Sun Network QDR InfiniBand Gateway Switch

Administration Guide for Firmware Version 2.0

Part No.: E26708-01
November 2011, Revision A

Contents

Using This Documentation

This administration guide provides detailed procedures that describe administration of the Sun Network QDR InifiniBand Gateway Switch Firmware Version 2.0 from Oracle. This document is written for technicians, system administrators, and users who have advanced experience administrating InfiniBand fabric hardware.

Product Notes

For late-breaking information an known issues about this product, refer to the product notes at:

http://www.oracle.com/pls/topic/lookup?ctx=E26699-01

Related Documentation

Documentation	Links
All Oracle products	http://www.oracle.com/documentation
Sun Network QDR InfiniBand Gateway Switch Firmware Version 2.0	http://www.oracle.com/pls/topic/lookup?ctx=E26699-01
Oracle Solaris OS and other systems software	http://www.oracle.com/technetwork/indexes/documentation/index.html#sys_sw
Oracle Integrated Lights Out Manager (ILOM) 3.0	http://www.oracle.com/pls/topic/lookup?ctx=E19860-01

Feedback

Provide feedback on this documentation at:

http://www.oracle.com/goto/docfeedback

Support and Accessibility

Description	Links
Access electronic support through My Oracle Support	http://support.oracle.com For hearing impaired: http://www.oracle.com/accessibility/support.html
Learn about Oracle's commitment to accessibility	http://www.oracle.com/us/corporate/accessibility/index.html

Troubleshooting the Gateway

These topics help you resolve some basic problems that might occur with your gateway.

Related Information

Gateway Hardware Problems

This table lists situations that might occur with gateway hardware and corrective steps that you can take to resolve the problem.

Situation	Corrective Steps
The Attention LED on a power supply is lit or the power supply seems dysfunctional.	1. Check the power supply status. See "Display Power Supply Status" on page 24. 2. Unplug the respective power cord, wait 15 minutes, then reattach the power cord. 3. If the previous steps do not rectify the situation, replace the power supply. See *Gateway Service*, servicing power supplies.

Situation	Corrective Steps
There is no network communication with the management controller.	1. Verify that the management controller is powered without faults. See "Check Chassis Status LEDs" on page 8 . 2. Verify that there is a LinkUp condition at the network management port. See "Check Network Management Port Status LEDs" on page 9. 3. Verfiy that the DHCP server is providing the IP address you are using to access the management controller. See DHCP server documentation. 4. If you are able to access the management controller through the USB connector, restart the management controller. See *Gateway Installation*, accessing the management controller and "Restart the Management Controller" on page 31. 5. If you are unable to access the management controller through the USB connector, power cycle the gateway. See *Gateway Service*, powering off power supplies and *Gateway Service*, powering on power supplies. 6. If the previous steps do not rectify the situation, replace the gateway. See *Gateway Service*, removing the gateway from the rack. See *Gateway Installation*, installaing the gateway into the rack.
The Attention LED on the gateway chassis is lit or the management controller seems dysfunctional.	1. If you are unable to access the management controller, power cycle the gateway. See *Gateway Service*, powering off the power supply and *Gateway Service*, powering on the power supply. 2. If you are able to access the management controller, restart the management controller. See "Restart the Management Controller" on page 31. 3. Check overall gateway health. See "Display Gateway General Health" on page 24 4. Verify that the gateway is within operating temperatures and voltages. See "Display Gateway Environmental and Operational Data" on page 27. 5. If the previous steps do not rectify the situation, replace the gateway. See *Gateway Service*, removing the gateway from the rack. See *Gateway Installation*, installing the gateway into the rack.
The Attention LED on a fan is lit or the fan seems dysfunctional.	1. Check the fan speed. See "Display Fan Status" on page 26. 2. If the previous step does not rectify the situation, replace the fan. See *Gateway Service*, servicing fans. 3. If the previous step does not rectify the situation, install the fan into another available slot. See *Gateway Service*, servicing fans. 4. If no other slots are available, replace the gateway. See *Gateway Service*, removing the gateway from the rack. See *Gateway Installation*, installing the gateway into the rack.

Situation	Corrective Steps
After installation, no links are operational.	1. Verify that there is at least one Subnet Manager active on the InfiniBand fabric. See "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45. 2. If no Subnet Manager is active, start the Subnet Manager within the gateway. See *Gateway Installation*, staring the Subnet Manager. 3. If the previous steps do not rectify the situation, restart the Subnet Manager. See "Disable the Subnet Manager" on page 53 and "Enable the Subnet Manager" on page 52.
After installation, not all links are operational.	1. Determine which links are nonoperational. See "Display Link Status" on page 36. 2. For links that are "`Down`", disable and re-enable the respective ports. See "Disable a Switch Chip Port" on page 40 and "Enable a Switch Chip Port" on page 41. 3. If the previous steps do not rectify the situation, disable the respective port. See "Disable a Switch Chip Port" on page 40.
There was a power outage during a firmware update.	1. If you are able to access the management controller, restart the management controller. See "Restart the Management Controller" on page 31. 2. If you are unable to access the management controller, power cycle the gateway. See *Gateway Service*, removing the gateway from the rack. See *Gateway Installation*, installing the gateway into the rack. 3. Reperform the firmware upgrade. See *Gateway Remote Administration*, upgrading the gateway firmware.

