# Server and Network Configuration Guide

**Version 3.4**          **October 19, 2013**          **Edition 1**

## Copyright Information

## Trademarks

# About this Guide

The SeaMicro system is designed to replace existing equipment and functionality. This guide assumes that the user is familiar with servers and server load-balancing topics. The guide documents the system's functionality, but does not cover all its possible applications in a data center environment.

This guide is divided into four parts:

- Part 1 — Getting Started. This part contains the First-time System Bring-up procedure, which is a high-level list of tasks that you must complete when you are finished installing the chassis. Each step in the procedure directs you to the section in the document where you will find the configuration information that you need.

- Part 2 — Basic Software Features. This part contains the features that you will use first and most often, like interface configurations. In general, these features require a lower level of expertise with regard to the system and the operating environment.

- Part 3 — Advanced Software Features. This part contains the features that require more knowledge of the operating environment. While these features might be simple to execute from the system command line, they are designed for specific applications, and require a higher level of expertise.

- Part 4 — Appendices. This part contains supplementary information that you might need for system monitoring.

# Conventions

Commands are defined in this guide using the conventions in Table 1. For example, let's use the following command:

**interface {{gigabitethernet|tengigabitethernet}** *slot*/*port*} [**description** *string* ||
**gratuitous-arp** || **shutdown**.

**Table 1** Command Syntax Documentation Conventions

| Command Element | Description | Example |
|---|---|---|
| **Keyword** | Enter the word exactly as shown. | You must enter **gigabitethernet**. |
| *Variable* | Enter a valid string or value. The CLI help displays the available options. | *slot*/*port*  requires you to enter a number. |
| X \| Y | A bar between keywords or variables separated indicates that you may choose only one. | You may enter different type of interfaces, such as **gigabitethernet** or **tengigabitethernet**. |
| X \|\| Y | A double bar between keywords indicates that you if you enter the first keyword you may then still enter the next option. | You may enter **gratuitous-arp** before you enter **shutdown.** |
| [X] | Brackets around a keyword or variable indicates that entering the element is optional. | You may enter only **description** *string* and/or **gratuitous-arp** and/or **shutdown**. You may enter one, two, or three elements. |
| {X \| Y} | Braces around set of keywords or variables indicates that you must enter one of the elements. | You must enter either **gigabitethernet** or **tengigabitethernet** when issuing this command. |

# Related Documentation

- *Command Line Reference*
- *Hardware Installation Guide*
- *Quick Start Guide*
- *Release Notes*

# Technical Support

For support, log in to the SeaMicro Support page at: http://www.seamicro.com/support. If you do not have a SeaMicro Support account, you can request one at: http://www.seamicro.com/supportform. Support login credentials gives you access to hardware installation and software configuration documentation, Generally Available code, release notes, Frequently Asked Questions, and Technical Tips. You can also open, track, and administer SeaMicro Support cases.

To contact SeaMicro Support directly by E-mail or by phone:

- **E-mail:** support@seamicro.com
- **Phone:** International: 1-408-701-5077
        US: 1-888-522-8760

# Part 1 — Getting Started

# Chapter 1 — First-time System Bring-up

This chapter has a specific purpose, which is to help you bring up the system for the first time. The process of configuration begins immediately after you have completed chassis installation with instructions from the *Hardware Installation Guide,* and ends with the internal servers booting up an Operating System (OS).

There are two options for configuring your chassis:

- Zero touch provisioning, and/or
- Manual configuration of the chassis.

# Zero Touch Provisioning

Configuration is made simple with '**Zero Touch Provisioning,**' which allows accelerated and automatic deployment of a standard configuration template onto the chassis. After installation, the chassis can be configured automatically with zero touch provisioning, without an administrator logging into the system.

In order to execute zero touch provisioning, the chassis must be in a state of factory default configuration. As part of such provisioning, the DHCP client running on the SeaMicro chassis will:

- Procure management IP addresses and chassis configuration URL from the DHCP server.
- Fetch the configuration file from the specified URL, and apply it to the chassis.

Using zero touch provisioning does not preclude you from manual configuration of the chassis. You may configure the chassis manually at any time, and vice versa.

**Figure 1** Zero Touch Provisioning
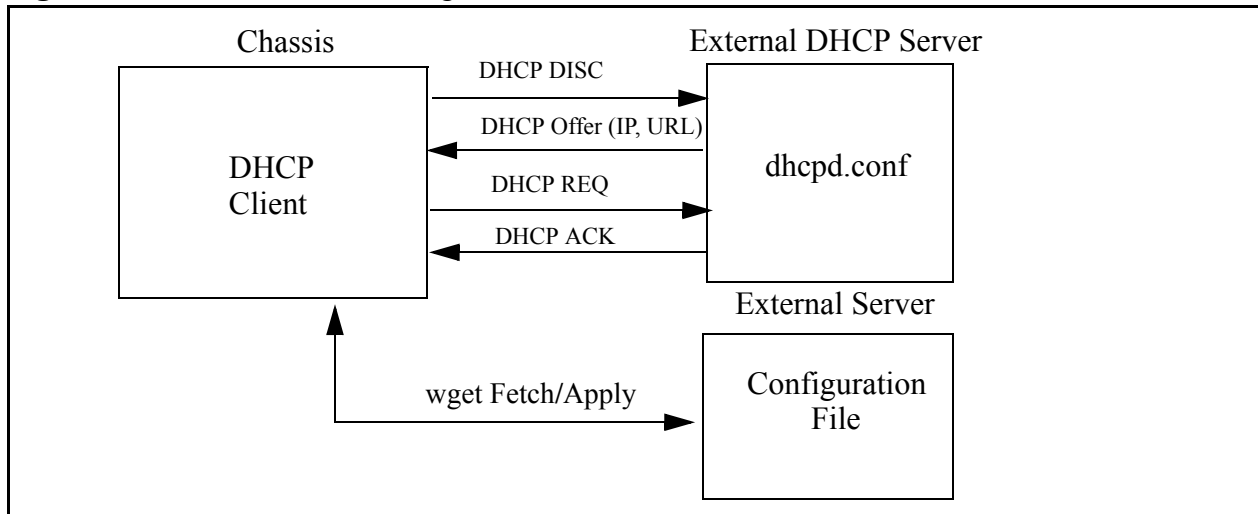


To execute zero touch provisioning:

| Step | Task | ☑ |
|:---:|:---|:---:|
| 1 | Ensure that the chassis is in factory default configuration state. | ☐ |
| 2 | If the chassis is not in factory default configuration state, or if it contains pre-existing configuration, clear the configuration using the command:<br><br>'**clear startup-config**' or '**clear startup-config force**' | ☐ |
| 3 | Before you power on the chassis or execute zero touch provisioning, save your configuration file in an external DHCP server from where the DHCP client can fetch your preferred configuration.<br><br>Make sure it includes the IP address and the configuration URL information. | ☐ |

| Step | Task | ☑ |
|------|------|---|
| 4 | In the dhcpd.conf file on the external DHCP server, specify the following in the order listed to send to the URL: | ☐ |

option space seamicro;
option seamicro.configfile_url code 1 = string;

  subnet 192.168.2.0 netmask 255.255.255.0 {
    host ch8 {
    hardware ethernet 00:21:53:91:03:80;
    fixed-address 192.168.2.209;
    vendor-option-space seamicro;
    option seamicro.configfile_url "http://192.168.2.1/chassisCfg";  }}

*(Replace the URL in the above sample with the URL that contains your saved configuration file, and confirm that the URL location actually contains the file.)*

| Step | Task | ☑ |
|------|------|---|
| 5 | Next, connect the chassis to an ethernet port. | ☐ |
| 6 | When in factory default configuration, the chassis will automatically come up with DHCP IP enabled. If not, in the configuration mode for mgmteth interface, enable DHCP IP: | ☐ |

**interface mgmteth**
**ip dhcp enable**

Note that when DHCP IP is enabled, no other static configurations, such as IP, default gateway, or DNS are allowed. You can always disable DHCP IP to configure static management IP if needed at a later time.

| Step | Task | ☑ |
|------|------|---|
| 7 | As soon as the chassis is powered on, the DHCP client will request and procure the IP address and the URL information from the DHCP server. Note: | ☐ |

- The chassis can procure either DHCP IP addresses or Static IP addresses, but not both.
- The command '**show interfaces mgmteth**' will display 'DHCP' if the IP address is received from DHCP.

| Step | Task | ☑ |
|------|------|---|
| 8 | If the DHCP response contains a URL, the saved configuration file from that URL will be fetched and applied to the chassis. | ☐ |
| 9 | Reload the chassis using the command '**reload**.' | ☐ |

| Step | Task | ☑ |
|------|------|---|
| 10 | At this point, the chassis will automatically come up with your specified configuration.<br><br>Congratulations, your chassis is now configured! | ☐ |

# Manual Configuration of Chassis

To manually configure the chassis without executing zero touch provisioning, complete each step in the order listed below. Not following some steps might result in the loss of data or degraded system operation.

| Step | Task | ☑ |
|:---:|:---|:---:|
| 1 | Connect to the system console. You may use either of the ports on the SM-card, or the console port on the U-card. The console port on the U-card takes a rollover cable; for the DB-9 port, the pinout is shown in the *Hardware Installation Guide,* and the console parameters are given below:<br><br>```BAUD: 9600```<br>```Data bits: 8```<br>```Stop bits: 1```<br>```Parity: None``` | ☐ |
| 2 | Log in using the default login credentials:<br><br>```username:``` **admin**<br>```password:``` **seamicro** | ☐ |
| 3 | Assign the inband interface an IP address and default-gateway, and enable SSH.<br><br>**interface inband**<br>**ip address \<ip-address>/\<mask>**<br>**ip default-gateway \<ip-address>**<br>**allow ssh** | ☐ |
| 4 | Configure an external DHCP server. | ☐ |
| 5 | Prepare the internal servers to install an OS via PXE. The system does not support KVM features, so some edits to the standard bootloader configuration file and answer file are required. Refer to SeaMicro Operating Systems and Environment Installation Guide. | ☐ |
| 6 | Assign volumes to the internal servers. See Configuring Storage on page 36. | ☐ |
| 7 | (Optional) Configure port-channels. See Port-channel Interfaces on page 89 in the configuration guide; this section includes information on static and dynamic (LACP) port-channels. | ☐ |
| 8 | If you did not configure port-channels in Step 8, configure gigabitethernet or tengigabitethernet interfaces. See Gigabit Ethernet Interfaces on page 83 or Tengigabit Ethernet Interfaces on page 91. | ☐ |

| Step | Task | ☑ |
|:---:|:---|:---:|
| 9 | Save your configuration.<br><br>`write memory`<br><br>**Important:** You must save your configuration or else your configuration data, and potentially all of the data, will be lost when you reload the system. | ☐ |
| 10 | Power on the internal servers.<br><br>`power-on server all` | ☐ |

# Part 2 — Basic Software Features

# System Management

The SeaMicro system has a simple software management environment that provides users with the ability to monitor and manage the servers, storage, and Ethernet infrastructure using an array of user interfaces including CLI, SNMP, Syslog, IPMI, and RESTful API.

This chapter covers the following system management topics:

# Remote System Connection

Using SSH or Telnet, you can remotely connect to the chassis, and you can connect to another device from the SeaMicro chassis.

| Task | Command | CLI Level |
|---|---|---|
| SSH to the chassis over the out-of-band management interface, or SSH to another device from the chassis.<br><br>**Note:** You can pass commands as a command-line argument to SSH. For example: **ssh admin@ch7 "show running-config; show clock"**. This includes configuration commands, for example: **ssh admin@ch7 "enable; configure; interface gigabitethernet1/0; no shut"**. | **ssh** *ip-address username* | Privileged |

| Task | Command | CLI Level |
|---|---|---|
| Telnet to the chassis over the management port, or Telnet to another device from the chassis. To Telnet to the chassis, the Telnet service must be enabled on the management port; see Management Ethernet Interface on page 86. | **telnet** *ip-address* | Privileged |
| (Optional) Set the terminal parameters. | **terminal {idle-timeout** *seconds* \| **length** *rows* \| **width** *columns*} Defaults: idle-timeout: 3600 seconds rows: 24 columns: 80 | Privileged |

# Key-based Authentication over SSH

**Note:** Only one user can be configured to access the chassis by this method.

To enable key-based authentication over SSH:

| Step | Task |
|---|---|
| 1 | Generate an RSA public/private key pair on the remote host or terminal server using the SSH-Keygen tool (**ssh-keygen**). Press enter at the dialog prompts to use the default values. This will generate *id_rsa* and *is_rsa.pub* in the location indicated.<br><br>```<br>[techpubs@CentOS_5.3 ~]$ ssh-keygen<br>Generating public/private rsa key pair.<br>Enter file in which to save the key (/home/techpubs/.ssh/id_rsa):<br>[Enter]<br>Created directory '/home/techpubs/.ssh'.<br>Enter passphrase (empty for no passphrase):[Enter]<br>Enter same passphrase again:[Enter]<br>Your identification has been saved in /home/techpubs/.ssh/id_rsa.<br>Your public key has been saved in /home/techpubs/.ssh/id_rsa.pub.<br>The key fingerprint is:<br>27:69:76:da:f1:cc:b6:0f:63:da:36:7e:17:41:15:cc techpubs@tacserver1<br>[techpubs@CentOS_5.3 ~]$ cd /home/techpubs/.ssh<br>[techpubs@CentOS_5.3 .ssh]$ dir<br>id_rsa  id_rsa.pub<br>[techpubs@CentOS_5.3 .ssh]$<br>``` |
| 2 | Login to the SeaMicro chassis. |

| Step | Task |
|------|------|
| 3 | Copy the public key to the system.<br><br>**copy authorized-key scp:** *ip-address filepath*/*filename* **system:**<br>CLI Level: Privileged<br><br>```<br>seamicro# copy authorized-key scp: 10.11.0.1 techpubs /home/techpubs/<br>.ssh/id_rsa.pub system:<br>Are you sure you want to overwrite the authorized key [no,yes] yes<br>The authenticity of host '10.11.0.1 (10.11.0.1)' can't be established.<br>RSA key fingerprint is<br>fd:51:2a:f3:fa:42:b0:65:38:2d:6d:4c:37:c8:bb:b8.<br>Are you sure you want to continue connecting (yes/no)? yes<br>Warning: Permanently added '10.11.0.1' (RSA) to the list of known<br>hosts.<br>techpubs@10.11.0.1's password:<br>id_rsa.pub<br>100%  400     0.4KB/s   00:00<br>seamicro#<br>``` |
| 4 | Exit your terminal session, and then you will be able to connect again (from the host with the private key) using **ssh admin@**{<ip-address > | hostname>}, without having to provide your login credentials.<br><br>```<br>[techpubs@CentOS_5.3 ~]$ ssh -b 10.11.0.1 admin@192.168.1.1<br>Copyright (c) 2009-2014 SeaMicro, Inc.<br>Welcome to SeaMicro OS!<br>admin connected from 10.11.0.1 using ssh on seamicro<br>seamicro><br>``` |

To delete the public key:

| Step | Task | Command |
|------|------|---------|
| 1 | Delete the public key from the system. | Command: **clear authorized-key**<br>CLI Level: Privileged |

# Reload, Reset, and Halt

You can reload the system or system components, and gracefully power down the chassis with a single command.

**Reload**—Reboot the system. The servers must be powered off in order to reload. If there are unsaved changes in the running-config, you are prompted to save them to the startup-config before the reload is executed.

| Step | Task | Command |
|------|------|---------|
| 1 | Power off all internal servers. | **power-off server all** [**force**]<br>CLI Level: Privileged |
| 2 | Reboot the system. You may enter the no-confirm option to avoid all of the dialogs (Steps 3 and 4). | **reload** [**no-confirm**]<br>CLI Level: Privileged |
| 3 | If there are unsaved changes to the running-config, you are prompted to save them. Enter "yes" or "no" in response to the dialog. | |
| 4 | If you did not enter the no-confirm option, you are prompted to confirm the reload action. Enter "yes" or "no" in response to the dialog. | |

The following is an example of a system reload:

```
seamicro(config)# do write mem
seamicro(config)# do reload
Are you sure you want to reboot the chassis (yes/no):no
seamicro(config)# hostname seamicro
seamicro(config)# do reload
Do you want to save the running-config (yes/no): yes
Are you sure you want to reboot the chassis (yes/no):no
seamicro(config)#
```

**Reset**—Reboot a component. You can reboot an internal server, MX-card, or an MX-card and its corresponding S-card.

| Task | Command |
|------|---------|
| Reset one or more servers. | **reset** {**server** {**all** \| *server-number* \| **assigned-to** *disk*}<br>CLI Level: Privileged |
| Reset the MX-card. | **reset mxcard** *number* [**force**]<br>CLI Level: Privileged |
| Reset the S-card. | **reset scard** *number* **force**<br>CLI Level: Privileged |

**halt**—Gracefully, stop all processes, and then power down the system.

| Step | Task | Command |
|------|------|---------|
| 1 | Halt the system. | **halt**<br>CLI Level: Privileged |

# Save Your Configuration

After you make configuration changes, you must save your changes to the startup-configuration so that they persist across a system reload.

| Task | Command | CLI Level |
|------|---------|-----------|
| Your current configuration is in the running-configuration file. To save it, you must write the running-configuration over the startup-configuration. | **write memory** | Configuration |

# File Management

Copy the startup-config file from the system to another machine:

| Command | CLI Level |
|---------|-----------|
| **copy config startup-config scp:** *ip-address username destination-file* | Privileged |

```
seamicro# copy config startup-config scp: 10.11.0.1 techpubs techpubs-config
techpubs@10.11.0.1's password:
seamicro-startupConfig                         100%   41KB  40.9KB/s   00:00
seamicro#
```

Copy a configuration file on another machine to the system, writing over the current startup-config:

| Command | CLI Level |
|---------|-----------|
| **copy config scp:** *ip-address username* **startup-config** | Privileged |

```
seamicro# copy config scp: 10.11.0.1 techpubs techpubs-config ?
Possible completions:
  startup-config   Startup configuration
```

```
seamicro# copy config scp: 10.11.0.1 techpubs techpubs-config startup-config
Are you sure you want to overwrite the config [no,yes] yes
techpubs@10.11.0.1's password:
techpubs-config                                    100%   28KB  27.8KB/s   00:00
Startup-config copied successfully
system: %SYSTEM-5-GENERAL-NOTICE: Startup-config successfully overwritten
seamicro#
```

# Change the System Hostname

The hostname appears as part of the CLI prompt for all levels and contexts, and is used to mnemonically identify the system.

| Task | Command | CLI Level |
|------|---------|-----------|
| Change the system hostname. | **hostname** *string*<br>Default: SeaMicro | Configuration |

# Set the System Clock

You can set the system clock manually, or synchronize the clock with an NTP server.

**Note:** On internal servers, you cannot execute the OS command "hwclock --systohc" to set the current system time to the hardware real time clock (RTC). This is because there is only one physical RTC in the chassis, which resides on the SM card.

-
-

## Set the System Clock Manually

The time is set manually from Privileged level. The default time is non-null, but arbitrary, and should be set to the current time upon first boot. When setting the system clock, you may optionally specify the time zone. The time zone is set manually from Configuration mode, and the setting is stored in the running-configuration.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Set the system date and time manually. | **clock set** *time date* | Privileged |

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 2 | Set the clock time zone using one of the following keywords:<br>• Central Standard Time: **CST6CDT**<br>• China Standard Time: **China**<br>• Eastern Standard Time: **EST5EDT**<br>• India Standard Time: **India**<br>• Japan Standard Time: **Japan**<br>• Mountain Standard Time: **MST7MDT**<br>• Mexico Standard Time: **Mexico** | **clock timezone** [**CST6CDT** \| **China** \| **EST5EDT** \| **GMT** \| **India** \| **Japan** \| **MST7MDT** \| **Mexico** \| **PST8PDT**]<br>Default: PST8PDT | Configuration |
| 3 | Reload the system. | **reload** | Privileged |
| 4 | (Optional) Verify the system time.<br><br>`seamicro# show clock`<br>`Sat Jun  5 19:22:36 PDT 2010` | **show clock** | Privileged |

# Set the System Clock using NTP

When a system is booted up:

- NTP will be initiated on all MX-cards and S-cards.
- The primary MX-card will sync with the configured external server, or if none is configured, it will sync with the local reference clock.
- All secondary MX-cards and all S-cards will sync with the primary MX-card.

You may specify more than one server, and the system will attempt to synchronize with each server in turn, beginning with the first configured server, until it is able to sync with one. A single server can be specified as the most preferred, and synchronization will be attempted with this server first.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Specify an NTP server. To specify multiple servers, enter a separate command for each server. | **ntp server** *ip-address* | Configuration |
| 2 | Manually set the time zone. | **clock timezone** [**CST6CDT** \| **China  EST5EDT** \| **India** \| **Japan  MST7MDT** \| **Mexico  PST8PDT**] | Configuration |
| 3 | (Optional) Specify an NTP server as the most preferred. | **ntp server** *ip-address* **prefer** | Configuration |

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 4 | (Optional) Specify the NTP version to use with a particular server. | **ntp server** *ip-address* **version** *number*<br>Default: 3<br>Range: 1 - 4 | Configuration |
| 5 | Reload the system. | **reload** | Privileged |
| 6 | Verify the NTP configuration. | **show ntp** | Privileged |

```
seamicro# show ntp
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
*192.168.1.145   192.168.10.21    3 u  364  512  377    0.229   15.929   0.921
```

# System Boot Parameters

There are two user-configurable parameters to influence system boot-up:

-
-

## Specify the Primary MX-card

Of all the MX-card, one card is the primary, and the rest are secondary.

- **Primary MX-card**—handles configuration, global Ethernet, and system management.
- **Secondary MX-card**— handles local Ethernet interfaces, and storage management.

By default the card in the lowest slot at boot-up is elected the primary. If the primary fails, the card in the next highest slot becomes the primary. You may prefer an MX-card as the primary.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Specify an MX-card as the preferred primary. | **system boot master-slot** *slot* | Configuration |
| 2 | Verify the primary MX-card setting. | **show run system boot master-slot** | Privileged |

# Choose the Boot Flash Partition

The boot flash has two partitions, each of which can contain a version of the SeaMicro software. You can select the software version that the system boots by pointing to the corresponding boot flash partition.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Specify the flash partition from which the system boots. | **system boot flash:** {**0** \| **1**} | Configuration |
| 2 | Verify the boot flash partition setting. | **show boot** | Privileged |

# System Auto Restoral

The system will automatically restore on power recovery. By default it takes 7 seconds for the system to auto restore. You can configure the time delay to auto restore between 1-60 seconds and then save the configuration. Configuring to 0 disables this feature.

The following shows an example of configuration.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Restore power automatically upon loss of power. | **system restore-power-state** *time-delay in seconds*<br><br>Default: 7 seconds.<br>Note: Setting delay to **0** disables this feature. | Configuration |

Example:

```
seamicro# show startup-config system restore-power-state
system restore-power-state 7
seamicro(config)# system restore-power-state ?
Possible completions:   Delay between 1-60 (default: 7). Setting delay to 0 disables
this feature[7]
seamicro(config)# system restore-power-state 47
seamicro# show running-config system restore-power-state
system restore-power-state 47
seamicro# show startup-config system restore-power-state
system restore-power-state 7
seamicro# write memory
seamicro# show startup-config system restore-power-state
system restore-power-state 47
seamicro#
```

# Power On/Power Off Servers

A server may be powered on or powered off from its serial console. The facility to power on or power off a server from its serial console is disabled by default.

To enable the ability to power on and power off a server:

1.  From the config prompt, enter the command:

    **config# system console-server key-sequence-reset enable**

2.  Once enabled, power-on a server by pressing:

    **CTRL + SHIFT + - I**

3.  To power off a server, press:

    **CTRL + SHIFT + - O**

4.  To see if this facility is enabled, enter the command:

    **config# show system console-server key-sequence-reset**

To disable the ability to power on or power off a server:

*   From the config prompt, enter the command:

    **config# no system console-server key-sequence-reset**

# Control Key Sequences

A list of control key sequences to power on or power off the system, and to access the console of the MX-card or interrupt the boot sequence is included in the table below.

| Control Key Sequences | Result |
| --- | --- |
| Ctrl-Shift-6 Shift-I | Power on system. |
| Ctrl-Shift-6 Shift-O | Power off system. |
| Ctrl-Shift-6 <#> | Takes the user to the console of the MX-card stipulated by #. |
| Ctrl-Shift-6 b | Interrupts Boot Sequence (after 'prepare to boot from flash'). |

# RESTful API

RESTful API commands may be used to get information on individual components in a chassis or system. They are accessible through an external terminal or browser using the http or https protocol via the chassis IP. The chassis IP can be either the inband interface IP or the management ethernet interface IP (chassis id).

Follow the steps outlined below to execute RESTful API commands:

1. In the CLI configuration mode, enable RESTful API with the following command: '**restserver enable**'

2. Save the configuration using the command '**write memory**.'

   Note that the '**restserver proxy'** is disabled by default. Restserver proxy is the interface between the restserver and the rest of the system.

3. Ensure that http and/or https protocol on the inband interface or the management ethernet interface is enabled.

   - By default, https is enabled on the management ethernet interface.
   - By default, http and https are disabled on the inband interface.

   Based on the type of protocol used to access RESTful API, enable http or https at the inband interface or the management ethernet interface.

4. To invoke and use RESTful API, refer to the *CLI and RESTful API Guide* for more information.

# Chapter 3 — Storage Management

This chapter covers storage overview and the following storage related-tasks:

# Storage Overview

The SM15K chassis offers flexible, shared, high capacity, high utilization, central storage management for large scale deployments. Storage management includes the following:

- Highly scalable storage with 5.3 PB capacity per SM15K.
- Data protection with integrated hardware RAID.
- Fully integrated management including provisioning, monitoring, and failure detection.
- Simple two-step provisioning of all storage attached to the SM15K.
- Flexibility to configure volumes or assign raw disks to servers depending on the application.
- Ease of operation with end-to-end fault monitoring with:
    - Alerts and Alarms through CLI to detect faulty devices.
    - Automatically activated LEDs to indicate bad disks.

Storage management is a way of managing disks within an S-Card that are internal to a chassis or managing disks within a storage enclosure that are external to a chassis. The chassis contains up to eight S-cards, and eight disks within each S-card. It can also have one or more storage enclosures physically connected to each of its S-Cards.

With the addition of storage enclosures, the storage capacity of a system can be expanded beyond the eight S-Cards. In addition, RAID levels may be set for each S-card. RAID is available in volume management mode, and Non-RAID is available with raw disk and volume management modes.

To understand storage management, it is important to be familiar with the storage components, such as the S-card, disks, storage enclosures, and the concept of RAID, Hot Spares, and storage management modes that include raw disk management and volume management.

# S-Cards

The SM15K chassis can hold a combination of the following HDDs or SSDs, SAS or SATA drives:

- Up to 8 S-cards per chassis, and up to 8 internal disks per S-card. Therefore, a total of 64 internal disks per chassis.
- Up to 1 Fabric Storage Enclosure per S-card, and therefore, up to 8 Fabric Storage Enclosures per chassis in the volume management mode.
- Up to 2 Fabric Storage Enclosures per S-card, and therefore, up to 16 Fabric Storage Enclosures per chassis in the raw management mode.

# Disks

As mentioned before, up to 8 disks reside in an S-card, and a total of 64 disks within a system. In addition, there are disks within a storage enclosure to extend the storage capacity of a system. All disks within an S-card are operational when the LED on each disk is green.

