration Completed screen is displayed.

To complete DAAD configuration:

1 Select **Finish**.



The DAAD system restarts and displays start-up messages. After a login prompt is displayed on the console, you can connect to DAAD for additional setup and configuration. An example screenshot is given here:

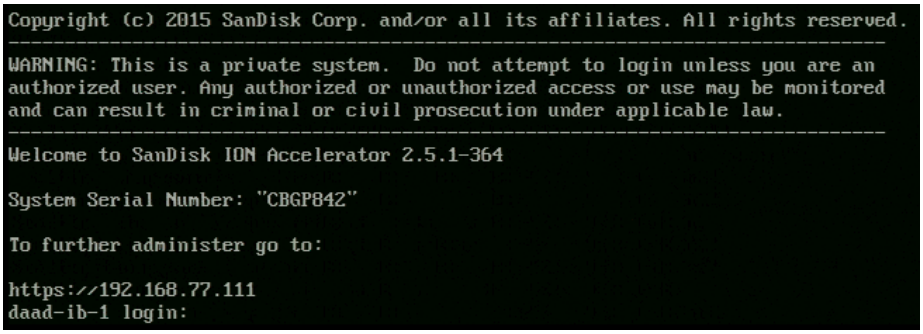


7

Connecting to DAAD

> **NOTE:** If you are setting up a permanent HA cluster, Dell recommends that you connect to each node. By connecting to the nodes, in case of a failback, the first-time setup on the second node is not necessary.

You can access and use the Dell Acceleration Appliance for Databases software by using two methods: the CLI or the GUI. The address of the DAAD system is displayed on the console.



Logging in through the CLI

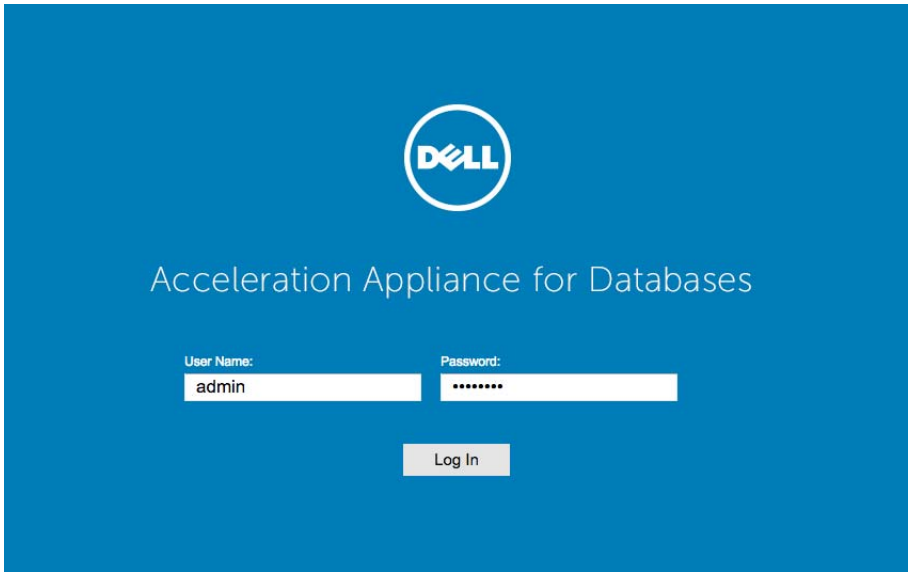
1. At the login prompt, type **admin** as the username.
2. At the password prompt, type the password you created in First Boot.
3. After you log in to the Dell Acceleration Appliance for Databases, use the *Dell Acceleration Appliance for Databases CLI Reference* for instructions about setting up and managing your DAAD storage.

Logging in through the GUI

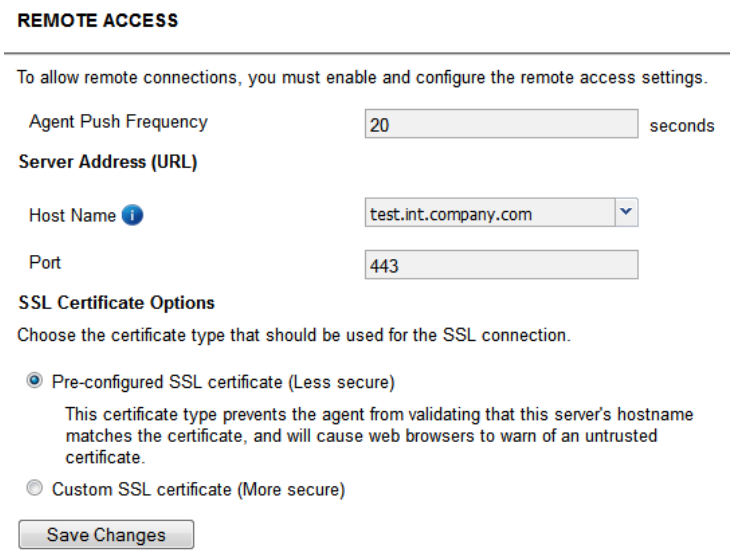
When you access the Dell Acceleration Appliance for Databases URL, the admin username and password login prompt is displayed:

1. If not pre-poulated, type **admin** in the Username box.
2. Type the password in the Password box. (This is the password you created in First Boot configuration.)

3 Click **Login**.

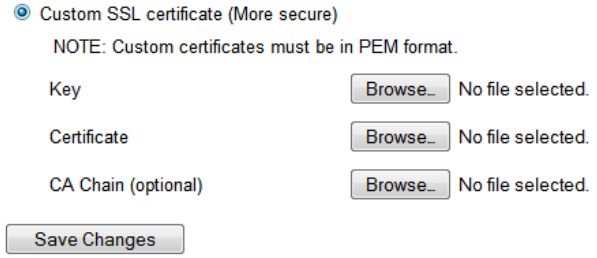


The SSL Certificate Options dialog box is displayed. By default, the pre-configured option is selected.



4 To use the less-secure pre-configured SSL certificate, click **Save Changes**.

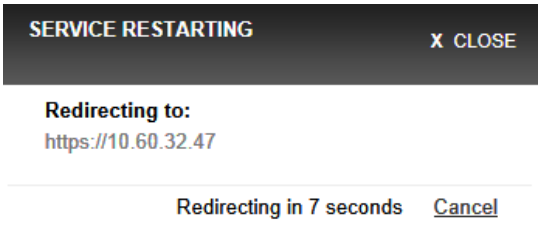
Or, to use a custom SSL certificate, click **Custom SSL certificate (more secure)**, and then click **Save Changes**.



NOTE: The default **Use pre-configured SSL Certificate**" option is less secure, as it causes untrusted certificate warnings to appear.

5 Click **Save Changes**.

NOTE: If the port, host name, or SSL information has changed, the service restarts in a few seconds, and the browser is redirected to the port you specified.



If the host name has changed, the service restarts, and the login process is initiated again. After re-direction, the login dialog box is displayed again.

6 Log in with the same username and password that you entered earlier.

The software is now configured so you can begin managing your Dell Acceleration Appliance for Databases. The Overview screen is displayed as shown in the *Basic Tasks and Overview Tab* section of the *Dell Acceleration Appliance for Databases GUI Guide*.

NOTE: For items not covered in the steps 1-6, see Remote Access in the *Dell Acceleration Appliance for Databases GUI Guide*.

If you want to change the admin password, see Changing Passwords in the *Dell Acceleration Appliance for Databases GUI Guide*..

NOTE: The admin account allows up to 10 concurrent sessions to run on the Dell Acceleration Appliance for Databases.

7 After you are logged in to the Dell Acceleration Appliance for Databases, see the *Dell Acceleration Appliance for Databases GUI Guide* for instructions on how to set up and manage your DAAD storage.

Backing up DAAD configuration

NOTE: It is a best practice to save and back up configurations of your Dell Acceleration Appliance for Databases system on a regular basis. As you set up storage pools, create volumes, and manage initiator groups, ensure that you back up your current configuration. For more information about backing up and restoring configurations, see the `config` commands in the *Dell Acceleration Appliance for Databases CLI Reference Guide*.

To back up the Dell Acceleration Appliance for Databases configuration:

1 Create a backup of the current configuration by running `config:backup`.

For example, the following commands create an XML backup that can be used later for a restore:

```
config:backup --host <hostname> --directory <path> --user <username> --
password <password>
```

2 Store the configuration in a location other than on the DAAD host.

8

Changing network settings

This section describes the tasks to change the network settings for the Dell Acceleration Appliance for Databases nodes after they have been configured in the First Boot process. For information about using the `system:setup` command referred to in this section, see the *Dell Acceleration Appliance for Databases CLI Reference*.

Changing a node host name in a cluster

NOTE: Changing a node host name requires cluster downtime.

1. Ensure that both nodes are turned on.
2. Close all active sessions by disabling or disconnecting all target ports.
3. At the CLI, run **`system:setup lan`** to view the Setup dialog box for LAN configuration.
4. Select the **Hostname/DNS** tab.
5. Change the name of the node in the dialog box.
6. Select **OK**. Both nodes will restart one at a time.
7. Repeat steps 5 and 6 for the second node in the cluster.

Changing the management IP address

NOTE: While changing the management IP address do not change any values for the gateway IP address. To change the gateway IP address perform the procedure Changing the gateway IP address on page 96 as a separate step.

1. Ensure that both nodes are online.
2. On the first node, at the CLI, enter **`maintenance on`**
3. Enter **`system:setup lan`** to view the Setup dialog box for LAN configuration.
4. Edit the Ethernet adapter configured with the management IP.

   The Network Card Setup dialog box is displayed.
5. Change the management IP address in the dialog box.

NOTE: Ensure that the host name that may be associated with the IP address in your network's DNS matches the host name set for the DAAD system.

6 Select **OK**. The node with the cluster IP address will fail over to the other node.
7 Enter **`maintenance off`**
8 On the second node, at the CLI, enter **`maintenance on`**
9 Repeat steps 3–7 on the second node.

Changing the management IP address to or from DHCP

1 Ensure that both nodes are online.
2 At the CLI, enter **`maintenance on`**
3 Enter `system:setup lan` to view the Setup dialog box for LAN configuration.
4 On the Network Card Setup screen, you can change the IP address setting to **Dynamic Address** or **Statically assigned IP Address**.

> **NOTE:** Ensure that the host name that may be associated with the IP address in your network's DNS matches the host name set for the DAAD system.

5 Select **OK**. The node with the cluster IP address will fail over to the other node.
6 Enter **`maintenance off`**

Changing the DAAD cluster name or IP address

1 Ensure that both nodes are online.
1 On the first node, at the CLI, enter **`maintenance on`**
2 Enter `system:setup cluster` to view the Setup dialog box for cluster configuration.
3 Change the cluster name or IP address in the dialog box.
4 Select **OK**.
5 Enter **`maintenance off`**
6 On the second node, at the CLI, enter **`maintenance on`**
7 Repeat steps 3–6 on the second node.

Changing the gateway IP address

1 Ensure that both nodes are turned on.
2 At the CLI, run `system:setup lan` to view the Setup dialog box for LAN configuration.
3 Select the Routing tab in the dialog box.
4 Change the gateway IP address.
5 Select **OK**.

Changing IP addresses for cluster interconnect ports

NOTE: You can change interconnect ports only when no volumes or LUNs are configured.

1. Ensure that both nodes are turned on.
2. At the CLI, run `system:setup lan` to view the Setup dialog box for LAN configuration.
3. In the dialog box, change the cluster interconnect IP address for the adapters configured as HA ports.
4. Select **OK**.
5. Restart both nodes.

Changing the iSCSI port IP address

1. Ensure that both nodes are turned on.
2. Enter `maintenance on`
3. At the CLI, run `system:setup lan` to view the Setup dialog box for LAN configuration.
4. Change the iSCSI port IP address in the dialog box.
5. Select **OK**.
6. For Linux initiators:
   a. Log out of all related target portals by using the following command:
      `iscsiadm -m node -U all`
   b. Remove old target ports by using the following command:
      `iscsiadm -m node -p <oldIPaddress> -o delete`
      Where
      oldIPaddress—is the IP address of one of iSCSI ports you want to change.
      Repeat this step for all the old target ports that you want to change.
   c. Add the new target session to the initiator by running the following command:
      `iscsiadm -m discovery -t st -p <new iSCSI IP address>`
   d. For Linux initiators, log in to the target (all sessions) from the initiator:
      `iscsiadm -m node –l`
   e. Run the CLI command `initiators –dt` to view the new initiator.
   f. Add this initiator to an existing group:
      `initiator:update –a <inigroup name> <initiatoriqn#target IP Address>`

7 For Windows initiators:

   a Disconnect and remove old target ports by clicking **iSCSI Initiator > Discovery** and removing old target ports with old IP addresses

   b Click **iSCSI Initiator > Targets** and disconnect old targets

   c Click **OK**.

   d For Windows initiators, connect and discover targets with new IP addresses by clicking **iSCSI Initiator > Targets** and entering new IP addresses.

   e Click **Quick Connect**.

   f Click **OK**.

8 From the console of the DAAD target, enter the following command to update the initiators displayed in step 9 to the initiator group:

   `initiator:update -a <initiatorGroup> <newInitiator>`

   Where

   initiatorGroup—is the existing initiator group.

   newInitiator—is one of the new initiators displayed by the `initiators -dt` command.

   Repeat this step for all the new initiators you want to include in the initiator group.

9 From the command line, enter `maintenance off`

NOTE: If the initiators do not rediscover the LUNS, run the echo command on the initiator console. The syntax of the commad is

"echo "- - -" > /sys/class/scsi_host/<host0>/scan"

where *host0* is the host number

Changing MTU settings for iSCSI

To change the MTU of the iSCSI NICs:

1 Log in to the DAAD console as user admin.

2 Enter `maintenance on`

3 Enter `setup lan`

   The Network Settings screen is displayed.

4 Edit the port where you want to change the MTU setting.

   The Network Card Setup screen is displayed.

5 On the Network Card Setup screen, press the left arrow key to select **General**.

6 Tab to the Set MTU field.

7 Set MTU to the new value.

8 Select **Next**.

9 Repeat steps 4-8 for any other iSCSI ports where you want to change the MTU settings.

10 Exit the Network Card Setup screen and the Network Setting screen.

11 From the command line, enter `maintenance off`

12 Enter `ports -dt` to ensure the new MTU setting.

9

HA and host configuration

NOTE: If you are not planning to implement the high-availability feature of the Dell Acceleration Appliance for Databases software, go to Application tuning on page 151.

About DAAD high availability

The Dell Acceleration Appliance for Databases enables a powerful and effective High Availability (HA) environment for your shared storage. HA clustering provides an important option for customers who prefer array-based HA over application-based mirroring. This can be especially useful if your application does not provide logical volume management, such as with all VMware environments and most implementations of Microsoft Clustering.

NOTE: Clustering relies on point-to-point connections. HA networking across geographically distributed sites is not supported.

Figure 9-1 shows both a simple HA clustering setup by using a 40Gb Ethernet connection between appliances and a mirroring configuration without clustering.

Figure 9-1. SImple HA Clustering

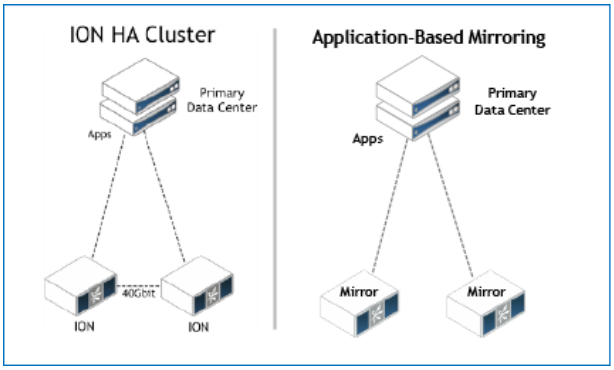
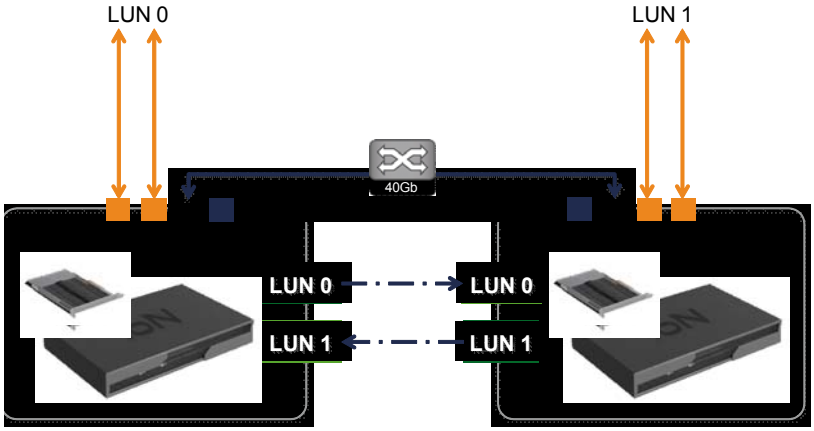


Figure 9-2 shows basic LUN access (exported volumes) in an HA configuration.