Related Information

- "InfiniBand Fabric Problems" on page 3
- "Identifying LEDs" on page 6
- "Understanding Routing Through the Gateway" on page 12
- "Switch GUIDs Overview" on page 14

InfiniBand Fabric Problems

This table lists situations that might occur with the InfiniBand fabric and corrective steps that can be taken to resolve the problem.

Situation	Corrective Steps
Performance of the InfiniBand fabric seems diminished.	1. Determine if there are errors or problems with the InfiniBand fabric. See: "Perform Comprehensive Diagnostics for the Entire Fabric" on page 67 "Find 1x, SDR, or DDR Links in the Fabric" on page 68 "Determine Which Links Are Experiencing Significant Errors" on page 69 2. Locate the affected nodes by the GUID provided in the output of the `ibdiagnet` command. See "Locate a Switch Chip or Connector From the GUID" on page 35. 3. If the problem is at a cable connection, swap the suspect cable with a known good cable or reconnect the cable to a known good remote port and repeat Step 1. See *Gateway Service*, servicing data cables. 4. If the problem still remains at the cable connection, disable and re-enable the respective port and repeat Step 1. See "Disable a Port" on page 72 and "Enable a Port" on page 73. Temporary solution: • If the problem still remains, disable the affected port. See "Disable a Port" on page 72. Permanent solution: • If the problem still remains, replace the affected component or the gateway. See *Gateway Service*, servicing data cables. See remote port's documentation for replacement procedures. See *Gateway Service*, removing the gateway from the rack. See *Gateway Installation*, installing the gateway into the rack.
An InfiniBand Link LED is blinking.	1. Disconnect and properly reconnect both ends of the respective InfiniBand cable. See *Gateway Service*, servicing the data cables. 2. If the LED is still blinking, determine the significance of the errors through use of the `ibdiagnet` command. See "Determine Which Links Are Experiencing Significant Errors" on page 69. 3. Determine which connectors map to the affected link by deconstructing the node's GUID and port. See "Locate a Switch Chip or Connector From the GUID" on page 35. 4. If some of the links are running at 1x or SDR, use that situation elsewhere in this table to rectify the problem. 5. Disable and re-enable the respective ports. See "Disable a Port" on page 72 and "Enable a Port" on page 73. 6. If the errors are still significant, swap the cable with a known good one or reconnect the cable to a known good remote port, and repeat from Step 2. 7. Depending upon what does or does not rectify the problem, replace that component. See *Gateway Service*, servicing the data cables. See remote port's documentation for replacement procedures.

Situation	Corrective Steps

Some InfiniBand links are running at 1x or SDR.	For a temporary solution:
	1. Identify the suspect links using the `ibdiagnet` command.
	See "Find 1x, SDR, or DDR Links in the Fabric" on page 68. Look for text like this:
	`-W- link with SPD=2.5 found at direct path "1,19"`
	`From: a Switch  PortGUID=0x00066a00d80001dd Port=19`
	`To:  a Switch  PortGUID=0x00066a00d80001dd Port=24`
	2. Determine which connectors map to the affected link by deconstructing the node's GUID and port.
	See "Locate a Switch Chip or Connector From the GUID" on page 35.
	3. Verify the cable connection at both ends.
	See *Gateway Service*, servicing the data cables.
	4. Disable and re-enable the respective ports.
	See "Disable a Port" on page 72 and "Enable a Port" on page 73.
	5. If the previous steps do not rectify the problem, disable the port.
	See "Disable a Port" on page 72.
	For a permanent solution:
	1. Perform the steps for a temporary solution, Steps 1 to Step 4.
	2. Swap the cable with a known good cable or reconnect the cable to a known good remote port, and repeat from Step 1.
	3. Depending upon what does or does not rectify the problem, replace that component or the gateway.
	See *Gateway Service*, servicing the data cables.
	See the remote port's documentation for replacement procedures.
	See *Gateway Service*, removin gthe gateway from the rack.
	See *Gateway Installation*, installing the gateway into the rack.

There are errors on some InfiniBand links.	1. Clear the error counters.
	See "Clear Data and Error Counters" on page 70.
	2. Start a fabric stress test.
	3. Identify the suspect links using the `ibdiagnet` command.
	See "Determine Which Links Are Experiencing Significant Errors" on page 69. Look for text like this:
	`-W- lid=0x0006 guid=0x0021283a8816c0a0 dev=48438 Port=34`
	`Performance Monitor counter : Value`
	`link_recovery_error_counter : 0x1`
	`symbol_error_counter : 0x25 (Increase by 3 during ibdiagnet)`
	4. For links that are experiencing recovery errors or substantial symbol errors, refer to other parts of this table to help identify the cause and rectify the problem.