Disks are physical disks present in the system in one of the following classifications:

- Raw disk: A physical disk assigned to a server in raw access mode.
- Pool disk: A disk that is configured to be a part of a pool in volume management mode.
- Hot Spare disk: A disk that is designated as RAID 'hot spare' with the RAID S-card.
- Free disk: A disk that is not assigned to a server or not configured in a pool.
- Missing disk: A physical disk not currently discovered, but one that was previously discovered.
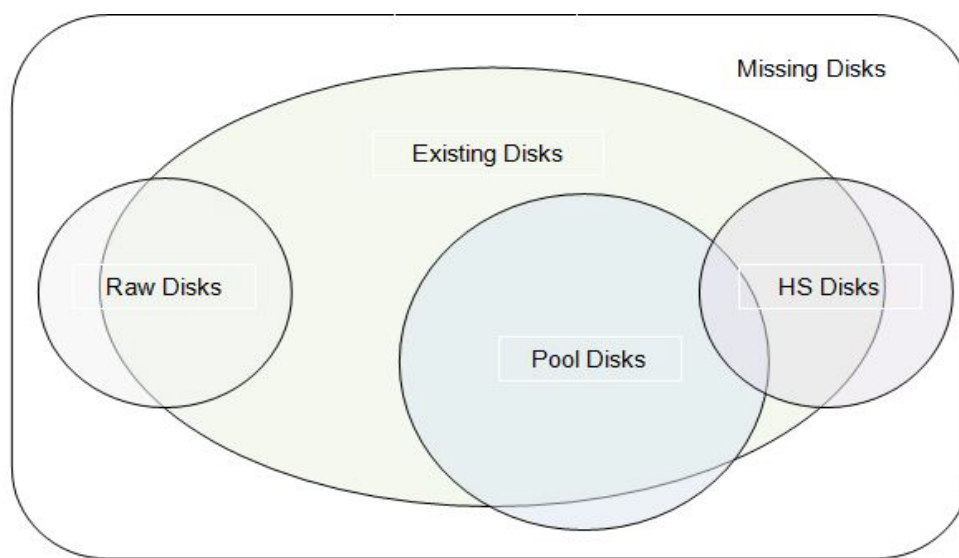
## Disk Identification

Disks may be identified by location or by Unique Identification of Disks (UUIDs).

- Location identification is available for raw disk management. If the location of a disk changes, the identifier will reflect the new location.
- Internal disks are identified in CLI as <bay> or <slot/<bay>
- External disks are identified in CLI as:
    - <device-name> device <device-id > slot <slot>
    - <scard slot>/<SAS port#>/<shelf #>/< disk slot>
        - For 2RU12 and 2RU24, the shelf# will be 0.
        - For 5RU84, the shelf# will be 0 or 1 as it has two drawers.

- Disk UUID identification is available for raw disk and volume management. It is location-independent, and is not available on RAID-enabled S-cards.
- For using an external disk for configuration, assignment, or display purposes, the device name needs to be explicitly configured using the command:

    **storage external** *device-name* **device** *device-id* **slot** *slot / bay*

    On startup, the device name of all discovered devices will be the same as their device id, but their slot will need to be configured explicitly before attempting to use any disks of that device.

    Note: During configuration, when assigning external disks, a name for the external device needs to be explicitly configured on a server.

- A disk may be powered on or off using the **power on disk** or **power off disk** command, along with internal or external disk identification.

---

# Storage Enclosures

Storage enclosures extend the storage capacity of a SeaMicro system. Each controller on the enclosure connects to <u>one</u> S-card port through SAS ports on the MX-card. With fully integrated monitoring of PSUs, IO Controllers, fans, and temperature sensors, the storage enclosures operate at an optimum level. The storage enclosures and their specifications are listed below.

- Seamicro Freedom Fabric Enclosure FS 2012-L
- Seamicro Freedom Fabric Enclosure FS 2024-S
- Seamicro Freedom Fabric Enclosure FS 5084-L
- Seamicro Freedom Fabric Enclosure FS 4060-L

|  | FS 5084-L | FS 2012-L | FS 2024-S | FS 4060-L |
|---|---|---|---|---|
| Features | High Capacity | Low Cost | Optimized Performance | Optimized Density and Performance |
| Height (RU) | 5RU | 2RU | 2RU | 4RU |
| Disk Count | 84 | 12 | 24 | 60 |
| Disk Types | 2.5"/3.5" SAS/SATA | 3.5" SAS/SATA | 2.5" SAS/SATA | 3.5" SAS/SATA HDDs |
| Controller | Dual HA Storage Bridge Bay (SBB) 2.0 Compatible Controllers | | | Dual, hot-swappable, controllers. |
| Interfaces | Three x4 6 Gb mini-SAS connectors per controller | | | Four x4 Gb mini-SAS port per SIM |
| Max Storage per Enclosure* | 336 TB | 48 TB | 24 TB | 240 TB |
| Max Storage per SM15K | 5,376 TB (5.3 PB) | 768 TB | 384 TB | 3,840 TB (3.84 PB) |

\* Based on 4TB 3.5"and 1TB 2.5" hard disk drives (HDD).

## Release 3.4:

With Release 3.4, SeaMicro offers increased storage capacity per SM15K as follows:
- Extended support for Fabric Storage Enclosures from 8 to 16 per SM15K in 'raw disk' mode only.
- Expanded possibility to connect 16 different storage enclosures to existing 16 SAS ports across 8 MX-cards.

## Storage Enclosure Naming Convention

The Storage Enclosure is represented in CLI commands and alarms in the following format: **ScardID/EnclosureID**, where S-card = Slot 0 to 7; Enclosure ID = 0 or 1.

The left-most enclosure is numbered **0.**

- In the rear of the enclosure, the left-most power supply, or fan, or IO controller, or IO port is numbered **0.**
- Temperature sensors are numbered as **0** to **n-1**.

The storage enclosure can be managed with CLI commands, and alerts, alarms and CLI commands help to troubleshoot and isolate enclosure issues.

Note:
- Any time you remove a storage enclosure from a chassis or reinsert a storage enclosure into a chassis, make sure to stop all I/O operations to the affected storage enclosure.
- During a storage enclosure hotswap, when raw disks are mapped to a server, switching a storage enclosure I/O port will cause a logical disk id change, making previous storage enclosure disk assignments inactive. In such a case:
    - Unassign current assignments corresponding to all physical disks in the enclosure.
    - Reassign them to servers.

# RAID

Integrated Redundant Array of Independent Disks (RAID) technology allows the administrator to manage data loss and recovery in the event of disk failure. There are multiple RAID levels, and depending on the RAID level, the number of supported disks will vary. RAID can be fully managed through the SM15K user interface.

- The RAID level configuration is unique per S-card, and is persistent.
- The available RAID levels are 0, 1, 5, 6, and 10. The default is level 5.
- With Release 3.4, mirroring enables the coexistence of multiple RAID levels per S-card. The two configurations supported are:
    - RAID 1 and RAID 5
    - RAID 1 and RAID 6
- The same chassis can have S-cards in different RAID levels.
- The RAID level cannot be changed if a pool is defined on an S-card. Note that the RAID level you want to change to should be a super set of the RAID level(s) of the configured pool(s).
- Before you change the RAID level of an S-card:
    - All volumes should be unassigned.
    - All pools should be removed.
- With integrated RAID, server data is protected on the S-card, and all *pools* and *vdisks* are automatically protected per the RAID level on that S-card.
- Backward and Forward Compatibility of RAID across Releases:
    - If Release 3.3 is replaced with Release 3.4 or newer, the existing RAID level will be inherited.
    - If Release 3.4 is replaced with Release 3.3 or older, the first RAID level of the RAID levels configured in Release 3.4 will be applied.

## RAID Considerations

- SeaMicro supports multiple RAID levels across one S-card, and all enclosures attached to that S-card. All pools and volumes, therefore, will inherit the specified RAID levels.
- Once you create a pool at a certain RAID level, making any changes to the RAID level will result in an error message. If you need to change the RAID level, it is best to delete the pool, and start again with the setting of a new RAID level.
- You can create multiple RAID pools in a given S-card, and these pools will share the 'global Hot Spare (HS) disks' across all pools on the S-card. However, as with pools, HS disks cannot be shared across S-cards.
- If a RAID daughter card is installed in the S-card, the S-card disk hotswap behavior will be different.
- The RAID/non-RAID mode is global to a given S-card hardware configuration. It does not change dynamically at runtime.
- The RAID stripe size is global to the S-card. All pools in the S-card inherit the stripe size. The default stripe size is 256KB. Set up stripe size based on the application running on your C-card.

- Note that it is not possible to move a storage enclosure from a RAID S-card to a non-RAID S-card.

To set a RAID level per S-card, enter the following command. The default level is '**5**' for RAID-capable S-cards.

| Task | Command | CLI Level |
|------|---------|-----------|
| Set RAID level. | **storage set raid level** {0\|1\|5\|6\|10\|1,5\|1,6} **slot** *slot* | Privileged |

## RAID Levels

The SeaMicro system can be configured in RAID levels 0, 1, 5, 6, and 10. See table below for a description of each RAID level.

| RAID Level | Description |
|:----------:|-------------|
| **0** | If RAID level **0**, called disk striping, is configured, the number of elements in the disk-list should be at least two, but more commonly three or more. The specified disks will be grouped for the RAID function. This command can be invoked multiple times as long as there are spare disks to add. |
| **1** | If RAID level **1**, called disk mirroring, is configured, the number of elements in the disk-list should be two. The two specified disks will be grouped for the RAID function. This command can be invoked multiple times as long as there are spare disks to add. |
| **5** | If RAID level **5**, which uses disk striping with striped parity, is configured, the number of the list elements should be at least three. This level is the most commonly used level, because it achieves a good balance between performance and availability. The specified disks will be grouped together for the RAID function. This command can be invoked multiple times with the same or different number of list elements. |
| **6** | If RAID level **6**, which uses disk striping with two dedicated parity disks, is configured, the number of the list elements should be at least four. This level is good for storing data for long periods of time. The specified disks will be grouped together for the RAID function. This command can be invoked multiple times with the same or different number of list elements. |

| RAID Level | Description |
|:---:|:---|
| 10 | If RAID level **10** is used, which is a combination of RAID levels **0** and **1**, an even number of drives must be selected. This level requires a minimum of four disks to complete disk striping and mirroring to provide maximum reliability and redundancy. |

# Hot Spares

Combined with the RAID function, Hot Spare (HS) is a failover mechanism through which a damaged pool can be dynamically detected and repaired to optimal status by removing a failed disk from the pool, and joining it instead with one of the HS disks. Through the process of mirroring and rebuilding, data is copied onto another disk, thereby preserving data and enabling data recovery.

In the event of a disk failure, HS disks become active disks. HS configuration is per S-card slot, and persistent. By default, the HS function is **disabled**. At least one hot spare disk per pool is recommended.

- A user may designate one or more HS disks.
- HS disks are global to each S-card, with no specific limits on the number of HS disks.
- HS disks automatically replace failed disks protecting assigned volumes and reducing the risk of data loss and recovery.
- HS disks rebuild volume assignments without manual intervention upon drive failure.
- New HS disks may be added to the pool to bring the hot spare count to the original number of disks.
- If there is no free HS disk whose size is greater than or equal to that of a failed disk, the replacement will not occur, and all data on the failed disk will be lost.

| Step | Task | Command<br>Keyword and Variables |
|:---|:---|:---|
| 1 | Configure Hot Spares.<br><br>This command allows the user to specify one or more disks as hot spare disks within an S-card in a specified slot. | **storage set raid hot-spares** (**none**\| (**list** *disk-list*) \| (**count** *disk-count*) \| *auto*) **slot** *slot* |
|  | The default **'None'** disables the hot spare function. | Default: **None**. |

# Storage Management Mode

The chassis allows two modes of storage management - volume management and raw disk management.

- Volume Management, which offers logical storage access for flexible capacity and re-provisioning for web hosting and Cloud Iaas. This mode is the default storage management mode.
- Raw disk management, which offers raw, physical storage access for applications requiring low overhead, such as Hadoop and cloud storage.

The difference between the two modes is shown in the table below:

| Volume Management | Raw Disk Management |
| --- | --- |
| Flexible capacity with pools and volumes. | Low overhead disk access with no metadata written to disk. |
| Volumes are transparently assigned to a server as direct-attach SATA drives. | Whole disk assigned to server as direct-attach SATA drives. |
| Storage portable across SeaMicro SM15K fabric servers only. | Storage portable across SM15K and any industry standard server. |
| Both modes are agnostic to OS running on the C-card. | |
| Note: Raw disk management mode is not supported on RAID S-cards | |

## Display Management Mode

The output for '**show storage scard** will display '**disk**' for raw disk management mode, and '**volume**' for volume management mode.

Example:

```
seamicro# show storage scard brief
slot      mgmt. status   mgmt. mode   raid level   stripe size   hot-spare count
----------------------------------------------------------------------------
0         up             disk         none
1         up             volume       none
```

## Display Volume/Raw Disk Assignments

To check if volumes or raw disks have been assigned to a server, run '**show server detail**.' The output will show '**disk**' if mapped to a raw disk, and '**volume**' if mapped to a volume.

Example:

```
seamicro# show server detail
Server Number 0/0 is  down
 CPU Type : Intel Atom N570
 UUID : 7787b70e-d3b1-3f76-ae71-c29a662f9f74
         Description:
         Port80 Code: 0xff
 NIC(s) : 2
         NIC Number: 0
                 MAC Address: 00:22:99:03:80:00
         NIC Number: 1
                 MAC Address: 00:22:99:03:80:01
 Vdisk(s): 1
         Vdisk: 0 Disk: 0/0 Size: 256GB Serial number: 5VJ2P2J4

Server Number 0/1 is  down

 CPU Type : Intel Atom N570
 UUID : 7afae3e4-9491-3976-adba-b81fd83401a0
         Description:
         Port80 Code: 0xff
 NIC(s) : 2
         NIC Number: 0
                 MAC Address: 00:22:99:03:80:02
         NIC Number: 1
                 MAC Address: 00:22:99:03:80:03
 Vdisk(s): 1
         Vdisk: 0 Volume: 1/nicepool/vol-2 Size: 256GB Serial number:
tCHrorgKeSR4foKr0qFX
```

## Set DISK I/O Mode

Set disks in write-through mode or write-back mode based on specific performance or data protection needs. The command to set disk I/O mode is as follows:

| Task | Command | CLI Level |
|------|---------|-----------|
| Sets disk input/output mode. | **storage set disk-io mode {write-through \| write-back \| nofsync-write-back} slot** *slot* | Privileged |
| | Default mgmt-mode: **write-back** | |

**write-through:** Enter this option to filter fsync, and disable vdisk write caching. This mode ensures data is always written to the disk by disabling the write-cache feature on the S-card and disabling write-cache on the physical disks attached to the S-card.

This mode is most suitable for applications that need the S-card to ensure data consistency. Fsync commands are dropped in this mode, since data is always written to the physical disk.

**write-back**: Enter this option to honor fsync, and enable vdisk write-back. This mode improves I/O performance by caching write-data on the S-card and by enabling disk write-cache.

This mode is set when applications can manage data consistency at the application layer. In this mode, applications can use the fsync command to flush data in the cache to the physical disk. This mode is the **default mode** of operation.

**nofsync-write-back**: Enter this option to filter fsync, but enable vdisk write-back. This mode improves I/O performance by caching write-data on the S-card and by enabling disk write-cache.

This mode is set when applications need weak data consistency. When the S-card is set in this mode, all fsync commands are ignored. This mode is the same as write-back, except the fsync commands are ignored in this mode.

To verify the current configured mode, use the '**show storage scard**' command as shown in the example below:

```
seamicro(config)# show storage card
mgmt.status: up
disk io mode: write-back
```

**Note:**

- This command will cause the specified S-card to restart.
- A warning message will display, and user confirmation will be required.
- For a RAID S-card, only write-back mode will be allowed, and a disk I/O mode change to any other will be denied.

The above sections contained general information on storage, storage components, and terms. Now, proceed to the appropriate section to configure, delete, display, and troubleshoot storage.

# Configuring Storage

All storage configuration is saved in a central configuration file. This configuration is applied to subsystems and/or software modules during system initialization. Only storage configuration is managed and saved by the MX Card or the management plane. Storage configuration involves using CLI commands in both the **Privileged** and **Configuration** modes.

Before you configure and manage storage in a system, decide if you will work with **raw disk** or **volume** management, or both.

- The system by default is in volume management mode.
- Use the **storage set mgmt-mode** command as listed below to switch from volume management to raw disk management, and vice versa.
- For raw disk management, the S-card must be a non-RAID S-card.
- Note that each S-card can be in a different storage management mode.

| Task | Command | CLI Level |
|---|---|---|
| Select storage management mode. | **storage set mgmt-mode {disk \| volume} slot** *slot* | Privileged |
| | Default mgmt-mode: **volume** | |

In order for the switch from volume management to raw disk management to be successful on the S-card:

- There should be no active assignments of disks or volumes on the S-card.
- If there is an active assignment, this command will result in an error unless issued with the '*force*' option.
- User confirmation will be required as there is a possibility of data loss on disks.
- Once the switching is complete, note that any existing pool, volume, or data may or may not be recoverable.
- Assignments will be activated or deactivated based on the availability of associated disk or volume.

# Raw Disk Management

Use the following command to assign raw disks to a server.

| Task | Command | CLI Level |
|---|---|---|
| Assign raw disks to a server. | Single disk:<br>**storage assign** *server-id vdisk-number* **disk** *disk-id* [**read-only**]<br><br>Multiple disks:<br>**storage assign-range** *server-range {vdisk-range \|* **all***} vdisk-number {***disk** *disk-list \|* **internal-disks** \| **external-disks** \| **all***}* [**share**] [**read-only**]<br><br>Note:<br>*<slot>/<bay>* for an internal disk.<br>*<configured device name>/<bay>* for an external disk. | Configuration |

**Note**: Disks assigned as raw disks through CLI will be seen as a '**raw**' device on the C-card .

# Volume Management

Volume management is the default mode in the system. It involves creating a pool of disks, creating a volume in a pool, and assigning volumes to servers as shown in Figure 2, Volume Management. In this storage management mode, you may provision storage automatically or configure storage manually in the system.

**Figure 2** Volume Management
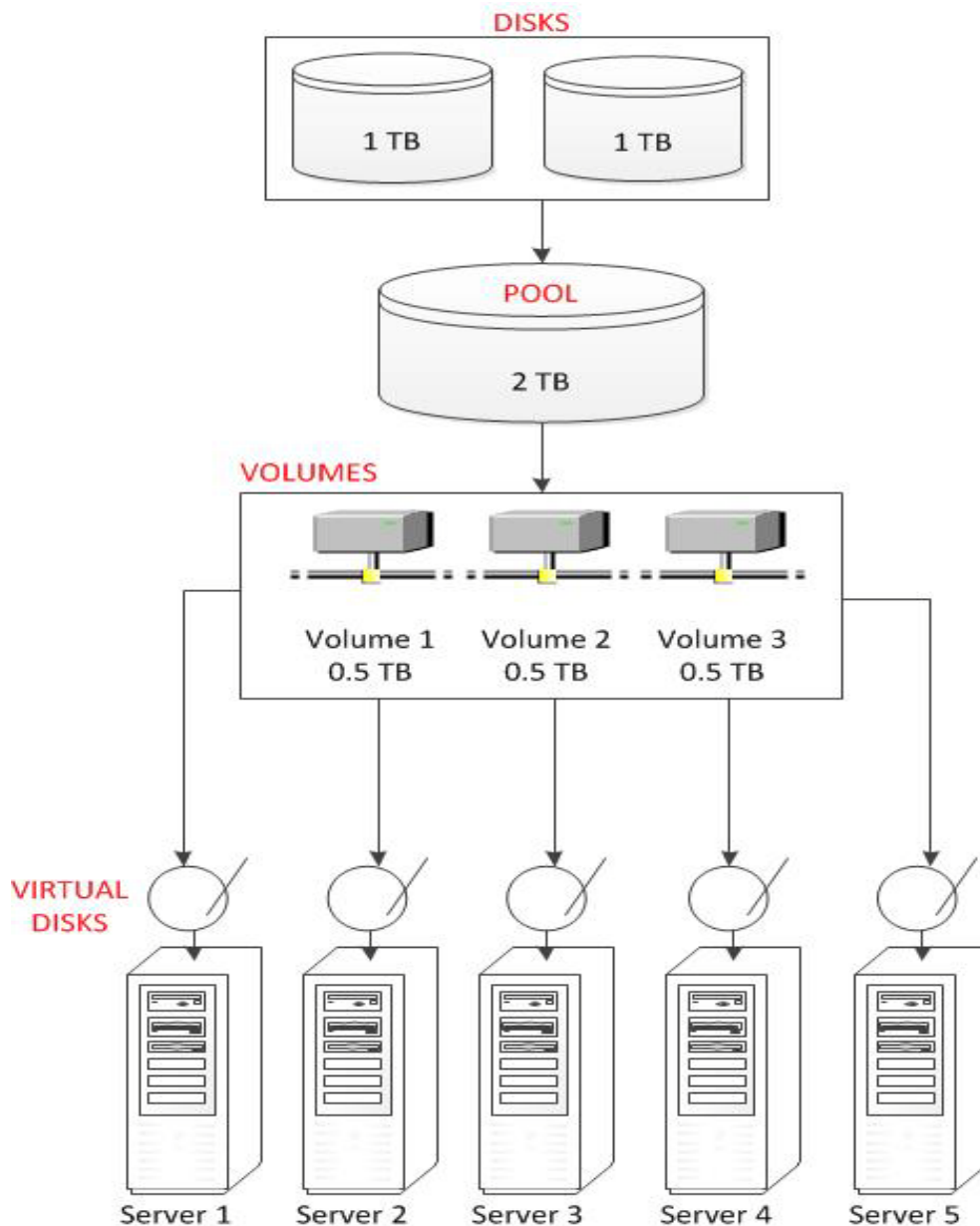
# Pools and Volumes

## Pools

A pool is simply a group of disks pooled together. An administrator needs to be familiar with storage pools to perform failure and performance isolation. Pool and volume information are always saved as meta data on the disk. In the event of a system reboot, the pool and volume information will be automatically retrieved from the meta data on the disk.

There are three ways to create a pool:

- Create a pool using the disks within an S-Card.
- Create a pool exclusively using disks within a storage enclosure.
- Create a pool with a combination of a disk or disks within an S-Card and a disk or disks within a storage enclosure.

Note: A pool can be created only within an S-Card, and not across S-Cards. However, *vdisks* assigned to servers can belong to one or more S-Cards. Assigning a *vdisk* to a server is the same as assigning a volume within a pool.

In Release 3.3, the capability to configure pools spanning across multiple physical disks in a non-RAID S-card is deprecated, and is strongly discouraged as it may not be available in future releases.

**Figure 3**  Creating Pools
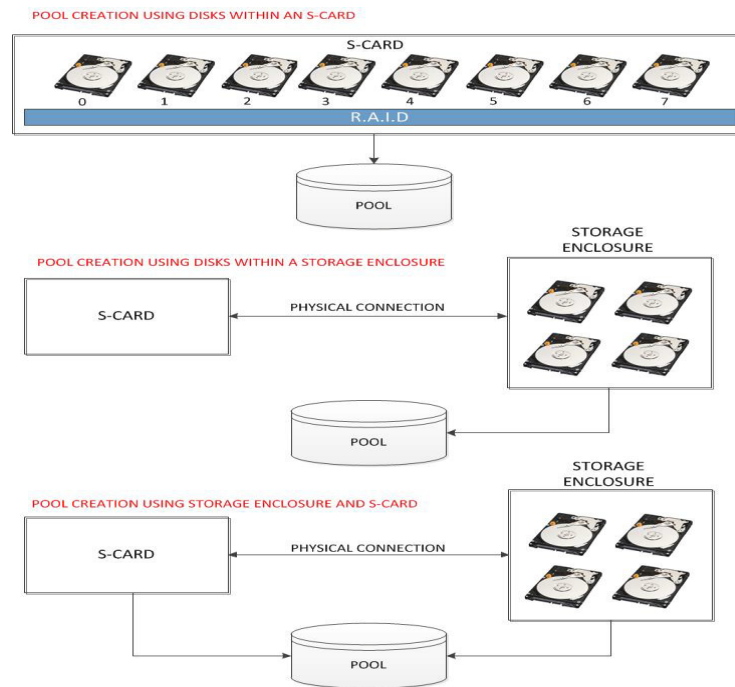
# Volumes

Assigning a volume is a way of assigning space within a pool of disks to be accessed by various applications. Volumes show up as individual *vdisks* in a server. The storage management model allows a server to be associated with up to 32 *vdisks*. A volume can be shared by multiple servers, and data within the volumes may be distributed in a pool across multiple disks. Note, a volume cannot be deleted if it is assigned to a server.

In Release 3.3, the following assignments are possible:

- A maximum of 64 assignments per disk or volume.
- A maximum of 512 disk or volume assignments per S-card.

**Figure 4**   Assigning and Sharing Volumes among Multiple Internal Servers
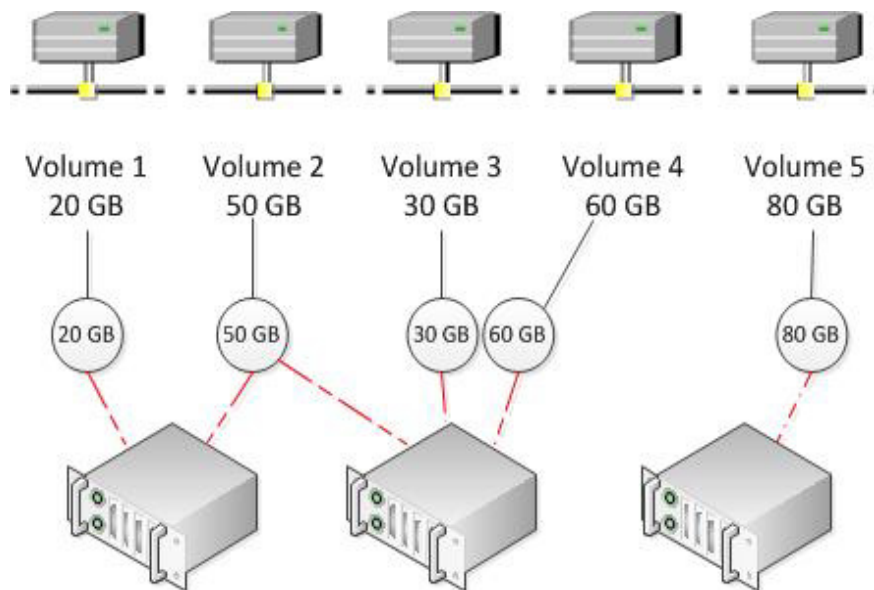


Note: When disks in volume mode are re-used in the raw disk mode:

- Creating pools/volumes on disks using the LVM on S-card will show up as pools/volumes on C-cards in the raw disk mode.
- Creating pools/volumes on disks using LVM on C-card will show up as pools/volumes on S-card in the volume mode.

# Provisioning Storage

To provision storage is to automatically configure storage for internal or external drives.

To automatically create pools and volumes or to provision storage in a pre-defined manner, use the **storage provision** command. This command allows for one-step completion of system-wide or slot-wide provisioning of storage. The maximum number of volumes that **storage provision** will create per S-card is **384**. In the case of storage enclosures, storage provisioning can occur even when an enclosure is added after the system is up and running.

With the **storage provision** command, you can:

- Create a 1:1 or many:1 mapping of physical drives to *vdisks*.
- Assign RAID level with Hot Spares on all RAID capable S-Cards.
- Run storage provision multiple times.
- Provision storage on storage enclosures.
- Assign volumes across a server or a range of servers.

**Note**:

- As storage provision is used mainly to create pools and volumes, it is not applicable for raw disks.
- A stripe size configured by the '**storage set raid stripe-size**' command will be used for all pools created by the '**storage provision**' command on a RAID S-card.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Automatically create pools and volumes within a pool. <br><br>With RAID: | **storage provision** *size* (*gigabytes*\|*max*) **slot** (*slot-range*\|**all**) **disk** (*internal-disks*\|*external-disks*\|**all**) **method** (**pool-per-disk**\|**pool-per-slot**) [**raid-level** (0\|1\|5\|6\|10) [**stripe-size** (16 \| 32 \| 64 \| 256 \| 512 \| 1024)] \|\| [**hot-spares** *number-of-hot-spare-disks*]] **pool-name** (*pool-name*\|*pool-name-prefix*) **volume-name-prefix** *volume-name-prefix* [no-confirm [**background**]] | Privileged |
| | Without RAID: | **storage provision** *size* (*gigabytes*\|*max*) **slot** (*slot-range*\|**all**) **disk** (*internal-disks*\|*external-disks*\|**all**) **method** (**pool-per-disk**\|**pool-per-slot**) **pool-name** (*pool-name*\|*pool-name-prefix*) **volume-name-prefix** *volume-name-prefix* [no-confirm [**background**]] | |

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 2 | Assign a single disk or volume to a single server or multiple servers. | Single Server<br><br>**storage assign** *server-id vdisk-number* **volume** *volume-specifier* [**uuid** *volume-uuid*] [read-only]<br><br>**storage assign** *server-id vdisk-number* **disk** *disk-id* [read-only]<br><br>Multiple Servers<br><br>**storage assign-range** *server-range* {*vdisk-range*|**all**} **volume** *range-volume-specifier* **[start** *volume-start-number*] **[share] [read-only]**<br><br>**storage assign-range** *server-range* {*vdisk-range* **\| all**} **disk** *disk-list* \| *i***nternal-disks** \| **external-disks** \| **all**} **[share] [read-only]** | Configuration |

# Configuring Storage Manually

If you choose to manually configure storage, you may create pools, create a volume or volumes with a pool, and assign volumes to servers.