Figure 9-2. LUN access in HA configurations



In this simplified example, Node 1 presents LUN 0 to the application, while Node 2 presents LUN 1 to the application. All "write" operations to LUN 0 are synchronously replicated to Node 2. All Writes to LUN 1 are synchronously replicated to Node 1. When the replication is complete, the original node acknowledges the write by the application.

Replication occurs over the 40Gb Ethernet interconnect between the two nodes, which consists of two dual-ported ConnectX-3 adapters. In the event a node stops functioning, all data will be available on the remaining active unit.

InfiniBand/SRP and iSCSI connections for HA

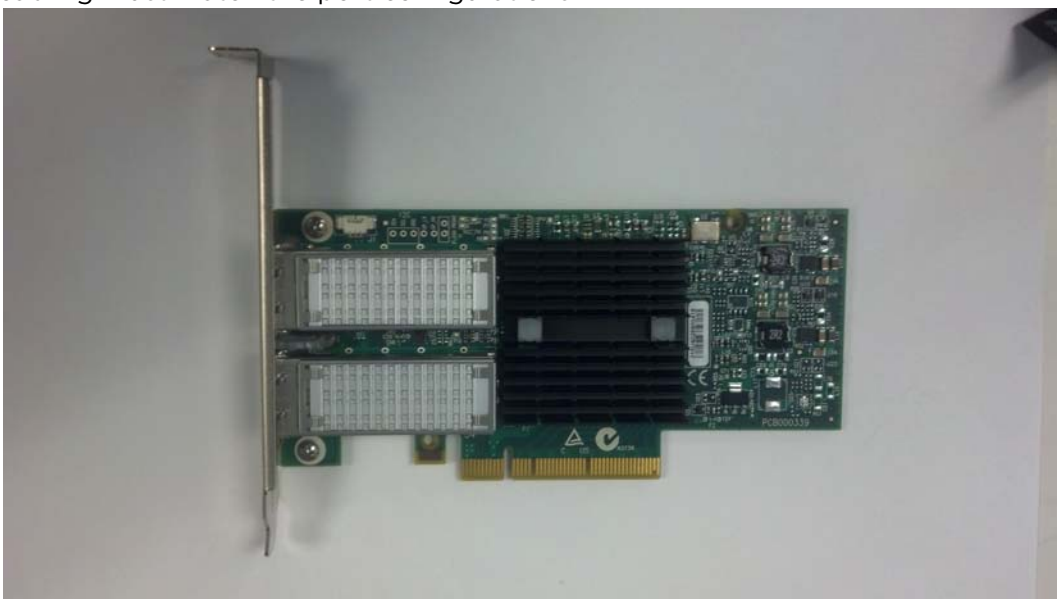
Both iSCSI and InfiniBand/SRP systems use dual-port ConnectX-3 cards for cluster interconnections.

For HA systems, the two ports on the interconnect card are split — one port on each of the two ConnectX-3 cards in slot 1 and slot 3 is used for the cluster interconnect, and one is used for the fabric connection. For InfiniBand/SRP systems, two cards in slot 1 and slot 3 are automatically split at boot time by the Dell Acceleration Appliance for Databases software, with the first port being put in InfiniBand mode and the second port being put in Ethernet mode. (Mellanox requires that the first port on a split card be in InfiniBand mode.)

See About DAAD first boot on page 13 for details and dialog boxes regarding cluster setup, and refer to Split-function ports for InfiniBand/SRP and iSCSI on page 164 for details on split-port configuration.

The Dell Acceleration Appliance for Databases does not enforce port-splitting for iSCSI HA configurations (unlike InfiniBand/SRP HA configurations). To ensure maximum availability for HA systems, your iSCSI configurations should split the ports on the two cards in slot 1 and slot 3, using one from each card for the cluster interconnect.

When cabling InfiniBand/SRP HA systems, configure the ports as explained here. The cabling must match the port configurations.



ConnectX-3 card in slot 1 (if HA)

- Port 1: InfiniBand/SRP
- Port 2: Cluster Interconnect

ConnectX-3 card in slot 2

- Port 1: InfiniBand/SRP
- Port 2: InfiniBand/SRP

ConnectX-3 card in slot 3

- Port 1: InfiniBand/SRP
- Port 2: Cluster Interconnect (if HA, otherwise InfiniBand/SRP)

Fabric and cluster interconnect cards

The photos here show common fabric and cluster interconnect cards, with ports labeled. As you configure your Dell Acceleration Appliance for Databases HA system refer to the port locations shownin DAAD configuration on page 163. This will ensure that the optimal port connections are made for a high-performance HA configuration.

Figure 9-3. QLogic QLE2562 Fibre Channel card (Dual-port)

Figure 9-4. ConnectX-3 cluster interconnect card

Multipathing overview

NOTE: For information about managing initiators, such as creating initiator groups, or adding or deleting initiators, refer to the *Dell Acceleration Appliance for Databases GUI Guide* or the *Dell Acceleration Appliance for Databases CLI Reference*.

Supported network fabric and hosts

The Dell Acceleration Appliance for Databases supports Fibre Channel, InfiniBand/SRP, and iSCSI network fabrics for initiators. Refer to the *Dell Acceleration Appliance for Databases Compatibility Guide,* available at dell.com/support/home for complete details on hardware compatibility, multipathing support, and the list of all operating systems supported on the hosts connecting to DAAD.

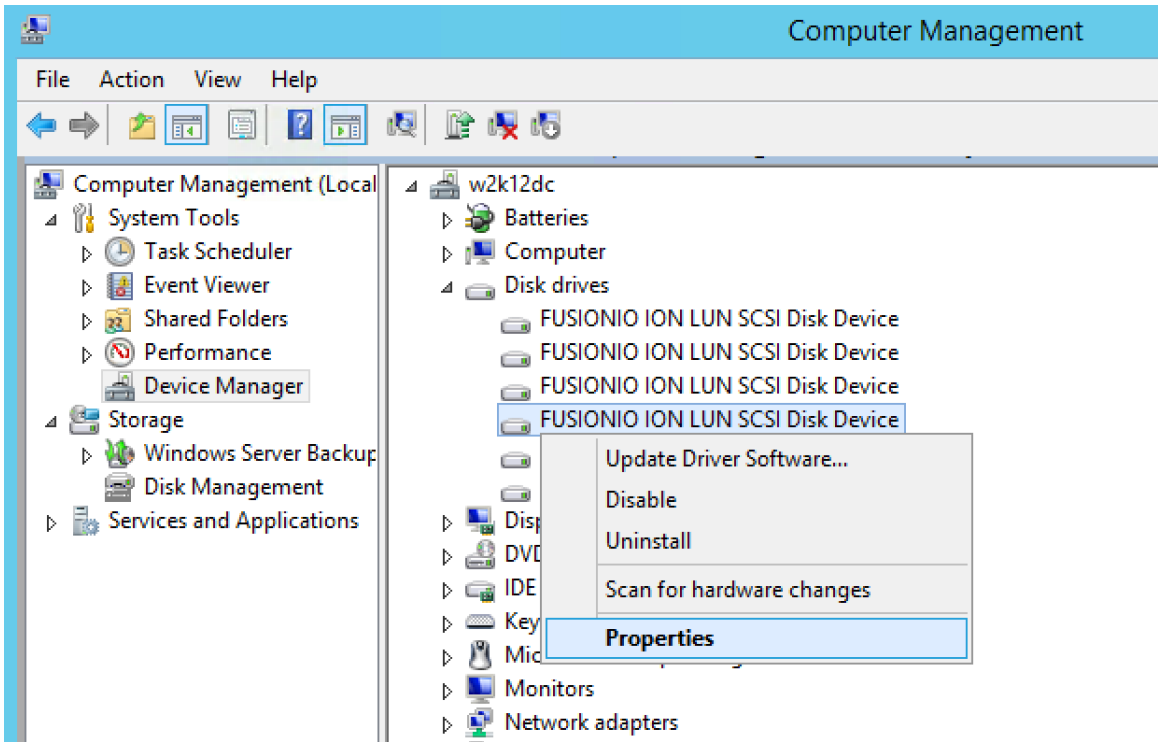
About multipathing for Windows

Multipath I/O (MPIO) establishes multiple routes and connections to a storage array, by using redundant physical paths (adapters, cables, switches). That way, when a component fails, an alternate I/O path is used. Multipathing provides redundancy of I/O paths and can improve overall system performance.

Before multipathing is installed, you see one drive for each path setup (so with two paths, you see two drives). The Drive Properties dialog box now contains a multipath tab, where you can set the load-balance policy.

To open the Driver Properties dialog box:

1. Click **Start > Administrative Tools > Computer Management > Device Manager**,
2. Right-click the disk that you want to add multiple paths to.
3. Select **Properties**.



NOTE: If you are using Fibre Channel, ensure that you have the latest HBA driver installed before configuring multipathing.

Load-balance policies for MPIO

Here are some of the basic load-balance policies typically used with multipathing:

- Round-robin, subsets: Standby paths are used only if all primary paths fail.

  NOTE: For Windows MPIO, the “round-robin with subsets” method is required for HA configurations; for Windows standalone configurations, round-robin or dynamic least queue depth methods may be used, with the latter generally preferred.

- Failover: No balancing. Standby paths are used.
- Failback: I/O is rerouted to preferred path when available.
- Round-robin: All available paths are used for balanced I/O.
- Dynamic least queue depth: I/O to path with fewest outstanding requests (in Linux, this is queue-length). This enables multipathing to compensate for an unbalanced load on the fabric. This may be advantageous for standalone configurations.
- Weighted path: Paths are assigned priority weights.

Configuring multipathing on Windows

Installing multipath on windows

When MPIO is installed, the Microsoft device-specific module (DSM) and the MPIO Control Panel are also installed.

For more information about Windows MPIO settings, refer to http://blogs.msdn.com/b/san/archive/2011/12/02/updated-guidance-on-microsoft-mpio-settings.aspx

> **NOTE:** In Windows 2008 R2 SP2, disk I/O operations may fail even when valid failover paths exist. To resolve this issue, apply the Microsoft hotfix available at http://support.microsoft.com/kb/2752538/en-us

Installing MPIO in Windows Server 2008 R2

To add MPIO on a server running Windows Server 2008 R2,

1. Start Server Manager: **Start > Administrative Tools > Server Manager**

2. In the Server Manager tree, click **Features**.

3. In the Features section, click **Add Features**.

4. In the Add Features Wizard, on the Select Features page, select **Multipath I/O**, and then click **Next**.

5. On the Confirm Installation Selections page, click **Install**.

6. After the installation completes, on the Installation Results page, click **Close**.

7. When prompted to restart the computer, click **Yes**. After restarting, the computer finalizes the MPIO installation.

8. Click **Close**.

Installing MPIO in Windows Server 2012 R2

To add MPIO on a server running Windows Server 2012 R2,

1. Start Server Manager: **Start > Administrative Tools > Server Manager**

2. In the Server Manager tree, click **Add roles and Features**.

3. Continue in the wizard to the **Select Features** section.

4. In the Features window, select **Multipath I/O**, and then click **Next**.

5. On the Confirm Installation Selections page, click **Install**.

6. After the installation completes, on the Installation Results page, click **Close**.

7. When prompted to restart the computer, click **Yes**. After restarting, the computer finalizes the MPIO installation.

8. Click **Close**.

More information is available in the *Microsoft Multipath I/O (MPIO) User's Guide for Windows Server 2012* which is available at http://www.microsoft.com/en-us/download/details.aspx?id=30450.

Configuring multipath on Windows

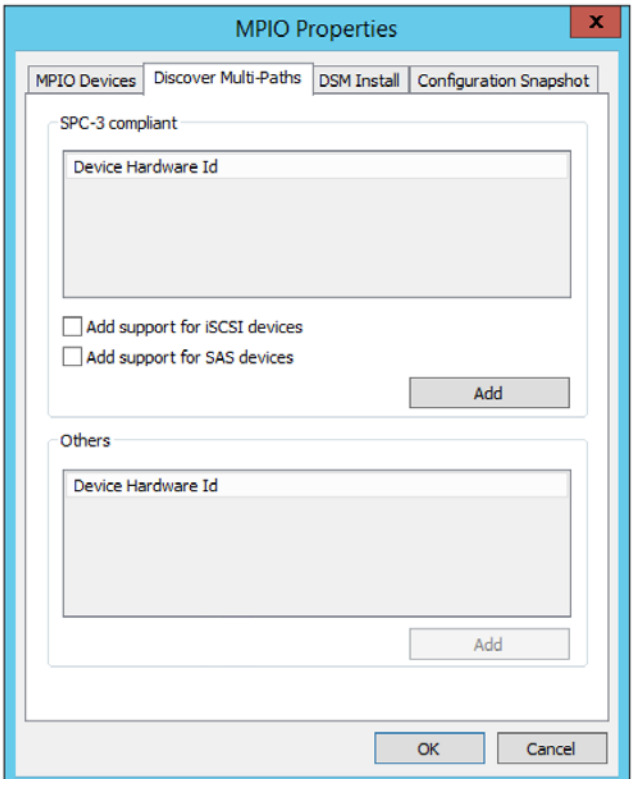
The MPIO Control Panel enables you to:

- Configure MPIO functionality
- Install additional storage DSMs
- Create MPIO configuration reports

To configure Windows multipathing:

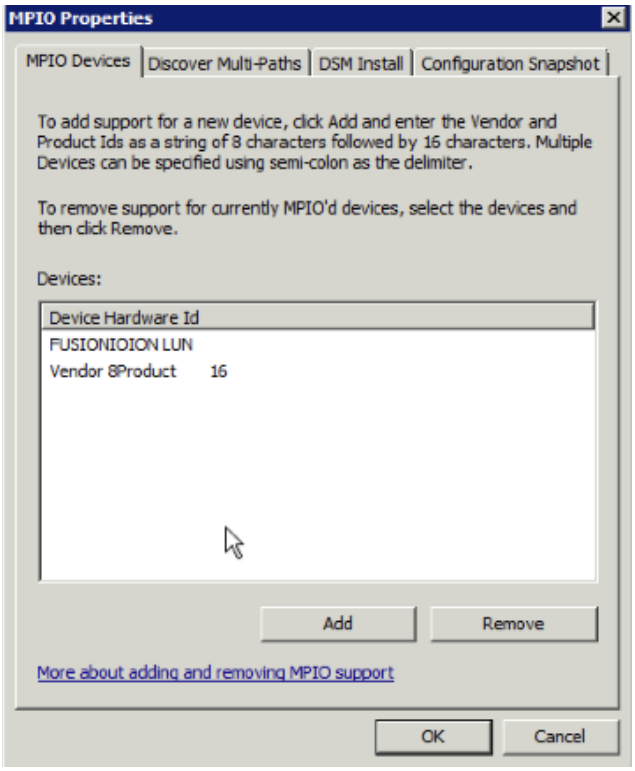
1. Click **Start > Control Panel > Views list > Large Icons > MPIO**.
2. On the User Account Control page, click **Continue**.

   The MPIO Properties dialog box is displayed.
3. If you are using iSCSI, click the **Discover Multi-Paths** tab, and ensure that the **Add support for iSCSI devices** is selected.



4 Click the **MPIO Devices** tab.

This tab displays the hardware IDs (for example, FUSIONIOION LUN) of the devices managed by MPIO, whenever they are present.



5 To add a hardware ID,

a Click **Add**.

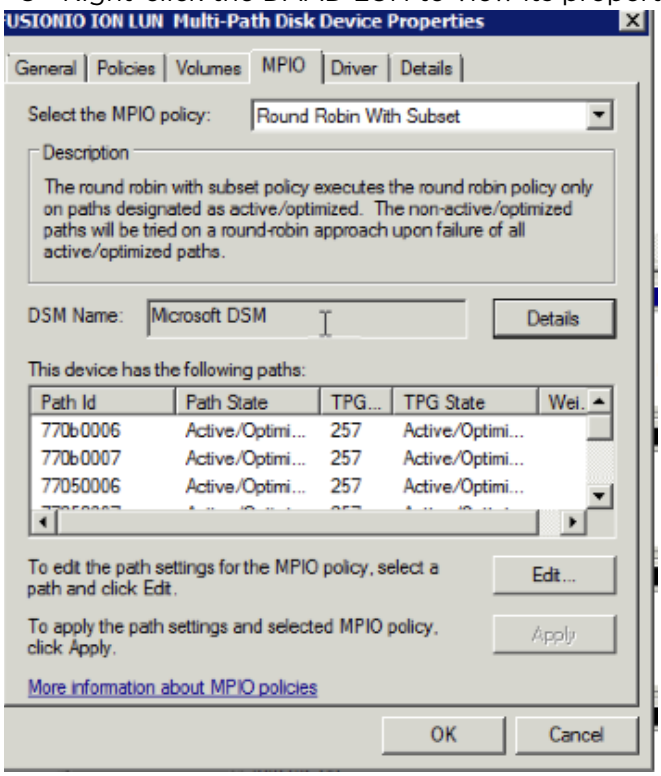
b Type the hardware ID, which is Vendor ID (eight characters), and type the Product ID (16 characters).

c Click **OK**.