Situation	Corrective Steps
Output of InfiniBand commands provides only GUID and port, not switch chip or QSFP connectors.	1. You can find the location of a node in the gateway by deconstructing the node's GUID and port. See "Locate a Switch Chip or Connector From the GUID" on page 35. 2. Use the `dcsport` command to provide port-to-connector and connector-to-port mapping. See "Display the Switch Chip Port to QSFP Connector Mapping" on page 34.

Related Information

- "Gateway Hardware Problems" on page 1
- "Identifying LEDs" on page 6
- "Understanding Routing Through the Gateway" on page 12
- "Switch GUIDs Overview" on page 14

Identifying LEDs

Status LEDs are used on many components of the gateway chassis as a means of indicating the component's state.

- "Front Status LEDs" on page 7
- "Rear Status LEDs" on page 8
- "Check Chassis Status LEDs" on page 8
- "Check Network Management Port Status LEDs" on page 9
- "Check Link Status LEDs" on page 10
- "Check Power Supply Status LEDs" on page 10
- "Check Fan Status LEDs" on page 11

Related Information

- "Gateway Hardware Problems" on page 1
- "InfiniBand Fabric Problems" on page 3
- "Understanding Routing Through the Gateway" on page 12
- "Switch GUIDs Overview" on page 14

Front Status LEDs

FIGURE: Front Status LEDs

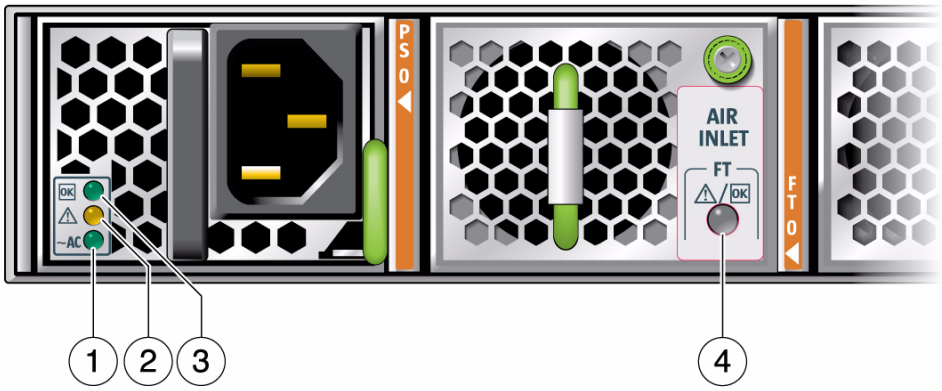


Figure Legend

1	Power supply AC LED
2	Power supply Attention LED
3	Power supply OK LED
4	Fan Attention LED

Related Information

- "Check Power Supply Status LEDs" on page 10
- "Check Fan Status LEDs" on page 11

Rear Status LEDs

FIGURE: Rear Status LEDs

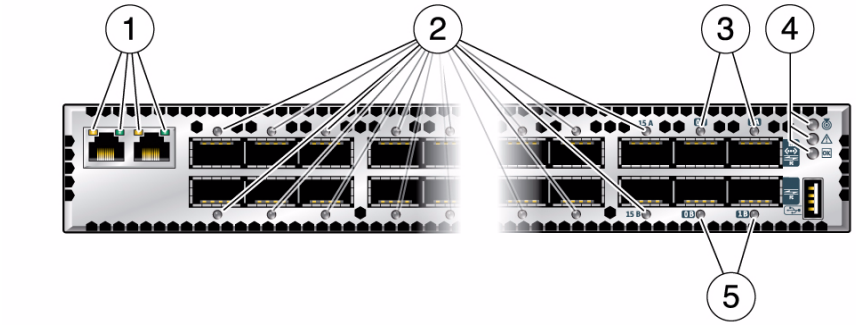


Figure Legend

1	NET status LEDs
2	InfiniBand Link status LEDs
3	Ethernet/Fibre Channel link status LEDs
4	Chassis status LEDs
5	Fibre Channel link status LEDs

Related Information

- "Check Chassis Status LEDs" on page 8
- "Check Network Management Port Status LEDs" on page 9
- "Check Link Status LEDs" on page 10

▼ Check Chassis Status LEDs

The chassis status LEDs are located on the left side of the rear panel. See "Rear Status LEDs" on page 8.

1. **Visually inspect the chassis status LEDs.**
2. **Compare what you see to this table.**

Glyph	Location	Name	Color	State and Meaning
	Top	Locator	White	On – No function. Off – Disabled. Flashing – The gateway is identifying itself.
	Middle	Attention	Amber	On – Normal fault detected. Off – No faults detected. Flashing – No function.
	Bottom	OK	Green	On – Gateway is functional without fault. Off – Gateway is off or initializing. Flashing – No function.

Related Information

- "Display Gateway Environmental and Operational Data" on page 27

▼ Check Network Management Port Status LEDs

The network management port status LEDs are located on the network management connector of the rear panel. See "Rear Status LEDs" on page 8.