- If you have RAID enabled S-cards, set the **RAID** level, required **stripe size**, and **hot spare count** before you create pools and volumes within a pool.
- Or, set the RAID level within a pool with the '**storage create pool'** command as shown in the table below.
- Note that the provided RAID level should be one of the RAID levels supported by the S-card in the slot. This command will fail:
    - If a RAID level is provided for a non-RAID S-card.
    - If a RAID level is not provided for a RAID S-card.
    - If the specified RAID level is not one of the configured RAID levels.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Create a pool within a physical disk or disks with or without RAID. | Non-RAID<br>**storage create pool** *slot/pool-name* **disk** *disk-list*<br><br>RAID<br>**storage create pool** *slot/pool-name* **disk** *disk-list* [raid-level (0\|1\|5\|6\|10)] | Privileged |
| 2 | Create a single volume or multiple volumes within a pool. | Single Volume<br><br>**storage create volume** *slot/pool-name/volume-name* **size** {*gigabytes* \| **max**} [**background**]<br><br>Multiple Volumes<br><br>**storage create volume-prefix** *slot/pool-name/volume-name-prefix* **size** {*gigabytes* \| **max**} [**start** *volume-start-number*] [**count** *number-of-volumes*] [**background**] | Privileged |

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 3 | Assign a single disk or volume to a single server or multiple servers. | Single Server<br><br>**storage assign** *server-id vdisk-number* **volume** *volume-specifier* [**uuid** *volume-uuid*] [read-only]<br><br>**storage assign** *server-id vdisk-number* **disk** *disk-id* [read-only]<br><br>Multiple Servers<br><br>**storage assign-range** *server-range* {*vdisk-range*|**all**} **volume** *range-volume-specifier* [**start** *volume-start-number*] [**share**] [**read-only**]<br><br>**storage assign-range** *server-range* {*vdisk-range* | **all**} **disk** *disk-list* | *internal-disks* | **external-disks** | **all**} [**share**] [**read-only**] | Configuration |

Note: Alerts are generated when volumes or disks are shared between servers. For example:

```
2013-06-02 09:22:47   Disk      1/0            Disk assignment      this disk is shared by servers
2013-06-02 03:13:35   Volume  1/Mixed/vol-0  Volume assignment    this volume is shared by servers
```

## Adding or Removing Disks

After configuring storage, if there is a need to add or remove disks, follow the steps outlined below:

To add new disks to a pool:

- Create a new pool with new disks, or add disks to an existing pool to increase its capacity.
- Then, create volumes within the pool, and assign the volumes to servers.

Starting with Release 3.4, when adding a disk, the RAID level of the existing pool will be used rather than the RAID level of the S-card. To add disks, use the following command:

| Task | Command | CLI Level |
|------|---------|-----------|
| Add disks to a pool. | **storage add disk** *disk-list* **pool** *slot/pool-name* [**no-confirm**] | Privileged |

To remove disks from a pool:

- Un-assign any volumes from the server.

- Un-mount the pool using the **storage umount pool** command.
- Remove the disks from the pool.

**Note:**

Any time you remove or insert a disk, make sure to stop all I/O operations to the affected storage disk.

## Renaming a Pool

To rename an existing pool, follow the steps outlined below:

- Un-assign any volumes from the server.
- Rename the pool.

To rename the pool, use the following command:

| Task | Command | CLI Level |
|---|---|---|
| Rename a pool. | **storage rename pool** slot/pool-name **to** new-pool-name | Privileged |

## Assigning a Unique Name to a Storage Enclosure

To assign a unique name to an external disk device address, and indicate which slot the external device is physically connected to, use the following command.

| Task | Command | CLI Level |
|---|---|---|
| Assign a unique name to an external storage enclosure. | **storage external name** *external-disk-device-name* **device** *external-disk-device-address* **slot** *slot* | Configuration |

The "external disk device" refers to a storage enclosure, and the "external disk device address" refers to a storage enclosure address. Typically, an enclosure address is a string of numbers, such as 11:22:33:44:55:66:77:88. With the storage external command, the string of numbers can be replaced with a name for all discovered storage enclosures.

Also, you can always pre-configure names of storage enclosures in a particular slot so users can verify if the storage enclosures are inserted correctly in the configured slot.

## Moving an S-card between Management Modes

- When an S-card is moved from volume management to raw disk management, pools and volumes will show up as pools and volumes on the C-card.
- When an S-card is moved from raw disk management to volume management, the pools and volumes will show up as pools and volumes on the S-card.

## Moving Storage across the Chassis

**Volume Management - To move storage across the SM15K:**

- Power down the existing SM15K.
- Move the storage enclosure to the new SM15K.
- Verify that pools and volumes have carried over from one chassis to another.
- Re-assign volumes on servers or issue secure copy (SCP) command to copy assignments from one chassis to another.

**Raw Disk Management - To move raw disks from one chassis to another:**

- Make sure the S-card on the other chassis where the disk is inserted is in raw disk mode. If it is in volume mode, you will not be able to access the data if any on the raw disk.
- If you do not want the server on the new chassis to be able to see the data, make sure the disk is formatted before the move.

**Raw Disk Management - To move the S-card from chassis to another:**

- When the entire S-card is moved from one chassis to another, the S-card will be present in raw disk mode. Assign the raw disks to the intended servers.
- If there are any partitions/data present on those disks, the server which has the raw disk assignment will be able to still access that data.
- If you do not want the server on the new chassis to be able to see the data, make sure the disk is formatted before the move.

## Moving a Storage Enclosure

**Volume Management: To move a storage enclosure from one S-Card to another S-card:**

- Run **show storage** command to identify configured volumes.
- Review assignments, and delete all old assignments.
- Re-assign volumes to servers with the **storage assign-range** command.

Note: Even after swapping a storage enclosure from one slot to another, the <previous slot>/pool may still show up in '**show storage pool brief**' marked as 'failed.' You may ignore it, or delete the entry using the '**delete pool**' command options.

**Raw Disk Management: To move a storage enclosure from one chassis to another:**

- Make sure the storage card on the other chassis where the storage enclosure is moved is already in raw disk mode. If it is in volume mode, you will not be able to access the data if any on the storage enclosure disks
- If you do not want the server on the new chassis to be able to see existing data, please make sure all the disks are formatted before the move.

# Powering off or Removing S-cards, Disks, or vdisks

A few important considerations when S-cards and disks are powered off or removed, and when raw disks fail:

## Powering off or Removing the S-card

- In a managed S-card removal, when you power-off the S-card, it will first power down the physical disks causing the assignments to become inactive, and the C-card OS will see a disconnect event.
- Once the S-card is re-inserted and powered on, the disks will be automatically discovered, and the C-card OS will see a disk connection event.
- In an un-managed S-card failure or removal, such as when an S-card is removed without powering off first, the management system will detect the S-card removal in about 60 seconds, and will mark all disks as inactive, and the OS will see a disconnect event.

## Powering off or Removing Disks

- When a disk is removed, powered off or fails, it will appear to the C-card server and the Operating System (OS) as a SATA disconnect. The associated interruption is communicated to the OS, and will be signaled back to the OS when the disk is re-inserted.
- When a disk is replaced, the new disk will be automatically mapped back to a server as a raw disk assignment
- If the same disk is powered back on, the server will see the disk reappear

## Powering off or Removing the Vdisk

- Vdisks are virtual disks assigned to servers.
- Its assignments are managed by system management on the MX card.
- Vdisks on the C-card are mapped to the actual physical disks.
- When a vdisk assignment is removed, it will be signaled to the C-card OS as a disk power-off/removal.
- When a vdisk assignment is reconfigured. the C-card OS will see the event as a power-on/insertion of the physical raw disk.

## Upgrading Disk Firmware

To upgrade specified disk firmware within an S-card, enter the following command:

| Task | Command | CLI Level |
|------|---------|-----------|
| Upgrading disk firmware by model. | **upgrade disk scp:** *ip-address user fiirmware_file disk_model* | Privileged |

- The firmware file is stored on a remote host. When the upgrade starts, the chassis executes a secure copy of the file from the specified IP address, and applies it to target disks.
- When the firmware update is complete, the new firmware version will display in various storage-related '**show**' commands.
- However, if disk firmware is updated using third party tools for specific disk vendors via debug commands, then restart the disk management process on the related S-card, so that firmware information is correctly displayed on disk and S-card-related '**show**' commands.
- To restart the disk management process on related S-cards, enter the following command:

| Task | Command | CLI Level |
|------|---------|-----------|
| Restarting S-card Process. | **restart scard-process {config | cpu-temp-writer | dataproc | ntp-daemon | pericom-monitor} slot** *slot number* | Privileged |

## Upgrading Storage Enclosure Firmware

To upgrade storage enclosure firmware, enter the following command:

| Task | Command | CLI Level |
|------|---------|-----------|
| Upgrade storage enclosure firmware. | **upgrade storageEnclosure scp:** *ip-address user-filename scardID IenclosureID* | Privileged |

- Enter the IP address of the node.
- Enter the name of the file you want to copy.
- Enter the id of S-card and the storage enclosure.

Note: The enclosure will reboot as required, possibly affecting disk operations. This command may take up to 10 minutes to execute completely.

# Deleting Storage

Disks, pools, and volumes can be deleted or removed from the system with simple CLI commands. Deleting storage is different in each mode.

| Volume Management | Raw Disk Management |
|---|---|
| First, delete pools and volumes. | Simply, un-assign disks. |
| Next, mount or un-mount pools. | |

## Delete Pools and Volumes

Using the following commands, you may:

- Delete an individual pool or volume or delete all pools and volumes from an S-card.
- Remove all assignments corresponding to the deleted pools or volumes on an S-card.

Note: You have to remove all assigned volumes from servers using the "**storage assign clear**" command before you delete any pools and volumes.

| Task | Command | CLI Level |
|---|---|---|
| Clear disks or assigned volumes from servers. | **storage assign-clear** *server-range vdisk-range* | Configuration |
| Delete a pool. | **storage delete pool** *slot/pool-name* | Privileged |
| Delete a volume or multiple volumes. | Single Volume:<br><br>**storage delete volume** *slot/pool-name/volume-name* [**background**]<br><br>Multiple Volumes:<br><br>**storage delete volume-prefix** *slot/pool-name/volume-name-prefix* [**start** *volume-start-number*] [**count** *number-of-volumes*] [**background**] | Privileged |
| Remove all pools and volumes. | **storage clear** | Privileged |
| Clear metadata on disk.<br><br>**Note:** This command is only available in Release 3.2.2 and future releases. | **storage clear-metadata disk** *<number>* **slot** *<number>* | Privileged |

## Mount or Un-mount Pools

You may mount or un-mount pools and volumes from an S-card using the following commands:

| Task | Command | CLI Level |
|---|---|---|
| Mount pools | **storage mount pool** [**slot** *slot*] | Privileged |
| Un-mount pools | **storage umount pool** *slot/pool-name* | Privileged |

## Unassign Disks

To delete storage in raw management mode, simply un-assign disks from servers.

# Displaying Storage

With specific "**show storage**" commands, you can display storage information on S-cards, disks, pools, volumes, assignment to servers, and external devices.

The "**show alerts"** command is only available in **Release 3.2.3 and future releases**. With the "**show alerts**" command, you can display alerts on faulty storage components, such as down S-cards, disks, expanders, and storage enclosures.

All storage-related "**show "**commands are listed in the table below. Refer to the *CLI Guide* for more information on all commands, options, and variables.

| Function | Command | CLI Level |
|---|---|---|
| Displays pool, volume, and disk information. | **show storage** | Privileged |
| Displays disk or volume assignments of servers and their status.<br><br>The '**type**' column indicates if a disk or volume is assigned to the server. The '**id**' column indicates either an internal or external disk id, or a volume id.<br><br>The non-brief version of this command displays the uuid if a volume is assigned, or the raw disk signature if one is assigned. | **show storage assign** | Privileged |
| Displays disk information. It will also display whether the disk is assigned as a raw disk. The string '**<raw>**' will be printed in the pool column or line.<br><br>It will also display the external disk location. | **show storage disk\*** | Privileged |
| Displays an alias table for external disk device addresses. | **show storage external** | Privileged |
| Displays a list of CLI commands in progress in the background. | **show storage jobs** | Privileged |
| Displays storage pool information. | **show storage pool** | Privileged |
| Displays S-card status and information. It will also display the configured management mode for each S-card. | **show storage scard** | Privileged |

| Function | Command | CLI Level |
|---|---|---|
| Displays statistics for volumes, servers, and disks. | **show storage statistics** | Privileged |
| Displays storage volume information for a single volume or multiple volumes. | <u>Single Volume:</u><br><br>**show storage volume**<br><br><u>Multiple Volumes:</u><br><br>**show storage volume prefix** | Privileged |
| Displays all storage-related configurations. | **show running-config storage** | Privileged |
| Display alerts on the S-card, disk, expander, and storage enclosures.<br><br>**Note:** This command is only available in Release 3.2.3 and future releases. | **show alerts** [**begin** *date-time*] [**end** *date-time*] [**reverse**]<br><br>**show alerts scard** (*slot-range* \| **all**) [**begin** *date-time*] [**end** *date-time*] [**reverse**]<br><br>**show alerts disk** *slot/ (ba*y \| **all**) [**begin** *date-time*] [**end** *date-time*] [**reverse**]<br><br>**show alerts disk** *external-disk-device-name/ (bay* \| **all**) [**begin** *date-time*] [**end** *date-time*] [**reverse**]<br><br>**show alerts disk** *external-disk-device-id/ (bay* \| **all**) [**begin** *date-time*] [**end** *date-time*] [**reverse**] | Privileged |
| Displays storage enclosure disk location map. | **show chassis storageEnclosure** {*scardID/ enclosureID*} **diskMap** | Privileged |

**\*As of Release 3.2.3:**

The command '**show storage disk**' for a storage enclosure will display the disk location in the following format:  <scard slot>/<SAS port#>/<shelf #>/< disk slot>

This display allows for easy identification of an external disk with its logical disk id, and is supported on Freedom Fabric Enclosures FS 5084-L, FS 2024-S, and FS 2012-L.

| Command | Values |
| --- | --- |
| S-card slot | **0 - 7** |
| SAS port # on the MX-card | **1** = Left Port<br>**0** = Right Port |
| Shelf# | **0** = Upper Shelf<br>**1**= Lower Shelf |
| Disk slot in Storage Enclosure | **0-83** (5RU)<br>**0-11** (2RU-12 disks)<br>**0-23** (2RU-24 disks) |

Also, the command **'activate/deactivate led disk external-location** *<physical disk location>'* will now allow users to activate or deactivate disk led using the logical or the physical disk location in the above mentioned format. See examples below.

seamicro# activate led disk 0/1 << internal disk

seamicro# activate led disk JBOD1/24  << JBOD disks

Possible completions:

```
DISK-ID in the format SLOT#/BAY# or DEVICE-NAME/BAY# e.g., 1/3, JBOD2/3
external-location
DISK-LOCATION in the format SCARD-SLOT#/SAS-PORT#/SHELF#/BAY#
```

**Physical Location**

```
seamicro# activate led disk external-location ?
Possible completions:
DISK-LOCATION in the format SCARD-SLOT#/SAS-PORT#/SHELF#/BAY#

seamicro# activate led disk external-location
seamicro# show storage disk JBOD1/24
disk JBOD1/24
          pool:
          size: 1863.00GB
         model: SEAGATE ST2000NM0001
        serial: Z1P0KR650000C2126UHF
 firmware rev.: XQB6
        status:
      location: 1/0/0/3
          note:

seamicro# activate led disk external-location 1/0/0/3
```

# Troubleshooting Storage

Storage troubleshooting involves identifying, removing, and replacing storage components, such as down or defective disks, S-cards, or S-card expanders, and verifying that storage is back to normal. With alerts, alarms, and LEDs, it is easy to identify the failed component, which may then be corrected as needed to bring the system back to its optimum level.

*   Alerts are saved in flash and are persistent across chassis reloads.
*   They will be synced across MX-cards so that the information is preserved during failovers.
*   A maximum of 10,000 alerts will be saved and managed in a circular buffer.

Most common storage errors occur due to:

*   Disk failures
*   S-Card failures
*   Storage Enclosure failures
*   CLI Errors for pools and volumes

This section includes information on alerts, alarms, LEDs, and storage errors. Refer to Appendix B on page 156 for common storage errors and steps you need take to correct those errors.

To identify the source of the problem, use the show alerts and show alarms commands for disks, S-cards, and external disks. In R3.2.3 and R3.3, the show alerts command displays storage-related information only.

| Component | Command | Level |
|---|---|---|
| Disks, S-cards, and Storage Enclosures. | **show alerts [begin** time **[end** time**]]** [reverse] | Privileged |
| | **show alarms** | Privileged |

Example:

```
seamicro# show alerts
Alert time             object         identifier       description
----------------------------------------------------------------------
04-Feb-2013@12:21:48    scard              0           <problem description>
04-Feb-2013@12:21:48    disk              2/1          <problem description>
04-Feb-2013@12:21:48    disk            JBOD1/2        <problem description>
04-Feb-2013@12:21:48     disk 11:22:33:44:55:66:77:88/2 <problem description>
```

# Disk Failures

Disks may fail for any of the following reasons:

- Disk I/O errors
- Missing or removed disks
- Unresponsive disks
- Dead disks

To identify the type of disk failure, use the following command.

| Component | Command | Level |
|---|---|---|
| Disk | **show alerts disk** slot/(bay\|**all**) [**begin** time [**end** time]] [reverse] | Privileged |

The table below describes the alerts and alarms, and the type of disk error to be resolved.

| Alert/ Alarm | Indicates | Description |
|---|---|---|
| Alert | **Disk I/O Error** | The disk I/O error will be logged in the Syslog. If the number of errors on a disk is greater than **3,** the disk will be marked as '**Failed,**' and the affected pool will be taken down.<br><br>The S-Card's data plane maintains disk I/O error collecting data structure per vdisk. Each vdisk is mapped to a volume or disk. The data structure records the following error information:<br><br>• **Server/Port Number**: Volume affected, which may span across multiple disks.<br>• **Error Count**: Number of disk I/O errors in a volume.<br>• **Error flags**: Read, write, or fsync error.<br>• **Offset**: Location of last error in a volume.<br>• **Error Code**: Linux-provided errno of the last error. |
| Alarm | **Removed Disk** (Internal) | When an **internal disk** is removed, an alarm will be raised.<br><br>If a removed disk is part of a pool, the pool will be taken down, un-configuring all assignments associated with volumes in that pool. |

| Alert/ Alarm | Indicates | Description |
|---|---|---|
| Alert | **Removed Disk** (External) | An alert will be raised when:<br><br>• An **external disk** is removed.<br>• A pool goes into failed/degraded state.<br><br>If a removed disk is part of a pool, the pool will be taken down, un-configuring all assignments associated with volumes in that pool. |
| Alert | **Unresponsive Disk** | An unresponsive disk is a transient condition.<br><br>It may not respond to some control/management commands.<br><br>The S-card management will not retrieve disk-related information. |
| Alert | **Dead Disk** | A dead disk is a disk that is physically present but does not respond to disk I/O requests, and is not seen on CLI.<br><br>It will be reported as an alert or as a removed disk. |

## Disk Failure and Hotswap Behavior

Refer to the table below for LED indicators of an active or failed disk.

| LED | Indicates | Correction |
|---|---|---|
| Solid Red | Internal Disk failed to power on. | To recover from an internal disk failure:<br><br>• Re-seat the disk.<br>• If the LED is still 'red,' remove and replace the disk.<br>• Verify that the LED turns 'green.'<br>• Verify that the disk is operational using the **'show disk'** command. |

Depending on the RAID level of the disk, data protection is provided to the volumes in an S-card. In the non-RAID mode, no data protection is provided to the volumes in an S-card.

**RAID Mode**

The fault tolerance of a pool depends on its RAID level. If the number of failed disks exceed the number that can be handled by a RAID level, the pool will be marked as failed, and the user will not see the pool in CLI. The user will then have to reconfigure the pool.

Best Practice: In general, be careful when hot-swapping any disks, S-cards, and storage enclosures that contain RAID pools. Before removing any such disks, S-cards, or storage enclosures, power off the S-card, and then hot-swap the required component.

## Hot Spare Disk Status, Alarms and Alerts

When a pool is damaged, a failed disk is automatically replaced with one of the free HS disks. Right after replacement, the RAID function will rebuild the new pool member disk, and repair the pool to optimal status.The following values indicate the HS disk status:

| HS Disk Status | Message |
|---|---|
| Missing | HS disk is not found. |
| Available | HS disk is not yet selected to repair a damaged pool. |
| In-Use | HS disk is being used to repair a damaged pool, to create a pool, or to add a disk to a pool. |

| Alarms/Alarms | Indicators |
|---|---|
| HS Alarm | Will be raised by the **sysmgr** only if the HS function is **enabled**. Will be raised when the number of HS disks gets to **zero**. Will be cleared when the number of HS disks is **positive**. |
| HS Alert | Will be raised by the **sysmgr** if the HS function is **enabled**. Will be raised whenever a failed disk is found by the RAID function. |
| Auto HS Function | With the '**auto**' option, there will be no need to explicitly configure HS disks. Any currently free disk will be treated as an available HS disk. If auto HS function is configured, '**in-use** and "missing" statuses will not be applicable. |

## Hot-Swapping Disks and Storage Enclosure with RAID

Note the following important considerations when hot-swapping storage enclosures with RAID or when hot-swapping disks with RAID pool information into an S-card without RAID capability:

| IF | THEN |
|---|---|
| If you hot-swap a storage enclosure from one S-card to another: | You run the risk of losing all of its pools, and the enclosure being marked as '**Failed**' by the RAID controller. |
| If you move your disks from one S-card to another: | Pools will not be automatically created on the new S-card. Re-create the pools on the S-card. |
| If you connect a storage enclosure that contains RAID pools to an S-card that is on a different RAID level: | An alarm will be raised, and the pool on the storage enclosure will be placed in an un-mounted mode.<br><br>You must match the S-card's RAID level or move the storage enclosure to an S-card that contains the same RAID level. |
| If you insert a disk that contains a RAID pool into an S-card that does not have RAID: | You may see incorrect pools mounted on the S-card.<br><br>You will then have to delete the incorrect pools and disks from the S-card. |

**Non-RAID Mode (Volume Management)**

In the non-RAID mode, no data protection is provided to the volumes in an S-card. As volumes do not cross storage pool boundaries, the impact of a disk going bad is limited to the storage pool that uses the disk.

For example, in the following S-card configuration, where Pool1 contains disk 1, 2, and 3 and Pool2 contains disk 4, 5, and 6, if disk 1 goes bad or is unplugged, all the volumes in pool1 will be affected, since volumes stripe across multiple disks. Therefore, when disk 1 is unplugged:

- The "**show storage pool**" command will show only pool2, and pool1 will disappear.
- The "**show storage disk**" command will show disk 2, 3, 4, 5, 6, and disk 2 and 3 will contain no associated pool.

Note:

- Using one disk to create one pool is recommended. Starting with Release 3.3, pools spanning across disks are not supported.
- All disks can only be in one of two states, disks assigned to a pool, or disks free to be assigned. As a result, when a disk is hot-swapped, the disks remaining in the pool will not have any pool associated with them, and the user will have to recreate the pool.

**Non-RAID Mode (Raw Disk Management)**

Assignment of pools and volumes is not applicable in the Raw Disk Management mode.

# S-card Failures

- S-card down or degraded.
- Expander failure.

An S-card failure might affect server access to the disks on the affected S-card. If the S-card experiences a complete or partial failure, the S-card can be replaced or hot-swapped while the chassis is powered on without affecting the internal server's access to the disks. In the event of a card failure and during a hot-swap procedure, traps, Syslogs, console messages, alarms, and LCD updates will be generated.

The table below illustrates the S-card error that may need to be resolved.

| Component | Command | Level |
|---|---|---|
| S-card | **show alerts scard** (slot-range\|**all**) [**begin** time [**end** time]] [reverse] | Privileged |
| | **show storage scard** | Privileged |
| | **show chassis** | Privileged |
| | **show alarms** | Privileged |

The table below displays possible S-card errors.

| Alert/Alarm/ LED | Indicates | Description |
|---|---|---|
| **Solid Red** LED | S-card Error | Faulty S-card or S-card may not be powered on. All eight LEDs should be green. |
| **show storage scard** displays 'not present' | Missing S-card | S-card may not be present or may have been removed from the chassis. |
| **show storage scard** displays 'down' | Lost management connection. | S-card may need to be powered on, or may need to be reset, re-seated, or replaced to function properly. |
| **show chassis** displays 'degraded' | Management connection is degraded. | S-Card management connection may be broken while the data plane of the S-Card may still work. Hence, the card is considered 'degraded' and not 'down.' |

| Alert/Alarm/ LED | Indicates | Description |
|---|---|---|
| **show alarms** displays 'card degraded' | S-card degraded. | S-card may need to be powered on, or may need to be reset, re-seated, or replaced to function properly. |

Note: Currently, there is no way to know whether an S-Card is powered '**On**' or '**Off**.'

In the examples below, the chassis contains an S-card in slot 1, 2, and 4, and the management interface of S-card 1 and 4 are connected.

```
seamicro# show storage scard brief
slot   mgmt. status   mgmt. mode   raid level   stripe size   hot-spare count
-------------------------------------------------------------------------------
   0     not present
   1        up          volume        none
   2     not present
   3        up          volume        none
   4        up          volume          6          256KB
   5     not present
   6        up          volume          6          256KB              0
   7        up           disk         none
* 8 entries

seamicro# show chassis
.
 --- SCards ---
Slot    Status          Internal Disks  External Disks  Uptime
-------------------------------------------------------------------------------
1       active          8               0               2 hrs, 30 mins, 11 secs
2       degraded        0               0               2 hrs, 30 mins, 11 secs
4       active          8               0               2 hrs, 29 mins, 31 secs
.
.
.

seamicro# show alarms
Alarm time            Class        Resource       Description
-------------------------------------------------------------------------------
2013-06-28 12:51:38   critical     Disk- 6/6      diskMissing
2013-06-28 12:45:47   critical      System         systemDegraded
```

# Storage Enclosure Failures

Storage enclosures may experience external disk failure, an S-card expander failure, or S-card expander management interface failure.

A new application process called the Storage Enclosure Manager (SEMgr) will establish a connection with all the S-cards in the chassis, and query the S-card for storage enclosure-related information. The SEMgr runs only on the primary MX-card. In case of process restarts and failovers, the SEMgr will re-establish all connections with the S-card and retrieve data once again.

The SEMgr will get the following information from the Scard:

- Enclosure vendor name, and the number of connected Storage Enclosures.
- IO controller and IO port information.
- Number of connected power supplies and their status.
- Number of connected fans, their RPM value, and their status.
- Number of voltage sensors and voltage sensor readings.
- Number of current sensors, and current readings.
- Number of temperature sensors, their threshold value, and current temperature readings.

The SEMgr will send the gathered storage enclosure data to the Sysmgr.

- The Sysmgr will cache this data.
- The Sysmgr compares the changes in storage enclosure data received from SEMmgr with its cached values.
- If there are any failures, it will raise/clear alarms/alerts.
- When a user issues a CLI command to get storage enclosure-related information, the Sysmgr will display the cached values.
- The Sysmgr will send commlib message to SNMPproxy application to raise/clear any SNMP traps.
- With the updated data, the SNMPproxy will use the information for responding to all SNMP walk queries.