6 Click **OK**.

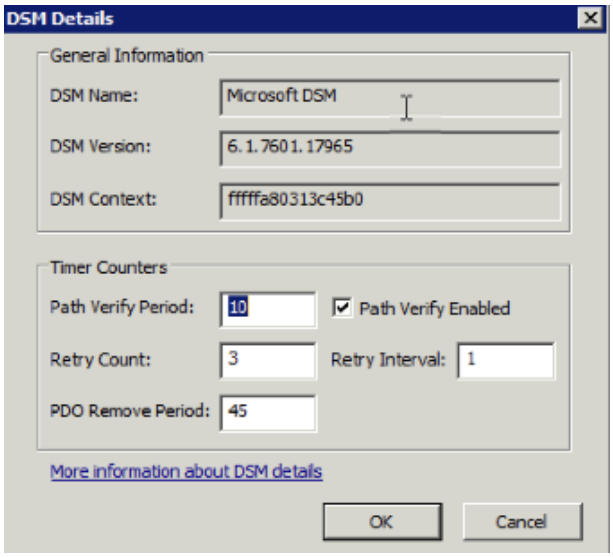
7 Click **Storage >Server Manager >Disk Management**.

8 Right-click the DAAD LUN to view its properties.



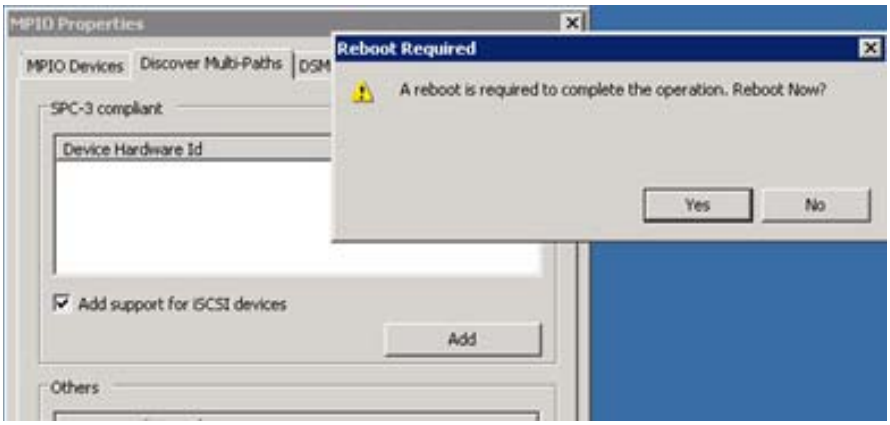
9 Click **MPIO Policy** on the MPIO tab, ensure that **Round Robin With Subset** is selected.

10 Click **Details** to view the DSM Details dialog box.



11 Ensure that the **Path Verify Enabled** check box is selected.

12 Click **OK** to close each dialog box, until prompted to restart the system.



13 If you are running in standalone mode, click **Yes** to restart, and then go to Discovering multipaths on page 112.

14 If you are running in HA mode, click **No** at the reboot prompt, as there more tasks to be completed.

15 Use `regedit` to edit the Registry with the following changes:

```
HKLM\SYSTEM\CurrentControlSet\services
\Disk\TimeOutValue = 60 secs
\mpio\Parameters\UseCustomPathRecoveryInterval = 1
```

NOTE: If you are using Fibre Channel, use the 30-second interval shown below; if you are using iSCSI, set the interval to 5:

```
\mpio\Parameters\PathRecoveryInterval = 30 secs
```

16 For the QLogic 25xx HBA (Fibre Channel), make the following changes in the driver, by using the QLogic CLI or GUI in Windows:

```
QLogic Port Link/Down Timeout = 15 secs
QLogic Port Retry/Down Timeout = 15 secs
```

Or if you are using Emulex drivers with Fibre Channel, make the following changes:

```
Emulex LinkTimeout = 15 secs
Emulex NodeTimeout = 60 secs
```

17 Restart the Windows system.

Discovering multipaths

To discover multipaths that exist in your configuration:

1 Ensure that multiple instances represent the same Logical Unit Number (LUN) through different paths. Hardware IDs for those devices are displayed for use with MPIO.

2 Add a second initiator to a volume.

3 With multipath running, open a disk management tool, such as Windows Disk Manager.

4 View the volume from the initiator and ensure that only one volume shows up. (Before installing multipath, the volume would show up twice).

Configuring iSCSI initiators for Windows

The instructions in this section apply to multipathing for Windows Server 2008 R2 and Windows Server 2012.

NOTE: Each initiator iSCSI port should be in a separate subnet that is the same as that of the target iSCSI port.

NOTE: Ensure that the default port 3260 for iSCSI is not blocked by a firewall.

Standalone mode

1. Ensure that the following Windows hot fix module and settings are applied: http://support.microsoft.com/kb/2752538/en-us
2. Use `regedit` to disable the delayed ACK in the registry:

   ```
   HKLM\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters\Interfaces\<Interface GUID>\TcpAckFrequency = 1
   ```

   To create this registry setting:

   a. Right-click in the interface you are using for iSCSI.

   b. Select a new DWORD 32-bit value.

   c. Name the new value `TcpAckFrequency`.

   d. Set the value to 1.
3. Increase the reconnect retries value in the registry as follows:

   ```
   HKLM\SYSTEM\CurrentControlSet\Control\Class\{4D36E97B-E325-11CE-BFC1-08002BE10318}\<Instance Number>\Parameters\PortalRetryCount = ffffffff
   ```
4. Set the registry values and restart the system.

HA mode

1. Complete the tasks 1 through 3 in the procedure for Standalone mode on page 113.
2. Click **Control Panel > MPIO > Add Device Hardware ID**.
3. Specify the device as **FUSIONIOION LUN**.
4. Use regedit to set the timeout value to 60 seconds:

   ```
   HKLM\SYSTEM\CurrentControlSet\services\Disk\TimeOutValue = 60 secs
   ```
5. Use regedit to set the custom path recovery interval to 1:

   ```
   HKLM\SYSTEM\CurrentControlSet\services\mpio\Parameters\UseCustomPathRecoveryInterval = 1
   ```
6. Use regedit to set the path recovery interval to 5 seconds:

   ```
   HKLM\SYSTEM\CurrentControlSet\services\mpio\Parameters\PathRecoveryInterval = 5 secs
   ```
7. Set the registry values and restart the system.

Setting up iSCSI in Windows

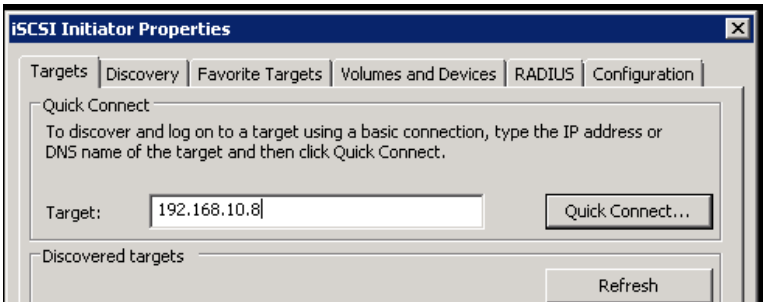
For the example used in this configuration, the following items must be noted:

- The Windows 2008 R2 initiator has an iSCSI card that uses IP addresses 192.168.10.125 and 192.168.11.125.
- The target is the Dell Acceleration Appliance for Databases in HA mode. Each node has a card installed and configured for iSCSI. The IP addresses are 192.168.10.8 and 192.168.11.8 for the first node, and 192.168.10.9 and 192.168.11.9 for the second node.
- The Dell Acceleration Appliance for Databases volume of 100 GB is used, with LUNs already created and presented to the Windows initiator.

To set up iSCSI in Windows:

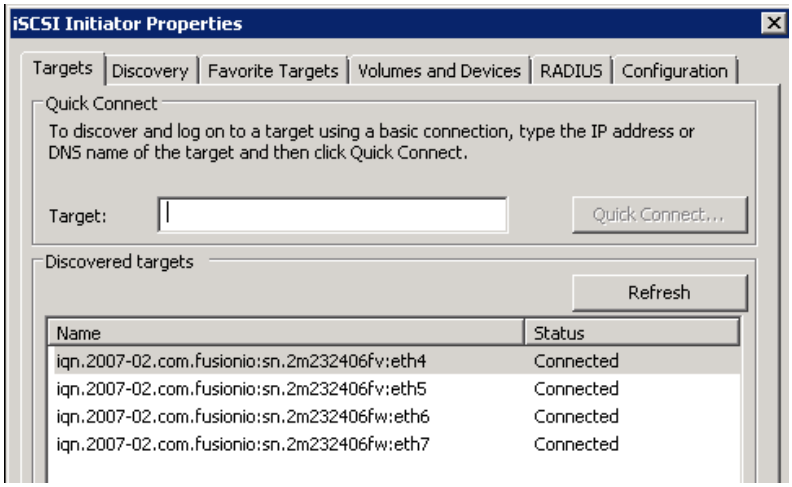
1 Click **Start > Administrator Tools > iSCSI Initiator**.

2 Log in to each target port and click **Quick Connect**.

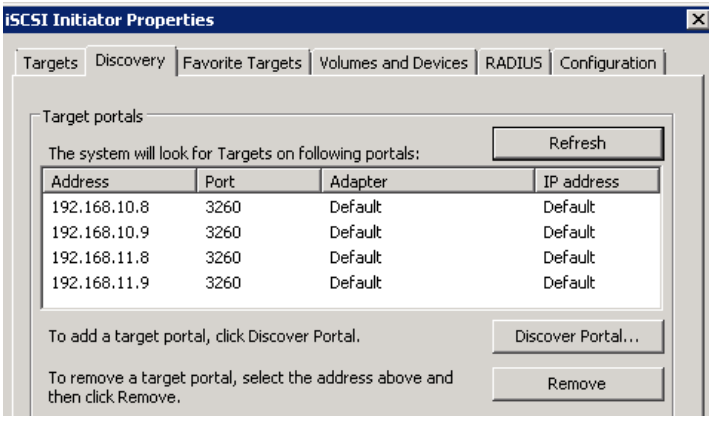


After you successfully log in to all four targets, the discovered targets are displayed.

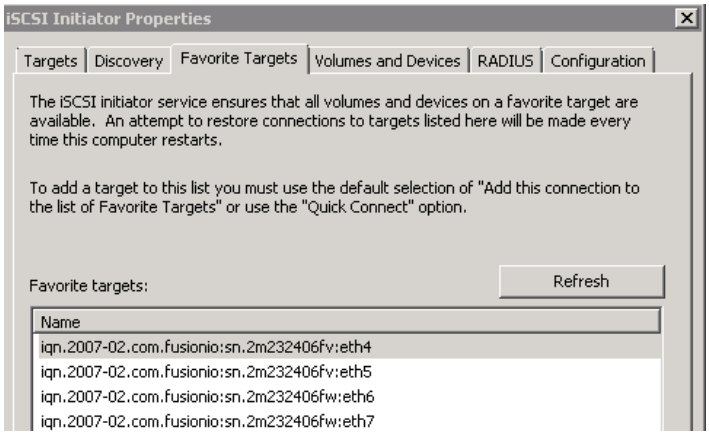
Discovered Targets (Targets Tab)



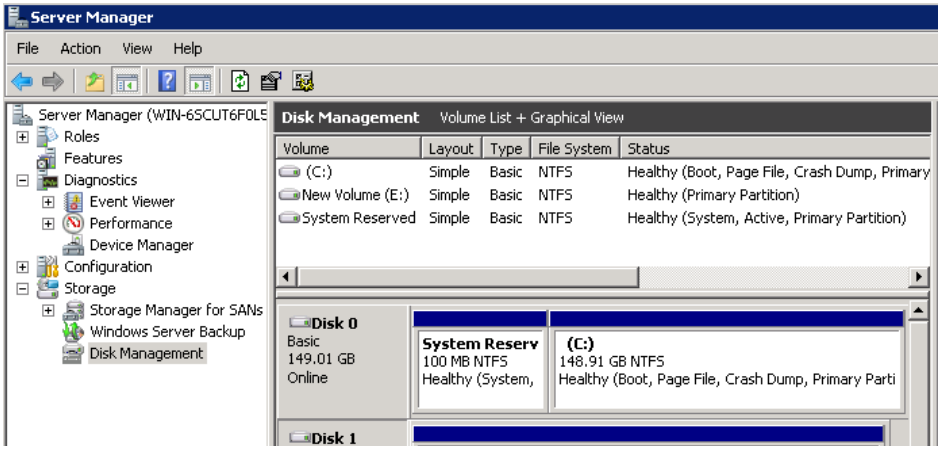
Target Portals (Discovery Tab)



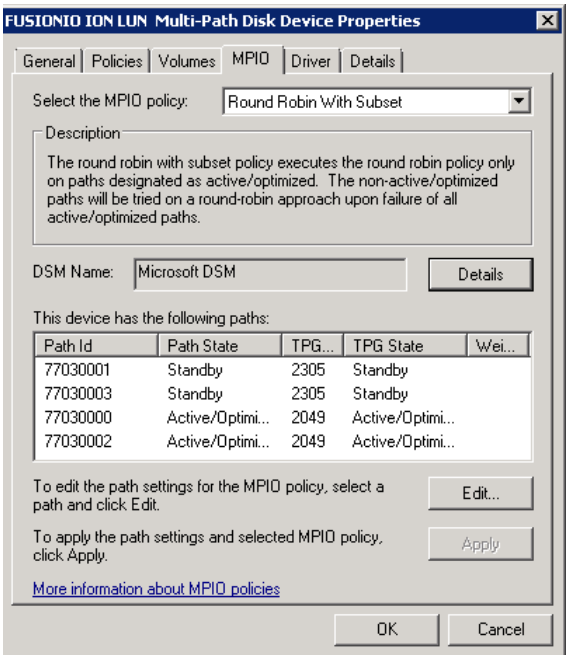
Favorite Targets Tab



3 To view the LUN discovery in Server Manager, open the **Disk Management** volume list. Disk 1 in the example here shows the discovered volume formatted and partitioned as drive E:



4 To see the discovered LUN as an MPIO multipath configuration disk device, with two active and two standby paths, right-click Disk 1 and click **Properties**.



Creating Windows initiators in DAAD

1 From DAAD (Stand-Alone or HA), create an initiator group:

```
inigroup:create W2K8
Id  W2K8
Parent
Initiators  []
UUID  79f7b844-5bbf-11e3-acce-0015178fbc10
```

2 For each target, create an initiator. The examples below are for four iSCSI targets in the Dell Acceleration Appliance for Databases HA:

```
initiator:create -a W2K8 iqn.1991-05.com.sqa.microsoft:win-
ru3kmomq0ov#192.168.10.8 win_1
id  win_1
UUID  iqn.1991-05.com.sqa.microsoft:win-ru3kmomq0ov#192.168.10.8
Protocol  iSCSI
Discovered  false
Initiator Group  79f7b844-5bbf-11e3-acce-0015178fbc10


initiator:create -a W2K8 iqn.1991-05.com.sqa.microsoft:win-
ru3kmomq0ov#192.168.11.8 win_2
id  win_2
UUID  iqn.1991-05.com.sqa.microsoft:win-ru3kmomq0ov#192.168.11.8
Protocol  iSCSI
Discovered  false
Initiator Group  79f7b844-5bbf-11e3-acce-0015178fbc10


initiator:create -a W2K8 iqn.1991-05.com.sqa.microsoft:win-
ru3kmomq0ov#192.168.10.9 win_3
id  win_3
UUID  iqn.1991-05.com.sqa.microsoft:win-ru3kmomq0ov#192.168.10.9
Protocol  iSCSI
Discovered  false
Initiator Group  79f7b844-5bbf-11e3-acce-0015178fbc10


admin@ionr8i48/> initiator:create -a W2K8 iqn.1991-
05.com.sqa.microsoft:win-ru3kmomq0ov#192.168.10.9 win_4
id  win_4
UUID  iqn.1991-05.com.sqa.microsoft:win-ru3kmomq0ov#192.168.11.9
Protocol  iSCSI
Discovered  false
Initiator Group  79f7b844-5bbf-11e3-acce-0015178fbc10
```

Configuring multipathing on Linux

Installing MPIO on Linux

The basic components for Linux multipathing are:

- `dm_multipath` kernel module — Reroutes I/O, supports failover for paths & path groups.
- `mpathconf` utility — Configures and enables device mapper multipathing
- `multipath` command — Lists & configures multipath devices. Start with `/etc/rc.sysinit` or `udev` when a block device is added.
- `multipathd` daemon — Monitors paths; as paths fail and come back, it may initiate path group switches. Provides for interactive changes to multipath devices. This must be restarted for any changes to the `/etc/multipath.conf` file
- `kpartx` command — Creates device mapper devices for the partitions on a device. Use this command for DOS-based partitions with DM-MP. The `kpartx` is provided in its own package, but the device mapper-multipath package depends on it.