1. **Visually inspect the NET status LEDs.**
2. **Compare what you see to this table.**

Name	Location	Color	State and Meaning
Link speed	Left	Amber or green	Amber on – 100BASE-T link. Green on – 1000BASE-T link. Off – No link or link down. Flashing – No function.
Activity	Right	Green	On – No function. Off – No activity. Flashing – Packet activity.

▼ Check Link Status LEDs

The link status LEDs are located at the data cable connectors of the rear panel. See "Rear Status LEDs" on page 8.

1. **Visually inspect the link status LEDs.**
2. **Compare what you see for a particular link to this table.**

Name	Color	State and Meaning
Link	Green	On – Link established. Off – No link or link down. Flashing – Symbol errors.

Related Information

- "Display Link Status" on page 36
- "Display the Link Status of a Node" on page 60
- "Display Counters for a Node" on page 61

▼ Check Power Supply Status LEDs

The power supply status LEDs are located on the power supply at the front of the chassis. See "Front Status LEDs" on page 7.

1. **Visually inspect the power supply's status LEDs.**
2. **Compare what you see on the power supply to this table.**

Glyph	Location	Name	Color	State and Meaning
	Top	OK	Green	On – 12 VDC is supplied. Off – No DC voltage is present. Flashing – No function.
	Middle	Attention	Amber	On – Fault detected, 12 VDC shut down. Off – No faults detected. Flashing – No function.
	Bottom	AC	Green	On – AC power present and good. Off – AC power not present. Flashing – No function.

Caution – If a power supply has shut down because of a thermal or overcurrent condition, signified by the amber Attention LED lighting, remove the respective power cord from the chassis. Allow the power supply to completely cool for at least

15 minutes. A shorter cooling time might cause damage to the power supply when the power cord is reattached. If the Attention LED lights amber upon reattaching the power cord, replace the power supply.

Related Information

- "Display Power Supply Status" on page 24
- "Check Board-Level Voltages" on page 25

▼ Check Fan Status LEDs

The fan status LEDs are located in the lower right corner of the fans at the front of the gateway chassis. See "Front Status LEDs" on page 7.

1. **Visually inspect the fan status LEDs.**
2. **If the LED is lit, there is a problem with that fan.**

Related Information

- "Display Fan Status" on page 26

Understanding Routing Through the Gateway

The tables in these topics describe the routing through the gateway. The first table maps the switch chip port to a QSFP connector. The second table provides a reverse map. When command output provides a switch chip port, you can use these tables to determine the route that link is following.

- "Switch Chip Port to QSFP Connectors and Link LED Routes" on page 12
- "QSFP Connectors and Link LEDs to Switch Chip Port Routes" on page 13
- "Signal Route Through the Gateway" on page 14

Related Information

- "Gateway Hardware Problems" on page 1
- "InfiniBand Fabric Problems" on page 3
- "Identifying LEDs" on page 6

- "Switch GUIDs Overview" on page 14

Switch Chip Port to QSFP Connectors and Link LED Routes

Port	Connector	Port	Connector	Port	Connector	Port	Connector
1	1A-ETH-1(P1) 1A-ETH-2(P2)	10	13B	19	0B	28	4A
2	1A-ETH-3(P3) 1A-ETH-4(P4)	11	12A	20	0A	29	5B
3	0A-ETH-1(P1) 0A-ETH-2(P2)	12	12B	21	1B	30	5A
4	0A-ETH-3(P3) 0A-ETH-4(P4)	13	9B	22	1A	31	8A
5	15A	14	9A	23	2B	32	8B
6	15B	15	10B	24	2A	33	7A
7	14A	16	10A	25	3B	34	7B
8	14B	17	11B	26	3A	35	6A
9	13A	18	11A	27	4B	36	6B

Related Information

- "Display the Switch Chip Port to QSFP Connector Mapping" on page 34
- "Display Link Status" on page 36
- "Display Switch Chip Port Status" on page 37
- "QSFP Connectors and Link LEDs to Switch Chip Port Routes" on page 13
- "Signal Route Through the Gateway" on page 14

QSFP Connectors and Link LEDs to Switch Chip Port Routes

Connector Group	Connector A Port	Connector B Port	Connector Group	Connector A Port	Connector B Port
0	20	19	9	14	13
1	22	21	10	16	15
2	24	23	11	18	17
3	26	25	12	11	12
4	28	27	13	9	10
5	30	29	14	7	8
6	35	36	15	5	6
7	33	34	0	3 (ETH-1, ETH-2) 4 (ETH-3, ETH-4)	
8	31	32	1	1 (ETH-1, ETH-2) 2 (ETH-3, ETH-4)	

Related Information

- "Display the Switch Chip Port to QSFP Connector Mapping" on page 34
- "Display Link Status" on page 36
- "Display Switch Chip Port Status" on page 37
- "Switch Chip Port to QSFP Connectors and Link LED Routes" on page 12
- "Signal Route Through the Gateway" on page 14

Signal Route Through the Gateway

By combining the information from the tables in "Understanding Routing Through the Gateway" on page 12, it is possible to determine a route through the gateway. This topic describes a sample situation that might occur.