## Monitoring Storage Enclosure

Use the following commands to monitor the health of a storage enclosure.

| Task | Command | Level |
|---|---|---|
| Display fan speed, temperature values, and voltage readings. | **show environment**<br><br>**Option: [storageEnclosure** *enclosureID***]** | Privileged |
| Display enclosure model number and unique ID. | **show inventory** | Privileged |
| Display brief storage enclosure elements. | **show chassis storageEnclosure** {*scardID/ enclosureID*} | Privileged |

## Identifying Storage Enclosure Failures

To identify any storage enclosure failures, use the following commands:

| Component | Command | Level |
|---|---|---|
| Storage Enclosure | **show alerts disk external-disk-device-name/ (bay|all) [begin time [end time]] [reverse]** | Privileged |
| | **show alarms** | Privileged |
| | **show chassis storageEnclosure** {*scardID |* *enclosureID*} **diskMap** | Privileged |

Refer to the table below to identify, troubleshoot, and resolve problems with a storage enclosure and/or S-card expanders.

| Alert/Alarm/ LED/CLI Command | Indicates | Description |
|---|---|---|
| Solid Amber | External disk failed. | Disk failure in a storage enclosure. |
| Alarm | S-card expander management interface disconnected. | Faulty SAS port on the MX-card or faulty Expander. |

| Alert/Alarm/ LED/CLI Command | Indicates | Description |
|---|---|---|
| Alert | S-card Expander failed. | Expander failure indicated as:<br><br>`Expander <problem area> -`<br>`<worrisome line>`<br><br>OR<br><br>`Expander <problem area> -`<br>`<worrisome line> -`<br>`<interpretation>` |
| **'show alarms'** | Display faults on storage enclosures, indicating severity as 'Missing' or 'Critical.'<br><br>Display over and under temperature warnings, and over and under temperature failures on storage enclosures, indicating the severity as 'Missing' or 'Critical.' | `Example:`<br>`encFan Failure/Missing - Critical`<br><br>`encPS Failure/Missing - Critical`<br><br><br>`Example:`<br>`encTempSensor Failure/Missing -`<br>`Critical`<br><br>`Storage Enclosure Missing - Critical` |

# CLI Errors

In the volume management mode, CLI errors indicate a failure to create pools and/or volumes, and a failure to assign volumes to servers.

| Alerts/Alarms | Indicates | Description |
| --- | --- | --- |
| Alerts | **RAID or Non-RAID**: Failed to create pool or volume. | Failure in creating pools and volumes.<br><br>A validation message will display, such as "could not initialize disk" or "error creating pool." |
| Alerts | Failed to assign volumes to servers. | Failure in assigning volumes to servers. |

- For more information on common storage errors and remedies, refer to Appendix B on page 156 of this Guide.
- For more information on troubleshooting storage-related components that are internal to the chassis, refer to the *Hardware Guide*
- For more information on troubleshooting storage-related components that are external to the chassis, refer to the following:
    - Freedom Fabric Enclosure Guide FS 2024-S & FS 2012-L Installation and User Guide
    - Freedom Fabric Enclosure Guide FS 4060-L Installation and User Guide
    - Freedom Fabric Enclosure Guide FS 5084-L Installation and User Guide

# Chapter 4 | **Layer 2 and MAC Addresses**

This chapter covers information on Layer 2, MAC Addresses, and Load Balancing.

# Layer 2

SeaMicro supports Layer 2 (L2) mode switching that forwards traffic based on MAC layer addresses. The MX-Card now supports the following:

- MAC table size 6K.
- MAC address aging out of the MAC table.
- MAC learning independent of other MX-Cards.

When non-IP, IPv4 or IPv6 traffic enters the chassis, the MX-card will spread the traffic across all 16 links between the MX-card and the fabric.

All traffic detected in the network will be switched based on MAC addresses and VLANs. With Release 3.2, all MAC learning will occur based on ingress and egress data packets from the chassis.

## MAC Addresses

Initially, all servers are assigned MAC addresses with the SeaMicro OUI. These addresses are statically installed during system initialization, and cannot be cleared or aged out. However, all non-SeaMicro OUI MAC addresses learned from external Ethernet ports or internal servers are dynamic addresses, and can be cleared or aged out.

The default internal server MAC addresses are derived through a hashing mechanism, and they include the SeaMicro OUI (00:22:99). You can configure another MAC address with any OUI for any internal server from the server itself. For example, on a RedHat Linux, use **ifconfig** or change the MAC in the ifcfg script.

You may configure, display, and clear MAC address tables, aging interval, and port limits. See the table below for tasks and commands associated with MAC address tables.

| Task | Command | CLI Level |
|------|---------|-----------|
| Configure MAC address table aging interval. | **seamicro(config)# mac address-table aging-time** | Configuration |

| Task | Command | CLI Level |
|------|---------|-----------|
| Configure MAC address table port limit. | **seamicro(config)# mac address-table port-limit** | Configuration |
| Display MAC address table aging interval. | **seamicro# show mac address-table age** | Privileged |
| Display MAC address table port limit. | **seamicro# show mac address-table port-limit** | Privileged |
| Display MAC addresses on the NIC, the NIC bond, the Ethernet interface, or the VLAN. | **seamicro# show mac address-table** | Privileged |
| Clear MAC addresses on the NIC, the NIC bond, the Ethernet interface, or the VLAN. | **seamicro# clear mac address-table** | Privileged |

*Refer to the SeaMicro Release 3.2.0.0 CLI Guide for more information on each CLI command.*

# Configure MAC Addresses

You may configure the MAC address aging interval or the port limit.

| Task | Command | CLI Level |
|------|---------|-----------|
| Configure MAC address table aging interval. | **seamicro(config)# mac address-table aging-time** | Configuration |
| Configure MAC address table port limit. | **seamicro(config)# mac address-table port-limit** | Configuration |

## MAC Address Aging Interval

When devices enter and leave a network, or are not active anymore, you can remove them from the network through MAC address aging. Static addresses that are pre-installed in the system, will never be aged out. However, dynamic address that are learned, can be aged out of the system by setting the MAC address aging interval. The aging interval can be short or long, ranging anywhere between 30 seconds to 6000 seconds. If there are not a lot of fluctuations on the device, the aging interval can be set up for a longer interval.

1. To configure MAC address aging, enter the following command:

```
seamicro(config)# mac address-table aging-time <time in seconds>
```

2. Enter aging time in seconds ranging from 30 to 6000. The default value is 300. In the example below, the configured value is 567.

   Example:

   ```
   seamicro(config)# mac address-table aging-time 567
   ```

3. To display MAC address aging interval, enter the following command:

   ```
   seamicro# show mac address-table aging-time
   ```

   Example:

   ```
   seamicro# show mac address-table aging-time
   300 seconds
   ```

## MAC Address Port Limit

Setting MAC address port limit is a security measure. When packets pass through the network, many MAC addresses are learned by the system. To avoid a MAC address glut in the system, you can limit the number of MAC addresses that are learned on the port.

1. Use the following command to configure the MAC address port limit:

   ```
   seamicro(config)# mac address-table port-limit
   ```

2. Enter a value from 0 to 6000 for port limit. The default value is 1200. In the example below, the configured value is 4500.

   Example:

   ```
   seamicro(config)# mac address-table port-limit 4500
   ```

3. To show MAC address port limits, enter the following command:

   ```
   seamicro# show mac address-table port-limit
   ```

   Example:

   ```
   seamicro# show mac address-table port-limit
   1200
   ```

# Display MAC Addresses

The **show mac address-table** command displays all learned MAC addresses on the various filters, such as the NIC, the NIC bond, the ethernet interface, and the table count.

| Task | Command | CLI Level |
|------|---------|-----------|
| Display MAC address table aging interval. | **seamicro# show mac address-table age** | Privileged |
| Display MAC address table port limit. | **seamicro# show mac address-table port-limit** | Privileged |
| Display MAC addresses on the NIC, the NIC bond, the Ethernet interface, or the VLAN. | **seamicro# show mac address-table** | Privileged |

1. To display MAC addresses, enter the following command:

```
seamicro# show mac address-table
```

2. Enter one of the possible completions for this command:

```
address       MAC Address, e.g. 00:11:22:aa:bb:cc
aging-time    Show MAC address aging time
bond          Show MAC address on the bond
count         Show MAC address table count
interface     Show MAC address on the interface
nic           Show MAC address on the NIC
port-limit    Show MAC address table port limit
server        Show MAC address on the server
vlan          Show MAC address on the VLAN
```

3. See sample output for **show mac address-table** below:

```
seamicro# show mac address-table
    MAC address        VLAN   Server/NIC   If Name    Type
------------------     ----   -----------  --------   -------
 00:21:53:11:03:80      -            n/a   inband     static
 00:22:99:04:00:00      -          0/0/2      n/a     static
 00:22:99:04:00:01      -          0/0/5      n/a     static
 00:22:99:04:00:02      -          0/0/0      n/a     static
 00:22:99:04:00:03      -          0/0/6      n/a     static
 00:22:99:04:00:04      -          0/0/1      n/a     static
 00:22:99:04:00:05      -          0/0/7      n/a     static
 00:22:99:04:00:06      -          0/0/3      n/a     static
```

```
      -------------------------------------------------------------
      00:22:99:04:01:00    -   not assigned      n/a    static
      00:22:99:04:01:01    -   not assigned      n/a    static
      00:22:99:04:01:02    -   not assigned      n/a    static
      -------------------------------------------------------------
      00:22:99:04:02:10    -   not assigned      n/a    static
      00:22:99:04:02:11    -   not assigned      n/a    static
      00:22:99:04:02:12    -   not assigned      n/a    static
      00:22:99:04:02:13    -   not assigned      n/a    static
      -------------------------------------------------------------
      00:22:99:04:07:00    -         4/0/0       n/a    static
      00:22:99:04:07:01    -         4/0/1       n/a    static
      00:22:99:04:07:02    -         4/1/0       n/a    static
      00:22:99:04:07:03    -         4/1/1       n/a    static
      -------------------------------------------------------------
      00:11:22:33:44:04    -       1/0/bo 8      n/a    dynamic
      00:22:99:04:00:17    -       1/0/bo 8      n/a    static
      00:11:22:33:44:03    -       1/0/bo 8      n/a    dynamic
      00:22:99:04:00:16    -       1/0/bo 8      n/a    static
      00:22:99:04:00:10    -       1/0/bo 8      n/a    static
      00:11:22:33:44:02    -       1/0/bo 8      n/a    dynamic
Number of addresses: 536
```

4.  See sample output for **show mac address-table** on the various filters below:

**On the Bond:**

```
seamicro# show mac address-table bond 8

    MAC address       VLAN    Server/NIC    If Name    Type
------------------    ----    -----------   --------   -------
 00:11:22:33:44:04    -        1/0/bo 8      n/a    dynamic
 00:22:99:04:00:17    -        1/0/bo 8      n/a    static
 00:11:22:33:44:03    -        1/0/bo 8      n/a    dynamic
 00:22:99:04:00:16    -        1/0/bo 8      n/a    static
 00:22:99:04:00:10    -        1/0/bo 8      n/a    static
 00:11:22:33:44:02    -        1/0/bo 8      n/a    dynamic
Number of addresses: 6

seamicro# show mac address-table count bond 8
Number of addresses: 6
```

**On the Tengigabitethernet:**

```
seamicro# show mac address-table interface tengigabitethernet 0/0
    MAC address       VLAN    Server/NIC    If Name    Type
------------------    ----    -----------   --------   -------
Number of addresses: 0
seamicro# show mac address-table interface tengigabitethernet 0/0
    MAC address       VLAN    Server/NIC    If Name    Type
------------------    ----    -----------   --------   -------
 00:00:00:11:00:18    20           n/a     te 0/0    dynamic
```

```
00:00:00:11:00:17    20                n/a    te 0/0    dynamic
00:00:00:11:00:16    20                n/a    te 0/0    dynamic
00:00:00:11:00:15    20                n/a    te 0/0    dynamic
00:00:00:11:00:14    20                n/a    te 0/0    dynamic
00:00:00:11:00:13    20                n/a    te 0/0    dynamic
00:00:00:11:00:12    20                n/a    te 0/0    dynamic
00:00:00:11:00:11    20                n/a    te 0/0    dynamic
00:00:00:11:00:10    20                n/a    te 0/0    dynamic
00:00:00:11:00:0f    20                n/a    te 0/0    dynamic
-------------------------------------------------------------------
Number of addresses: 20
```

## On the VLAN:

```
seamicro# show mac address-table vlan 20
   MAC address       VLAN   Server/NIC   If Name    Type
-------------------  ----   -----------  --------  -------
 00:00:00:11:00:18   20                n/a    te 0/0    dynamic
 00:00:00:11:00:17   20                n/a    te 0/0    dynamic
 00:00:00:11:00:16   20                n/a    te 0/0    dynamic
 00:00:00:11:00:15   20                n/a    te 0/0    dynamic

Number of addresses: 20
seamicro#

seamicro# show mac address-table count vlan 20
Number of addresses: 20
seamicro# show mac address-table interface tengigabitethernet 0/0
   MAC address       VLAN   Server/NIC   If Name    Type
-------------------  ----   -----------  --------  -------
 00:00:00:11:00:18   20                n/a    te 0/0    dynamic
 00:00:00:11:00:17   20                n/a    te 0/0    dynamic
 00:00:00:11:00:16   20                n/a    te 0/0    dynamic
 00:00:00:11:00:15   20                n/a    te 0/0    dynamic
 00:00:00:11:00:14   20                n/a    te 0/0    dynamic
 00:00:00:11:00:13   20                n/a    te 0/0    dynamic
 00:00:00:11:00:12   20                n/a    te 0/0    dynamic
 00:00:00:11:00:11   20                n/a    te 0/0    dynamic
-----------------------

Number of addresses: 20
```

## On the Server/NIC:

```
seamicro# show mac address-table server 1/0
   MAC address       VLAN   Server/NIC   If Name    Type
-------------------  ----   -----------  --------  -------
 00:22:99:04:00:11   -          1/0/5      n/a    static
 00:22:99:04:00:12   -          1/0/0      n/a    static
 00:22:99:04:00:13   -          1/0/6      n/a    static
 00:22:99:04:00:14   -          1/0/1      n/a    static
```

```
00:22:99:04:00:15   -          1/0/7      n/a   static
00:11:22:33:44:04   -      1/0/bo 8       n/a   dynamic
00:22:99:04:00:17   -      1/0/bo 8       n/a   static
00:11:22:33:44:03   -      1/0/bo 8       n/a   dynamic
00:22:99:04:00:16   -      1/0/bo 8       n/a   static
00:22:99:04:00:10   -      1/0/bo 8       n/a   static
Number of addresses: 10
```

**On the NIC:**

```
seamicro# show mac address-table nic 1/0 6
   MAC address      VLAN   Server/NIC   If Name   Type
------------------  ----  ------------  --------  -------
 00:22:99:04:00:13   -        1/0/6      n/a   static
Number of addresses: 1
seamicro#
seamicro# show mac address-table address 00:00:00:11:00:09
   MAC address      VLAN   Server/NIC   If Name   Type
------------------  ----  ------------  --------  -------
 00:00:00:11:00:09   20        n/a     te 0/0   dynamic
Number of addresses: 1
```

# Clear MAC Addresses

MAC addresses can be cleared on the bond, the interface, the NIC, the server, and the VLAN.

| Task | Command | CLI Level |
|------|---------|-----------|
| Clear MAC addresses on the NIC, the NIC bond, the Ethernet interface, or the VLAN. | **seamicro# clear mac address-table** | Privileged |

1. To clear MAC addresses, enter the following command. The default is '**all'** addresses.

```
seamicro# clear mac address-table
```

**Note:** Clear MAC addresses with caution as it removes all dynamic addresses. If cleared by accident, the addresses will be re-learned, but will also result in temporary flooding.

2. Enter one of the following filters as a possible completion for the command:

```
address      MAC Address, e.g. 00:11:22:aa:bb:cc
no-confirm   No confirmation needed
bond         Clear MAC address on the bond
interface    Clear MAC address on the interface
nic          Clear MAC address on the NIC
server       Clear MAC address on the server
vlan         Clear MAC address on the VLAN
```

Note: VLAN can be an option on **show mac address-table interface**. For example:

**show mac address-table interface tengigabitethernet 0/0**

Or

**show mac address-table interface tengigabitethernet 0/0 vlan 20**

The first command shows all MAC addresses learned on interface tengigabitethernet 0/0. and the second show command only shows MAC addresses learned on VLAN 20 on interface tengigabitethernet 0/0.

3. As confirmation is needed for **clear** commands, use the **show mac address-table** command to confirm that the mac address table has been cleared.

See examples below of clearing mac address tables, mac addresses on the interface and on the VLAN.

**Example 1:**

```
seamicro# clear mac address-table
Clear MAC addresses (yes/no)? no
```

**Example 2:**

```
seamicro# clear mac address-table interface tengigabitethernet 0/0
Clear MAC addresses (yes/no)? yes

seamicro# show mac address-table interface tengigabitethernet 0/0
    MAC address       VLAN   Server/NIC   If Name    Type
------------------  ----  ------------  --------  -------
Number of addresses: 0
```

**Example 3**:

```
seamicro# clear mac address-table vlan 20
Clear MAC addresses (yes/no)? yes
seamicro# show mac address-table vlan 20
    MAC address       VLAN   Server/NIC   If Name    Type
------------------  ----  ------------  --------  -------

Number of addresses: 0
```

# Common MAC for Multiple VLANs

The SM15K supports ethernet packet forwarding directly between two servers (through the fabric) based on the MAC address table resident within each server's forwarding ASICs. The ASICs look only at the destination MAC address when making a forwarding decision. Note, that the system does not support two internal servers using the same MAC address in different VLANs.

Release 3.3 enables using the same MAC address for multiple external destinations in different VLANs. An external destination is one that exists outside of the SM15K chassis, such as an upstream router or host. This support is possible because the MX-card uses both destination MACs and VLANs when making forwarding decisions.

# Server to Server Load Balancing

This section provides information on internal server to server load balancing between a SeaMicro MAC to a SeaMicro MAC, and a SeaMicro MAC to an arbitrary MAC. Each forwarding ASIC in

the system has two different tables that are used to forward frames -- the fabric table and the MAC address table.

- The fabric table is used to map SeaMicro MAC addresses to destination servers.
- The MAC address table is used to map arbitrary MAC addresses to destination servers.

Each ASIC has a MAC address table size of 1K entries, and these entries are programmed on demand when a server needs to send traffic to a particular MAC. When establishing NIC bonds to servers, SeaMicro recommends using an arbitrary MAC address rather than a SeaMicro MAC to represent the bond in order to maximize system scalability.

### SeaMicro MAC to SeaMicro MAC Forwarding

- When a NIC becomes part of a NIC bond with a SeaMicro MAC address, the ASIC corresponding to that NIC will disable its fabric table, and put both arbitrary and SeaMicro MAC addresses into the MAC address table, limiting its total capacity.
- Since the MAC address table for each ASIC is programmed on demand, load balancing is achieved by programming each ASIC's MAC address table with different bonded destinations in a round-robin fashion.

### SeaMicro MAC to Arbitrary MAC Forwarding

- When a NIC becomes part of a NIC bond with an arbitrary MAC address, the ASIC corresponding to that NIC will continue to use the fabric table to forward traffic to SeaMicro MACs, using the MAC address table only for arbitrary MACs.
- Since the MAC address table for each ASIC is programmed on demand, load balancing is achieved by programming each ASIC's MAC address table with different bonded destinations in a round-robin fashion.

# IPv4, IPv6, and non-IP Traffic Load Balancing

The SM15K performs load balancing for IPv4, IPv6, and non-IP traffic entering the chassis from the outside that are destined for internal servers. This load balancing is performed by hashing on the MX-card to one of the 16 links between the MX-card and the fabric, and across any server NICs in a bond.

- For non-IP traffic, the hash is chosen based on the ethernet source and destination, and incoming port number.
- For IPv4 traffic originating outside of the chassis, the hash is chosen based on the ethernet source and destination, IP source and destination, and TCP/UDP source and destination port.
- For IPv6 traffic originating outside of the chassis, the hash is chosen based on the ethernet source and destination, IPv6 source and destination, IPv6 flow label, and TCP/UDP source and destination port.
- For traffic originating at the server, the hash is chosen based only on the IPv4 source and destination address.

# Multicast

The SeaMicro system will forward IPv4 multicast streams in the incoming and outgoing directions, and between internal servers. It will prune the received traffic or IPv4 multicast packets, process the packets, and route the packets to the appropriate group or destination.

The system will not act as an L2 transit switch for forwarding multicast streams between systems outside of the SeaMicro system. Within the system, multicast packets will be treated as broadcast packets. In the multicast mode, IGMP snooping is implemented to avoid unnecessary flooding of multicast traffic.

SeaMicro recommends that IGMP snooping is enabled on all LAN-facing chassis.

## Multicast and IGMP Snooping

The system supports IGMP v2, but not IGMP v3. If snooping is enabled, IGMP v3 packets will be dropped. If snooping is disabled, IGMP v3 packets will be broadcast within the VLAN domain.

A total of 1K unique VLAN and multicast group pairs are supported.

- When IGMP snooping is enabled, IGMP control packets will be built, and multicast data packets will be forwarded to a multicast group cache and a router cache.
- When IGMP is disabled, multicast control and data packets will be flooded in the VLAN. By default, IGMP snooping is **disabled**.

For IGMP snooping to work correctly, external queriers are required.

# Multicast and VLANs

Multicast packets come into the VLAN as tagged or untagged packets. The VLAN will contain a valid ID for tagged packets, and will be '**Zero'** for untagged packets.

- Registered multicast packets will be forwarded to group subscribers and multicast routers on the VLAN.
- Unregistered multicast packets will be forwarded to multicast routers by default
- If no multicast routers exist, all unregistered multicast packets will be dropped by default.
- If '**flooding**' is configured, all unregistered multicast packets will be flooded in the VLAN.
- Multicast data streams will not be forwarded across VLAN boundaries.
- When the system resource is full and IGMP snooping is enabled, multicast traffic from unknown VLANs will be dropped.
- When a new VLAN is added, multicast groups will be learned on the new VLAN.
- When a VLAN is removed, all group cache entries that are related to the deleted VLAN will be removed.
- When a port or a node is moved from one VLAN to another, multicast groups will be learned on the port or the node of the new VLAN. The old entries related to the old VLAN will be removed by the group cache aging mechanism.

# Multicast and Port Channels

Multicast traffic will not be load balanced among members of a port channel.

- Multicast data and control traffic will be forwarded only to a designated port of the Port Channel.
- The multicast process will listen to the port channel designated port change event, and move the multicast group data traffic from the old designated port to the new designated port.
- When a port channel is added, multicast groups will be learned on the new port channel.
- When a port channel is deleted, the group cache entries that are learned on the deleted port channel will be removed by the group cache aging mechanism.

# Multicast and NIC Bonds

- When a NIC is added to a NIC bond on an internal server, the group cache entries for that NIC will be removed as that NIC is now a part of the NIC bond. NIC bond is learned as a member when IGMP packets are received from the NIC bond interface.

# Multicast and Servers

- When a node is up, multicast groups will be learned on the new 'up' node.
- When a node is down, the group cache entries learned on the port will be removed by the group cache aging mechanism.

# Multicast and System Events

- During and after a multicast process restart, the system will continue to forward multicast traffic. However, during the process restart, IGMP snooping will not happen as the application is not active.
- During and after a primary MX-card switch over, the system will continue to forward multicast traffic. However, IGMP snooping will not happen during the switchover, and therefore, new group membership will not be learned.
- When a new MX card is inserted, it will bulk sync from the primary MX-card to construct a group cache, a router cache, and a group member table.
- When an MX card is removed, the group cache entries related to the removed card will be removed by the group cache aging mechanism.
- When an MX-card is inserted again, it will become a member of all the existing multicast groups only if that MX-card is already participating in that group, i.e. if a router port or a receiver is connected to that MX-card port.

# IGMP Snooping Control Packets

When IGMP control packets are snooped, the IGMP snooping switch will build an IGMP snooping cache. All subsequent multicast data packet forwarding will take place per this cache.

Each MX card will run its own instance of IGMP snooping, and all IGMP control packets will be forwarded to all MX cards for snooping purposes. Note that there will be an exchange of IGMP snooping cache only when an MX-card is inserted.

## Types of IGMP Control Packets

- **IGMP Query and PIM Hello**: General and group specific IGMP query and PIM hello packets are snooped to construct a VLAN mrouter cache. IGMP queries will be forwarded on the VLAN, and PIM hello packets will also be forwarded.
- **IGMP Report**: IGMP Reports, v1 and v2, are snooped and used to construct the group cache. All IGMP report packets will be forwarded to the mrouter list learned on the VLAN. Report suppression is not supported. The IGMP group cache will consist of internal nodes, external ports, and mrouter ports only.
- **IGMP Leave**: IGMP leave messages are snooped. If the '*leave*' is from an internal node, the node will be removed from the group cache. Also, the IGMP leave message will be forwarded to the mrouter cache learned on the VLAN.

## Group and Mrouter Cache Aging

**Group Cache Aging**: Each IGMP snooping cache entry will be aged out with an aging timer. Incoming IGMP report packets will reset the aging timer for the corresponding entries. The default aging time for each entry will be 260 seconds. Use the CLI command to configure aging time.

**Mrouter Cache Aging**: Each mrouter cache entry/port will be aged out with an aging timer. Incoming IGMP query or PIM hello packet will reset the aging timer for the corresponding entry/port. The default aging time for each entry will be 260 seconds. Use the CLI command to configure aging time.

**Router Alert Options**: IGMP packets with or without a router alert option will be snooped. IGMP packets generated, such as group queries, will not have a router alert option.

# IGMP Querier

The IGMP querier within the system facilitates the proper routing of multicast traffic. This traffic is internal, where the sender and receiver of multicast traffic exists within the system. Only the primary MX-card can initiate IGMP general queries. Upon the primary MX-card failure, the new primary MX-card will become the querier if enabled. The system supports a maximum of 64 IGMP queriers.

Using CLI commands, you may configure the interval between general queries, the maximum response time for general queries, and queriers for VLANs.

- IGMP has reserved space for 32K subscribers.
- IGMP querier can be enabled only when IGMP snooping is enabled.
- Even if the querier is configured, it will be disabled when snooping is disabled.
- The system does not allow querier election, and does not generate group-specific queries.

**IP Address**:

- Source IP address is not pingable.
- The querier source IP address is optional, and may be configured per VLAN.
- If source IP is not configured, the default source IP address will be 0.0.0.0.

**VLAN:**

- There is no dependency on system VLAN CLI.
- IGMP querier can be enabled per configured or un-configured VLAN independent of system VLAN.
- However, IGMP queries will be sent out for a VLAN only when system VLAN is configured.
- If system VLAN is un-configured and the querier is configured on VLAN, queries will not be sent out unless system VLAN is configured back again.
- IGMP enabled queries will be broadcast only to the internal network in configured VLANs..
- Note: VLAN range is not supported.