For specific installation information about Oracle Linux 6.4 and RHEL 6.4+ see the MPIO installation sections later in this guide.

The basic steps for installing DM Multipath are listed here:

1. Install `device-mapper-multipath` RPM.
2. On OL 6.4 or RHEL, run `mpathconf --enable` to enable multipathing.
3. If necessary, edit the `multipath.conf` configuration file to modify default values and save the updated file.
4. Start the multipath daemon.

Common mpathconf settings

Here are some common multipathing configuration settings:

```
mpathconf --find_multipaths y
mpathconf --with_module y
mpathconf --with_module y
```

Configuring multipathing in HA mode

To configure multipathing on the host node in the Dell Acceleration Appliance for Databases HA environment, use the `multipath.conf` file:

> **NOTE:** These instructions apply to OL 6.4 and RHEL 6.4 and later.

1. Complete the tasks in the first line of the code about copying the file.
2. Follow the uncomment instructions near the end of the file.

3 Restart `multipathd` as indicated on the first line of the `multipath.conf` file:

```
# copy this file to /etc/multipath.conf and restart multipathd

defaults {
   user_friendly_names     yes
   queue_without_daemon    no
}
devices {
   device {
           vendor                  "FUSIONIO"
           features    "3 queue_if_no_path pg_init_retries 50"
           hardware_handler        "1 alua"
           path_grouping_policy    group_by_prio
           path_selector           "queue-length 0"
           failback                immediate
           path_checker            tur
           prio                    alua
      # Uncomment if using FC. Do not use for SRP and iSCSI
      #fast_io_fail_tmo       15
      #dev_loss_tmo           60
   }
}
```

Configuring standalone multipathing with host-based mirroring

To configure multipathing with RAID/LVM/ASM (host-based mirroring) in a standalone Dell Acceleration Appliance for Databases environment, use the `multipath.conf` file:

NOTE: These instructions apply to OL 6.4and RHEL 6.4 and later.

1 Complete the tasks in the first line about copying the file.

2 Follow the uncomment instructions at the end of the file.

3 Restart multipathd as indicated on the first line.

```
# copy this file to /etc/multipath.conf and restart multipathd

defaults {
   user_friendly_names     yes
   queue_without_daemon    no
}
devices {
   device {
           vendor                  "FUSIONIO"
           features                "0"
           hardware_handler        "1 alua"
           path_grouping_policy    group_by_prio
           path_selector           "queue-length 0"
           failback                immediate
           path_checker            tur
           prio                    alua
           no_path_retry           3

           # Uncomment if using FC. Do not use for SRP and iSCSI
           #fast_io_fail_tmo       15
           #dev_loss_tmo           60
   }
}
```

Configuring multipathing for Oracle VM Server

To configure multipathing to run on Oracle VM Server, add a device entry in the `multipath.conf` file. Near the bottom of the `multipath.conf` file, in the `devices` section append the following lines:

```
device {
   vendor                  "FUSIONIO"
   features                "3 queue_if_no_path pg_init_retries 50"
   hardware_handler        "1 alua"
   path_grouping_policy    group_by_prio
   path_selector           "queue-length 0"
   failback                immediate
   path_checker            tur
   prio                    alua

   # Uncomment if using FC. Do not use for SRP and iSCSI
   #fast_io_fail_tmo       15
   #dev_loss_tmo           60
   }\
```

NOTE: Ensure that the closing bracket "}" for the `devices` section follows the closing bracket of the section you just added.

Restarting multipathing

If your multipath configuration changes, you may need to restart multipathing, as in this example:

```
# multipathd -k
multipathd> reconfigure
ok
multipathd> show config
multipathd> exit
```

When using the show config command, search through the output and ensure that it matches what was entered in `/etc/multipath.conf`.

If the restart is not successful, you can try running the following commands:

```
# multipath -F
# service multipathd stop
# service multipathd start
// confirm the new settings are used for existing multipath devices
# multipath –ll
```

(See Creating RHEL initiators for DAAD on page 124 for sample `multipath –ll` output.)

Setting the node session timeout for iSCSI

If you are using the iSCSI protocol, you need to set the `node.session.timeo.replacement_timeout` value. This should be done for both HA and standalone systems.

> **NOTE:** This should not be done while applications are using the iSCSI or dm-multipath devices.

The `node.session.timeo.replacement_timeout` setting controls the time a path is tried before it is failed. When using dm-multipath, a value of 15 seconds is safe. If your applications require faster failovers, you may need to set the value lower, such as to 5 seconds.

To set this value so it can be used for both current and new sessions:

1. Edit the `/etc/iscsid.conf` file and set an appropriate value for `node.session.timeo.replacement_timeout`.
2. Log out of current sessions by running `iscsiadm -m session –u`.
3. Set the replacement timeout for currently discovered portals, as follows:

   ```
   iscsiadm -m node -o update -n node.session.timeo.replacement_timeout -v 15
   ```
4. Log in to the targets again:

   ```
   iscsiadm -m node -T target -p IP -l
   ```

Linux SCSI initiator notes

When using Linux SCSI initiators with the Dell Acceleration Appliance for Databases consider the following:

- The SCSI H:C:I:L address assigned by a Linux initiator system consists of four components: Host, Channel, Target ID, and LUN. Of these components, only the LUN number is assigned by the Dell Acceleration Appliance for Databases system. The other three components are assigned by the initiator system.
- An HBA reset (caused by the host being reset or the `sg_reset` command being run) can cause the first three components of an H:C:I:L address to change.
- An HBA reset can cause a new `/dev/sd*` device node to be assigned to a path.
- A LUN rescan initiated by the rescan-scsi-bus.sh script changes neither the H:C:I:L SCSI address nor the `/dev/sd*` device node assigned to a path.

Setting up Linux iSCSI initiators

For the example used in this configuration, the following items must be noted:

- The Red Hat Linux initiator has an iSCSI card that uses IP addresses 192.168.10.119 and 192.168.11.119.
- The target is the Dell Acceleration Appliance for Databases in HA mode. Each node has a card installed and configured for iSCSI. The IP addresses are 192.168.10.8 and 192.168.11.8 for the first node, and 192.168.10.9 and 192.168.11.9 for the second node.
- The Dell Acceleration Appliance for Databases volume of 100GB is used, with LUNs already created and presented to the Linux initiator.

NOTE: Each initiator iSCSI port should be in a separate subnet that is the same as that of the target iSCSI port.

Discovering DAAD volume

By running the targets command in the CLI, you can view the targets as seen from the HA side of the Dell Acceleration Appliance for Databases system.

```
targets -dt --cluster
ionr8i48:
Id  |UUID                    |Protocol|Enabled|State
-----------------------------------------------------
eth4|iqn.2007-02.com.fusioni\|iSCSI   |true   |Online |o:sn.2m232406fv:eth4 |||
eth5|iqn.2007-02.com.fusioni\|iSCSI   |true   |Online |o:sn.2m232406fv:eth5 |||
ionr8i49:
Id  |UUID                    |Protocol|Enabled|State
-----------------------------------------------------
eth6|iqn.2007-02.com.fusioni\|iSCSI   |true   |Online |o:sn.2m232406fw:eth6 |||
eth7|iqn.2007-02.com.fusioni\|iSCSI   |true   |Online |o:sn.2m232406fw:eth7 |||
```

Discovering DAAD target portals from RHEL initiators

To see all the target portals that are discovered, you can run the `iscsiadm` command at the CLI:

```
# iscsiadm -m discovery -t st -p 192.168.10.8
192.168.10.8:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fv:eth4
# iscsiadm -m discovery -t st -p 192.168.10.9
1    92.168.10.9:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fw:eth6
# iscsiadm -m discovery -t st -p 192.168.11.8
192.168.11.8:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fv:eth5
# iscsiadm -m discovery -t st -p 192.168.11.9
192.168.11.9:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fw:eth7
```

```
# iscsiadm -m node -P1
Target: iqn.2007-02.com.fusionio:sn.2m232406fw:eth6
        Portal: 192.168.10.9:3260,1
                Iface Name: default
Target: iqn.2007-02.com.fusionio:sn.2m232406fw:eth7
        Portal: 192.168.11.9:3260,1
                Iface Name: default
Target: iqn.2007-02.com.fusionio:sn.2m232406fv:eth5
        Portal: 192.168.11.8:3260,1
                Iface Name: default
Target: iqn.2007-02.com.fusionio:sn.2m232406fv:eth4
        Portal: 192.168.10.8:3260,1
                Iface Name: default
```

The following example logs in to each target portal:

```
# iscsiadm -m node -T iqn.2007-02.com.fusionio:sn.2m232406fw:eth6 -p
192.168.10.9  -l
# iscsiadm -m node -T iqn.2007-02.com.fusionio:sn.2m232406fw:eth7 -p
192.168.11.9  -l
# iscsiadm -m node -T iqn.2007-02.com.fusionio:sn.2m232406fv:eth5 -p
192.168.11.8  -l
# iscsiadm -m node -T iqn.2007-02.com.fusionio:sn.2m232406fv:eth4 -p
192.168.10.8  -l
# iscsiadm -m session
tcp: [5] 192.168.10.9:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fw:eth6
tcp: [6] 192.168.11.9:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fw:eth7
tcp: [7] 192.168.11.8:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fv:eth5
tcp: [8] 192.168.10.8:3260,1 iqn.2007-02.com.fusionio:sn.2m232406fv:eth4
```

Creating RHEL initiators for DAAD

1. At the CLI, run the following to create an initiator group:

   ```
   inigroup:create RHEL
   Id  RHEL
   Parent
   Initiators  []
   UUID  29f1007e-1751-11e3-8482-0015178fbc10
   ```

2. For each target, create an initiator. The examples here are for four iSCSI targets from the Dell Acceleration Appliance for Databases HA:

   ```
   initiator:create -a RHEL iqn.1994-05.com.redhat:595862111da4#192.168.10.8
   rhel_1
   id  rhel_1
   UUID  iqn.1994-05.com.redhat:595862111da4#192.168.10.8
   ```

```
Protocol  iSCSI
Discovered  false
Initiator Group  29f1007e-1751-11e3-8482-0015178fbc10

initiator:create -a RHEL iqn.1994-05.com.redhat:595862111da4#192.168.11.8
rhel_2
id  rhel_2
UUID  iqn.1994-05.com.redhat:595862111da4#192.168.11.8
Protocol  iSCSI
Discovered  false
Initiator Group  29f1007e-1751-11e3-8482-0015178fbc10
initiator:create -a RHEL iqn.1994-05.com.redhat:595862111da4#192.168.10.9
rhel_3
id  rhel_3
UUID  iqn.1994-05.com.redhat:595862111da4#192.168.10.9
Protocol  iSCSI
Discovered  false
Initiator Group  29f1007e-1751-11e3-8482-0015178fbc10

initiator:create -a RHEL iqn.1994-05.com.redhat:595862111da4#192.168.11.9
rhel_4
id  rhel_4
UUID  iqn.1994-05.com.redhat:595862111da4#192.168.11.9
Protocol  iSCSI
Discovered  false
Initiator Group  29f1007e-1751-11e3-8482-0015178fbc10
```

The DAAD LUN is now discovered with a MPIO multi-path configuration disk device, with two active and two enabled paths:

```
# multipath -ll
mpathag (26538623635336336) dm-3 FUSIONIO,ION LUN
size=93G features='3 queue_if_no_path pg_init_retries 50' hwhandler='1 alua'
wp=rw
|-+- policy='queue-length 0' prio=130 status=active
| |- 11:0:0:0 sdd 8:48 active ready running
| `- 12:0:0:0 sde 8:64 active ready running
`-+- policy='queue-length 0' prio=1 status=enabled
  |- 9:0:0:0  sdb 8:16 active ready running
  `- 10:0:0:0 sdc 8:32 active ready running
```

Utilizing full bandwidth of a 10 Gbs Ethernet link

To utilize the full bandwidth of a 10 Gbs Ethernet link, the iSCSI protocol requires you to set up multiple sessions. Follow these steps to set up multiple sessions:

1. Configure `/etc/multipath.conf` so that `multipathd` controls at least two iSCSI LUNs.
2. Start `multipathd`.
3. Change the parameter `node.session.nr_sessions` in `/etc/iscsid.conf` from 1 to **4**.
4. Trigger an iSCSI logout and login.
5. Verify the `multipath -ll` output and look up which `/dev/dm-<n>` node corresponds with which iSCSI LUN.
6. Use the `/dev/dm-<n>` device nodes for communication over iSCSI with the ION system.

Mounting a file system on a LUN during boot

In order to mount a file system on top of the Dell Acceleration Appliance for Databases LUN during Linux system boot time, complete the following tasks:

1. Determine the SCSI ID of the Dell Acceleration Appliance for Databases LUN. This is the number shown between parentheses in the output of multipath -l that starts with the digit 2. For example:

   ```
   # multipath -l
   26164613638323832 dm-0 FUSIONIO,ION LUN
   size=34G features='3 queue_if_no_path pg_init_retries 50' hwhandler='0' wp=
   rw
   `-+- policy='queue-length 0' prio=0 status=active
   `- 17:0:0:0 sdc 8:32 active undef running
   23630313437393135 dm-1 FUSIONIO,ION LUN
   size=31G features='3 queue_if_no_path pg_init_retries 50' hwhandler='0' wp=
   rw
   `-+- policy='queue-length 0' prio=0 status=active
   `- 17:0:0:1 sdd 8:48 active undef running
   ```

2. Add an entry in `/etc/fstab` for that LUN and use the `noauto` option so the entry is skipped during boot time. For example:

   ```
   /dev/disk/by-id/dm-uuid-mpath-23165353937623137 /mnt auto noauto 0 0
   ```

3. Add a script in `/etc/init.d` that waits until multipathd has detected the LUN and then mounts the filesystem. See Script for Mounting the file system on page 127.

4. Make the system run during startup and shutdown. For example:

   ```
   chkconfig mount-ion-filesystems on
   ```

Script for Mounting the file system

```
#!/bin/sh
### BEGIN INIT INFO
# Provides: mount-ion-filesystems
# Required-Start: multipathd
# Required-Stop: multipathd
# Default-Start: 2 3 4 5
# Default-Stop: 0 1 6
# Description: Mount ION File Systems
### END INIT INFO
### BEGIN CHKCONFIG INFO
# chkconfig: 2345 13 87
# description: Mount ION File Systems
### END CHKCONFIG INFO

timeout=120

usage() {
echo
```

```
echo "Usage: `basename $0` {start|stop|restart|status}"
echo
return 2
}

is_ion_lun() {
[ "${1#/dev/disk/by-id/dm-uuid-mpath-}" != "$1" ]
}

luns_exist() {
grep -v '^#' /etc/fstab |
while read dev mountpoint vfstype options freq passno; do
is_ion_lun "$dev" && [ ! -e "$dev" ] && return 1
done
return 0
}

mount_luns() {
grep -v '^#' /etc/fstab |
while read dev mountpoint vfstype options freq passno; do
is_ion_lun "$dev" && fsck -n "$dev" && mount "$dev"
done
}

unmount_luns() {
grep -v '^#' /etc/fstab |
while read dev mountpoint vfstype options freq passno; do
is_ion_lun "$dev" && umount "$dev"
done
}

start() {
for i in $(seq $timeout); do
luns_exist && break
sleep 1
done
mount_luns
}

stop() {
unmount_luns
}

case "$1" in
start) start;;
stop) stop;;
restart) stop; start;;
status) ;;
*) usage;;
esac
```

Configuring DAAD with ESXi initiators

The following sections describe how to configure ESXi initiators to work with DAAD.

Configuration procedure: Mixed sub-storage

This section explains how to set up ESXi 5.1 and ESXi 5.5 initiators to work with the Dell Acceleration Appliance for Databases, if your system has other storage subsystems that do not allow MPIO to run in Round-Robin mode.

> **NOTE:** If your only storage system is Dell Acceleration Appliance for Databases, skip to *Configuration Procedure: ION Accelerator Only*.

To change the MPIO policy for each ESX LUN (DataStore) in the system:

1. Select the ESXi host you want to modify and click the **Manage** tab at the upper-right corrner.
2. Click **Storage > Storage Devices** and click the LUN you want to modify.
3. In the Properties pane, click **Edit Multipathing** to obtain the paths that are in use.
4. Under the Policy section, select **Round Robin** from the drop-down menu.
5. Click **Change** to apply the change in path policy.
6. Click **OK** to close the dialog box
7. Repeat steps 2-6 for each LUN.

Complete the configuration by performing the steps found in Completing the configuration on page 130.