1. A route is initiated at connector 2A. The LED blinks.
2. Using "QSFP Connectors and Link LEDs to Switch Chip Port Routes" on page 13, it is determined that the link routes to the switch chip through port 24.
3. The Subnet Manager instructs the switch chip to use port 5 to forward the link.

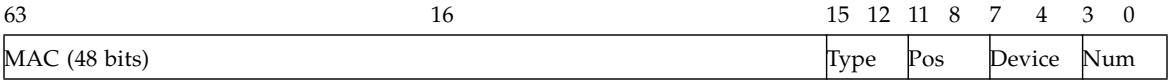
4. Using "Switch Chip Port to QSFP Connectors and Link LED Routes" on page 12, it is determined that the link routes to connector 15A. The LED blinks.

Related Information

- "Switch Chip Port to QSFP Connectors and Link LED Routes" on page 12
- "QSFP Connectors and Link LEDs to Switch Chip Port Routes" on page 13
- "Display the Switch Chip Port to QSFP Connector Mapping" on page 34
- "Display a Route Through the Fabric" on page 59

Switch GUIDs Overview

Global unique identifiers (GUIDs) are unique 64-bit strings that identify nodes such as switches and channel adapters. For the Sun Datacenter InfiniBand switches and this gateway, GUIDs are modified to identify the node's role and location. This table describes the GUID's structure.



The five fields of the GUID are described as follows:

- **MAC** – Bits 63 through 16 are the Machine Allocation Code (MAC) address. A standard for network components, the MAC address is typically provided by manufacturers in a 6-byte, colon delimited string. For example, 00:11:22:33:44:55.
- **Type** – Bits 15 through 12 is the type of board on which the node resides.
- **Pos** – Bits 11 through 8 identify the position of the board within the gateway.
- **Device** – Bits 7 through 4 identify which device on the board has that node.
- **Num** – Bits 3 through 0 are numbers reserved for the programs which modify the GUID. In most occurrences, the value is 0x2.

This table provides values for Type, Pos, Device.

Board	Type	Position	Device
Fabric card	0xF	0x0 (Fabric card 0)–0x8 (Fabric card 8)	0xA (I4 chip 0)–0xB (I4 chip 1)
Line card	0x1	0x0 (Line card 0)–0x8 (Line card 8)	0xA (I4 chip 0)–0xD (I4 chip 3)

Board	Type	Position	Device
Gateway	0xC	0x0	0xA (I4 chip 0), 0x0 (BX chip 0), 0x4 (BX chip 1)
36-Port	0xA	0x0	0xA (I4 chip 0)
72-Port	0xB	0x0	0xA (I4 chip 0)–0xF (I4 chip 5)

For example, given this output from the `ibnodes` command:

```
Ca      : 0x0021283bad45c000 ports 2 "SUN IB QDR GW switch gw-2 Bridge 0"
```

The GUID is 0x0021283bad45c000 or 0021283BAD45 C 0 0 0. Using the information provided in this topic:

- The MAC address is 0x0021283BAD45 or 00:21:28:3B:AD:45.
- The type is 0xC, or a gateway.
- The position is 0x0.
- The device is 0x0, or BridgeX chip 0.
- The number is 0.

Related Information

- *Gateway Reference*, `ibnetdiscover` command
- *Gateway Reference*, `ibnodes` command
- *Gateway Reference*, `ibswitches` command
- *Gateway Reference*, `ibhosts` command
- "Identify All Switches in the Fabric" on page 56
- "Gateway Hardware Problems" on page 1
- "InfiniBand Fabric Problems" on page 3
- "Identifying LEDs" on page 6
- "Understanding Routing Through the Gateway" on page 12

Understanding Administrative Commands

These topics provide an overview of administrative tasks and the command sets to perform those tasks. Administering the gateway requires accessing the management controller.

Related Information

- "Troubleshooting the Gateway" on page 1
- "Administering the Chassis" on page 23
- "Administering the I4 Switch Chip" on page 33
- "Administering Gateway Resources" on page 99
- "Administering the InfiniBand Fabric" on page 55
- "Administering the Subnet Manager" on page 43

Configuration Overview

After you have installed the gateway, you can configure its functionality. This table lists functions of the gateway to configure, in the order of dependency. If you do not want to configure a particular functionality, skip that step.