# Configure IGMP Snooping

Using CLI commands, you can configure IGMP snooping, group aging time, and mrouter aging time. See table below for the CLI commands.

| Task | Command Syntax | CLI Level |
|---|---|---|
| Enable/Disable IGMP snooping system-wide | **ip igmp snooping enable**<br>**no ip igmp snooping**<br><br>Default: disabled. | Configuration |
| Set IGMP group aging time system-wide. | **ip igmp snooping group-aging-time** *time in seconds*<br><br>Default: 260 sec. Range: 20 - 1000 sec. | Configuration |
| Set the IGMP mrouter aging time system-wide. | **ip igmp snooping router-aging-time** *time in seconds*<br><br>Default: 260 sec. Range: 20 - 1000 sec. | Configuration |
| Enable flooding of unregistered multicast data packets. | **ip igmp snooping flood**<br><br>Default: flooding disabled. | Configuration |
| Configure the interval between general queries. | **ip igmp snooping querier query-interval** *interval*<br><br>Default: 125 sec. Range: 10 - 450 sec. | Configuration |
| Configure maximum response time for IGMP general queries to tune burstiness of IGMP traffic on a subnet.<br><br>**Note:** The Syslog will display a warning if '**group-aging time**' is not correct. The query '**response**' must be smaller than the query '**interval**.' | **ip igmp snooping querier query-response-interval** *interval*<br><br>Default: 10 sec. Range: 2 - 100 sec.<br><br>Default and Range are based on group aging time range of 20 - 1000 seconds. | Configuration |
| Enable IGMP queriers for VLANs, if VLANs are configured. | **ip igmp snooping querier vlan** [*vlan-id* [**source-address** *ip-address*]] | Configuration |

# Display IGMP Snooping

Using CLI commands, you can display IGMP snooping group caches, router caches, and snooping parameters.

| Task | Command Syntax | CLI Level |
|------|----------------|-----------|
| Display IGMP snooping group cache. | **show ip igmp snooping group-cache** *[group <IP> vlan <vlanId>]*<br><br>**show ip igmp snooping group-cache** *[group <IP>]*<br><br>**show ip igmp snooping group-cache** *[vlan <vlanId>]* | Privileged |
| Display IGMP snooping router cache. | **show ip igmp snooping router-cache** *[vlan <vlanId>]* | Privileged |
| Display IGMP snooping parameters. | **show ip igmp snooping params** | Privileged |
| Display IGMP querier information for VLAN.<br><br>The output will be sorted by VLAN. | **show ip igmp snooping querier** [*vlan <vlanId range* \| default>] | |

# Clear IGMP Snooping

Using CLI commands, you can clear IGMP snooping group caches, router caches, and snooping statistics.

| Task | Command Syntax | CLI Level |
|---|---|---|
| Clear IGMP snooping group cache. | **clear ip igmp snooping group-cache**<br><br>**clear ip igmp snooping group-cache** *[group <IP> vlan <vlanId>]*<br><br>**clear ip igmp snooping group-cache** *[group <IP>]*<br><br>**clear ip igmp snooping group-cache** *[vlan <vlanId>]* | Privileged |
| Clear IGMP snooping router cache. | **clear ip igmp snooping router-cache**<br><br>**clear ip igmp snooping router-cache** *[vlan <vlanId>]* | Privileged |
| Clear IGMP snooping statistics.<br><br>This command will also clear queries for VLANs. | **clear ip igmp snooping statistics** | Privileged |

*Refer to the CLI Guide for details on each command.*

# Chapter 6    Interface Configurations

This chapter covers the following configuration topics:

SeaMicro MX-cards provide uplinks from the chassis to the external data center network via Gigabit Ethernet or Ten Gigabit Ethernet interfaces. The interfaces support link aggregation (IEEE 802.3ad) both within and across cards, providing redundancy against link and card failure. Inter-server communication traffic is on the interconnect fabric and does not consume any uplink bandwidth. Nor does the management Ethernet interface, which is the out-of-band interface that provides remote access to the system for management and monitoring.

There are five kinds of interfaces:

- **Gigabit Ethernet**—each 1 gigabit MX-card has eight gigabit speed ports for ingress and egress production traffic. They are identified in the CLI using the MX-card slot number followed by the interface number, in the form *slot/port.*
- **Inband Interface**—the system has an internally-facing interface to perform some DHCP and monitoring functions within the chassis. This interface must have its own IP address and default-gateway.
- **Management Ethernet**—this is the out-of-band port through which the administrator accesses the system to control system functionality.
- **Port-channel**—Gigabit Ethernet ports can be statically or dynamically grouped to function as a single port in order to increase bandwidth and provide redundancy; the group is called a *port-channel* or *LAG.*
- **Ten Gigabit Ethernet**—each 10 gigabit MX-card has two, ten gigabit speed ports for ingress and egress production traffic. They are identified in the CLI using the MX-card slot number followed by the interface number, in the form *slot/port.*

# Gigabit Ethernet Interfaces

Aside from entering a description, the other configuration option available for gigabitethernet interfaces is Gratuitous ARP. Gratuitous ARP is an ARP reply that is not sent in response to a request. It is instead sent to:

- inform other nodes of its presence,
- inform other nodes a hardware address change, or
- detect an IP address conflict

**Note:** 1 Gbps interfaces only operate at 1Gbps and not at 10 Mbps or 100 Mbps.

To configure a gigabitethernet interface:

| Step | Task | Command |
| --- | --- | --- |
| 1 | Move to the Interface CLI context for a Gigabit interface.<br>**Note:** You can configure multiple interfaces with a single command by using the **range** keyword. | **interface gigabitethernet** *slot*/*port*<br>CLI Level: Configuration |
| 2 | (Optional) Enter a description of the interface. | **description** *string*<br>CLI Level: Interface |
| 3 | (Optional) Enable gratuitous ARP. | **gratuitous-arp**<br>Default: Disabled<br>CLI Level: Interface |
| 4 | Enable the interface. | **no shutdown**<br>CLI Level: Interface |

# Inband Interface

The system has an internally-facing interface to perform some DHCP and monitoring functions within the chassis. This interface must have its own IP address and default-gateway, and the interface must be configured before any other IP commands are accepted. The specified address and default-gateway may, but do not need to be in any of the internal server subnets.

| Task | Command |
|---|---|
| Assign an IP address and default-gateway to the inband interface. | **interface inband** {**ip-address** *ip-address subnet-mask* \| **ip default-gateway** *ip-address*} <br> CLI Level: Configuration |

## Inband Management Services

If the SM-card is removed or fails completely, the out-of-band management and console ports become unavailable. You can preserve management access to the chassis in such an event by enabling SSH and any other desired management services on the inband interface.

**Note:** You may also use the inband interface as the primary management interface. Whether you are using the inband interface as the primary or backup management interface, you must at least enable SSH, and you must enable it *before* an SM-card failure or hot-swap.

| Step | Task | Command |
|---|---|---|
| 1 | Move to the inband interface CLI context. | **interface inband** <br> CLI Level: Configuration |
| 2 | Enable SSH. | **allow ssh** <br> CLI Level: Interface |
| 3 | (Optional) Enable one or more services. | **allow {http \|\| https \|\| ipmi \|\| ntp \|\| radius \|\| snmp \|\| tacacs \|\| termserv}** <br> CLI Level: Interface |

Table 2 shows the services that are available on the inband management interface as compared to the out-of-band management interface.

**Table 2:** Out-of-band Management Services vs. Inband Management Services

| Service | Out-of-band Management | Inband Management |
|---|---|---|
| IP Address | Supported | Supported |
| Gateway | Supported | Supported, but it is only for DHCP and DDNS. |
| HTTP | Supported | Supported, but you must enable through the CLI. |
| HTTPS | Supported | Supported, but you must enable through the CLI. |
| IPMI | Supported | Supported, but you must enable through the CLI, and a host can run IPMI only through the same subnet the inband IP. |
| SNMP | Supported | Supported, but you must enable through the CLI, and a host can run SNMP only through the same subnet as the inband IP. |
| Syslog | Supported | Supported |
| SSH | Supported | Supported, but you must enable through the CLI, and a host can run SSH only through the same subnet as the inband IP. |
| Telnet | Supported, but you must enable through the CLI. | Not Supported |
| NTP | Supported | Supported, but you must enable through the CLI. |
| copy image with SCP | Supported | SCP is enabled when SSH is enabled. |
| RADIUS | Supported | Supported, but you must enable through the CLI. |
| TACACS+ | Supported | Supported, but you must enable through the CLI. |

# Management Ethernet Interface

This section discusses configuration of management interfaces.

## Configuring the Management Interface

To configure the management interface.

| Step | Task | Command |
|:---:|---|---|
| 1 | Move to the Management Interface CLI context. | **interface mgmteth**<br>CLI Level: Configuration |
| 2 | Assign an IP address to the interface. | **ip address** *ip-address*<br>CLI Level: Interface |
| 3 | Specify the default-gateway. | **ip default-gateway** *ip-address*<br>CLI Level: Interface |
| 4 | (Optional) Enable the Telnet service. | **allow telnet**<br>Default: disabled<br>CLI Level: Interface |
| 5 | (Optional) Specify a one or more domain names to complete unqualified host names. Enter a separate command for each domain name. Domains are tried in the order configured. | **ip domain-list** *domain*<br>CLI Level: Interface |
| 6 | (Optional) Define the default domain name. | **ip domain-name** *domain*<br>CLI Level: Interface |
| 7 | (Optional) Specify a name-server. | **ip name-servers** *ip-address*<br>CLI Level: Interface |
| 8 | (Optional) The management interface is enabled by default. Issue a **no shutdown** if the interface is disabled. | **no shutdown**<br>CLI Level: Interface<br>Default: enabled |

# Configuring the Management Interface for DHCP

To configure the management interface for DHCP:

| Step | Task | Command |
|------|------|---------|
| 1 | Move to the Management Interface CLI context. | **interface mgmteth**<br>CLI Level: Configuration |
| 2 | Enable the DHCP client on the chassis. The DHCP client will provide the IP address and default gateway information.<br><br>Note:<br>• A shutdown of the management interface will release the DHCP IP, and bring the interface down.<br>• On 'no shutdown,' when the management interface is brought up by CLI, it will renew its IP address.<br><br>To renew DHCP leases, issue the following command: | **ip dhcp enable**<br>CLI Level: Configuration<br><br><br><br><br><br><br><br><br><br>**restart process dhclient-daemon** |
| 3 | (Optional) Enable the Telnet service. | **allow telnet**<br>Default: disabled<br>CLI Level: Interface |
| 4 | (Optional) The management interface is enabled by default. Issue a **no shutdown** if the interface is disabled. | **no shutdown**<br>CLI Level: Interface<br>Default: enabled |

## Display DHCP

The following show commands display IP addresses received from DHCP, DHCP parameters, and the status of dhclient-daemon.

| Function | Command | CLI Level |
| --- | --- | --- |
| Displays 'DHCP' if IP addresses were received from DHCP. | **show interfaces mgmteth** | Privileged |
| Displays DHCP parameters for mgmteth lease. | **show interfaces mgmteth dhcp-params** | Privileged |
| Displays status of dhclient-daemon. | **show process summary dhclient-daemon** | Privileged |

**Note:**
When DHCP IP is enabled, no other static configurations, such as IP, default gateway, or DNS are allowed. You can always disable DHCP IP to configure static management IP if needed at a later time.

# Port-channel Interfaces

Gigabit or Ten Gigabit Ethernet ports can be statically or dynamically grouped in order to function as a single port called a *port-channel* or *LAG*. Port-channels increase bandwidth and provide interface redundancy.

Dynamic port-channels are created using Link Aggregation Control Protocol (LACP). LACP is a control protocol that automatically detects multiple links between two LACP-enabled devices and trunks them together. Up to eight ports can be aggregated into single static or LACP port-channel, and up to 16 port-channels (static or LACP) can be configured. Port-channels can also have members that are on different MX-cards.

**Note:**

- Moving channel-members in and out of port-channel interface might result in incorrect ARP address resolution. SeaMicro recommends that you shut down gigabitethernet or tengigabitethernet channel-members before re-cabling them or adding/removing them from an LAG.
- In Release 3.2, slow LACP is the only option supported at the server NICs.

To create a static port-channel:

| Step | Task | Command |
|---|---|---|
| 1 | Create a port-channel. | **interface port-channel** *number* <br> CLI Level: Configuration |
| 2 | (Optional) Enter a description of the interface. | **description** *string* <br> CLI Level: Port-channel Interface |
| 3 | Add one or more interfaces as port-channel members. | **channel-member** *slot*/*port* <br> CLI Level: Port-channel Interface |
| 4 | Enable the port-channel interface. | **no shutdown** <br> CLI Level: Port-channel Interface |
| 5 | Enable the gigabitethernet or tengigabitethernet interfaces that are channel members. | **no shutdown** <br> CLI Level: Gigabitethernet Interface or Ten Gigabitethernet Interface |

To create a dynamic port-channel using LACP:

**Note:** LACP configurations are accepted only when there are no channel-members configured on the port-channel. If there are channel-members configured, un-configure them, configure LACP, and then reconfigure the channel-members.

| Step | Task | Command |
|------|------|---------|
| 1 | Create a port-channel. | **interface port-channel** *number*<br>CLI Level: Configuration |
| 2 | (Optional) Enter a description of the interface. | **description** *string*<br>CLI Level: Port-channel Interface |
| 3 | Set the port-channel mode to LACP. This command can only be issued after the port-channel has been created and before any member are added. | **mode lacp**<br>Default: static<br>CLI Level: Port-channel Interface |
| 4 | (Optional) Change the LACP mode from active to passive. | **lacp-options lacp-mode** {**passive** \| **active**}<br>Default: active<br>CLI Level: Port-channel Interface |
| 5 | (Optional) Select the frequency at which LACP-active interfaces send LACPDUs. **fast** sends one every second, **slow** sends one every 30 seconds. | **lacp-options periodic** {**slow** \| **fast**}<br>Default: slow<br>CLI Level: Port-channel Interface |
| 6 | Add an interface as a port-channel member. | **channel-member** *slot*/*port*<br>CLI Level: Port-channel Interface |
| 7 | Enable the port-channel interface. | **no shutdown**<br>CLI Level: Port-channel Interface |
| 8 | Enable the gigabitethernet or tengigabit interfaces that are channel members. | **no shutdown**<br>CLI Level: Gigabitethernet Interface or Ten Gigabitethernet Interface |

# Tengigabit Ethernet Interfaces

Aside from entering a description, the other configuration option available for tengigabitethernet interfaces is Gratuitous ARP. Gratuitous ARP is an ARP reply that is not sent in response to a request. It is instead sent to:

- inform other nodes of its presence,
- inform other nodes a hardware address change, or
- detect an IP address conflict.

To configure a tengigabitethernet interface:

| Step | Task | Command |
|---|---|---|
| 1 | Move to the Interface CLI context for a tengigabit interface.<br>**Note:** You can configure multiple interfaces with a single command by using the **range** keyword. | **interface tengigabitethernet** *slot*/*port*<br>CLI Level: Configuration |
| 2 | (Optional) Enter a description of the interface. | **description** *string*<br>CLI Level: Interface |
| 3 | (Optional) Enable gratuitous ARP. | **gratuitous-arp**<br>Default: Disabled<br>CLI Level: Interface |
| 4 | Enable the interface. | **no shutdown**<br>CLI Level: Interface |

# Interface Redundant Groups

An Interface Redundant Group (IRG) is a group of two or more (up to 16) port-channels. Only one port-channel is active at any given time, while the remaining port-channels in the group are in a stand-by state; if one port-channel goes down, the IRG switches over to a stand-by port-channel. IRGs can be used to connect to a Layer 2 switch where both increased bandwidth and link redundancy are required. You can configure up to eight IRGs per chassis.

The lowest numbered port-channel becomes the active port-channel within the IRG. When all of the interfaces in the active port-channel go down, the next lowest numbered port-channel becomes the active port-channel. After a port-channel failover, ARP entries must be re-learned, which means that traffic stops until an external or internal host ARPs the IP address.

An IRG is created when you place a port-channel in it and give it a number. This is done from Interface Port-channel context.

| Task | Command |
|------|---------|
| Assign a port-channel to an IRG. | **redundant-group** *number*<br>CLI Level: Interface |
| Display the configured IRGs and the status of an IRG. | **show interface brief**<br>CLI Level: Privileged |

```
seamicro# show interfaces brief
last run: 29.751 seconds ago
time now: Mon Oct 18 14:48:50 2010
last run: Mon Oct 18 14:48:20 2010
Name    MAC Address     Admin  Oper      PCinfo   IRG  Grat ARP  Rx Pkts  Tx Pkts
-----   --------------  -----  ----      ------   ---  --------  -------  ------
po 3    0021.5311.0943  up     down      down/0   7    n/a             0        0
po 7    0021.5311.0947  up     up        up 4/4   7    n/a         34553    39743
po 13   0021.5311.094d  up     blocked   up 2/2   0    n/a         33304      780
```

| Display the IRG to which a port-channel belongs. | **show interface port-channel**<br>CLI Level: Privileged |
|------|---------|

```
seamicro# show interfaces port-channel 10
last run: 14.784 seconds ago
time now: Mon Oct 18 14:48:20 2010
last run: Mon Oct 18 14:48:06 2010
port-channel 10 is up, line protocol is up
 Hardware is Gigabit Ethernet, address is 0021.5311.094a
 port-channel 10 has 5 channel-members, 5 channel-members are up
  Channel members: gi 0/2(up), gi 0/3(up), gi 0/4(up), gi 0/5(up), gi 1/2(up)
 Interface redundant group is 0
[...]
```

## Minimum-link LAG

You can specify the number of links in a LAG that must be up in order for the LAG to become Operationally Up. If the count of up interfaces becomes less than configured minimum, then the LAG is blocked.

If the LAG is part of an interface redundant group (IRG) when it transitions to blocked mode, then the LAG becomes a stand-by in the IRG, and another eligible LAG becomes the primary. If there is no other eligible LAG, then the current active LAG remains the primary, but it is still blocked.

When the count of up interfaces becomes greater than or equal to the minimum link requirement it is eligible to come back up from a minimum-link perspective. However, if the LAG is part of an IRG, and another LAG is active, then the LAG will remain blocked until it can transition to primary. The output of **show interfaces** command indicates that reason that the LAG is blocked. The operational status may be:

- "blocked-minimum-link" or "blocked-IRG."

| Task | Command |
|------|---------|
| Specify the minimum number of links that must be up for the LAG to be Operationally Up. | **minimum-links** *links*<br>Range: 1-8<br>Default: 1<br>CLI Level: Interface |

# Display Configurations and Status

| Task | Command |
|------|---------|
| Display the configuration, status, and statistics for an interface. | **show interfaces** [**gigabitethernet** \|<br>**inband** \| **port-channel** \| **mgmteth**]<br>CLI Level: Privileged |
| Display the configuration, status, and statistics for an interface. | **show interfaces** [**tengigabitethernet** \|<br>**inband** \| **port-channel** \| **mgmteth**]<br>CLI Level: Privileged |

# Internal Servers

This chapter contains the following sections:

# Connect Directly to Servers

You can Telnet directly to an internal server using the management IP and a port number that represents the internal server. The port number is a base number plus the server number. By default, the base number is 2000. For example, default port number for Server 1 is 2001, as 2000 plus 1 is 2001.

You can have open connections to all internal servers at the same time.

**Note:** The requirement for multiple concurrent sessions (up to 512) is that only 20 of these sessions are sending high character output rates (e.g. 9600 baud per second), while the rest are in a mostly quiescent state sending no more than 10 characters per second on average. If the number of sessions that generate high character output rates is higher than 20, then the CPU utilization will rise significantly and could lead to sluggish terminal session character input/output behavior.

The base number is configurable:

| Task | Command Syntax |
| --- | --- |
| Select a port base number for terminal server connections to internal servers. | **system console-server portbase** *number*<br>CLI Level: Configuration<br>Default: 2000 |

You can also connect to the console of an internal server from the CLI:

| Task | Command Syntax |
| --- | --- |
| Connect to the console of an internal server. | **server console connect** *server-number*<br>CLI Level: Privileged |

| Task | Command Syntax |
|---|---|
| Disconnect from the console of an internal server. | Press the Ctrl and "]" keys together, and you are taken to the "telnet>" prompt. Then, enter "quit." |

## SSH Local Port Forwarding

You can access a server console through SSH port forwarding.

| Step | Task |
|---|---|
| 1 | Forward incoming traffic on the local (client-side) port to a remote port. For connecting to an internal server use **127.0.0.1:***port-number*, as the remote port. *port-number* is a base number plus the server number, as described in . In the example below, incoming traffic on port 2329 on the local machine is forwarded to server 329 on the remote machine, ch4. <br><br> ``` $ ssh  -L 2329:127.0.0.1:2329 admin@ch4 Copyright (c) 2009-2014 SeaMicro, Inc. Welcome to SeaMicro OS! admin connected from 192.168.10.207 using ssh on seamicro seamicro> ``` |
| 2 | At this point, the remote port is available to other programs on the local host, outside of the SSH session. From another window, you can telnet to the server 329 console as shown below: <br><br> ``` $ telnet localhost 2329 Trying ::1... Connected to localhost.localdomain. Escape character is '^]'. Connecting to server 329... Success! ``` |

# Power On/Off Servers

There are two ways to power servers:

-
-

## Related Configuration Tasks

-

# Power Servers Manually

By default, as part of the system boot, internal servers are not powered on; after the system is booted, you must explicitly power on servers.

| Task | Command |
|---|---|
| Power on some or all servers. | **power-on server** {**all** | *server-number* | **assigned-to** *disk-number*}<br>CLI Level: Privileged |
| Power off some or all servers. | **power-off server** {**all** | *server-number* | **assigned-to** *disk-number*}<br>CLI Level: Privileged |

# Power Servers during System Boot

By default, as part of the system boot, internal servers are not powered on; after the system is booted, you must explicitly power on servers. However, you can configure servers to power on as part of the boot sequence.

| Step | Task | Command |
|---|---|---|
| 1 | Enter the Server CLI context. | **server id** *number*<br>CLI Level: Configuration |
| 2 | Configure the system to power on servers as part of the system boot-up process. | **power-on**<br>CLI Context: Server |

# Staggered Boot Settings

When the system boots, internal servers that are configured to power on (using the command **power-on** under the Server context) are powered on. However, all internal servers are not powered on at the same time. They are powered on in sets, and there is a delay between powering sets. This process is called *staggered boot*.

| Task | Command | CLI Level |
|---|---|---|
| You can configure the system to power on servers randomly or sequentially. | **server poweron-algorithm** {**random** | **sequential**}<br>Default: random | Configuration |

| Task | Command | CLI Level |
|------|---------|-----------|
| Specify the number of servers in a set for a staggered boot. | **server poweron-count**<br>Default: 32<br>Range: 1-512 | Configuration |
| Specify the number of seconds between powering sets during a staggered boot. | **server poweron-delay**<br>Default: 10 seconds<br>Range: 5-20, in 5 second increments | Configuration |

# BIOS Parameters

Three BIOS settings are configurable:

- **Boot order**: By default, each server will attempt to boot from the following locations, in the following order: vdisk 0, vdisk 1, vdisk 2, vdisk 3. If the OS is not found in the first location, vdisk 0, the server will attempt to boot from the next location, vdisk 1. You can change the order of preference, and also include PXE in the list. When you specify PXE in the boot order, PXE is enabled. To disable PXE, remove it from the list. Also, PXE is supported only on the first NIC of the two NICs on each N570 internal server or each Xeon server.
- **Hyperthreading**: Intel Atom CPUs are capable of hyperthreading, which is the ability to complete multiple instructions simultaneously. The benefits of hyperthreading depends on the application, and so the SeaMicro system provides the ability to turn hyperthreading on and off through the CLI.
- **C-states**: C-states are power-saving modes that the CPU can enter based on demand. When the CPU is idle, parts of the core are shut down until interrupts increase. The longer the CPU is idle, the more parts of the core are shut down. However, there is a power cost to transitioning between states, and so if the CPU is not idle for long, power consumption is negatively impacted. SeaMicro provides the ability to turn C-states on and off through the CLI to suit your application.
- **CPU Frequency Scaling**—In the SeaMicro system, internal server CPUs have CPU frequency scaling disabled in the BIOS by default. CPU Frequency Scaling is a feature that adjusts the CPU frequency up or down based on usage, in order to conserve power. A command is provided to enable and disable CPU Frequency Scaling.

    - If the command is configured for a server, the CPU will be governed by the server OS through the acpi-CPUfreq scaling_governor. On Xeon internal servers, this command
    - enables "Intel Turbo Boost Technology 2.0."
    - If the command is not configured, then the CPU is controlled by the BIOS and set at the CPU's maximum frequency.

| Step | Task | Command |
|------|------|---------|
| 1 | Enter the Server CLI context. | **server id** *number*<br>CLI Level: Configuration |

| Step | Task | Command |
|------|------|---------|
| 2 | (Optional) Change the order of the locations servers look for the OS during boot-up. | See BIOS: Boot Order on page 80. |
| 3 | (Optional) Enable CPU Frequency Scaling. | **bios cpufreq-scaling**<br>CLI Context: Server<br>Default: disabled |
| 4 | (Optional) Enable C-States for system processors. | **bios c-states**<br>CLI Context: Server<br>Default: enabled |
| 5 | (Optional) Enable/disable hyperthreading. | **bios hyperthreading off**<br>CLI Context: Server<br>Default: enabled |

| Chapter 8 | Security |
| --- | --- |

System security involves two components: authorization and authentication. You can configure one or both.

This chapter covers the following configuration topics:

# Enable Administrator Authorization

The command line interface has three levels, and each level provides greater control over the system.

- **Unprivileged**—this level has only a small subset of commands that enable you to navigate the CLI and display the system configuration.
- **Privileged**—the command set at this level enables you to control system operation, trouble-shoot, and basic system settings.
- **Configuration**—configuration-level commands enable you to control the software features which provide the main functionality of the system. From here you can customize system, network, and server settings according to your requirements.

Administrators access Privileged level using the command **enable**. As a basic method of system security, you can require administrators to enter a password in order to access Privileged level. Unauthorized users are then restricted to Unprivileged level, where system operation cannot be controlled, and settings cannot be changed.

| Step | Task | Command | CLI Level |
| --- | --- | --- | --- |
| 1 | Create a password that must be entered before an administrator may access a Privileged CLI level. | **enable password** *string* | Configuration |

# Enable Administrator Authentication

You can authenticate an administrator in two ways.

- **local authentication**—the username and password given by the administrator attempting to login is compared against the set of usernames and passwords stored on the system.
- **method list including remote authentication**—a method list specifies alternate authentication methods in case the primary method fails. The available methods are local username and password, RADIUS, and TACACS+.

## Local Authentication

Local authentication uses a username and password pair stored in the configuration file to authenticate the user, and it is enabled by default. Currently there can be only one username and password which is by default "admin" and "seamicro," respectively. The username is non-configurable, but the password is configurable.

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Create a password that must be entered before an administrator may access the system at Unprivileged level. | **passwd** *string* <br> Default: "seamicro" | Unprivileged |

When configuring the enable password, the system initially rejects passwords that are all lowercase. Upon rejection you are prompted to configure a different password, but at this prompt it will accept an all-lowercase password.

```
seamicro# passwd
Are you sure you want to change the password [no,yes] yes
Changing password for admin.
New Password:
Retype New Password: <------------------all lowercase password rejected
Please don't use an all-lower case password.
Unusual capitalization, control characters or digits are suggested.
New Password:
Retype New Password: <------------------all lowercase password accepted
```

# Authenticate using a Method List

Authentication using a method list requires you to specify one or more authentication methods; only one configurable method list is available, the "default" method list.

You can specify in any order, any or all of three login methods: Local, RADIUS, and TACACS+. If authentication fails using the first method, authentication will be attempted using the remaining methods in succession until the administrator is authenticated. If the user is not authenticated using the final method, the connection is terminated.

Enabling authentication using a method list is a two-step process:

1.  Configure an Authentication Method List on page 101.
2.  Configure the parameters for the authentication methods specified in the method list.
    *   Local Authentication on page 100
    *   RADIUS Authentication Parameters on page 102
    *   TACACS+ Authentication Parameters on page 102

**Note:** RADIUS and TACACS+ authentication over an external network is unavailable from secondary MX-cards. Authentication to secondary MX-cards is only through local authentication via the primary MX-card, or through the console connection.