Configuration procedure: DAAD only

This section describes the tasks to set up ESXi 5.1 and ESXi 5.5 initiators to work with the Dell Acceleration Appliance for Databases if it is the only storage subsystem.

1. Before presenting DAAD volumes to the ESXi initiator, set a storage global policy on each ESXi initiator in-use by running the following command. This ensures that the default path policy for ALUA based storage is set to Round-Robin.

   ```
   ~ # esxcli storage nmp satp set --default-psp=VMW_PSP_RR
   --satp=VMW_SATP_ALUA
   ```

   The default PSP for VMW_SATP_ALUA is now VMW_PSP_RR.
2. Restart the ESXI initiators for the policy to take effect.

3 Ensure the settings by running the following command:

```
~ # esxcli storage nmp satp list

Name            Default PSP  Description
--------------  -----------  ----------------------------------------
VMW_SATP_ALUA    VMW_PSP_RR  Supports non-specific arrays that use the ALUA
protocol
```

Completing the configuration

Following the recommendations in the *Setting Up Storage* section in the *Dell Acceleration Appliance for Databases GUI Guide*, create the storage profile, volumes, and initiator groups. The initiators will come from ESXi hosts.

1 Present the volumes to the initiator group.

> **NOTE:** Dell recommends that all the initiators from each of the ESXi cluster nodes are part of one Initiator Group. Also, users can dynamically add initiators to an already existing initiator group. This results in the newly added initiators gaining access to the volume(s).

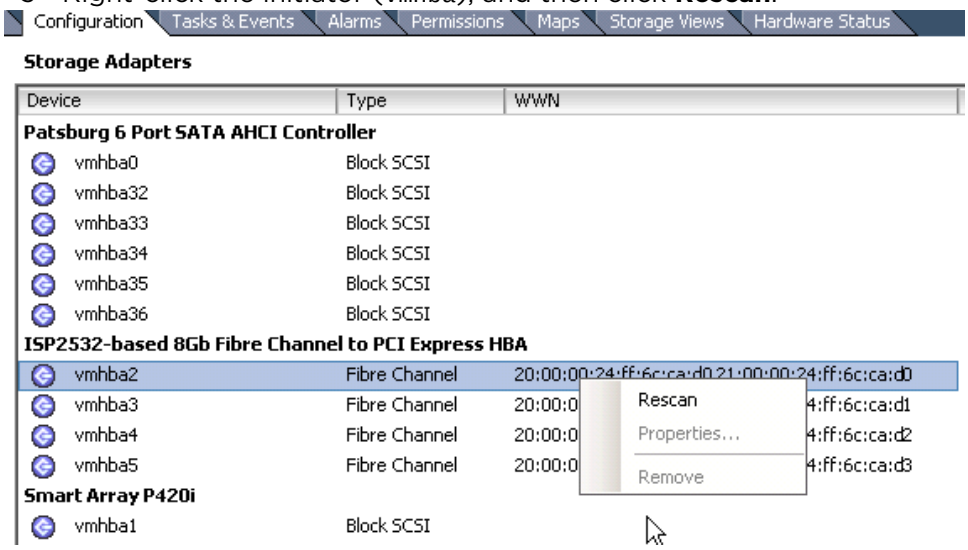
After the configuration is complete on the Dell Acceleration Appliance for Databases appliance and the volumes are presented to the ESXi nodes, you need to run a rescan to discover the FUSIONIO volumes.

2 In the vSphere GUI, log in to an ESXi host.

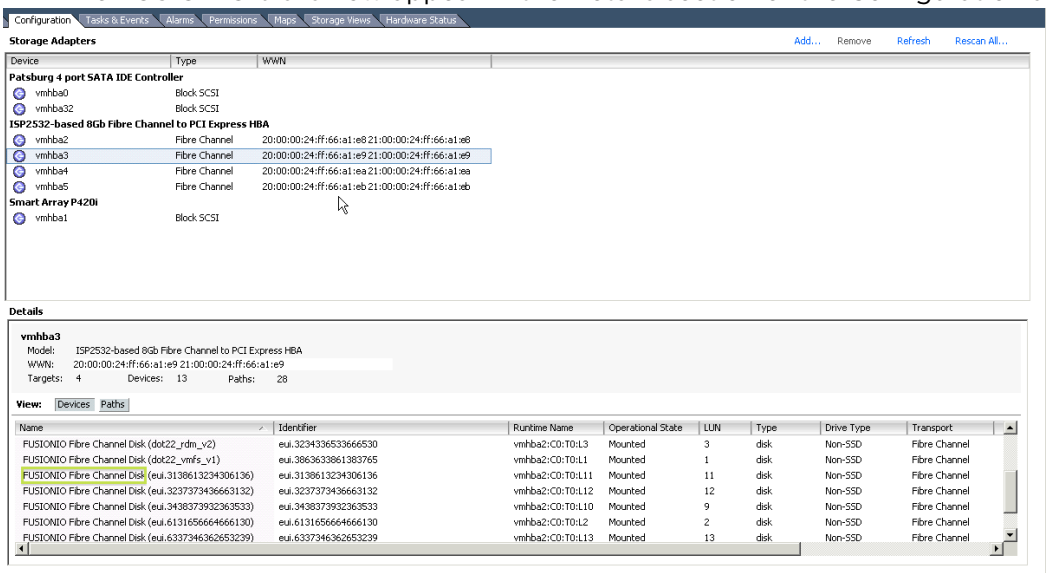
3 Click **ESXi Host > Configuration > Storage Adapters**.

4 Select the initiator .

5 Right-click the initiator (`vmhba`), and then click **Rescan**.

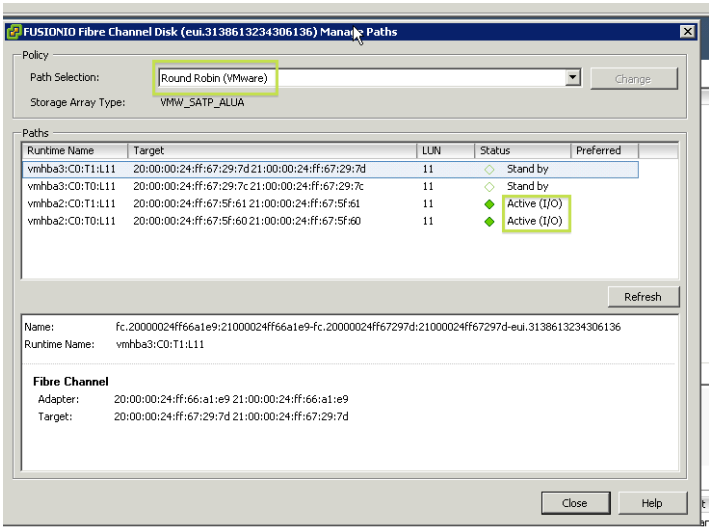


The FUSIONIO disks now appear in the Details section of the Configuration screen:



6 Repeat this procedure (steps 5–7) for all HBAs.

7 When the FUSIONIO volumes are discovered, ensure that the path policy is set to **Round Robin**.



8 Repeat this verification procedure for every ESXi host in the configuration.

To obtain the EUI (Extended Unique Identifier), run the following commands:

```
# esxcli storage core device list
# esxcli storage nmp satp generic deviceconfig set -d eui.3863633861383765 -c=
disable_alua_followover
```

At this point, the configuration is complete, and the volumes are ready to be used as VMFS datastores or raw device mappings.

> **NOTE:** To view the EUI in the Dell Acceleration Appliance for Databases, run `volumes -dt --display-flavor vmware`

Issues with automatic failback

When all initiator links or all target links fail, workload failover occurs. If the initiator links or the target links are then restored, the Round Robin policy does not perform an automatic failback. Because of the failure of the automatic failback, the system load is unbalanced.

The following explanation about failback policy is taken from the VMware ESX manual (https://pubs.vmware.com/vsphere-50/index.jsp?topic=%2Fcom.vmware,vcli.examples.doc_50%2Fcli_manage_storage.6.5.html):

From the VMware ESX manual (vSphere 5 Command Line Documentation > vSphere Command-Line Interface Documentation > vSphere Command-Line Interface Concepts and Examples > Managing Storage > Managing Path Policies – http://pubs.vmware.com/vsphere-50/index.jsp?topic=%2Fcom.vmware.vcli.examples.doc_50%2Fcli_manage_storage.6.5.html):

Path Policy Effects

Policy	Active/Active Array	Active/Passive Array
Most Recently Used	Administrator action is required to fail back after path failure.	Administrator action is required to fail back after path failure.
Fixed	VMkernel resumes using the preferred path when connectivity is restored.	VMkernel attempts to resume by using the preferred path. This action can cause path thrashing or failure when another SP now owns the LUN.
Round Robin	No fail back.	Next path in round robin scheduling is selected.

To manually restore the configuration of volumes back to its original path, the Manual failback procedure on page 132 must be followed as described here to rebalance the load.

Manual failback procedure

> **NOTE:** Before proceeding with manual failback, ensure that all paths have been re-established.

1. Determine the LUNs that do not have the same Current and Preferred Owner by running the CLI command:

   `luns –dt —-cluster`

   The Owner/Pref column must display the same node for each LUN. The example here shows different results in this column.

   ```
   admin@/> luns -dt
   ```

```
Id                  |Block|Target/Device          |Num|Conn                 |Ini Group|Initiators             |Owner/Pref|
-------------------------------------------------------------------------------------------------------------------------------
4f8b9efa-d856-11e2-87\|512b |21:00:00:24:ff:67:29:7c |0  |21:00:00:24:ff:66:a1:e9 |dot22_esx|21:00:00:24:ff:66:a1:e9 |IONr1i44  |
a6-90b11c06e928-LUN0  |     |dot22_rdm_v1           |   |                     |         |21:00:00:24:ff:66:a1:e8 |IONr1i45  |
```

2 Determine the volume name for that particular LUN, and determine the USN for the volume by running the CLI `volumes -dt` command.

```
Id          |UUID           |Capacity  |Device     |T10 Id              |USN     |Pool      |Status    |Read   |Written|Nodes   |
-------------------------------------------------------------------------------------------------------------------------------
dot22_rdm_v1 |bR1eXB-sZMx-zmY\|200.00 GB  |/dev/drbd0 |cd765853-dot22_rdm_v1 |cd765853|RAID0_POOL|Connected |0.00 GB|0.00 GB|IONr1i44|
            |5-fGMH-4rTa-eYz\|          |           |                    |        |          |          |       |       |IONr1i45|
            |3-03cYL1       |          |           |                    |        |          |          |       |       |        |
```

3 List the EUIs for FUSIONIO volumes on ESXi host:

```
~ # esxcli storage nmp device list
eui.6364373635383533
   Device Display Name: FUSIONIO Fibre Channel Disk (dot22_rdm_v1)
   Storage Array Type: VMW_SATP_ALUA
   Storage Array Type Device Config: {implicit_support=on;explicit_support=
on; explicit_allow=on;alua_followover=on;{TPG_id=257,TPG_state=STBY}{TPG_id=
513,TPG_state=AO}}
   Path Selection Policy: VMW_PSP_RR
   Path Selection Policy Device Config: {policy=rr,iops=1000,bytes=
10485760,useANO=0;lastPathIndex=2: NumIOsPending=13,numBytesPending=851968}
   Path Selection Policy Device Custom Config:
   Working Paths: vmhba3:C0:T0:L0, vmhba3:C0:T1:L0
   Is Local SAS Device: false
   Is Boot USB Device: false
```

4 Map the USN of the volume to the EUI in the VSphere GUI. You can use the CLI command `volumes -dt --display-flavor vmware` to do this. For example:

```
admin@IONr1n45/>  volumes -dt --display-flavor vmware
Id      |Capacity   |T10/USN/EUI         |Pool     |Status    |Read/Write|Nodes
-------------------------------------------------------------------------------
v12      |300.00 GB  |T10|ab0c4aa3-v12      |raid0_pool|Connected |R|0.00 GB
|IONr1n44
         |           |USN|ab0c4aa3          |          |          |W|0.00 GB
|IONr1n45 !!
        |          |EUI|6162306334616133  |         |         |         |
admin@IONr1n45/> luns -dt
Id                   |Status|Device|Target                     |Ini
Group|Initiators(*=Active)     |Owner/Pref
-------------------------------------------------------------------------------
```

```
4d929f3c-e5dd-11e3-b7\|Active|oel_vm|21:00:00:24:ff:67:29:7c |esx_clus
|21:00:00:24:ff:66:a1:e8  |IONr1n45
98-90b11c06e928-LUN0  |512b  |#0    |                        |
|21:00:00:24:ff:6c:ca:d3  |
                        |      |      |                        |
|21:00:00:24:ff:6c:ca:d2* |
                   |      |      |                        |
|21:00:00:24:ff:66:a1:ea* |
```

5 On the vSphere GUI, navigate to the ESXi node and then click **Configuration > Storage Adapters**.

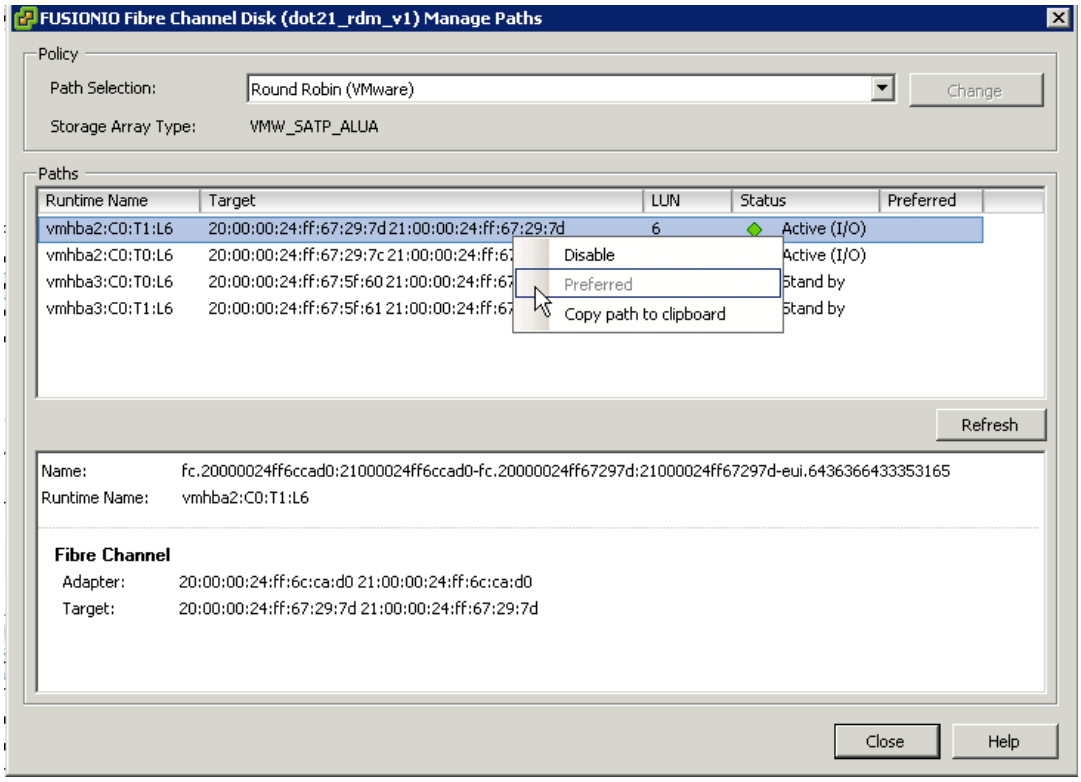
6 Select the desired `vmhba` number.

7 In the View section, click **Devices** to view the FUSIONIO volumes.

8 Select the volume and right-click **Manage Paths**.

9 Click **Disable Current Active Path**.

10 Re-enable the disabled path.



Disabling VAAI functionality in ESXi

VAAI (vStorage APIs for Array Integration) should be disabled when running the Dell Acceleration Appliance for Databases with ESXi. This avoids issues with responses to VAAI primitives.

To disable VAAI, complete the tasks outlined in this article:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1033665

Configuring ESXi initiators with iSCSI

HA configuration

1. Set the default PSP for ALUA SATP to `VMW_PSP_RR` (round robin):

   ```
   esxcli storage nmp satp set --default-psp VMW_PSP_RR --satp VMW_SATP_ALUA
   ```

2. To apply this setting to LUNs that have already been discovered, restart the system.

Setting up an ESXi initiator

For this example, the following assumptions are made:

- The ESXi 5.1 initiator has a dual-port, 10GbE Emulex NC552SFP card installed and configured for iSCSI, with IP addresses of 192.168.10.100 and 192.168.11.100.
- The target is the Dell Acceleration Appliance for Databases HA system.
- Each HA node has a dual-port 10GbE Intel 82599EB card installed and configured for iSCSI at 192.168.10.45, 192.168.11.45 first node, and 192.168.10.46 and 192.168.11.46 second node.
- The Dell Acceleration Appliance for Databases volume of 50GB with LUNs has been created and presented to the ESXi 5.1 initiator.