Step	Functionality	Description	Links
1.	Configure the Subnet Manager.	Configure the priority.	"Set the Subnet Manager Priority" on page 48
		Configure the prefix.	"Set the Subnet Manager Prefix" on page 49
		Enable controlled handover.	"Enable Subnet Manager Controlled Handover" on page 50
		Configure the M_Key.	"Set the Subnet Manager Management Key" on page 50
2.	Configure the InfiniBand fabric.	Create a fabric configuration.	"Create a Fabric Configuration" on page 77
3.	Partition the InfiniBand fabric.	Collect GUIDs.	"Identify All Switches in the Fabric" on page 56 "Identify All HCAs in the Fabric" on page 57 "Display Gateway Port Information" on page 133
		Create the `smnodes` file.	"Use the `smnodes` Command" on page 88
		Create the partition information file.	"Determine Partitions and P_Keys" on page 89
		Configure partitions and P_Keys.	"Create a User Partition" on page 91
		Add GUIDs to the partitions.	"Add or Remove a Port From a Partition" on page 93
4.	Configure LAGs.	Enable LAG mode.	"Enable LAG Mode" on page 137
		Create LAGs.	"Create LAGs" on page 138
5.	Enable host support of the gateway.	Acquire BXOFED software.	"Acquire the BXOFED Software (Linux)" on page 100
		Install BXOFED software.	"Install the BXOFED Software (Linux)" on page 101
6.	Determine the VNIC mode.	Use Gateway Manual Mode or Host Manual Mode.	"Gateway Manual Mode Overview (Linux)" on page 103 "Host Manual Mode Overview (Linux)" on page 108
7.	Configure VNICs.		
	Gateway Manual Mode	Create the MAC address file.	"Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104
		Append the MAC address file.	"Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105
		Create VLANs.	"Create VLANs" on page 143
		Create VNICs.	"Create VNICs" on page 145

Step	Functionality	Description	Links
	Host Manual Mode	Determine the configuration file.	"Central Configuration File (Linux)" on page 108 "VNIC-Specific Configuration File (Linux)" on page 110
		Create the MAC address file.	"Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
		Create the configuration file.	"Create the Central Confguration File (Linux)" on page 113 "Create the VNIC-Specific Confguration Files (Linux)" on page 114
		Enable VLANs and VNICs.	"Configure and Create VNICs for Host Manual Mode (Linux)" on page 117
8.	Save the configuration.	Back up the configuration for restoration at a later time, should it become necessary.	*Gateway Remote Administration*, backing up the configuration.

Related Information

- "Hardware Command Overview" on page 19
- "InfiniBand Command Overview" on page 20
- "ILOM Command Overview" on page 20

Hardware Command Overview

The management controller uses a simplified Linux OS and file system. From the # prompt on the management controller, you can type hardware commands to perform some administrative and management tasks. Alternatively, you can use the Oracle ILOM restricted Linux shells, which are the prefferred method of issuing hardware commands. See *Gateway Remote Administration*, Linux restricted shells.

Hardware commands are user-friendly and can perform some testing upon the switch chip, enabling greater control of the gateway and its operation.

After you log in to the `root` account, the shell prompt (#) appears, and you can enter shell commands. Enter the hardware commands in this format:

`#` *command* `[`*arguments*`][`*arguments*`]...`

Related Information

- *Gateway Reference*, understanding hardware commands
- "InfiniBand Command Overview" on page 20
- "ILOM Command Overview" on page 20

InfiniBand Command Overview

The InfiniBand commands are a means of monitoring and controlling aspects of the InfiniBand fabric. These commands are also installed on and run from the management controller, which is also the host of the Subnet Manager. Use of these commands requires thorough knowledge of InfiniBand architecture and technology.

After you log in to the `root` account, the shell prompt (`#`) appears, and you can enter shell commands. Alternatively, you can use the Oracle ILOM restricted Linux shells, which are the prefferred method of issuing InfiniBand commands. See *Gateway Remote Administration*, Linux restricted shells.

Enter the InfiniBand commands in this format:

```
# command [option][option] ...
```

Related Information

- *Gateway Reference*, understanding InfiniBand commands
- "Hardware Command Overview" on page 19
- "ILOM Command Overview" on page 20

ILOM Command Overview

The Oracle ILOM CLI, web, SNMP, and IPMI interfaces enable additional administration features and capabilities. The Oracle ILOM CLI, SNMP, and IPMI interfaces use remote clients for command-line administration of many Oracle ILOM targets. Alternatively, the Oracle ILOM web interface permits point-and-click administration of the Oracle ILOM components and services.

After you log in to the `ilom-admin` account, the Oracle ILOM prompt (`->`) appears, and you can enter Oracle ILOM commands. Enter the Oracle ILOM commands in this format:

-> *command* [*option*][*target*][*property=value*]...

Information about ILOM support of the gateway is available in the *Oracle Integrated Lights Out Manager (ILOM) 3.0 Supplement for the Sun Network QDR InfiniBand Gateway Switch*, available online at:

http://www.oracle.com/pls/topic/lookup?ctx=E19671-01&id=homepage

Overall information about Oracle ILOM 3.0 is available online at:

http://www.oracle.com/pls/topic/lookup?ctx=E19860-01&id=homepage

Related Information

- *Gateway Remote Administration*, understanding Oracle ILOM commands
- "Hardware Command Overview" on page 19
- "InfiniBand Command Overview" on page 20

Administering the Chassis

These topics describe the administration of the gateway.