## Configure an Authentication Method List

| Task | Command | CLI Level |
|------|---------|-----------|
| Select one or more authentication methods in order of preference. To enter more than one method, you must enter "[ (space)" before the first method. | **authentication login default** {**[** ‖ **local** ‖ **radius** ‖ **tacacs+** ‖ **]**} | Configuration |

## RADIUS Authentication Parameters

If you selected RADIUS as an authentication method, you must at least specify the RADIUS server address and the shared secret. Optionally, you can change the port number.

| Step | Task | Command | CLI Level |
|---|---|---|---|
| 1 | Authenticate administrators using RADIUS. | **radius-server host** *ip-address* | Configuration |
| 2 | Specify the shared secret used for encryption/decryption during RADIUS authentication. | **radius-server host** *ip-address* **secret** *key* | Configuration |
| 3 | (Optional) Change the port that RADIUS uses to authenticate users. RADIUS uses UDP port 49 by default. | **radius-server host** *ip-address* **auth-port** *port-number*<br>Default: 49 | Configuration |

## TACACS+ Authentication Parameters

If you selected TACACS+ as an authentication method, you must at least specify the TACACS+ server address, and the shared secret. Optionally, you can change the port number.

| Step | Task | Command | CLI Level |
|---|---|---|---|
| 1 | Authenticate administrators using TACACS+. | **tacacs-server host** *ip-address* | Configuration |
| 2 | Specify the shared secret used for encryption/decryption during TACACS+ authentication. | **tacacs-server host** *ip-address* **secret** *key* | Configuration |
| 3 | (Optional) Change the port that TACACS+ uses to authenticate users. TACACS+ uses TCP port 65 by default. | **tacacs-server host** *ip-address* **auth-port** *port-number*<br>Default: 65 | Configuration |

# Display Authentication Configuration

| Step | Task | Command | CLI Level |
|------|------|---------|-----------|
| 1 | Display the local and remote authentication configuration. | **show authentication** | Configuration |

# System Monitoring

There are three ways that the system can be configured to provide event-driven feedback. Additionally, you can display system information from the CLI.

# Monitor Thresholds

You can define high-usage thresholds for various system components, and log messages when the thresholds are crossed.

| Task | Command | CLI Level |
|---|---|---|
| Set a high CPU usage threshold for all of the MX-cards. If any of the cards crosses the threshold in either direction, a message is logged. | **system threshold cpu-usage** | Configuration |
| Set an rotations-per-minute threshold for the fan trays. If either of the fan trays crosses the threshold in either direction, a message is logged. | **system threshold fan-rpm** | Configuration |
| Set an upper flash memory usage limit in percentage, which after crossed (in the increasing direction) generates a system message. | **system threshold flash-memory** | Configuration |
| Set a high memory usage threshold for the MX-cards. If any of the cards crosses the threshold in either direction, a message is logged. | **system threshold memory-usage** | Configuration |

| Task | Command | CLI Level |
|------|---------|-----------|
| Set a high chassis temperature threshold. If the chassis temperature crosses the threshold in either direction, a message is logged | **system threshold sensor-temp** | Configuration |
| Display the system threshold settings. | **show run system threshold** | Privileged |

# System Logs

The SeaMicro system uses Syslog to inform administrators about system events, conditions, and behaviors. Syslogs are categorized according to severity. There are eight severity levels:

- **info**—an informational message about an event, condition, or system behavior.
- **notice**—a system event, condition or system behavior of which the user should be aware.
- **warning**—indicates a condition that might lead to an error.
- **err**—indicates an abnormal condition.
- **crit**—a condition exists that might limit system functionality.
- **alert**—immediate action is required to prevent an emergency condition.
- **emerg**—an event ocurred that rendered the system unusable.
- **debug**—debug messages describe system events at a granular level for troubleshooting. These messages are for SeaMico use only.

A list of Syslog messages and their descriptions is given in for troubleshooting.

You can choose whether to store and/or display each type of message:

| Task | Command | CLI Level |
|------|---------|-----------|
| Save Syslog messages to a buffer. | **logging buffered** *severity-level* Default: **info** | Configuration |
| Display messages on the console. | **logging console** *severity-level* Default: **err** | Configuration |
| Assign a facility number to a severity level. | **logging facility** {*local* \| **user**} | Configuration |
| Send all messages to another node on the network. | **logging host** *ip-address* | Configuration |

| Task | Command | CLI Level |
|---|---|---|
| Send messages with particular severity level to a Syslog server. | **logging trap** *severity-level* Default: **debug** | Configuration |
| Display the logging configuration. | **show run logging** | Privileged |

```
seamicro# show running-config logging
logging trap info
logging console warning
logging buffered info
logging host 192.168.1.129
logging host 192.168.1.146
```

| Task | Command | CLI Level |
|---|---|---|
| Display the logged messages. | **show logging** [**reverse**] | Privileged |

```
seamicro# show logging
Syslog logging: enabled
       logging facility: local5
       Console logging: level warning
       Buffer logging: level info
       Trap logging: level info
               Logging to 192.168.1.129
               Logging to 192.168.1.146

Log Buffer:
Jun  7 14:23:05 CHASSIS-19_seamicro syslogd: restart
Jun  7 14:23:05 CHASSIS-19_seamicro /netbsd: SeaMicro M-Card
Jun  7 14:23:05 CHASSIS-19_seamicro /netbsd: total memory = 1023 MB
```

# SNMP

**Note:** Running snmpwalk on a Windows server returns some invalid values; this issue is not SeaMicro-related. In particular, the processor entry in the device list might return "Unknown processor type," and the value returned for the disk size might not match the actual partition size.

To read MIB object values on the system from an SNMP agent, set up an SNMP community:

| Step | Task | Command | CLI Level |
|---|---|---|---|
| 1 | Setup an SNMP community. | **snmp-server community** *name* Default: public | Configuration |

To send SNMP traps to a particular host:

| Step | Task | Command | CLI Level |
|---|---|---|---|
| 1 | Send SNMP traps to a particular host | **snmp-server host** *ip-address* [**port** *number*] [**version** {**1** | **2c**}] | Configuration |

# Monitor from the CLI

You can use the following commands to monitor the system and access high level system information:

| Task | Command | CLI Level |
|---|---|---|
| Display the software version. | **show version** | Privileged |
| Display hardware information about the installed components. | **show inventory** | Privileged |
| Display uncleared alarms. | **show alarms** | Privileged |
| Collect system information for use by SeaMicro technical support. | **show tech-support detail** | Privileged |

# Part 3 — Advanced Software Features

# Advanced Internal Server Features

This chapter covers the following topics:

# NIC Bonding

The internal servers support only mode 4 (802.3ad) for interface bonding. If any other mode is configured, the bonded interface will behave as if in mode 3 (broadcast) from a chassis perspective.

**Note:** In order to change the MAC address on a bond, you must unbundle the NIC bond, change the MAC, and then re-bundle it.

## NIC Teaming for Windows

| Step | Task |
|------|------|
| 1 | Extract the Intel drivers to a folder.<br><br>Prowinx64.exe /s /e /f "C:\temp" |
| 2 | Right-click on the network adapter in device manager and select "Update Driver Software..." |
| 3 | Select "Browse my computer for driver software." |
| 4 | Select "Let me pick from a list of device drivers on my computer." |
| 5 | Select "Have Disk." |
| 6 | Select "Browse." |
| 7 | Browse to *c:\temp\pro1000\winx64\ndis61\e1g6032e.inf* |
| 8 | Click "Next" to install the driver. |
| 9 | Repeat steps 2-8 for each network adapter. |

| Step | Task |
|---|---|
| 10 | Once you have all the adapters updated, run the Intel setup program, which will install the advanced network services software with VLANs, etc. |

# Configuring Internal Server as a PXE

An internal server can be configured as a PXE server for deployment.

- A PXE server may be configured as one or more internal servers.
- A PXE server can process the request of multiple internal servers at the same time.
- Only one PXE server can be configured per DHCP pool, but multiple such DHCP pools are supported.
- PXE requests coming from other internal servers inside the chassis or other external servers in the network are supported.

# Configuring LACP with Internal Servers

No chassis configuration is required for LACP to work with internal servers, as the system will automatically configure LACP bonds. The SeaMicro chassis:

- Supports only '**slow**' mode for LACP bond with internal servers.
- Supports only '**passive**' mode for LACP bond with internal servers.

To create a port channel, enable LACP at the OS level, and:

- Set LACP to '**active**' mode.
- Set LACP to '**slow**' mode.

Use the following "**show**"commands to display information on LACP:

| Task | Command | CLI Level |
|---|---|---|
| Display LACP info. | **show lacp info** <bond \| port-channel \| server> | Privileged |
| Display LACP counters. | **show lacp counters** <port-channel \| server> | Privileged |

# Preventing ARP Flux

When a Linux machine is connected to a network segment by multiple network cards, a problem with the MAC-to-IP-address mapping can occur. The target host might respond to ARP requests

from multiple Ethernet interfaces. As a result the source host might associate the target IP address with the wrong interface. This behavior can also lead to the target IP address non-deterministically migrating in the ARP cache between multiple MAC addresses. This behavior is called "ARP flux," and SeaMicro internal servers, because they have dual NICS, are exposed to this issue.

You can use ARP options to avoid the ARP flux issue. Set arp_ignore=1 and arp_announce=2 for all interfaces by appending the lines "net.ipv4.conf.all.arp_ignore=1" and "net.ipv4.conf.all.arp_announce=2" to */etc/sysctl.config.* You can write these lines to the file by entering the commands:

```
sysctl -w net.ipv4.conf.all.arp_ignore=1
sysctl -w net.ipv4.conf.all.arp_announce=2
```

Setting arp_ignore=1 directs hosts to reply to ARP requests only out of the interface that is configured with the target IP address. Setting arp_announce=2, directs hosts to, if possible, use a source IP address that is in the same subnet as the target IP address when sending an ARP request.

Another way to avoid the ARP Flux issue is to set the arp_filter=1 by appending the line "net.ipv4.conf.all.arp_filter=1" in */etc/sysctl.conf.* You can write this line to the file by entering the command:

```
sysctl -w net.ipv4.conf.all.arp_announce=1
```

This option directs the target host to perform a route lookup and respond only from the interface that 1) received the ARP request, and 2) can be used to reach the source IP.

**Note:** Setting arp_filter=1 on the SeaMicro internal servers might cause the health-check to fail.

# Chapter 11    Virtual Local Area Network

## Overview

A Virtual Local Area Network (VLAN) is a logically segmented network, mapped over physical hardware, without regard to the physical location of the users. It is an administered and configured LAN or broadcast domain. Typically, a router creates the broadcast domain, but with VLANs, a switch can create its own broadcast domain.

When a network or traffic gets too large, a VLAN is created to effectively manage collision, increase performance, and segment traffic to appropriate ports. VLANs have the same attributes as LANs, and can include end stations (even if they are not located in the same LAN).

Configuration of different native VLAN domains allows you to segregate traffic going to different C-card servers. By configuring separate native VLAN domains on C-card servers, you can effectively manage traffic separation, without having to re-configure the server OS interface.

By default, all ports are part of a VLAN called the default VLAN. The default VLAN cannot be created or modified by the user.   The users, however, can configure non-default, untagged native VLAN or tagged VLANs on 1G/10G ports, port-channels, or server-NICs.

The SM15K allows you to configure different native VLAN domains within the system. VLANs can be configured on server-NICs, external-facing gigabitethernet ports, tengigabitethernet ports, and port-channels.

Until Release 3.2.4, SM15K supported 64 VLANs on the external interfaces and 8 VLANs per NIC. Beginning Release 3.3, SM15K will support 4094 VLANs (4K VLANs) on the internal server, NICS, and external gigabitethernet, or tengigabitethernet interfaces.

To support cloud hosting environments, you must configure 4K VLANs:

*   On all servers and each NIC, interface, port channel, and NIC bond.
*   To specify VLAN ranges, and to isolate traffic amongst different applications.

## VLAN Membership

By default all ports and server-NICs allow untagged traffic and drop tagged traffic.

When a tagged VLAN is configured on a port, the default native VLAN is still active on the ports. External servers on default native VLANs can communicate with internal servers on the default

VLAN. At the same time, external servers on a different tagged VLAN can communicate with internal servers on both tagged or untagged VLANs.

Configure a non-default untagged VLAN on a VLAN member to allow communication within the untagged VLAN domain. The untagged VLAN domain is separated from the default native VLAN domain.

The following VLAN memberships are supported:

- Gigabitethernet ports
- Tengigabitethernet ports
- Port channels
- Server NICs

Note: Configure VLANs only on port-channels.   Configuration of VLANs on an individual channel member is not allowed. Configuration of VLANs is not allowed on the management port or on the inband channel.

All ports and server-NICs are part of default VLAN.

# VLAN Configuration

Configure VLANs on a system to switch tagged or untagged traffic. You can create or delete VLANs, and add or remove memberships from VLANs. Note that native VLAN is always available, and cannot be deleted.

You will have to create a global VLAN, using the **system-vlan** command, on the system before it can be used by any other entities.

## Creating Global VLANs

To create a global VLAN, issue the following commands:

1. Enter the Privileged configuration mode:

    **configure terminal**

2. Enter the VLAN configuration command and specify a VLAN ID.
   The VLAN ID must be a unique number, or a range, such as 10-20,35,300 and so on.

    **switch system-vlan** *vlan-id*

    Example:

    seamicro(config)# **switch system-vlan 12,15-18**

---

3.  Confirm creation of a VLAN with the following command:

    **do show vlan**

    Example:

    ```
    seamicro(config)# do show vlan
    Default Vlan                  : 0
    Number of User Configured Vlans : 1
    Number of Default Vlans       : 1
    Flags : T = Tagged            U = Untagged
          : I = Incomplete bond in a VLAN
    Vlan    Port Members  62
    ```

# Deleting Global VLANs

To delete a global VLAN, issue the following commands:

1.  Enter the configuration mode.

    **configure terminal**

2.  Enter the following command:

    **no switch system-vlan**

    Example:

    ```
    seamicro(config)# no switch system-vlan 1-10
    ```

3.  Confirm deletion of a VLAN with the following command:

    **do show vlan**

    Example:

    ```
    seamicro(config)# do show vlan
    seamicro(config)# interface gigabitethernet 0/1
    seamicro(config-gigabitethernet-0/1)# untagged-traffic-drop
    seamicro(config-gigabitethernet-0/1)# end
    seamicro# show vlan
    Default Vlan                  : 0
    Number of User Configured Vlans : 7
    Number of Default Vlans       : 1
    Flags : T = Tagged            U = Untagged
    ```

```
                : I = Incomplete bond in a VLAN

                : D = interface configured for untagged traffic drop


        Vlan    Port Members
        ----- -----------------------------------------------------------------------
        ----------------------------------
        10        srv 1/0/0      (T ) ,
        20    po 1      (U ) ,srv 1/0/0   (T ) ,srv 1/0/1   (T ) ,
        30    te 1/0     (T ) ,po 1      (T ) ,srv 1/0/1   (T ) ,
        40       te 1/0          (T ) ,
        50    te 1/0      (T ) ,srv 3/0/0    (T ) ,
        60    gi 0/1       (TD) ,te 1/0      (T ) ,
        70      gi 0/1          (TD) ,te 1/0        (T ) ,srv 3/0/0      (U ) ,
```

Note: You cannot delete a global, system VLAN without deleting the VLANs from all interfaces, such as the gigabitethernet, tengigabitethernet, port channels, server-ids, and associated NICs. Global VLANs can only be cleared if it has no members, NICs, or interfaces. If you need to see what interfaces are members of a VLAN, use the **show vlan** command. Then delete or clear the VLAN from that specific interface.

# Adding VLAN Membership

During initial bootup, all gigabitethernet ports, tengigabitethernet ports, and server NICs are initialized to default native VLAN.   A maximum of 4094 global, unique VLANs can be configured per system:

- Server-NICs may become a member of up to eight VLANs.
- External gigabit ports and port channels may become a member of 4094 unique VLANs.

## Adding Tagged or Untagged VLAN Memberships

Use the following command to add VLAN or tagged VLAN membership on the following interfaces, internal servers, server-NICs, and static IP subnets.

Example

```
seamicro(config-tengigabitethernet-5/1)#no tagged-vlan ... | untagged-vlan<>


Seamicro(config)# server id { server-id }
Seamicro(config)# nic-0
Seamicro(config)# no vlan <>
Seamicro(config)# interface gigabitEthernet0/1
Seamicro(config)# no vlan <>
```

```
Seamicro(config)# interface port-channel 12
Seamicro(config)# no vlan <>
```

## Configuring and Confirming VLAN Configuration

The examples below show configuration of VLANs, and confirmation of the configuration.

**On MX-cards:**

```
interface gigabitethernet 0/1 vlan { vlan-id } tagged | untagged

Example:

seamicro(config)# interface gigabitethernet 1/4
seamicro(config-gigabitethernet-1/4)# untagged-vlan 12

interface tengigabitethernet 5/1
Seamicr(config)# tagged-vlan <id-or-range> | untagged-vlan <id> [no range]

Example:
seamicro(config)# interface tengigabitethernet 2/3
seamicro(config-tengigabitethernet-2/3)# tagged-vlan 14


interface port-channel 0/1
Seamicro(config)# tagged-vlan <id-or-range> | untagged-vlan <id> [no range]

Example:
seamicro(config)# interface port-channel 0
seamicro(config-if)# tagged-vlan 14
```

**On Server-NIC cards:**

```
server id 0 { server-id } nic 0 { nic-number }
tagged-vlan <id-or-range> | untagged-vlan <id> [ no range]

Example:
seamicro(config)# server id 95
seamicro(config-id-95)# untagged-vlan 12
```

# Removing VLAN Membership

Use the following command to remove VLAN membership from an interface:

```
seamicro(config-tengigabitethernet-5/1)# no tagged-vlan <>
```

```
Seamicr(config)# server id { server-id }
Seamicr(config)# nic-0
Seamicr(config)# no tagged-vlan <>
Seamicr(config)# interface gigabitEthernet0/1
Seamicr(config)# no vlan <>


Seamicr(config)# interface port-channel 12
Seamicr(config)# no vlan <>
```

# Managing Untagged Traffic

Traffic may be untagged for the following reasons:

- When you want to capture only tagged traffic.
- When you want to drop untagged traffic for security reasons.

Use the following command to drop all untagged traffic into the default VLAN. Tagged traffic will pass through, and untagged traffic will be dropped.

To drop incoming and outgoing untagged traffic per interface, use the following commands:

```
Seamicro(config)# interface gigabitethernet 0/1
Seamicro(config)# untagged-traffic-drop
Seamicro(config)# no untagged-traffic-drop

Seamicro(config)# interface tengigabitethernet 0/1
Seamicro(config)# untagged-traffic-drop
Seamicro(config)# no untagged-traffic-drop

Seamicro(config)# interface port-channel 0/1
Seamicro(config)# untagged-traffic-drop
Seamicro(config)# no untagged-traffic-drop

Seamicro(config)# server id { server-id }
Seamicro(config)# nic-0
Seamicro(config)# untagged-traffic-drop
Seamicro(config)# no untagged-traffic-drop
```

There is no VLAN pass-through mode support on the external gigabitethernet, tengigabitethernet, or port-channel interfaces.

# VLAN Pass-Through Mode

The system supports 4K VLANs on MX-card interfaces and on internal server NICs. You must configure the port VLAN membership on the internal server NICs, the MX-card interfaces, and port-channels.

To configure pass-through mode on a NIC, use the following command:

```
seamicro(config-nic-1)# vlan-pass-through-mode
```

VLAN-pass-through-mode supports the following:

- Allows you to configure 4094 VLANs on the NIC.
- Permits you to configure respective VLAN tagging on OS interfaces.
- Requires you to remove all prior VLAN configurations on the NIC.
- Classifies untagged traffic in the default VLAN.
- Displays the Server NICs with a "P' flag for the relevant NICs.
- Supports untagged-traffic-drops and untagged VLANs on external ports.

# Non-VLAN Pass-Through Mode

Non VLAN-pass-through-mode supports the following:

- VLAN-based traffic filtering is available on NICs.
- Untagged-traffic-drop CLI command is not supported.
- Untagged VLANs are not supported.

# VLAN and NIC Bonds

NIC Bonds are formed by bundling one or more NICs on a server in a NIC Bond. The server cards, or C-Cards, support bundling of up to eight NIC members in a Bond. When NICs are configured with VLANs and are part of a Bond, that Bond inherits all VLANs configured on those NIC members in the Bond.

For uninterrupted data traffic between internal servers or external servers, the VLAN configuration on each of the NIC members in a Bond must be identical. In order to put the NIC in a NIC Bond, make sure that all the interfaces that are part of the NIC Bond are members of the same VLAN configuration. When all the NIC members in a Bond do not have a symmetrical or identical VLAN configuration:

- The data traffic over the respective VLAN on the Bond will be interrupted.
- The Bond state will be declared incomplete.
- A **syslog** message will be generated indicating that the VLAN configuration is not symmetric for this bond.

---

In the example below, nic 0 and nic 1 on server 25 are members of Bond 0, which is also configured on the same server.  The configuration for these NICs is as follows:

```
server id 25
nic 0
untagged-vlan 10
tagged-vlan 20
tagged-vlan 30

nic 1
untagged-vlan 10
tagged-vlan 20
tagged-vlan 30
```

In this case, as all the NIC members indicate symmetric VLAN configuration, the Bond is complete in all the configured VLANs.

In another example, the two NIC members are configured in VLAN 10 and 20. However, VLAN 30 tagged is configured on nic 0 but not on nic 1.

```
server id 25
nic 0
vlan 10 untagged
vlan 20 tagged
vlan 30 tagged

nic 1
vlan 10 untagged
vlan 20 tagged
```

In such a case, the Bond will show up as incomplete for VLAN 30, and the following message will display:

```
"Incomplete Bond.All bond members are NOT in the same set of VLANs."
```

Such VLAN will be shown by flag (I) in the show VLAN commands

# VLAN Pass-Through Mode and NIC Bonds

When a NIC is part of a NIC-bond and it is in VLAN pass-through mode, all the bond members must be in the same VLAN pass-through mode and configured as VLAN members.

• Any of the NIC bond members not having the same VLAN pass-through mode configuration or VLAN configuration, will be flagged as "Incomplete" for that particular VLAN in the **show vlan** command output.
• When the bond is placed in an "Incomplete" state in a VLAN, the broadcast traffic will not be forwarded to the VLAN.

---

- Configuring the same VLAN and VLAN pass-through mode on the NICs will enable the traffic to pass on to the bond members.

# VLAN Show Commands

VLAN show commands are used to show VLANs, VLAN servers, VLAN interfaces, tagged and untagged VLANs, their membership, and the interface counters. Refer to the `show` commands below for each task.

## Show VLAN

Use the following command to display a range of VLANs, gigabit ethernet and tengigabit ethernet ports, and server NIC membership information:

```
show vlan [id | range]
```

Example:

```
seamicro# show VLAN 1-3
Default Vlan                    : 0
Number of User Configured Vlans : 4094
Number of Default Vlans         : 1
Flags : T = Tagged               U = Untagged
      : I = Incomplete bond state because of difference in the bond mem-
ber configuration.
      : D = interface configured for untagged traffic drop
      : P = Vlan pass through enabled
Vlan    Port Members
-----  -------------------------------------------------------
1       srv 1/0/2      (TP)
2       srv 1/0/2      (TP)
3       srv 1/0/2      (TP)
```

## Show Running Switch System VLAN

Use the following command to show VLANs configured in a system:

> seamicro# **show running switch system-vlan**

Example:

```
seamicro# show running switch system-vlan
system-vlan 101
system-vlan 201
```

## Show VLAN Server:

Use this command to display VLAN servers in the system.

seamicro# **show vlan server** *<server-id>*

Example:

```
Show vlan server <server-id>
    seamicro# show vlan ?
optional vlan-id: 1 to 4094
Possible completions:
  default    Show default VLAN membership
  interface
  server     Show server vlan information
  tagged     Show VLAN tagged membership
  untagged   Show VLAN un tagged membership
  |          Output modifiers
<cr>
seamicro# show vlan server 0/0

 Flag    : D = interface configured for untagged traffic drop
 Server/Nic-          Vlan(s)
 Untagged-Traffic
 Drop
 -----------  --------------------------------------------------------
srv 0/0/0        2       (U ) ,2121    (T )
srv 0/0/1        default
srv 0/0/2(D)     default
srv 0/0/3        default
srv 0/0/4        default
srv 0/0/5        default
srv 0/0/6        default
srv 0/0/7        default
seamicro#
```

## Show VLAN Interface

Use this command to display VLAN gigabit ethernet interfaces, VLAN tengigabit ethernet interfaces, and VLAN port-channels:

seamicro# **show vlan interface** ?

Example:

```
seamicro# show vlan interface ?
Possible completions:
  gigabitethernet     Show VLAN membership and configuration information
  port-channel        Show VLAN membership and configuration information
  tengigabitethernet  Show VLAN membership and configuration information


seamicro# show vlan interface gigabitethernet 4/1
Untagged Traffic Drop : Drop
 IfName              Vlan(s)
---------------- --------------------------------------------------
gi 4/1              default ,10      (T ) ,20       (T )
seamicro#


seamicro# show vlan interface tengigabitethernet 4/1
Untagged Traffic Drop : Drop
IfName               Vlan(s)
---------------- --------------------------------------------------
te 4/1              default ,10      (T ) ,20       (T )
seamicro#


seamicro# show vlan interface port-channel 1
Untagged Traffic Drop : Drop
IfName               Vlan(s)
---------------- --------------------------------------------------
Po 1                default ,10      (T ) ,20       (T )
seamicro#
```

## Show Tagged VLAN

Use this command to display all tagged VLANs and their memberships:

seamicro# **show vlan tagged**

Example:
```
seamicro# show vlan tagged
Default Vlan                       : 0
Number of User Configured Vlans : 22
Number of Default Vlans       : 1
Flags : T = Tagged                 U = Untagged
```

```
             : I = Incomplete bond in a VLAN

Vlan Port Members----- --------------------------------------------
1    po 14 ,  srv 32/0 ,  srv 33/0 ,  srv 34/0 ,   srv  35/0      ,
2    gi 6/5, srv 32/1 ,   srv 33/1 ,  srv 34/1 ,   srv  35/1      ,
seamicro#

seamicro# show vlan tagged 2
Vlan  Port Members
--------------------------------------------------------------------
2    gi 6/5  ,srv 32/1  ,srv 33/1   ,srv 34/1   ,srv  35/1       ,
seamicro#
```

## Show Untagged VLAN

Use the following command to display all untagged VLANs and their memberships:

seamicro# **show vlan untagged**

Example:
```
seamicro# show vlan untagged
Default Vlan                  : 0
Number of User Configured Vlans : 22
Number of Default Vlans       : 1
Flags : T = Tagged              U = Untagged
        : I = Incomplete bond in a VLAN

Vlan  Port Members
----- ----------------------------------------------------------------
1
2
3
4
5    gi 6/6     ,srv 32/4   ,srv 33/4   ,srv 34/4   ,srv  35/4       ,
6    gi 6/7     ,srv 32/5   ,srv 33/5   ,srv 34/5   ,srv 35/5        ,
7    po 1       ,srv 32/6   ,srv 33/6   ,srv 34/6   ,srv 35/6        ,
seamicro#
```

## Show Untagged VLAN and Membership

Use the following command to display a specific untagged VLAN and its membership:

seamicro# **show vlan untagged 5**

Example:
```
seamicro# show vlan untagged 5
Vlan Port Members
--------------------------------------------------------------------
5    gi 6/6     ,srv 32/4   ,srv 33/4   ,srv 34/4   ,srv  35/4       ,
seamicro#
```

## Show Interface Counters (Deprecated)

**This command is deprecated, and is not available in R3.3.**

Use the following command to show the number of packets in a particular VLAN.

> seamicro# **show interfaces counters vlan** *<id >*

Note: VLAN counters are supported only on gigabitethernet and tengigabitethernet ports, but not on server-NICs.