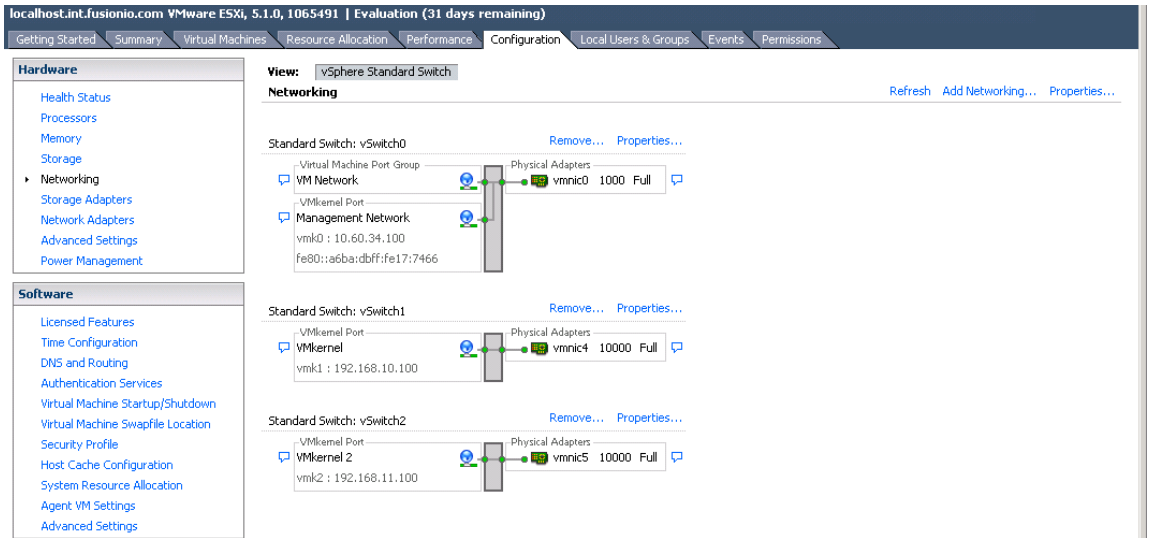
The following iSCSI information is reported from the ESXi host, after successful installation of the proper driver for the iSCSI adapter:

```
~ # esxcli network nic list |grep Emulex
vmnic4  0000:007:00.0  be2net  Up    10000  Full    44:1e:a1:17:49:18  1500
Emulex Corporation NC552SFP 2-port 10Gb Server Adapter
vmnic5  0000:007:00.1  be2net  Up    10000  Full    44:1e:a1:17:49:1c  1500
Emulex Corporation NC552SFP 2-port 10Gb Server Adapter
```

To set up the initiator:

1. Open the vSphere Client/VMware Infrastructure (VI) Client or vCenter Server.
2. Ensure that you have a separate vSwitch for iSCSI. (All iSCSI entries can be grouped into one vSwitch.)

3 Add VMKernel and Management Console ports and identify them with appropriate IP addresses.



4 Ensure that the iSCSI targets are reachable from the ESX host by using `ping` or `vmkping`:

```
~ # ping 192.168.10.45
PING 192.168.10.45 (192.168.10.45): 56 data bytes
64 bytes from 192.168.10.45: icmp_seq=0 ttl=64 time=0.254 ms
64 bytes from 192.168.10.45: icmp_seq=1 ttl=64 time=0.133 ms

~ # ping 192.168.10.46
PING 192.168.10.46 (192.168.10.46): 56 data bytes
64 bytes from 192.168.10.46: icmp_seq=0 ttl=64 time=0.271 ms
64 bytes from 192.168.10.46: icmp_seq=1 ttl=64 time=0.138 ms
```

Assuming that the Dell Acceleration Appliance for Databases has been set up with initiator groups, initiators, volumes, and LUNs available to the ESX initiator, you can log in to each target portal.

5 Log in to vCenter Server.

6 Click the ESX host and click **Configuration**.

7 Click **Storage Adapters**.

8 Click the iSCSI VMHBA to be modified.

9 Click **Properties**.

10 Click **Advanced**.

11 On the dialog box that is displayed, disable **DelayedAck** and set **Login Timeout** to 30 seconds.

12 On the Storage Adapters section, click the iSCSI Software Adapter.

13 On the Details pane, click **Properties**.

14 On the General tab, click **Configure** and click **Enabled**.

15 Click **OK**. The Status is enabled, and the iSCSI name has a valid IQN.

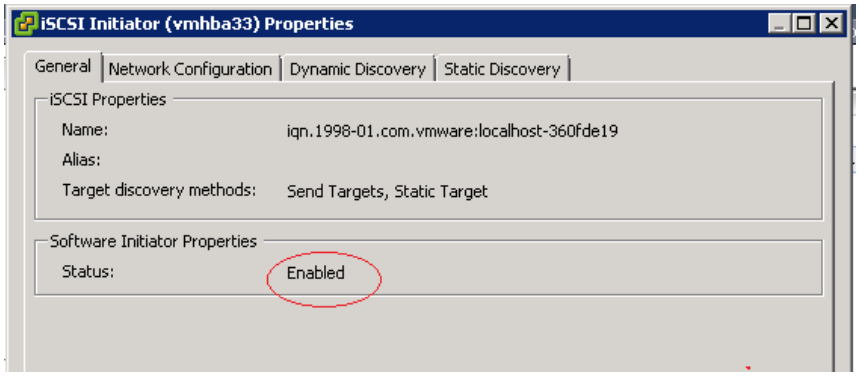
You can now configure the iSCSI array to use the IQN you have just created. Generally, you also need to assign storage to the ESX host from the array:

1 Click the **Dynamic Discovery** tab and click **Add**.

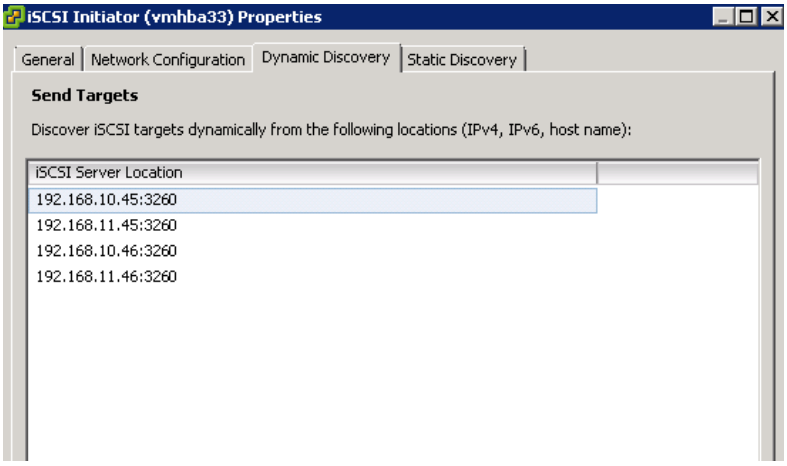
2 In the Add Send Targets Server section, add the iSCSI array IP addresses to the iSCSI Server field and click **OK**.

3 After the iSCSI Server is added, click **Close**.

The General tab for the iSCSI Initiator shows the iSCSI properties and initiator status:



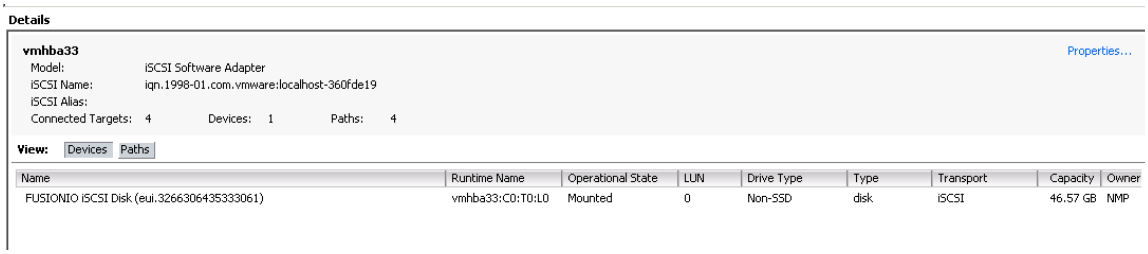
The Dynamic Discovery tab shows locations from where the iSCSI targets can be discovered.



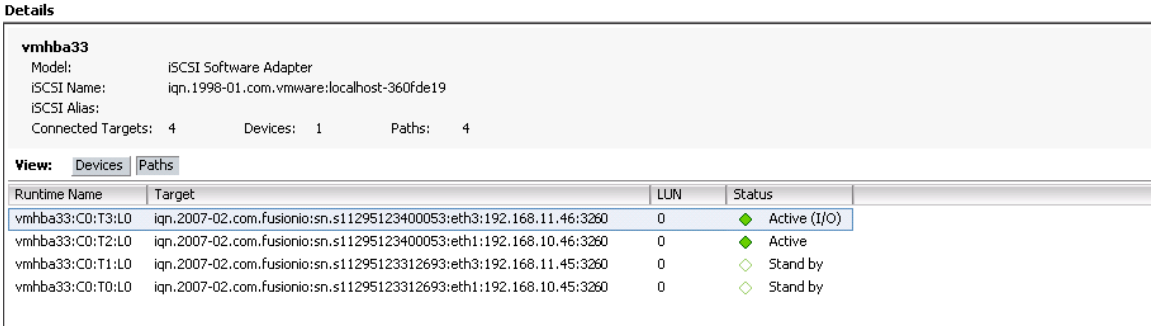
4 Rescan the ESX host storage in the Storage Adapters section.

5 After the rescan completes, click the iSCSI initiator to see information about the SAN array, and one or more targets.

The disk device is discovered:



Multipath information about the two Active and two Standby paths is also displayed:



Setting up InfiniBand/SRP initiators

NOTE: Although this section is primarily intended for HA mode, you can also use these instructions to configure standalone systems.

Installing the multipath daemon

In order to set up InfiniBand/SRP initiators, you must install and configure the multipath daemon. For assistance, contact support at dell.com/support/home.

Installing the InfiniBand/SRP software stack

After having installed and configured the multipath daemon, you need to install the InfiniBand/SRP software packages. The names of some of these packages vary between Linux distributions. Examples are provided for RHEL 6 and OL 6

Table 9-1. Infinband packages that need to be installed

Package	RHEL 6 / OL 6
InfiniBand core	`rdma`
Mellanox ConnectX user space driver	`libmlx4`
Mellanox firmware tool	`mstflint`
InfiniBand Subnet Manager	`opensm`

InfiniBand diagnostic tools	`infiniband-diags`
	`ibutils`
	`perftest`

SRP tools	`srptools`
PCI tools	`pciutils`
SCSI tools	`lsscsi` `sg3_utils`

The InfiniBand/SRP configuration files can be found in the following locations:

Table 9-2. Locations of InfiniBand/SRP configuration files

	RHEL6 / OL 6
InfiniBand/SRP core	`/etc/rdma/rdma.conf`
opensm	`/etc/rdma/opensm.conf`
SRP daemon	`/etc/srp_daemon.conf`
SRP kernel module	`/etc/modprobe.d/ib_srp.conf`

Configuring InfiniBand/SRP

Here are the configuration changes you need to make for InfiniBand/SRP:

1. Set `SRP_LOAD=yes` and `MLX_LOAD=yes` in the InfiniBand/SRP core configuration file `/etc/rdma/rdma.conf`.
2. When using an unmanaged InfiniBand/SRP switch, or when not using an InfiniBand/SRP switch at all, configure the GUID of the port over which opensm should run in the `opensm` configuration file. (An unmanaged switch is a switch without subnet management functionality.)

   NOTE: A subnet manager in an InfiniBand switch and opensm provide the same functionality: InfiniBand subnet management. Running a single subnet manager is sufficient. If multiple subnet managers are active in the same subnet one of them will be the master and the others slaves. If the master fails one of the slaves will become the new master. In other words, enabling multiple subnet managers provides redundancy. However, there is also a risk associated with enabling multiple subnet managers, namely that an incorrect configuration can cause one subnet manager to reassign all the LIDs. Reassigning LIDs is a disruptive operation in an IB network because it causes ongoing communication to fail until all IB connections have been reestablished by the software that uses the IB network.

3. If the SRP daemon must not automatically log in to all SRP targets, modify the SRP daemon configuration file accordingly.
4. Performance can be improved by increasing the maximum SRP scatter/gather list size. To do this, create the file `/etc/modprobe.d/ib_srp.conf`. This configuration will go into effect the next time the SRP initiator is loaded.

   ```
   echo options ib_srp cmd_sg_entries=255 > /etc/modprobe.d/ib_srp.conf
   modprobe -r ib_srp
   modprobe ib_srp
   cat /sys/module/ib_srp/parameters/cmd_sg_entries
   ```

5. Ensure that the InfiniBand/SRP HCAs are visible on the PCI bus:

   ```
   # lspci | grep Mellanox
   02:00.0 Network controller: Mellanox Technologies MT27500 Family [ConnectX-
   3]
   ```

6. Ensure that the number of PCIe lanes assigned to the HCA will allow it to operate at full speed. To do this, compare the `LnkCap` and `LnkSta` parameters in the `lspci` output:

   ```
   # lspci -vv | sed -rne '/Mellanox/,/^$/ s/^[^[:blank:]]|LnkSta:|LnkCap:/&/p'
   02:00.0 Network controller: Mellanox Technologies MT27500 Family [ConnectX-
   3]
   LnkCap: Port #8, Speed 8GT/s, Width x8, ASPM L0s, Latency L0 unlimited, L1
   unlimited
   LnkSta: Speed 5GT/s, Width x8, TrErr- Train- SlotClk+ DLActive- BWMgmt-
   ABWMgmt-
   ```

7 Ensure that the InfiniBand/SRP HCA firmware is up to date:

```
# lspci | grep Mellanox | while read a b; do echo ==== $a; mstflint -d $a
query; done
==== 0000:06:00.0
Image type:      ConnectX
FW Version:      2.11.500
Rom Info:        type=PXE  version=3.4.142 devid=4099 proto=VPI
Device ID:       4099
Description:     Node             Port1            Port2            Sys image
GUIDs:           0002c90300a05de0 0002c90300a05de1 0002c90300a05de2
0002c90300a05de3
MACs:                                 0002c9a05de0     0002c9a05de1
Board ID:        (MT_1090120019)
VSD:
PSID:            MT_1090120019
==== 0000:07:00.0
Image type:      ConnectX
FW Version:      2.11.500
Rom Info:        type=PXE  version=3.4.142 devid=4099 proto=VPI
Device ID:       4099
Description:     Node             Port1            Port2            Sys image
GUIDs:           0002c90300a04250 0002c90300a04251 0002c90300a04252
0002c90300a04253
MACs:                                 0002c9a04250     0002c9a04251
Board ID:        (MT_1090120019)
VSD:
PSID:            MT_1090120019
```

8 Ensure that the InfiniBand/SRP HCAs are recognized by their kernel driver. For example:

```
# (cd /etc/init.d; if [ -e ./rdma ]; then ./rdma start; else ./openibd start;
fi)
# ls /sys/class/infiniband
mlx4_0  mlx4_1
```

9 If necessary, start `opensm` and wait until it has configured the fabric. For a small fabric, configuration finishes less than twenty seconds after `opensm` starts.

```
(cd /etc/init.d; if [ -e ./opensm ]; then ./opensm start; else ./opensmd
start; fi)
```

10 Check which nodes can communicate with each other over the fabric, as follows:

```
# iblinkinfo
CA: ionr7i43 mlx4_2:
      0x0002c90300a4f4d1  6  1[  ] ==( 4X  10.0 Gbps Active/  LinkUp)==>  1
35[  ] "MF0;switch-62f0f2:SX60XX/U1" ( Could be FDR10)
      0x0002c90300a4f4d2  7  2[  ] ==( 4X  10.0 Gbps Active/  LinkUp)==>  1
36[  ] "MF0;switch-62f0f2:SX60XX/U1" ( Could be FDR10)
[ ... ]
```

11 Check the fabric status:

```
# ibdiagnet -r
[ ... ]
-I- Stages Status Report:
    STAGE                                    Errors Warnings
    Bad GUIDs/LIDs Check                     0      0
```

```
    Link State Active Check                 0      0
    General Devices Info Report             0      0
    Performance Counters Report             0      0
    Partitions Check                        0      0
    IPoIB Subnets Check                     0      0
    Subnet Manager Check                    0      0
    Fabric Qualities Report                 0      0
    Credit Loops Check                      0      0
    Multicast Groups Report                 0      0
```

12 Ensure that the initiator system can communicate with the InfiniBand/SRP target systems in the same subnet:

```
# for d in /dev/infiniband/umad*; do echo
==== SRP target ports reachable via $d; srp_daemon -oacd$d; done
==== SRP target ports reachable via /dev/infiniband/umad0
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300fc3211, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300fc3221, pkey=ffff,service_id=0002c90300a4f4d0
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300a34a31, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300a38491, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300a38492, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300a4f4c1, pkey=ffff,service_id=0002c90300a4f4d0
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300a4f4d1, pkey=ffff,service_id=0002c90300a4f4d0
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300a4f4d2, pkey=ffff,service_id=0002c90300a4f4d0
==== SRP target ports reachable via /dev/infiniband/umad1
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300fc3211, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300fc3221, pkey=ffff,service_id=0002c90300a4f4d0
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300a34a31, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300a38491, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a38490,ioc_guid=0002c90300a38490,dgid=
fe80000000000000002c90300a38492, pkey=ffff,service_id=0002c90300a38490
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300a4f4c1, pkey=ffff,service_id=0002c90300a4f4d0
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300a4f4d1, pkey=ffff,service_id=0002c90300a4f4d0
id_ext=0002c90300a4f4d0,ioc_guid=0002c90300a4f4d0,dgid=
fe80000000000000002c90300a4f4d2, pkey=ffff,service_id=0002c90300a4f4d0
```

13 After the entire IB stack is working, run `chkconfig` to make the InfiniBand/SRP core, `opensm`, and SRP tools start automatically.

```
# for s in rdma openibd opensm opensmd srpd; do [ -e /etc/init.d/$s ] && echo
==== $s && chkconfig $s on; done
# chkconfig --list | grep -E 'rdma|openib|opensm|srpd'
```

```
opensm          0:off   1:off   2:on    3:on    4:on    5:on    6:off
rdma            0:off   1:off   2:on    3:on    4:on    5:on    6:off
srpd            0:off   1:off   2:on    3:on    4:on    5:on    6:off
```

Installing the IB_SRP-BACKPORT package

After you have installed the InfiniBand/SRP components, you need to install the `ib_srp-backport.RPM`.