Related Information

- "Troubleshooting the Gateway" on page 1
- "Understanding Administrative Commands" on page 17
- "Administering the I4 Switch Chip" on page 33
- "Administering Gateway Resources" on page 99
- "Administering the InfiniBand Fabric" on page 55
- "Administering the Subnet Manager" on page 43

Monitoring the Chassis

These topics enable you to display and check the operation and status of the gateway.

Related Information

- "Monitoring the I4 Switch Chip" on page 33
- "Monitoring Gateway Resources" on page 128
- "Monitoring the InfiniBand Fabric" on page 55
- "Monitoring the Subnet Manager" on page 43
- "Controlling the Chassis" on page 30

▼ Display Gateway General Health

An easy way to perform a quick check of the gateway's operation is with the `showunhealthy` command.

- **On the management controller, type.**

```
# showunhealthy
OK - No unhealthy sensors
#
```

Related Information

- *Gateway Reference*, `showunhealthy` command
- "Display Gateway Environmental and Operational Data" on page 27

▼ Display Power Supply Status

The `checkpower` command performs a simple pass-fail test on the power supplies.

1. **On the management controller, type.**

```
# checkpower
PSU 0 present status: OK
PSU 1 present status: Alert
#
```

2. **If you see the words `Alert` or `Not Present` in the command output, perform these steps for the affected power supply.**

   a. **Remove the power cord and affected power supply from the gateway chassis.**

      See *Gateway Service*, powering off the power supply and *Gateway Service*, removing the power supply.

   b. **Wait one minute.**

c. **Reinstall the power supply.**

   See *Gateway Service*, installing the power supply.

d. **Reattach the power cord.**

   See *Gateway Service*, powering on the power supply.

e. **If the power supply's Attention LED lights or the `checkpower` command still reports `Alert` or `Not Present` for the power supply, replace the power supply.**

   See *Gateway Service*, servicing the power supplies.

Related Information

- *Gateway Reference*, `checkpower` command
- "Check Board-Level Voltages" on page 25
- "Display Gateway Environmental and Operational Data" on page 27

▼ Check Board-Level Voltages

When you want to know if the voltages used by the various components within the gateway are within nominal values, use the `checkvoltages` command. If a voltage deviates more than 10% of its nominal value, there is a problem. This check also reports if a battery has failed.

- **On the management controller, type.**

```
# checkvoltages
Voltage ECB OK
Measured 3.3V Main = 3.30 V
Measured 3.3V Standby = 3.42 V
Measured 12V = 12.06 V
Measured 5V = 5.03 V
Measured VBAT = 3.17 V
Measured 1.0V = 1.01 V
Measured I4 1.2V = 1.22 V
Measured 2.5V = 2.51 V
Measured V1P2 DIG = 1.18 V
Measured V1P2 ANG = 1.18 V
Measured 1.2V BridgeX = 1.22 V
Measured 1.8V = 1.80 V
Measured 1.2V Standby = 1.20 V
All voltages OK
#
```

Related Information

- *Gateway Reference*, `checkvoltages` command
- "Display Power Supply Status" on page 24
- "Display Gateway Environmental and Operational Data" on page 27

▼ Display Internal Temperatures

If you are concerned with the internal temperature of the gateway, you can use the `showtemps` command to display nominal and measured temperatures. If there is a great deviation, there might not be enough cooling.

- **On the management controller, type.**

```
# showtemps
Back temperature 29
Front temperature 29
SP temperature 47
Switch temperature 40, maxtemperature 42
Bridge-0 temperature 40, maxtemperature 41
Bridge-1 temperature 45, maxtemperature 46
All temperatures OK
#
```

Related Information

- *Gateway Reference*, `showtemps` command
- "Display Fan Status" on page 26
- "Display Gateway Environmental and Operational Data" on page 27

▼ Display Fan Status

The output of the `getfanspeed` command can help you determine fan speed inconsistencies, which might indicate a future failure.

- **On the management controller, type.**

```
# getfanspeed
Fan 0 not present
Fan 1 running at rpm 11212
Fan 2 running at rpm 11313
```

```
Fan 3 running at rpm 11521
Fan 4 not present
#
```

Related Information

- *Gateway Reference*, `getfanspeed` command
- "Display Gateway Environmental and Operational Data" on page 27

▼ Display Gateway Environmental and Operational Data

The `env_test` command enables you to perform several investigative tasks, helping you determine the internal environment and operational status of the gateway.

- **On the management controller, type.**

```
# env_test
Environment test started:
Starting Voltage test:
Voltage ECB OK
Measured 3.3V Main = 3.28 V
Measured 3.3V Standby = 3.37 V
Measured 12V = 12.06 V
Measured 5V = 5.03 V
Measured VBAT = 3.25 V
Measured 1.0V = 1.01 V
Measured I4 1.2V = 1.22 V
Measured 2.5V = 2.51 V
Measured V1P2 DIG = 1.18 V
Measured V1P2 ANG = 1.18 V
Measured 1.2V BridgeX = 1.22 V
Measured 1.8V = 1.80 V
Measured 1.2V Standby = 1.20 V
Voltage test returned OK
Starting PSU test:
PSU 0 present
PSU 1 present
PSU test returned OK
Starting Temperature test:
Back temperature 30.5
Front temperature 30.6
SP temperature 34.1
Switch temperature 48, maxtemperature 49
Bridge-0 temperature 40, maxtemperature 41
Bridge-1 temperature 45, maxtemperature 45
```

```
Temperature test returned OK
Starting FAN test:
Fan 0 not present
Fan 1 running at rpm 12075
Fan 2 running at rpm 11960
Fan 3 running at rpm 12075
Fan 4 not present
FAN test returned OK
Starting Connector test:
Connector test returned OK
Starting Onboard ibdevice test:
Switch OK
Bridge-0 OK
Bridge-1 OK
All Internal ibdevices OK
Onboard ibdevice test returned OK
Environment test PASSED
#
```