## Show VLAN Count

Use this command to display the total number of configured system VLANs.:

```
show vlan count
```

 Example:

```
seamicro# show vlan count
```
Total Vlan(s) Configured : 4094

## Show VLAN Members

Use this command to display only VLANs that have configured members. VLANs that do not contain any configured port or NIC members will be excluded from this list:

```
show vlan members
```

Example:

```
seamicro# show vlan members

Default Vlan                    : 0
Number of User Configured Vlans: 4094
Number of Default Vlans        : 1
Flags: T = Tagged              U = Untagged
     : I = Incomplete bond state because of difference in the bond member
configuration.
     : D = interface configured for untagged traffic drop
     : P = Vlan pass through enabled
Vlan    Port Members
----- -------------------------------------------------------------------
---------------------------------------------
1 gi 0/2  (T), gi 0/3 (TD), gi 0/4 (TD), gi 0/6  (TD), gi 07  (TD)
   gi 3/5  (T), gi 3/6 (T), gi 3/7 (T), te 6/0    (T), te 2/0 (T)
   te 2/1  (T), po 0   (T), po 15  (T), srv 0/0/2 (TP), srv 0/0/5 (TP)
```

```
      srv 0/0/0 (TP), srv 0/0/6(TP), srv 0/0/1(TP), srv 0/0/7(TP),srv 0/0 3
  (TP)
      srv 0/0/4 (TP), srv 1/0/2 (TP), srv 1/0/5 (TP), srv 1/0/0 (TP), srv 1
  0/6 (TP)
```

## Show VLAN Server All

Use this command to display information on all VLANs or on a specified range of servers:

**show vlan server all** [*range*]

Example:

```
seamicro# show vlan server 1/0-24/0

seamicro# show vlan server all

 Flag    : D = Server/nic configured for untagged traffic drop
         : P = Server/nic configured for vlan pass-through mode
 Server/Nic-Vlan(s)
 Untagged-Traffic Drop
 ------------   --------------------------------------------------------
 srv 0/0/0(P)      default  1-49 (T)
 srv 0/0/1(P)      default  1-49 (T)
 srv 0/0/2(P)      default  1-49 (T)
 srv 0/0/3(P)      default  1-49 (T)
 -------
```

## Show VLAN Inband

Use this command to display the VLANs configured on the inband interface:

**show vlan inband**

Example:

```
seamicro# show vlan inband
Inband Interface Vlan = Default Vlan
```

## Clear Counters (Deprecated)

**This command is deprecated, and is not available in R3.3.**

Use the following command to clear all counters on the member interfaces of a VLAN:

seamicro# **clear counters vlan**

Example:

```
seamicro# show interfaces counters vlan 2
Vlan 2 counters:
  Interface Name     RxPkts      RxBytes     TxPkts      TxBytes
  gi 6/5             0           0           0           0
seamicro#
```

Confirm VLAN deletion with the following show VLAN command:

```
CH32_seamicro(config)# do show vlan
```

Example:

```
CH32_seamicro(config)# do show vlan
Default Vlan                  : 0
Number of User Configured Vlans : 0
Number of Default Vlans       : 1
```

# High Availability

This chapter contains the following sections:

# MX-card Redundancy

When multiple MX-cards are installed one card is elected the primary, and the remaining are secondary. The primary MX-card handles configuration, global Ethernet, and system management, while secondary MX-cards handle local Ethernet interfaces and storage management. By default, the card in the lowest slot at boot-up is elected the primary; or you may specify a preferred MX-card as the primary.

There is a heartbeat mechanism between the SM-card and each MX-card. If the primary card fails, the card in the next highest slot (the stand-by MX-card) is elected the new primary; this is called a *failover*. The heartbeat may be lost due to a card failure or software failure, a manual reboot of the card, or removal of the primary card. There is also a heartbeat between each MX-card so that the stand-by can detect the failure of the primary, and the primary can detect the failure of the stand-by. An MX-card failover is also be triggered if any process exceeds its maximum restart count, and you can manually trigger a failover through the CLI.

MX-card redundancy is enabled by default when the following system requirements are met:

- At least two MX-cards must be installed, and two storage-controllers must be installed at the same levels as the MX-cards.
- The CPP/BIOS, bootloader, eFPGAs, and sFPGA image versions on at least one secondary MX-card must match the image versions of the primary MX-card. If there is an image version mismatch with any secondary MX-card, that MX-card does not become operational, and the system notifies you of the mismatch.

**Note:** Disk Supervisor cards and MX-cards without corresponding S-cards are ineligible to become the primary MX-card.

| Task | Command Syntax |
|------|----------------|
| Trigger an MX-card failover. | **system switchover** *number*<br>CLI Level: Privileged |

# Resolve an MX-card Version Mismatch

All of the MX-cards in the chassis must have the same software version. If you install an MX-card into a chassis, and its software version does not match the version on the primary MX-card, the card comes up in a subdued state. In this state, the set of operations the card can perform is strictly limited to those that are required to clear the condition. In order to clear the condition and make the MX-card operational, you must copy the correct image to the card and then reset it.

| Step | Task | Command |
|------|------|---------|
| 1 | Determine which MX-card is the primary, and which MX-card has the version mismatch (target MX-card). | **show redundancy** <br> CLI Level: Privileged |
| 2 | Copy the primary MX-card flash to the target MX-card flash. | **copy image mxcard** *primary-card* **mxcard** *new-card* <br> CLI Level: Privileged |
| 3 | Determine from which flash partition the primary MX-card is booting. | **show boot** [**mxcard** *number*] <br> CLI Level: Privileged |
| 4 | Set the target MX-card to boot from the same partition as the primary. | **system boot flash** {**0** | **1**} <br> CLI Level: Configuration |
| 5 | Reset the target MX-card. | **reset mxcard** <br> CLI Level: Privileged |
| 6 | Check the status of the MX-cards. | **show redundancy** <br> CLI Level: Privileged |

The following is an example of identifying and resolving an MX-card version mismatch:

**show redundancy** displays the primary MX-card, and the card with the software mismatch.
**show boot mxcard** *number* displays which software version the specified card has booted.

```
seamicro# show redundancy
System redundancy information
        Primary slot: 0
        Configured Primary slot:
        Eligible Primary slot(s): 0, 1
        Ineligible Primary slot(s): 3, 4, 6, 7
        Version mismatched slot(s): 2
        POST Failure slot(s):
        My slot: 0
        Number of Secondary: 6
        Redundancy enabled
        Switchover Count: 0
        Initial selection: System Selection
```

```
seamicro# show boot mxcard 0
Last boot source is flash partition-0
Next boot source is flash partition-0
    Image version : 2.5.0.0
Backup boot source is flash partition-1
    Image version : 2.4.0.0

seamicro# show boot mxcard 2
Retrieving flash image versions from slot-2
Last boot source is flash partition-0
Next boot source is flash partition-0
    Image version : 2.4.0.0
Backup boot source is flash partition-1
    Image version : 2.4.0.0
```

**copy image** copies the contents of each of the partitions on the primary MX-card flash (partitions 0 and 1) to their respective partitions on the target MX-card flash.

```
seamicro# copy image ?
Possible completions:
  flash:    Flash partition information
  mxcard:   Management/Ethernet Card
  scp:      Scp host information
seamicro# copy image mxcard: 0 ?
Possible completions:
  mxcard:   Management/Ethernet card
seamicro# copy image mxcard: 0 mxcard: 2 ?
Possible completions:
  |  <cr>
seamicro# copy image mxcard: 0 mxcard: 2
Are you sure you want to overwrite the mxcard image. This will take a few min-
utes [no,yes] yes
Copying Image from MX Card - Slot 0 Flash: 0
---------------------------------------------
Reading from flash ....
Copying Image from MX Card - Slot 0 Flash: 1
---------------------------------------------
Reading from flash ....
Copying Image to   MX Card - Slot 2 Flash: 0
---------------------------------------------
Performing Release Compatibility Check
No Hardware List
******Release 2.5.0.0 is not ISSU compliant******
WARNING: Do not interrupt, halt or power-cycle the chassis till the copy com-
mand is complete!
Copying into flash ......
Verifying the installed image .
Image version 2.5.0.0 copied successfully
Copying image to BOOTROM...
Erasing BOOTROM: ..
Updating BOOTROM: ....
Verifying the installed BOOTROM .
```

```
system: %SYSTEM-6-GENERAL-INFO: Image version 2.5.0.0 copied to mxcard-2 flash
partition-0

Copying Image to   MX Card - Slot 2 Flash: 1
---------------------------------------------
Performing Release Compatibility Check
No Hardware List
******Release 2.4.0.0 is not ISSU compliant******
WARNING: Do not interrupt, halt or power-cycle the chassis till the copy com-
mand is complete!
Copying into flash ......
Verifying the installed image .
Image version 2.4.0.0 copied successfully
system: %SYSTEM-6-GENERAL-INFO: Image version 2.4.0.0 copied to mxcard-2 flash
partition-1

MX Card - Slot 2 copied successfully
```

Above, the output of **show boot mxcard 0** showed that the primary MX-card boots from flash
partition 0. Below, **show boot mxcard 2** shows that MX-card 2 will also boot from partition 0,
so there is no need to change the system boot pointer using **system boot flash**.

```
seamicro# show boot mxcard 2
Retrieving flash image versions from slot-2
Last boot source is flash partition-0
Next boot source is flash partition-0
    Image version : 2.5.0.0
Backup boot source is flash partition-1
    Image version : 2.4.0.0
```

Resetting the MX-card will allow the card to come back up fully operational. **show
redundancy** displays that the mismatch condition is cleared.

```
seamicro# reset mxcard 2
Are you sure you want to reset the mxcard (yes/no):yes
seamicro# show redundancy
System redundancy information
        Primary slot: 0
        Configured Primary slot:
        Eligible Primary slot(s): 0, 1, 2
        Ineligible Primary slot(s): 3, 4, 6, 7
        Version mismatched slot(s):
        POST Failure slot(s):
        My slot: 0
        Number of Secondary: 6
        Redundancy enabled
        Switchover Count: 0
        Initial selection: System Selection
seamicro#
```

# Application Hitless Restart

The SeaMicro system is able to restart processes after a process crash or an MX-card failover. The process restart capability provides the system with a way to automatically recover from software process problems.

The system is minimally impacted due to a process restart that is caused by a process crash, and the restart is generally not perceived by the user beyond the appearance of a log message. When a MX-card failover occurs, some of the processes on each of the MX-cards must be restarted. Like process restarts after a process crash, process restarts due to MX-card failover generally do not impact the system, with the exception of the Ethernet application.

Because some data is synchronized across all MX-cards during normal chassis operation, and the way that the Ethernet application is restarted, data-plane traffic is hitless. That is, data traffic on hardware path continues to flow, and there is no packet loss. The control-plane and management-plane however are not hitless.

During a failover, the following functions remain up:

- **ARP**—ARP packet processing for known hosts continues to work.
- **Disks**—Disk access is unaffected.
- **Data Traffic**—Data traffic on the hardware path continues to flow and does not experience packet loss.
- **LACP Packet Processing**—LACP maintains peer connectivity.
- **Virtual Server Packet Processing**—Virtual server packet processing continues to work with the current virtual server state on each secondary MX-card.

During a failover, the following functions are down until the new primary MX-card takes control:

- **ARP/DHCP Packet Processing**—New ARPs for unknown IP addresses and DHCP packets will not be serviced for a short period of time (seconds) until the new primary MX-card takes control.
- **Disks and S-cards**—the attached S-card, all of the disks on the S-card, and all of the servers using those disks will remain functional. However, since the terminal server has failed over to another MX-card, all existing server console connections are lost. You can reconnect to a server console only after the failover completes.
- **IP Broadcast Processing**—IP broadcast processing will be down (for packets originating from internal servers) until the new primary MX-card takes control.
- **Management Connection**—The management connection will be down until the new primary MX-card takes control.
- **Terminal Server Connections**—All terminal server connections will be lost. New terminal server connections cannot be established until the new primary MX-card takes control.

**Note:** The only SeaMicro process that does not support process restart is ConfD. A ConfD process crash triggers an MX-card failover if MX-card redundancy is available or an chassis reload if MX-card redundancy is not available.

# Chapter 13      IPMI

The SeaMicro software supports Intelligent Programming Management Interface (IPMI) version 2 and Data Center Management Interface (DCMI) version 2.

This chapter contains the following sections:

# Enable IPMI/DCMI

You must create a user ID and password to enable IPMI/DCMI.

| Task | Command | CLI Level |
|---|---|---|
| Set an IPMI password. | **ipmi password** *string* | Configuration |

After creating a user ID and password, you can use ipmitool to retrieve data via APIs.

# Sensor Data Record

The Sensor Data Record (SDR) provides server status and access information.

## SDR Command

You can retrieve the SDR of all servers by using ipmitool as follows:

```
ipmitool -H <ip-address> -I lanplus  -U <userID> -P <password> sdr list
```

or

```
ipmitool -H <ip-address> -I lanplus  -U <userID> -P <password> sdr elist
```

where:

- <ip-address>: the IP address of the chassis.

- <userID>: the user ID created in the previous step, for example: "user1".
- <password>: the password created in the previous step, for example: "test".
- **list**: returns the basic list output format.
- **elist**: returns the extended list output format (user should use this format).

## SDR Example

```
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test sdr list
Server 000 Up  | disabled          | ns
Server 001 Up  | disabled          | ns
Server 002 Up  | disabled          | ns
[...]
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test sdr elist
Server 000 Up  | 00h | ns  | 23.1 | Disabled
Server 001 Up  | 01h | ns  | 23.1 | Disabled
Server 002 Up  | 02h | ns  | 23.1 | Disabled
[...]
```

## Output Description

- The first column "Server 000" indicates the server identification information.
- The second column of "list" (or last column of of "elist") shows "disabled". This is because SeaMicro does not support getting server state using ipmitool.
- "ns" indicates that some data formats are not supported.
- "00h" or "01h" is the server ID in hexadecimal format.
- "23.1" (17h.1): the SDR data is for chassis instance.

# System Event Log

The System Event Log (SEL) describes events similar to those that trigger SNMP traps. SeaMicro publishes only server events such as server-up or server-down.

## SEL Command

You can retrieve the SEL using ipmitool as follows:

```
$ ipmitool -H <ip-address> -I lanplus  -U <userID> -P <password> sel list
```

The "elist" format is more detailed:

```
$ ipmitool -H <ip-address> -I lanplus  -U <userID> -P <password> sel elist
```

## SEL Example

```
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test sel list
```

```
1 | 06/09/2010 | 23:16:06 | System Event | State Deasserted

$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test sel elist
1 | 06/09/2010 | 23:16:06 | System Event  Server 128 Down | State Asserted
2 | 06/09/2010 | 23:17:20 | System Event  Server 128 Down | State Deasserted
```

## Output Description

- The second and third columns show the date and time of the event.
- "Server 128 Down | State Asserted" indicates that the server 128 was powered-off.
- "Server 128 Down | State Deasserted" indicates that the server 128 was powered-on.

# raw Command

SeaMicro supports the "raw" command format in IPMI to execute server management functions, which include:

- Powering on a server.
- Powering off a server by force.
- Powering off a server through ACPI.
- Resetting a server.

The command syntax is as follows:

```
$ ipmitool -H <ip-address> -I lanplus -U <userID> -P <password> raw <hexstring>
```

Examples of commands to power-on, power-off, or reset a server using IPMI raw commands are shown below.

The following command powers on server 20:

```
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test raw 0x2E 1 0x00 0x7d 0xab 1 0 20
```

The following command powers off server 20 by force:

```
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test raw 0x2E 1 0x00 0x7d 0xab 5 0 20
```

The following command powers off server 20 by issuing ACPI shutdown:

```
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test raw 0x2E 1 0x00 0x7d 0xab 6 0 20
```

The following command resets server 20:

```
$ ipmitool -H 192.2.3.4 -I lanplus -U user1 -P test raw 0x2E 1 0x00 0x7d 0xab 2 0 20
```

# Part 4 — Appendices

# Appendix A Error Message Descriptions

This section lists the mnemonic/commands, the type/severity of the errors, and the error descriptions. Each error message has the following possible options for troubleshooting:

- No action may be required on your part. The situation may resolve itself over time. Or, you may try the command again later.
- You may have to call your technical support representative to report the problem.
- You may have to capture the screen output or 'tech support' output, and call your technical support representative.

The errors and error descriptions are listed under the following categories in the tables below:

- SYS Error Messages
- XENCLASS Error Messages
- CLI Helper
- CONFDSTARTER
- ROUTED
- STM2
- PMON
- ETHERNET
- SYSMGR
- FABMON
- MCMGR
- NNMT
- XENSRVR
- SMON

# SYS Error Messages

The following are possible SYS log error messages:

| No action required for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/Severity** | **Error Description** |
| GEN_DEBUG_LOG1 | DEBUG | Generic debugging log - type 1. No action required. |
| GEN_ERROR | ERROR | Generic debugging log - type 1. |
| GEN_ERROR1 | ERROR | Generic debugging log - type 1. |
| GEN_DEBUG_LOG2 | DEBUG | Generic debugging log - type 2. |
| GEN_DEBUG_LOG3 | DEBUG | Generic debugging log - type 3. |
| GEN_DEBUG_LOG4 | DEBUG | Generic debugging log - type 4. |
| GEN_INFO_LOG8 | DEBUG | Generic informational log - type 8. |
| GEN_INFO_LOG1 | INFO | Generic informational log - type 1. |
| GEN_INFO_LOG2 | INFO | Generic informational log - type 2. |
| GEN_INFO_LOG3 | INFO | Generic informational log - type 3. |
| GEN_INFO_LOG4 | INFO | Generic informational log - type 4. |
| GEN_INFO_LOG5 | INFO | Generic informational log - type 5.. |
| GEN_INFO_LOG6 | INFO | Generic informational log - type 6. |
| GEN_INFO_LOG7 | INFO | Generic informational log - type 7. |
| GEN_INFO_LOG9 | INFO | Generic informational log - type 9. |
| GEN_NOTICE_LOG7 | NOTICE | Generic notification log - type 7. |
| GEN_NOTICE_LOG1 | NOTICE | Generic notification log - type 1. |
| GEN_NOTICE_LOG2 | NOTICE | Generic notification log - type 2. |
| GEN_NOTICE_LOG3 | NOTICE | Generic notification log - type 3. |
| PROC_START | INFO | Process started. |
| CONFD_DATA_DEL | INFO | Application deleted object from config file. |
| ADD_SUB | INFO | Application received a subscription request from the specified connection. |
| REM_SUBS | INFO | Application is removing all subscriptions for the specified connection. |
| CL_CONNECT | INFO | This indicates that connection to server got accepted. |
| CL_CONNECT_NAME | INFO | This indicates that connection to server got accepted. |

| | | |
|---|---|---|
| CL_DISCONNECT | INFO | This indicates that connection to server got broken. |
| CL_DISCONNECT_NAME | INFO | This indicates that connection to server got broken. |
| CL_UNKN_MSG | DEBUG | Application got an unknown message. |
| CL_UNKN_CONN | INFO | Application got an unknown connection. |
| CL_GEN_LOG1 | DEBUG | Generic commlib log - type 1. |
| CL_GEN_ERROR1 | ERROR | Generic commlib error log - type 1. |
| CL_GEN_ERROR2 | ERROR | Generic commlib error log - type 1. |
| CL_GEN_ERROR3 | ERROR | Generic commlib error log - type 1. |
| CL_SERVER_ID_NOT_FOUND | ERROR | Requested server ID not found. |
| CL_CLIENT_ID_NOT_VALID | ERROR | Requested client ID not found. |
| PROC_EXIT1 | INFO | This indicates the specified process is starting its exit sequence |
| CONFIG_UPGRADE | ERROR | This indicates the configuration will be upgraded from the old release to the new release |
| NO_CONFIG_UPGRADE | INFO | This indicates that no configuration upgrade will be performed. |
| CONFIG_DOWNGRADE | ERROR | This indicates the configuration will be downgraded from the old release to the new release |
| CL_GEN_TRACE | ERROR | Generic commlib trace message. |
| *Call your technical support representative when you encounter the following error messages.* | | |
| DEBUG_ASSERT | ERROR | Assertion fired. When you see this message, report it. |
| ASSERT | CRITICAL | Assertion fired. When you see this message, report it. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| MEM_ALLOC_ERR | ERROR | Application encountered memory allocation error. |
| INTERNAL_ERR | ERROR | Application encountered internal error. |
| PANIC_ERR | ERROR | Application encountered unrecoverable internal error. |
| File_OPEN_ERR1 | ERROR | File is not valid. |
| File_OPEN_ERR2 | ERROR | File is not valid. |
| File_STAT_ERR1 | ERROR | File is not valid. |
| File_STAT_ERR2 | ERROR | File is not valid. |
| File_RENAME_ERR1 | ERROR | Renaming of file failed. |
| File_RENAME_ERR2 | ERROR | Renaming of file failed. |
| File_NAME_ERR | ERROR | Specified file name is too long. |
| CONFD_SESSION_ERR | ERROR | Application is not able to open a confd session. |

| | | |
|---|---|---|
| CONFD_NS_ERR | ERROR | This indicates the specified namespace was not set. |
| CONFD_DATA_RTRV_ERR | ERROR | Application could not get the value of object from config file... |
| CONFD_DATA_CREATE_ERR | ERROR | Application could not create object in config file.. |
| CONFD_DATA_SET_ERR | ERROR | Application could not create object in config file.. |
| CONFD_DATA_CD_ERR | ERROR | Application could not cd into object in config file.. |
| CONFD_GET_ELEM | ERROR | Application was unable to get a configuration element. |
| CONFD_DATA_DEL | INFO | Application deleted object from config file. |
| CONFD_FATAL | CRITICAL | Fatal error message from confD. |
| CONFD_ERROR | ERROR | Generic error from confD. |
| CONFD_SOCKET_CLOSED | ERROR | Generic error from confD. |
| FORK_FAILED | ERROR | CLI is not able to fork off a process to handle the command. |
| SIGACTION_FAILED | ERROR | Application encountered internal error. |
| PROC_KILL_FAILED | WARNING | Specified process couldn't be killed. |
| IOCTL_ERR | ERROR | Input-Output-Control failure. |
| IOCTL_ERR_INFO | INFO | Input-Output-Control failure. |
| TX_FAIL | WARNING | Input-Output-Control failure. |
| IOCTL_ERR2 | ERROR | Input-Output-Control failure. |
| IOCTL_ERR3 | ERROR | Input-Output-Control failure.. |
| IOCTL_ERR_INFO3 | INFO | Input-Output-Control failure. |
| DEVICE_OPEN_ERR1 | ERROR | Specified device couldn't be opened - type 1. |
| QUEUE_ERR | ERROR | Application passed a wrong queue ID to driver API. |
| DEVICE_OPEN_ERR2 | ERROR | Specified device couldn't be opened - type 2. |
| MEM_MAP_ERR | ERROR | This indicates that the device couldn't be memory-mapped. |
| DB_ERR1 | ERROR | This indicates that the database error - type 1. |
| DB_ERR2 | ERROR | This indicates that the database error - type 2. |
| MESSAGE_SUPPRESSED | ERROR | This indicates an earlier-occurring log message was suppressed because it occurred too frequently. |
| CL_INIT_ERR | ERROR | This indicates that the communication library initialization failed. |
| CL_SRVR_INIT_ERR | ERROR | This indicates that the communication server initialization failed. |

| | | |
|---|---|---|
| CL_SEND_MSG_ERR | ERROR | This indicates that send message failed. |
| CL_GET_ENTRY_ERR | ERROR | This indicates that get entry failed. |
| CL_COPY_MSG_ERR | ERROR | This indicates that copy message failed. |
| CL_CONN_ERR | ERROR | This indicates that connection to server failed. |
| CL_ADD_FD_ERR | ERROR | This indicates that adding the specified device to commlib poll failed. |
| CL_MALLOC_FAILED | ERROR | Communication library initialization failed. System is running with low memory. |
| CL_MSG_LENGTH_IS_SHORT | ERROR | This indicates that the communication library initialization failed. |
| CL_SERVER_REGISTRATION _FAILED | ERROR | Server already registered.. |
| NO_CONFIG_UPGRADE_PAT H | ERROR | This indicates that the configuration upgrade path doesnt exist from the old release to the new release |
| BAD_FAB_REG | WARNING | Inconsistent fabric register value. |
| IOCARD_POWER_TOGGLE | ERROR | Indicates power toggle state of IO card. |

# XENCLASS Error Messages

The following are possible XENCLASS errors

| No action required for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| OBJECT_GET_ERROR | INFO | Failed to get XEN object from database. |
| UPDATE_REC_FIELD_ERR | INFO | Failed to update record fields |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| GENERAL_ERROR2 | ERROR | General error condition. |
| VM_CREATED | ERROR | VM is created. |
| VM_DESTROYED | ERROR | VM is destroyed. |

# CLI Helper

The following are possible CLI Helper error messages:

| Capture the 'tech support' output and call representative for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| PROC_TERMINATE_ERR | ERROR | Specified process couldn't be terminated. |
| HOSTNAME_SET_ERR | ERROR | Hostname could not be set. |
| TIME_NOT_CHANGED | ERROR | Unable to change the system clock. |
| TIMEZONE_NOT_CHANGED | ERROR | Unable to change the timezone. |
| NTP_SYSTEM_RELOAD_REQUIRED | WARNING | To apply NTP configuration changes, save the configuration and reload the system. |
| PROC_RESTART | ERROR | Specified process could not be restarted. |

# CONFDSTARTER

The following are possible CONFDSTARTER error messages:

| No action required for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| CONFD_STOP | INFO | CONFD process was stopped. |
| INVALID_CFG | ERROR | Saved configuration is invalid and will be removed. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| DESC_RLIMIT | WARNING | The CONFD process was unable to increase one of its process limits potentially limiting the number of configuration sessions it can support. |

# ROUTED

The following are possible ROUTED error messages:

| No action required for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| CONNECT_OK | INFO | Successful connection to a peer process. |
| ROUTE_THROUGH_PATH | INFO | The routing system cannot find a path between two nodes without using route-through edges. This contains debugging information only. |
| ROUTE_COMPUTE_DONE | INFO | Route computing is done. Debugging information only. |
| FR_LINK_IN_SERVICE | INFO | Stable link state change. |
| FR_LINK_OUT_OF_SERVICE | INFO | Stable link state change. |
| CHASSIS_MIS_CONFIG | ERROR | Misconfiguration on the chassis. |
| PARTIAL_CHASSIS_DETECTED | ERROR | Not a production environment. Proceed with caution. |
| *Call your technical support representative when you encounter the following error messages.* | | |
| FR_LINK_DOWN | WARNING | Some links are down. Contact SeaMicro in your earliest convenience. |
| FR_DEGRADED | WARNING | Some links are down but no connectivity is lost. If alarm persists, contact SeaMicro immediately. |
| FR_ERROR | WARNING | Some fabric connectivity is lost. If alarm persists, contact SeaMicro immediately |

| Capture the 'tech support' output and call representative for the following errors. | | |
|---|---|---|
| DISCONNECT | NOTICE | This indicates connection to a peer process is broken. |
| CONNECTIVITY_LOSS | ERROR | Some fabric connectivity is lost. |
| UNREACHABLE_TARGET_SEEN | WARNING | ASIC register is not updated |
| ASIC_WRITE_ERROR_SEEN | ERROR | Writing SMOB register failed |
| LINK_WATCHDOG_SEEN | NOTICE | Watchdog interrupts have been seen. |
| LINK_HANG_SEEN | ERROR | Link hang discovered and not recovered. |

# STM2

The following are possible STM2 error messages:

| No action required for the following errors. | | |
|---|---|---|
| Mnemonic/Command | Type/ Severity | Error Description |
| SCARD_UP | NOTICE | This indicates that a connection is established with the specified Scard. |
| SCARD_DOWN | NOTICE | This indicates that a connection with the specified Scard gets disconnected. |
| SCARD_IMAGE_UPDATE | NOTICE | This indicates that Scard image update has started. |
| SCARD_IMAGE_UPDATE_SUCCESS | NOTICE | This indicates that Scard image update is successful. |
| ALERT | NOTICE | This indicates there is an alert on the specified resource. |
| NOTICE | NOTICE | Print a notice message. |
| INFO | INFO | Print an informational message. |
| DINFO | INFO | Print a debugging message. |
| TEST | INFO | Print a test message. |
| Capture the 'tech support' output and call representative for the following errors. | | |
| ERROR | ERROR | Print an error message. |
| SCARD_IMAGE_UPDATE_FAILURE | ERROR | This indicates that Scard image update has failed. |