About ib_srp-package.rpm

> **CAUTION!** Dell does not support the use of any InfiniBand OFED packages for Linux. You should only use the InfiniBand stack provided by your Linux vendor with all of the available kernel updates applied.

In addition, Dell recommends installing and using the following third-party `ib_srp-backport` package for faster failover times and higher performance:

https://github.com/bvanassche/ib_srp-backport

This package provides the source code for making two kernel modules: `scsi_transport_srp.ko` and `ib_srp.ko`. The updates in the `ib_srp-backport` package are in newer Linux kernel versions (3.11+). These changes should be available in updated releases of major Linux distributions such as Red Hat Enterprise Linux (RHEL) 6 and beyond.

The driver source code, including the procedure for generating a binary RPM is available on the website. The contents of the binary .RPM not only depend on the operating system version (for example, RHEL 6.4+) but also on the kernel updates that have been installed. Thus, as with many different kernel modules, a new build of the `ib_srp-backport .RPM` must be created and installed after each kernel update.

Installation

To install the ib_srp-backport .RPM

1. Run the following commands as a non-privileged user:

```
$ rpmbuild --rebuild ib_srp-backport-*.src.rpm
$ (cd $(rpm --eval '%{_topdir}')/RPMS/$(uname -m) && find -name 'ib_srp-backport*.rpm' )
./ib_srp-backport-3.10.0-123.20.1.el7.x86_64-2.0.29-1.x86_64.rpm
./ib_srp-backport-3.10.0-123.20.1.el7.x86_64-debuginfo-2.0.29-1.x86_64.rpm
```

2. Obtain root privileges and install the generated RPM. For example:

```
rpm -U $(rpm --eval '%{_topdir}')/RPMS/$(uname -m)/ib_srp-backport-*.rpm
```

Separating IPoIB from SRP traffic

Although a single HCA port can run IPoIB and SRP communication simultaneously, it is possible to separate IPoIB from SRP traffic. An effective way to separate these two communication protocols is to use a different IB partitions for each protocol. An IB partition is a set of ports that are allowed to communicate with each other. Each partition is identified by a 15 bit number which is called a P_Key. A single IB port can be a member of more than one partition at the same time. Communication between two IB ports is only allowed if both are member of the same partition. This means that defining disjoint partitions for the IPoIB and SRP ports will separate these two traffic types.

Since the IB subnet manager controls which IB port is member of which IB partition configuration of partitions happens through the subnet manager. Instructions for how to configure partitions when using OpenSM as subnet manager can be found in http://git.openfabrics.org/?p=~halr/opensm.git;a=blob;f=doc/partition-config.txt

10

Troubleshooting

The procedures in this section can be used to troubleshoot issues that may be encountered with the DAAD.

After node restore, initiators may not connect correctly

If a single node in a cluster is restored, the initiators that were previously connect to it do not connect correctly to the preferred node. This is because the UUID of the node volume is different after restore, and the initiators need to be reconfigured to see the volume correctly.

After the node has been restored, run the following command from any initiators that are not connecting to the correct preferred node: `multipath -r`

Login Failure With RHEL Initiators

When using RHEL's `qlogic qla2xxx` driver, the initiator may fail to log in to the target. When this happens, disk failure messages will appear in `/var/log/messages`.

To resolve this issue, use the `qla2xxx` module parameter on the initiator side:

```
modprobe qla2xxx ql2xasynclogin=0
```

To have this setting used automatically, create the `/etc/modprobe.d/qla2xxx.conf` file and add the following line:

```
options qla2xxx ql2xmaxqdepth=2
```

If `qla2xxx` is loaded during boot time, run `mkinitrd` (or `dracut –f` on RHEL 6.x) to pick up the changes.

Oracle Linux 6.3 issues

When running Oracle Linux 6.3, OpenSM and srptools may encounter problems. The following error may be displayed:

```
# srp_daemon -oaev >/dev/null
26/07/13 12:49:01 : umad_open_port failed for device mlx4_0 port 1
```

The reason is that neither the OpenSM package nor the `srptools` package works on an unmodified Oracle Linux 6.3 system.

To resolve this issue:

1. Add the following code to the end of the `/lib/udev/rules.d/50-udev-default.rules` file:

```
# InfiniBand
KERNEL=="umad*", NAME="infiniband/%k"
KERNEL=="issm*", NAME="infiniband/%k"
KERNEL=="uverbs*", NAME="infiniband/%k"
KERNEL=="ucm*", NAME="infiniband/%k"
KERNEL=="uat", NAME="infiniband/%k"
KERNEL=="ucma", NAME="infiniband/%k"
KERNEL=="rdma_cm", NAME="infiniband/%k"
```

2. Run the following command to restart `rdma`:

```
# /etc/init.d/rdma restart
```

3. Wait for a few seconds after `rdma` has restarted, and then run the following commands to ensure that the changes have solved the issue:

```
# srp_daemon -oac
# opensm
```

OpenSM and srptools will now start as expected.

Linux cleanup after LUN removal

Linux does not automatically clean up after the Dell Acceleration Appliance for Databases LUN removals. For RHEL or OL, the required steps for LUN removal are described here:

https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/removing_devices.html

For all distros, before running `multipath –f`, run the following command:

```
dmsetup message mpathXYZ 0 "fail_if_no_path"
```

where `mpathXYZ` is the name of the device.

LUN 0 deletion problems in Oracle VM

When deleting LUN 0, Oracle VM runs certain multipath commands that create a multipath device for the deleted LUN 0 device. This results in invalid multipath devices and Oracle VM HCL failures.

When using Oracle HCL:

1. Add the following to the `/etc/multipath.conf` file:

```
blacklist {
   wwid 23565323633313765
}
```

2 Restart `multipathd`, if it is running. This prevents the multipath tools that come with Oracle HCL from making a multipath device when there is no LUN 0.

Network manager and UDEV issues with RHEL 6.4+

Problem

When using the Mellanox ConnectX-3 driver, `mlx4_en`, in RHEL, the Network Manager and `udev` tools incorrectly reset and set up the network devices. This results in the network devices being renamed or incorrectly shut down, and possibly incapable of receiving frames.

Solution

Network Manager must be disabled, and one of the `udev` rules must be modified, as follows:

1 Complete the tasks listed at this Red Hat link to stop the Network Manager service: https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux_OpenStack_Platform/3/html/Installation_and_Configuration_Guide/Disabling_Network_Manager.html

2 Ensure that the Network Manager service is stopped by running the `service` command:

```
service NetworkManager stop
```

3 Ensure that the Network Manager service is disabled at startup by using the chkconfig command:

```
chkconfig NetworkManager off
```

4 Open each interface configuration file on the system in a text editor. Interface configuration files are found in the `/etc/sysconfig/network-scripts/` directory. They have names of the form `ifcfg-X`, where X is replaced by the name of the interface. Valid interface names include `eth0`, `p1p5`, and `em1`.

5 In each configuration file, ensure that the `NM_CONTROLLED` configuration key is set to no and the ON_BOOT configuration key is set to yes.

```
NM_CONTROLLED=no
ONBOOT=yes
```

6 Ensure that the network service is started by using the service command:

```
service network start
```

7 Ensure that the network service is enabled at startup by using the `chkconfig` command:

`chkconfig network on`

To prevent `udev` from creating duplicate eth-X entries in `/etc/udev/rules.d/70-persistent-net.rules`, modify the following `udev` rule:

In:

`/lib/udev/rules.d/75-persistent-net-generator.rules`

Add this line:

```
DRIVERS=="mlx4_core", ENV{MATCHDEVID}="", ENV{MATCHIFTYPE}=""
```

NOTE: When `udev` is updated, this line must be re-added.

Handling configuration changes

After volumes have been added, deleted, or resized, it is important to rescan LUNs on each affected initiator system. The most convenient way to rescan LUNs is by running the `rescan-scsi-bus.sh` script. For example:

```
# rescan-scsi-bus.sh --forcerescan
Scanning SCSI subsystem for new devices and remove devices that have disappeared
Scanning host 34 for  SCSI target IDs  0 1 2 3 4 5 6 7, all LUNs Scanning for
device 34 0 0 0 ...
OLD: Host: scsi34 Channel: 00 Id: 00 Lun: 00 Vendor: FUSIONIO Model: ION LUN Rev:
3243 Type: Direct-Access ANSI SCSI revision: 05 0 new device(s) found. 0
device(s) removed.
```

CAUTION! Failure to rescan LUNs after having reduced the size of a volume may lead to data loss.

Informing the multipath daemon of resized devices

After a multipath device is resized on Linux and a rescan is done, the multipath daemon needs to be informed of the change. To do this, run the following commands:

```
# multipathd -k'resize map <your_multipath_device>'
```

Tips for resizing devices with multipathing

To resize a device, you must first resize it on the target/Dell Acceleration Appliance for Databases side, and then on the initiator side. For RHEL 6.X and OEL 6.X, follow the steps out lined in the Red Hat Storage guide:

https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/online-iscsi-resizing.html

Resolving path failures during multipathing scans

If a path fails while `multipathd` is scanning the path UID, then `multipathd` does not recognize that path until it is rescanned explicitly. To ensure whether `multipathd` failed to add one or more paths, run the following command:

```
# echo show paths | multipathd -k | grep faulty
388:0:0:10 sdlt  68:432  1   undef  faulty running orphan
388:0:0:11 sdlu  68:448  1   undef  faulty running orphan
```

You can also force `multipathd` to reconsider faulty paths, as follows:

```
# echo reconfigure | multipathd -k
```

11

Application tuning

This section outlines best practices for using the Dell Acceleration Appliance for Databases with Oracle and Oracle ASM.

Oracle implementation best practices

- After configuring the Dell Acceleration Appliance for Databases, export the Dell Acceleration Appliance for Databases volumes to the Oracle nodes (standalone or RAC).
- Install Oracle with the “database binary” (software-only) option.
- Run the Oracle synthetic benchmark tool Orion to measure IOPS and bandwidth on DAAD volumes with various workloads (OLTP, DSS, RAID). You can also use the FIO tool to benchmark these performance categories.
- Ensure that the Orion/FIO performance statistics approximate the theoretically possible performance of the hardware setup (HBAs on the target, initiator, and ioMemory device).
- Install Oracle Grid Infrastructure. Configure the database and run the read only `ioCalibration` tool to determine the IOPS, MBPS, and latency of the setup.
- Record your performance statistics. These will serve as the baseline for the setup if application performance issues are noted during workload tests.
- Configure NUMA and HugePages for Oracle.

ASM best practices

- ASM limits the LUN size to 2TB or less, so the Dell Acceleration Appliance for Databases volumes should be created no larger than that. (This restriction applies to Oracle grid and database versions older than 12.1.0.1. Please refer to the administrative guide for your database version for further details.)
- An ASM diskgroup can be created using one or multiple Dell Acceleration Appliance for Databases volumes. For better performance, Dell recommends that you create fewer ASM diskgroups on more Dell Acceleration Appliance for Databases volumes.
- Use the 512B sector size when exporting volumes to initiators. If you need to export LUNs with a 4KB sector size, use Oracle or third-party-provided ASMLIB support packages. ASMLIB support is now available for RHEL 6.x. For details, click

  http://www.oracle.com/technetwork/server-storage/linux/asmlib/rhel6-1940776.html

- You may also need to use the `_disk_sector_size_override` parameter to overcome an Oracle ASM diskgroup creation error.
- ASM distributes data on all the ASM disks in the disk group. A diskgroup can be created with an option for NORMAL REDUNDANCY (two-way mirroring), HIGH REDUNDANCY (three-way mirroring) or EXTERNAL REDUNDANCY (no mirroring). For the Dell Acceleration Appliance for Databases, no significant performance differences have been observed between these ASM mirroring options.
- It is generally advisable to use the default 1MB allocation unit (AU) size. For larger (3TB+) databases, the AU size can be increased to 4M or 8M.
- If more than two DAADs are used (without HA options), then by carefully creating ASM diskgroup mirroring options can avoid outages due to the rare scenario of DAAD node/component failures. For example, if you have three sets of DAAD volumes from three different DAAD nodes and all part of a HIGH redundancy mirroring option, then losing two DAADs completely would still make the application function without errors (assuming there is enough free disk space for the database transactions).

Exporting LUNs to initiators

In RAC, Dell recommends that you have separate, small Dell Acceleration Appliance for Databases volumes for Oracle cluster voting disks, with Normal or High redundancy ASM mirroring. Optionally, versions 11gR2 and 12c support voting disks as part of regular ASM diskgroups. Careful planning is needed here to avoid Oracle RAC node evictions.

For Oracle RAC configuration, Dell recommends that you have one initiator group for all participating Oracle initiator nodes. This will greatly decrease the time spent in LUN-mapping.

If you are planning to assign LUNs to various initiators, you should create the initiator groups individually.

Because LUN mappings include the target/initiator port addresses, Dell recommends that you tag the cables and ports. This helps in reconnecting to the same ports if the direct-connected cables need to be removed or exchanged.

ASM storage configurations

The configurations that follow are variations that may serve your particular storage needs, using Oracle ASM and the Dell Acceleration Appliance for Databases.

HA mode, one diskgroup

One large ASM diskgroup can be created with External Redundancy, which means no mirroring is done by ASM. Here is the basic procedure:

1 Place all the Dell Acceleration Appliance for Databases storage from both the HA nodes into one ASM diskgroup called DATA.

2 Have Oracle write the OCR and Voting Disk files to DATA, which gets passed through to the underlying LUNs on the Dell Acceleration Appliance for Databases.

The HA implementation of the Dell Acceleration Appliance for Databases takes care of creating the redundant copies.

Standalone mode

In standalone mode, Oracle ASM can be used with Normal Redundancy. The procedure uses the first two steps in HA mode, one diskgroup on page 152.

A variation is to use ASM High Redundancy for the OCR and Voting files, and Normal Redundancy for the database files. To do this:

3 Create a 1 GB partition on *n* LUNs from each Dell Acceleration Appliance for Databases node (where *n* is 6 divided by the number of nodes).

4 Put those 6 partitions into an ASM diskgroup named CRS with High Redundancy.

5 Put all of the remaining storage into the DATA diskgroup with Normal Redundancy.

Standalone mode: Storing OCR and voting files

When using ASM (not Dell Acceleration Appliance for Databases HA) for redundancy, you do not have to store the OCR and Voting files in ASM. During Oracle Clusterware installation, the checkbox for this can be cleared so ASM is not used for these two types of files.

The basic procedure is to create a 1GB partition on the first *n* LUNs from each DAAD until there are six total partitions (two LUNs from each of three Dell Acceleration Appliance for Databases nodes, or one LUN from each of six nodes).

During installation, Oracle Clusterware prompts for the 6 devices on which Oracle stores the OCR and Voting Disk files. There are 6 OCR files and 5 Voting Disk files (the Voting Files are small extents stored in the headers of the OCR files, not separate files themselves).

It is very important to note that this count of files (6 OCR and 5 Voting) means one of the LUNs will not have a Voting File. It is critical that no single Dell Acceleration Appliance for Databases node has a majority of the Voting Disk files, because if a majority of them go offline, Oracle aborts the entire cluster.