# PMON

The following are possible PMON error messages:

| No action required for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| APPS_COMING_UP | ERROR | This indicates that system initialization is still in progress. |
| APPS_READY | ERROR | This indicates all the applications are ready. |
| APP_READY | ERROR | This indicates the specified application is ready. |
| WDOG_OP | INFO | This indicates the watchdog op. |
| CONFD_INIT | INFO | This indicates confd process was sent a message to finish initialization. |
| PROC_START | INFO | This indicates the process monitor is starting the specified process. |
| PROC_STOP | INFO | This indicates the process monitor is stopping the specified process. |
| PROC_EXIT | INFO | This indicates the specified process has exited by itself. |
| PROC_NOT_RUNNING | INFO | This indicates the specified process is not running. Process will restart automatically. |
| PROC_TERM | INFO | This indicates the specified process has been terminated. Process will restart automatically. |
| *Call your technical support representative when you encounter the following error messages.* | | |
| HB_RESP_FAILED_DEBUG | DEBUG | This indicates the process didn't respond to heartbeat. |
| HB_RESP_FAILED | INFO | This indicates the process didn't respond to heartbeat. |
| HB_SEND_FAILED | INFO | This indicates the process didn't receive the heartbeat. |
| *Capture the 'tech support' output and call representative for the following error messages.* | | |
| PROC_FORK_ERR | ERROR | This indicates the process was unable to start. |
| PROC_START_ERR | ERROR | This indicates the process was unable to start. |
| APPS_NOT_READY | ERROR | This indicates some of the applications may not be ready. |
| APP_NOT_READY | INFO | This indicates the specified application may not be ready. |
| WRONG_ARG_ERR | ERROR | One of the process argument for the specified app is wrong. |
| PROC_DIED_SIGNAL | DEBUG | Process monitor received process died signal. |

| PROC_STOPPED_HB_FAILURE | ERROR | This indicates the specified process has been stopped due to heartbeat failure. |
|---|---|---|
| PROC_STOPPED | INFO | This indicates the specified process has been stopped. |
| PROC_CRASH_FATAL | CRITICAL | This indicates reboot is initiated to handle the specified process crash. |
| PROC_CRASH | ERROR | This indicates the specified process crashed. |
| PROC_HB_FAILURE_FATAL | CRITICAL | This indicates reboot is initiated to handle the specified process heartbeat failure. |

# ETHERNET

The following are possible Ethernet error messages:

| No action required for the following errors. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| GENERAL_INFO | INFO | This indicates a general information log. |
| GENERAL_DEBUG | DEBUG | This indicates a general debug information log. |
| GENERAL_INFO2 | INFO | This indicates a general information log. |
| GENERAL_INFO_DEC | INFO | This indicates a general information log. |
| GENERAL_INFO_HEX | INFO | This indicates a general information log. |
| GENERAL_INFO_DEC2 | INFO | This indicates a general information log. |
| GENERAL_INFO_DEC3 | INFO | This indicates a general information log. |
| GENERAL_ERROR_DEC3 | ERROR | This indicates a general information log. |
| GENERAL_INFO_DEC4 | INFO | This indicates a general information log. |
| GENERAL_NOTICE | NOTICE | This indicates a general notice log. |
| GENERAL_NOTICE_DEC | NOTICE | This indicates a general notice log. |
| GENERAL_NOTICE_HEX | NOTICE | This indicates a general notice log. |
| GENERAL_NOTICE_HEX2 | NOTICE | This indicates a general notice log. |
| CHASSIS_INFO | INFO | This indicates IP packet debug tracing. |
| INIT_STATUS | NOTICE | Indicates a general informational initialization status. |
| EXIT_NOTICE | NOTICE | This indicates a note about why Ethernet is exiting. |
| CONNECT_INFO_DEC3 | INFO | This indicates connect information in log. |
| FABRIC_PORTS_WARNING | WARNING | This indicates a fabric link warning message. |

| CONFIG_INFO | INFO | This indicates config information. |
|---|---|---|
| CONFIG_INFO_DEC | INFO | This indicates config information. |
| GENERAL_INFO_HEX2 | INFO | This indicates a general information log. |
| GENERAL_INFO_HEX3 | INFO | This indicates a general information log. |
| INTERFACE_STATUS | ERROR | This indicates a interface specific message. |
| INTERFACE_INFO | INFO | This indicates a interface specific message. |
| INTERFACE_ID_INFO | INFO | This indicates a interface specific information log. |
| FABRIC_PORTS_INFO | INFO | This indicates a fabric link specific information log. |
| TAP_INTERFACE_INFO | INFO | This indicates information about tap interface. |
| UNKNOWN_SERVER_MESSAGE_INFO | INFO | This indicates unknown commlib message. |
| UNKNOWN_CLIENT_MESSAGE_INFO | INFO | This indicates unknown commlib message. |
| PKT_TX_DEBUG | DEBUG | This indicates the pkt write information. |
| PKT_RX_DEBUG | DEBUG | This indicates the pkt read information. |
| PKT_FWD_DEBUG | INFO | This indicates the pkt forward information. |
| TARGET_IP_INVALID | DEBUG | This indicates the specified node is down while processing packet. |
| ARP_REQ_IGNORED_DEBUG | DEBUG | This indicates why arp packet was ignored. |
| ARP_REQ | DEBUG | Arp request debug information. |
| ARP_REPLY | DEBUG | Arp reply debug information. |
| PING_HLTH_CHK_STATE_CHANGE | INFO | Ping Health check state change. |
| NPT_ERROR | ERROR | NPT error. |
| NPT_DEBUG | DEBUG | NPT information. |
| EMAP_DUMP | ERROR | EMAP dump. |
| EMAP_DEBUG | ERROR | EMAP debug. |
| NATMAP_DEBUG | ERROR | NATMAP debug. |
| HEALTH_CHECK_ERROR | ERROR | EMAP debug. |
| URL_HLTH_CHK_STATE_CHANGE | INFO | URL Health check state change. |
| IP_FCT_PROTO_EIP_EPORT_TRACE | DEBUG | Ingress FCT table debug. |
| IP_FCT_VIP_VPORT_TRACE | DEBUG | Ingress FCT table debug. |

| IP_FCT_NIP_NPORT_VALID_TRACE | DEBUG | Ingress FCT table debug. |
|---|---|---|
| IP_IFCT_PROTO_EIP_EPORT_TRACE | DEBUG | Egress IFCT table debug. |
| IP_IFCT_VIP_VPORT_TRACE | DEBUG | Egress IFCT table debug. |
| IP_IFCT_NIP_NPORT_VALID_TRACE | DEBUG | Egress IFCT table debug. |
| IP_TRACE | DEBUG | IP packet debug tracing. |
| DHCPD_PKT_FWD_TO_DHCPD_INFO | INFO | DHCPd packet forward information. |
| DHCPD_PKT_FWD_TO_DHCP_RELAY_INFO | INFO | DHCPd packet forward information. |
| DHCPD_PKT_FWD_TO_EFPGA_INFO | INFO | DHCPd packet forward information. |
| DHCP_MAC_FILTER_FAILED | ERROR | The reason why DHCP packet request packet was dropped. |
| DHCP_IGNORED | DEBUG | The reason why DHCP packet request packet was dropped. |
| DHCP_NON_IPv4_PACKET | ERROR | Non-IPv4 packet was received from tap interface and was dropped. |
| DHCP_NON_UDP_PACKET | ERROR | Non-UDP packet was received from tap interface and was dropped. |
| DHCP_IP_MULTICAST_PACKET | DEBUG | IP multicast packet was received from tap interface and was dropped. |
| NNMT_TRACE | DEBUG | SMOB NNMT table set/get debug. |
| NNMT_ERROR | ERROR | SMOB NNMT table set/get error. |
| EFPGA_DIAG_WARNING | WARNING | EFPGA diagnostics debug information. |
| *Resolve the following errors as recommended below.* | | |
| IP_ALREADY_IN_USE | ERROR | The user attempted to configure an IP address that is already in use. Use the show Ethernet IP command and release the IP on the corresponding server. |
| NNMT_RATE_CAPACITY_ERR | ERROR | The rate of external to internal node associations has been exceeded, possible data loss. Revisit your network architecture and server mappings to bring them within specification. |
| NODE_MACIP_CAPACITY_EXCEEDED | WARNING | An attempt has been made to exceed the number of IP addresses assigned to an internal node. Remove the excess IP address on the node. Currently, a maximum of 8 IP addresses are allowed. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |

| INTERNAL_CRITICAL | CRITICAL | Critical error. |
|---|---|---|
| GENERAL_ERROR | ERROR | General error condition. |
| MCAST_ERROR | ERROR | Mcast/ARP/IP subnet bcast general error. |
| MCAST_INFO | INFO | Mcast/ARP/IP subnet bcast general info message. |
| NICBOND_ERROR | ERROR | System has failed to learn/unlearn server NIC bond. |
| MAC_ERROR | ERROR | MAC address table general error. |
| GENERAL_ERROR2 | ERROR | General error condition. |
| GENERAL_ERROR3 | ERROR | General error condition. |
| GENERAL_ERROR4 | ERROR | General error condition. |
| GENERAL_WARNING | WARNING | General warning which may or may not be relevant. |
| GENERAL_MESSAGE | WARNING | General warning condition. |
| URL_HEALTH_CHECK_CONNECT_FAILED | DEBUG | URL health check connect() has failed. Check IP connectivity to the relevant IP. |
| NIP_TABLE_ERROR | ERROR | NIP table programming failed. |
| PCAP_CRITICAL | CRITICAL | PCAP error. |
| PCAP_CRITICAL | CRITICAL | PCAP error. |
| PCI_WRITE_ERR | ERROR | Mismatch in writing and reading a PCI offset. |
| VS_MMAP_CRITICAL | CRITICAL | VS critical error. |
| MAC_INFO | INFO | MAC address table general info message. |
| MAC_DEBUG | DEBUG | MAC address table debug info message. |

# SYSMGR

The following are possible SYSMGR error messages:

| No action required for the following error. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/Severity** | **Error Description** |
| NO_ELIGIBLE_MASTER | ERROR | This indicates that no eligible master found. |
| HIGH_TEMP_CONDITION_CLEARED | ERROR | This indicates that the high temperature condition is cleared and the system is decreasing the fan speed to handle the condition. |
| POWERING_OFF | EMERGENCY | This indicates that the system is being powered off. |
| FAN_PWM_CONFIG_FAILURE | ERROR | This indicates that the software failed to configure Fan PWM through SMCard interface. |

| REBOOT | EMERGENCY | This indicates that the management card is rebooted. |
|---|---|---|
| RELOADING_AFTER_AUTOPROVISION | EMERGENCY | This indicates that the system is reloaded. |
| RELOADING | EMERGENCY | This indicates that the system is reloaded. |
| SYSTEM_UNSTABLE | CRITICAL | This indicates that the system is in unstable state. |
| SYSTEM_SW_VERSION | ERROR | This indicates that MCard is running the specified software version. |
| CCARD_OPERATION_COMPLETE | ERROR | This indicates that the ongoing ccard operation is complete. |
| CCARD_OPERATION_COMPLETE_TIMEOUT | ERROR | This indicates that the ongoing ccard operation is forced to completed after a default time duration. |
| EMERGENCY_SHUTOFF | EMERGENCY | This indicates that the system is being shutdown as a precautionary measure. |
| CCARD_SHUTOFF | CRITICAL | This indicates that the ccard is being shutdown as a precautionary measure. |
| SERVER_SHUTOFF | CRITICAL | This indicates that the server is being shutdown as a precautionary measure. |
| MXCARD_SCARD_SHUTOFF | CRITICAL | This indicates that MXcard and Scard being shutdown as a precautionary measure. |
| SMCARD_HB_FAILURE | INFO | This indicates MXcard has failed to receive heartbeats from SMCard. |
| SMCARD_OPER_STATE | INFO | Operational state of SMCard. |
| SMCARD_UP | ERROR | SMCard is operational. |
| PS_DOWN_FOR_NODE | INFO | Not enough power to turn on the specified node. |
| PS_ON | NOTICE | Specified power supply is being turned on. |
| PRIMARY_USER_SELECTION | ERROR | User has designated the specified slot as primary. |
| PRIMARY_SYSTEM_SELECTION | ERROR | This indicates that the system has designated the specified slot as primary. |
| PRIMARY_SLOT_PICKED | ERROR | Specified slot has been picked to take over as primary. |
| BEGIN_ELECTION | ERROR | Primary/standby election has started. |
| SET_PRIMARY | ERROR | Slot has become primary. |
| SET_STANDBY | ERROR | Slot has become standby. |
| PS_OFF | NOTICE | Specified power supply is being turned off. |
| FAN_SET_RPM | NOTICE | Fan RPM is being set. |
| NODE_DOWN_FROM_SMOB | INFO | Specified node down event is received. |
| FAULT_RAISE | ERROR | Fault condition on the specified resource. |

| | | |
|---|---|---|
| FAULT_RAISE_WITH_INFO | ERROR | Fault condition on the specified resource. |
| ALERT_RAISE | ERROR | Alert on the specified resource. |
| ALERT_RAISE_WITH_INFO | ERROR | Alert on the specified resource. |
| FAULT_CLEAR | ERROR | Specified fault condition is cleared on the specified resource. |
| DIAG_MODE | ERROR | This indicates the specified card is in diagmode. |
| SMCARD_UPGRADE | ERROR | This indicates that SMcard is rebooted to upgrade its firmware. |
| SMCARD_UPGRADE_FAILURE | ERROR | This indicates that SMcard upgrade failed. |
| SET_ACTIVE_CONSOLE | ERROR | This indicates that sysmgr configured active console after SMcard reboot. |
| WAIT_SMCARD_REBOOT | ERROR | This indicates that sysmgr is waiting for SMcard reboot. |
| SET_CHASSIS_POWEROFF_REG | ERROR | This indicates that sysmgr configured power-off delay register on SMcard. |
| PCARD_EEPROM_NOT_ACCESSIBLE | INFO | This indicates that Pcard EEPROM read-write is failed. |
| SERVER_TAKEOVER | ERROR | This indicates that failed server is replaced with spare server. |
| SFP_VENDOR_NOT_SUPPORTED | ERROR | This indicates that failed server is replaced with spare server. |
| ERROR_CONFIG_LOW_POWER_MODE | ERROR | This indicates that software failed to configure prochot for given server. |
| PS_NOT_GOOD | ERROR | This indicates that software failed to configure prochot for given server. |
| MODULE_UP | ERROR | This indicates that a module is up now. |
| MODULE_DOWN | ERROR | This indicates that a module is up now. |
| MODULE_MISSING | ERROR | This indicates that a module is up now. |
| *Resolve the following errors as recommended below.* | | |
| SYSTEM_SW_VERSION_MISMATCH | ERROR | This indicates that there is a software version mismatch between this standby MCard and the primary. Use the; CLI command:copy image to upgrade. |
| SYSTEM_POST_FAILURE | ERROR | This indicates that one of the critical POST test is failed for this MXcard. Use the; CLI command:show tech-support system.It will print POST results for all MXcards |

---

| | | |
|---|---|---|
| SMCARD_BRCM_SWITCH_FAILU RE | ERROR | This indicates that the SMcard broadcom switch failed. Please replace SM card. |
| SYSTEM_DEGRADED | CRITICAL | This indicates that the System is going into degraded mode due to SMcard failure. Please replace SM card. |
| TOTAL_SMCARD_FAILURE | CRITICAL | This indicates SMcard is down. Please replace SM card. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| HIGH_TEMP_THRESHOLD_CROS SED | ERROR | This indicates that the high temperature threshold is crossed and the system is increasing the fan speed to handle the condition. |
| CCARD_CRIT_TEMP | ERROR | This indicates the temperature of ccard. |
| FANTRAY_CRIT_TEMP | ERROR | This indicates the temperature of fantray. |
| EFPGA_CRIT_TEMP | ERROR | This indicates the temperature of efpga. |
| SCARD_CRIT_TEMP | ERROR | This indicates the temperature of scard. |
| MCARD_CRIT_TEMP | ERROR | This indicates the temperature of mxcard. |
| NO_WATCHDOG_DETECTED | INFO | This indicates that the Fcard watchdog is not detected by SMcard. |
| WATCHDOG_DETECTED | INFO | This indicates that the Fcard watchdog is detected by SMcard. |
| WATCHDOG_DISABLED | ERROR | This indicates that the Fcard watchdog is disabled. |
| WATCHDOG_ENABLED | ERROR | This indicates that the Fcard watchdog is enabled. |
| WATCHDOG_I2C_FAILURE | ERROR | This indicates that the Fcard watchdog I2C is failed. |
| FANTRAY_I2C_FAILURE | ERROR | This indicates that the Fant ray controller I2C is failed. |
| FANTRAY_I2C_OK | ERROR | This indicates that the Fant ray controller I2C is working now. |
| PS_I2C_FAILURE | ERROR | This indicates that the Power supply I2C is failed. |
| PS_I2C_OK | ERROR | This indicates that the Powers upply I2C is working now. |

# FABMON

The following are possible FABMON error messages:

| No action required for the following error. | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| TRACE | DEBUG | This indicates code control has reached this point. |
| GENERAL_DEBUG | DEBUG | This indicates general debug message. |
| TIMER_CREATE | DEBUG | This indicates general debug message. |
| TIMER_DELETE | DEBUG | This indicates general debug message. |
| TIMER_EXPIRE | DEBUG | This indicates general debug message. |
| GENERAL_INFO | INFO | This indicates general info message. |
| GENERAL_NOTICE | NOTICE | This indicates general notice message. |
| GENERAL_WARNING | WARNING | This indicates general warning message. |
| GENERAL_ERROR | ERROR | This indicates general error message. |
| GENERAL_CRITICAL | CRITICAL | This indicates general critical message. |
| GENERAL_ALERT | ALERT | This indicates general alert message. |
| GENERAL_EMERGENCY | EMERGENCY | This indicates general emergency message. |
| **Resolve the following errors as recommended below.** | | |
| ROUTED_CONNECTION_BROKEN | ERROR | This indicates connection to routing-daemon is broken. Path configuration failed. Check if routing-daemon is running and if not start it. Wait for few minutes for routing-daemon to complete initialization before issuing configuration command. |
| ERROR_PATH_BROKEN | ERROR | User input cannot be processed as is. Enter correct input. |
| ERROR_SOURCE_NOT_PRESENT | ERROR | User input cannot be processed as is. Enter correct input. |
| ERROR_DESTINATION_NOT_PRESENT | ERROR | User input cannot be processed as is. Enter correct input. |
| ERROR_SOURCE_HAS_NO_IO | ERROR | User input cannot be processed as is. Enter correct input. |
| ERROR_DESTINATION_HAS_NO_IO | ERROR | User input cannot be processed as is. Enter correct input. |

# MCMGR

The following are possible MCMGR error messages:

| No action required for the following error. | | |
| --- | --- | --- |
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| GENERAL_INFO | INFO | This indicates a general information log. |
| GENERAL_INFO_DEC | INFO | This indicates a general information log. |
| GENERAL_DEBUG | DEBUG | This indicates a general debug information log. |
| GENERAL_NOTICE | NOTICE | This indicates a general notice log. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| INTERNAL_CRITICAL | CRITICAL | This indicates a critical error. |
| GENERAL_ERROR | ERROR | This indicates a general error condition. |
| GENERAL_WARNING | WARNING | This indicates a general warning which may or may not be relevant. |
| PCAP_CRITICAL | CRITICAL | This indicates a pcap error. |
| GENERAL_ERROR2 | ERROR | This indicates a general error condition. |

# NNMT

The following are NNMT error messages:

| No action required for the following error. | | |
| --- | --- | --- |
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| GENERAL_INFO | INFO | This indicates a general information log. |
| GENERAL_DEBUG | DEBUG | This indicates a general debug information log. |
| GENERAL_NOTICE | NOTICE | This indicates a general notice log. |
| NNMT_TRACE | DEBUG | This indicates SMOB NNMT table set/get debug. |
| resolve as recommended | | |
| NNMT_RATE_CAPACITY_ERR | INFO | The rate of external to internal node associations has been exceeded, possible data loss. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| INTERNAL_CRITICAL | CRITICAL | This indicates a critical error. |

| GENERAL_ERROR | ERROR | This indicates a general error condition. |
|---|---|---|
| GENERAL_WARNING | WARNING | This indicates a general warning which may or may not be relevant. |
| PCAP_CRITICAL | CRITICAL | This indicates a pcap error. |

# XENSRVR

The following is a XENSRVR error message:

| *No action required for the following error.* | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| XEN_OBJECT_GET_ERROR | INFO | Failed to get XEN object from database. |

# SMON

The following are SMON error messages:

| *No action required for the following errors.* | | |
|---|---|---|
| **Mnemonic/Command** | **Type/ Severity** | **Error Description** |
| CONNECT | INFO | This indicates a successful connection to a peer process. |
| DISCONNECT | INFO | This indicates a disconnect from a peer process. |
| *Resolve the following errors as recommended below.* | | |
| CONFIG_LINE_TOO_LONG | WARNING | This indicates a line from the configure File is too long. File a bug to report the error. |
| BAD_CONFIG_LINE | WARNING | This indicates a line from the configure File is malformed. File a bug to report the error. |
| CONFIG_BAD_CHIP_TYPE | WARNING | This indicates a line from the configure File specified a bad chiptype. File a bug to report the error. |
| BAD_REGISTER_DETAIL | WARNING | Detail of a register check failure. File a bug to report it. |
| *Capture the 'tech support' output and call representative for the following errors.* | | |
| BAD_REGISTER_VALUE | ERROR | This indicates a register value check failure. |

# Storage Errors

A few common errors, symptoms, and remedies in RAID and Non-RAID scenarios are listed below.

## RAID

---

**Symptom: Disk is physically present, but not seen on CLI.**

---

**Remedy:**

- If the disk is physically present, but not seen on CLI, run '**show alarms**.'
- If the alarm says '**disk missing**,' re-seat the disk.
- If the alarm does not say '**disk missing**,' run '**show alerts.**'
- If the alert says '**unresponsive disk**,' one of the following messages will display:
    - Could not initialize disk.
    - Could not read disk information.
    - Could not read disk rotation rate.
    - Could not read disk size.
- If the alert says '**dead disk**,' it means that the disk is not responding to any I/O requests.
- For any of the above disk errors, re-seat the disk and try again.
- Run '**show storage disk brief**.' If disk status says '**rebuilding**', wait a few minutes and re-check its status to make sure the disk is available for use.
- If the problem still persists, proceed as follows:
    - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
    - Copy '**show storage statistics**' output.
    - Copy '**show storage disk**' output for that particular disk.
- If '**show storage statistics**' displays uncorrectable errors, **replace the disk**.

---

**Symptom: S-Card down or degraded.**

---

**Remedy:**

- Verify if the problem of a down/degraded S-card occurred right after reloading the chassis.
- If yes, wait **15** minutes for firmware to reload.
- If not, follow the steps outlined below:
    - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
    - Power off and then power on the S-card.
    - If powering off and on does not fix the problem, reset the S-card.

---

- If the problem persists, re-seat the S-card.
- If the problem still exists, contact Customer Support.

---

**Symptom: Failed to create pool or volume.**

---

**Remedy:**

- Run '**show storage disk**' with disk ID.
- Check if the disk is rebuilding.
- If yes, create pools and volumes after the rebuild is complete.
- If not, follow the steps outlined below:
    - Power off and then power on the S-card.
    - If powering off and on does not fix the problem, reset the S-card.
    - If the problem persists, re-seat the S-card.
    - If the problem still exists, contact Customer Support.

---

**Symptom: Pool and volumes are created, but R/W errors are displayed after a few hours or days of runtime**

---

**Remedy:**

- Run '**show storage disk**.'
- Check if a disk in the pool is rebuilding.
- Wait **1 - 2 seconds** to see if the error disappears after the rebuild is complete.
- If yes, copy logs and syslog, and provide to Customer Support.
- If the error still persists, copy logs and syslog to provide to Customer Support.

# Non-RAID

**Symptom: LED is still red, and does not turn green even after disk insertion.**

- Check alarms to see if disk is missing.
- If the alarm shows '**disk missing**,' re-seat the disk.
- If LED is still red after re-seating the disk, replace the disk.
- If LED is still red after replacing the disk, replace the S-card.

**Symptom: Disk is physically present, but not seen on CLI.**

**Remedy:**

- If the disk is physically present, but not seen on CLI, run '**show alarms**.'
- If the alarm says '**disk missing**,' re-seat the disk.
- If the alarm does not say '**disk missing**,' run '**show alerts.**'
- If the alert says '**unresponsive disk**,' one of the following messages will display:
    - Could not initialize disk.
    - Could not read disk information.
    - Could not read disk rotation rate.
    - Could not read disk size.
- If the alert says '**dead disk**,' it means that the disk is not responding to any I/O requests.
- For any of the above disk errors, re-seat the disk and try again.
- Run '**show storage disk brief**.' If disk status says '**rebuilding**', wait a few minutes and re-check its status to make sure the disk is available for use.
- If the problem still persists, proceed as follows:
    - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
    - Copy '**show storage statistics**' output.
    - Copy '**show storage disk**' output for that particular disk.
- If '**show storage statistics**' displays uncorrectable errors, **replace the disk**.

**Symptom: Faulty S-card.**

After plugging in the S-card, all eight LEDs should turn green. If the LEDs are red, make sure the S-card is powered on.

**Remedy:**

- Power on the S-card.
- If LEDs are still red, re-seat the S-card.
- If LEDs are still red, it could mean that the S-card is not receiving power.
- Replace the S-card.
- Run '**show storage scard brief,'** and verify that the S-card status says '**Up.**'
- If the S-card status is not '**Up**,' contact Customer Support.

**Symptom: Faulty SAS port on MX-card or Faulty Expander.**

After a storage enclosure is plugged in, the LED on the SAS port should be green.

**Remedy:**

- If the LED on the SAS port is not green, re-plug the storage enclosure, and then power on.
- If the LED is still not green, try switching the storage enclosure from one SAS port to another.
- If it is still a problem, it could be a faulty SAS port on the MX-card.
- Replace the MX-card.
- Verify that the LED on the SAS port turns green.

**Symptom: Failed to create pools or volumes.**

**Remedy:**

- Copy the error message.
- Check '**show alerts**.'
- Check for '**unresponsive disk.**'
- Check for disk '**I/O Errors**.'
- If yes to any of the above, run '**show storage statistics**' or '**show storage disk**.'
- If the error is correctable:
  - Re-seat the disk.
  - Verify the disk on CLI.

- Recreate the pool.
- If the error is uncorrectable:
    - Replace the disk.
    - Recreate the pool.

---

**Symptom: S-card degraded.**

---

**Remedy:**

- Run '**show alerts**.'
- If the S-card is '**degraded**,' it means the management connection is lost. In such a case:
    - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
    - Restart STM.
    - Re-seat the S-card.
    - Make sure the S-card comes up.
    - Run '**show storage scard brief**,' and verify that the S-card status says '**Up**.'
    - If the S-card status is not '**Up**,' contact Customer Support.

---

**Symptom: S-card down.**

---

**Remedy:**

- Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
- Re-seat the S-card.
- If the S-card is still down, power off and then power on the S-card.
- Verify that the S-card status is '**Up**.'
- If the S-card status is not '**Up**,' contact Customer Support.

---

**Symptom: Read/Write Errors on C-card or application.**

---

**Remedy:**

- Run '**show alerts**.'
- Disk I/O Errors will display the disk ID (bay/port).
- Run '**show storage statistics**.'
- Storage statistics will display '**uncorrectable**' or '**correctable**.'
- If the error is uncorrectable:
    - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.

---

- Replace the disk.
- Recreate the pool and the volume with the same name and size.
- If you use the same name and size, no assignments will be necessary.
- If you use a different name and size, re-assign volumes to the server.
- If the error is correctable, it is probably a transient error. However, it is recommended that you still:
  - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
  - Power off and then power on the S-card.
  - Restart the application that failed.
  - Run '**show storage statistics'** to verify that the error has been fixed.

---

**Symptom: Expander Errors (RAID and Non-RAID).**

---

**Remedy:**

All missing disks, pool creation failure, disk I/O errors may be linked to an expander error. However, expander errors usually do not manifest into disk I/O errors.

- Run '**show alerts**.'
- Check if the problem is an '**unresponsive disk'** or one caused by disk '**I/O errors**.'
- If you notice 100s of expander errors, contact Customer Support.
- If the error is an **unreponsive disk**:
  - Re-seat the disk.
  - Copy logs, syslogs, '**show alerts**,' and '**show alarms'** output.
  - Copy **show storage statistics** output.
  - Copy **show storage disk** output for that particular disk.
- If '**show storage statistics**' displays uncorrectable errors, **replace the disk**.