For example, in a two-node configuration, there would be three voting files on the first node and two on the second node. Taking the second node offline for maintenance is not a problem, because the majority of Voting Files are still online. However, taking the first node offline causes the entire cluster to abort. The solution calls for a third storage node, which can be a third Dell Acceleration Appliance for Databases node or a legacy SAN. With three storage nodes, two voting files are placed on the first, two on the second, and one on the third. This way, no single node contains a majority of voting files, so the cluster is not aborted.

Best practices for MS SQL server

This section outlines best practices for using the Dell Acceleration Appliance for Databases with MS SQL Server.

MS SQL server implementation best practices

- Configure the Dell Acceleration Appliance for Databases and export its volumes to the SQL DB nodes (standalone or cluster)
- Run the SQLIO synthetic benchmark tool to measure the IOPS and bandwidth for volumes, with various workloads.
- Ensure that the SQLIO performance statistics approximate the theoretically possible performance of the hardware setup (HBAs on the target, initiator, and Fusion ioDrive devices).
- Record the performance statistics. These will serve as the baseline for the setup if application performance issues are noted during workload tests.

SQL server memory allocation

- Ensure that you allocate sufficient memory space for the OS processes. As a best practice, earmark 1GB for every 4 cores and allocate the rest to the *Max server memory (MB)* parameter.
- Use the following dynamic management views to troubleshoot memory issues that may occur:
  - `sys.dm_os_memory_brokers` provides information about memory allocations using the internal SQL Server memory manager. The information provided can be useful in determining very large memory consumers.
  - `sys.dm_os_memory_nodes` and `sys.dm_os_memory_node_access_stats` provides a summary of the memory allocations per memory node and node access statistics, grouped by type of page. This information can be used to quickly obtain a summary of memory usage, without running `DBCC MEMORYSTATUS`.
  - `sys.dm_os_nodes` provides information about CPU node configuration for SQL Server. This DMV also reflects software NUMA (soft-NUMA) configuration.
  - `sys.dm_os_sys_memory` returns the system memory information. The "Available physical memory is low" value in the `system_memory_state_desc` column is a sign of external memory pressure that requires further analysis.

Other best practices

- When formatting the partition that will be used for SQL server data files, you should use a 64 KB allocation unit size for data, logs and the TempDB database.
- Lock pages in memory. To reduce SQL server paging, grant the service account "Lock pages in Memory" privileges through the Windows Group Policy editor for both 32- and 64-bit servers.

- Adjust the “Degree of Parallelism” option to the number of cores in a single NUMA node, which is eight or fewer. Test the workload with various degrees of parallelism to arrive at the best combination.
- The default setting for Number of Worker Threads works well for most configurations.
- To ensure efficient operation, create multiple TempDB files (1 per physical CPU core). Pre-size the TempDB files and create them in equal sizes.
- For ease of maintenance, create more than one log file (there is no performance gain in configuring more files).

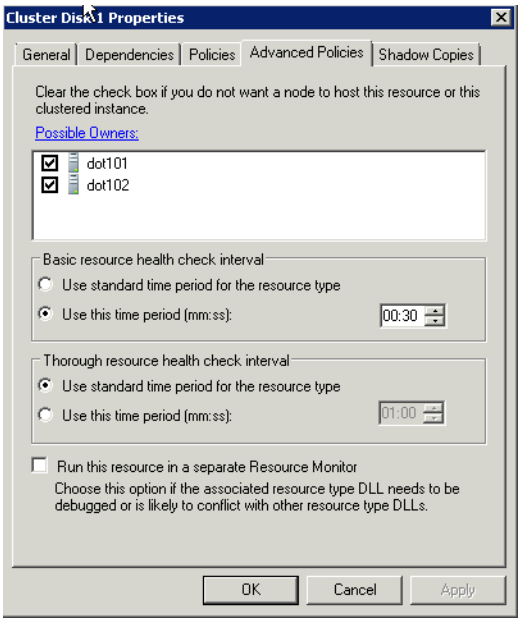
HA with Windows server failover cluster

If you are deploying Windows Server 2008 with the Dell Acceleration Appliance for Databases, configure your storage to use Windows Server Failover Cluster (WFSC). For MPIO settings, refer to Configuring multipath on Windows on page 108.

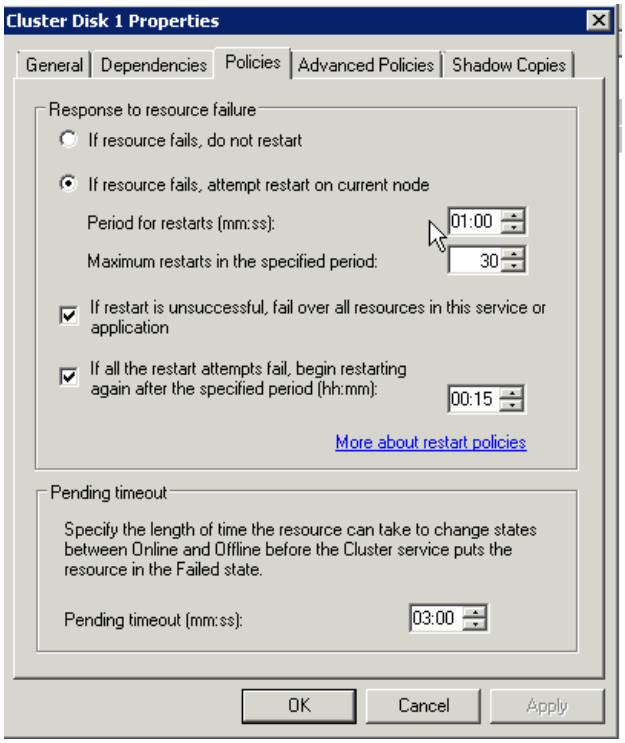
1 In WSFC, navigate to Storage.

2 Click the **Advanced Policies** tab.

3 Set **Use this time period** to **30** seconds.



4 Click the **Policies** tab.



5 Set **Period for restarts** to **1:00**.

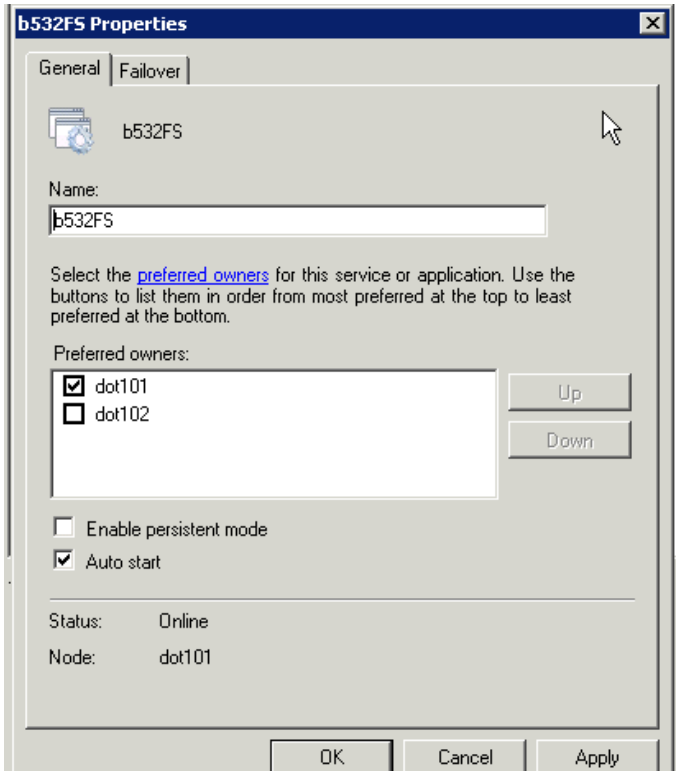
6 Set **Maximum restarts in the specified period** to **30**.

7 Set **If all restart attempts fail, begin restarting again after the specified period** to **15** minutes.

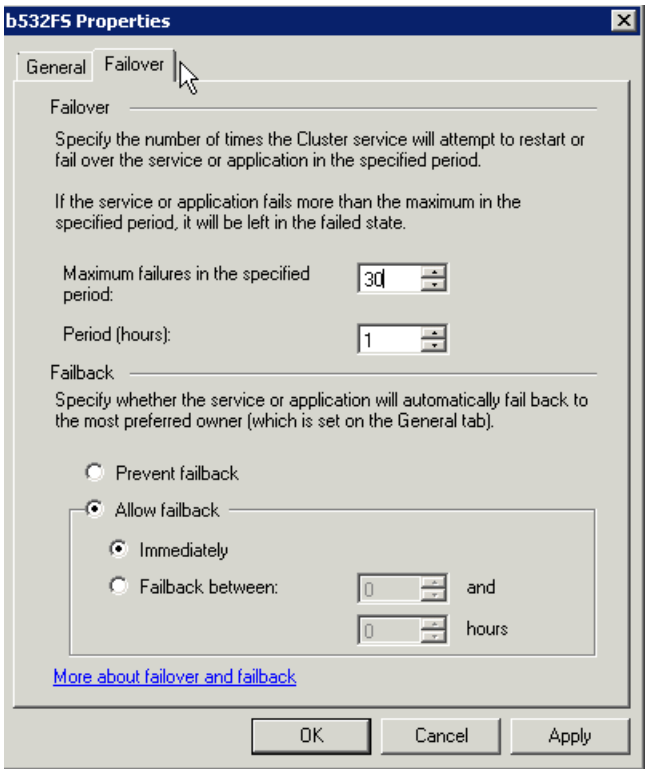
Configuring auto-failback settings

To specify auto-failback settings on a WSFC node for any service or application:

1 Ensure that the preferred owner is selected in the General tab. In the example, the preferred owner is **dot101**.



2 On the **Failover** tab, set the maximum number of failures as **30** in **1** hour.



3 In the Failback section, click **Allow failback** and then **Immediately**.

Windows clustering hot fixes

Several hot fixes for Windows Clustering are available on the Microsoft Support site. These may be useful in preventing or resolving issues with Windows clusters used in the Dell Acceleration Appliance for Databases system.

Hot Fix Reference	Description
http://support.microsoft.com/kb/2718576/en-us	This fixes an issue in which an MPIO disk is removed unexpectedly when a `PR_IN` command fails. This issue occurs when the Failover Clustering feature is configured in a Windows Server 2008 R2 environment.
http://support.microsoft.com/kb/2733575/en-us	SAN targets should reply to `PR_IN` commands within 3 seconds on Windows Server 2008 and Server 2008 R2.
http://support.microsoft.com/kb/2522766	The MPIO driver fails over all paths incorrectly when a transient, single failure occurs in Windows Server 2008 or in Windows Server 2008 R2.
http://support.microsoft.com/kb/2550886	A transient communication failure causes a Windows Server 2008 R2 failover cluster to stop working.

12

Contacting technical support

Dell Acceleration Appliance for Databases drivers, utilities, and related documentation are available at:

dell.com/support/home

Dell provides several online and telephone-based support and service options. Availability varies by country and product, and some services may not be available in your area. To get help with your Fusion ioMemory devices, contact your Dell Technical Service representative or access the Dell Support website.

Choose the method of contacting Dell that is convenient for you.

> **NOTE:** The safety information that shipped with your system provides important safety and regulatory information. Warranty information may be included within this document or as a separate document.

A

Appendix A: DAAD configuration

DAAD platform: Front panel

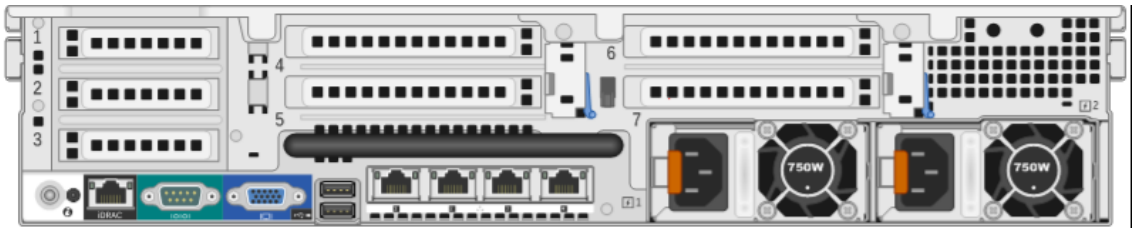
Figure A-1 shows the front view of the Dell Acceleration Appliance for Databases.

Figure A-1. Front view of DAAD

DAAD platform: Rear panel

Figure A-2 shows the rear-panel view of the Dell Acceleration Appliance for Databases.

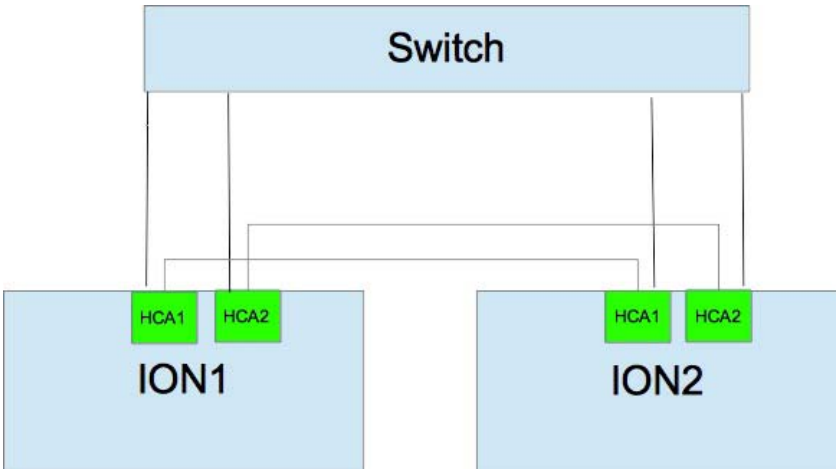
Figure A-2. Rear view of DAAD



Split-function ports for InfiniBand/SRP and iSCSI

For increased redundancy or high-availability, when using two Mellanox ConnectX-3 dual-port adapters for both connectivity between HA DAADs and fabric connectivity to the initiator nodes (as when running SRP/IB or iSCSI), it is best to split the functionality between the cards. This allows each card to be used for both a redundant connection to the partner node and to the fabric. That way, if one adapter fails, and both its ports are lost, the system can continue to run. The mode will be partially degraded, but not system-failed, with the remaining cluster interconnect and iSCSI or InfiniBand/SRP ports on the other adapter.

Figure A-3. HA configuration with split-port functionality



It is important to use the correct ports for fabric and DAAD interconnectivity in this case. The diagram that follows, along with the rear-view picture of the DAAD chassis shown previously, describes the ports to be used to connect DAADs, and the ports to be used when connecting to the fabric.

> **NOTE:** While split-port functionality is recommended for iSCSI, it is mandatory for InfiniBand/SRP.

In the example here, slots 1 and slot 3 use one port for HA cluster interconnect. The other ports on Slots 1, 2, and 3 are iSCSI ports.

NOTE: Note that the cluster interconnect protocol runs over the Ethernet ports on the InfiniBand HCA.

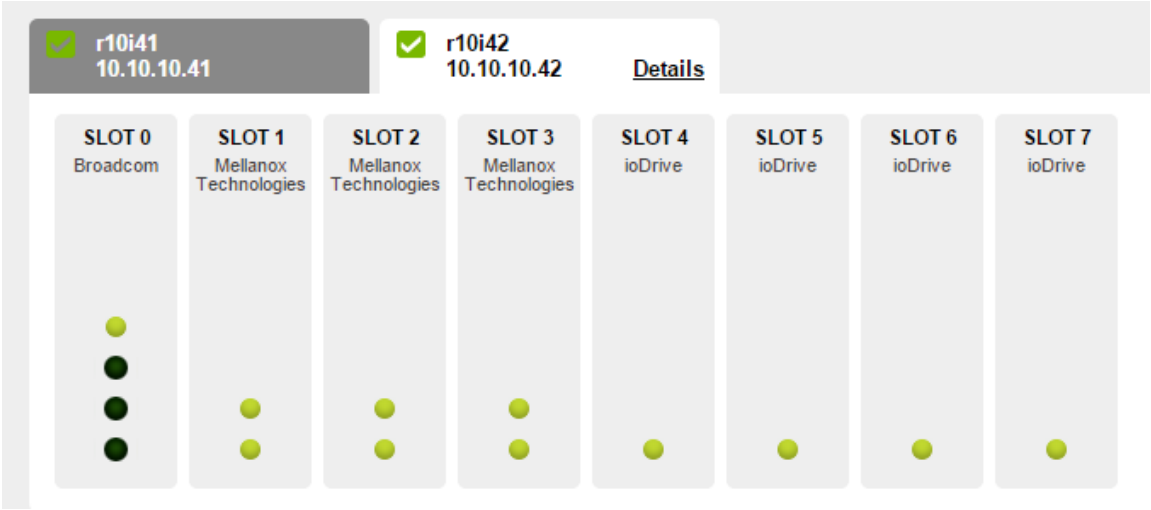


Figure A-4. ConnectX-3 cluster interconnect card