Related Information

- *Gateway Reference*, `env_test` command

▼ Display Chassis FRU ID

The `showfruinfo` command displays gateway chassis FRU ID information.

- **On the management controller, type.**

```
# showfruinfo
 UNIX_Timestamp32              :  Tue Mar 23 17:02:15 2010
 Sun_Fru_Description           :  ASSY,NM2-GW
 Vendor_ID_Code                :  11 E1
 Vendor_ID_Code_Source         :  01
 Vendor_Name_And_Site_Location :  4577 CELESTICA CORP. SAN JOSE CA US
 Sun_Part_Number               :  5111402
 Sun_Serial_Number             :  0110SJC-1010NG0045
 Serial_Number_Format          :  4V3F1-2Y2W2X4S
 Initial_HW_Dash_Level         :  03
 Initial_HW_Rev_Level          :  50
 Sun_Fru_Shortname             :  NM2 gateway
 Sun_Hazard_Class_Code         :  Y
 Sun_SpecPartNo                :  885-1655-01
#
```

Related Information

- *Gateway Reference*, `showfruinfo` command
- "Display Power Supply FRU ID" on page 29

▼ Display Power Supply FRU ID

The `showpsufru` command displays power supply FRU ID information.

- **On the management controller, type.**

```
# showpsufru slot
```

where *slot* is the power supply slot (`0` or `1`). For example:

```
# showpsufru 0
 Sun_SpecPartNo           :  885-1165-02
 UNIX_Timestamp32         :  Sun Jan  3 15:35:39 2010
 Sun_Part_Number          :  3002143
 Sun_Serial_No            :  BF15WA
 Vendor_ID_Code           :  02a2
 Initial_HW_Dash_Level    :  02
 PSU_Voltage 1            :  0x04b0 (1200)
 PSU_Current 1            :  0x186a (6250)
 PSU_Voltage 2            :  0x014a (330)
 PSU_Current 2            :  0x012c (300)
 Sun_Hazard_Class_Code    :  Y
 IPMI_Board_Manufacturer  :  EMERSON
 IPMI_Board_Product_Name  :  A237
 IPMI_Board_Serial_Number:  1357ZHO-0952BF15WA
 IPMI_Board_Part_Number   :  300-2143-02
#
```

Related Information

- *Gateway Reference*, `showpsufru` command
- "Display Chassis FRU ID" on page 28

▼ Display Gateway Firmware Versions

In time, newer or updated gateway firmware might become available. The `version` command displays the versions of the firmware within the gateway chassis.

- **On the management controller, type.**

```
# version
SUN DCS gw version: 2.0.5-1
Build time: Nov 25 2011 12:56:05
FPGA version: 0x33
SP board info:
Manufacturing Date: 2009.06.23
Serial Number: "NCD3R0527"
Hardware Revision: 0x0006
Firmware Revision: 0x0102
BIOS version: NOW1R112
BIOS date: 04/24/2009
#
```

Related Information

- *Gateway Reference*, `version` command
- *Gateway Remote Administration*, `version` command

Controlling the Chassis

These topics describe how you can manage the gateway and its components.

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell target of the Oracle ILOM CLI.

- "Restart the Management Controller" on page 31
- "Change the Administrator Password" on page 31

Related Information

- "Controlling the I4 Switch Chip" on page 40
- "Controlling Gateway Ports and Parameters" on page 151
- "Controlling the InfiniBand Fabric" on page 66
- "Controlling the Subnet Manager" on page 47
- "Monitoring the Chassis" on page 23

▼ Restart the Management Controller

Should the management controller enter an indeterminate state, you can reboot it. The Subnet Manager and other services are also restarted. Additionally, the links to the I4 switch chip are disabled and then re-enabled when the management controller reboots.

Note – The `reboot` command severs any management console link to the management controller. You must reaccess the management controller to regain administrative control.

- **On the management controller, type.**

```
# reboot
Broadcast message from root (pts/0) (Mon Nov 23 12:19:17 2009):
The system is going down for reboot NOW!
# Connection to 123.45.67.89 closed by remote host.
Connection to 123.45.67.89 closed.
#
```

The management controller power cycles and reinitializes.

Related Information

- "Change the Administrator Password" on page 31

▼ Change the Administrator Password

The default password for the `root` user is `changeme`. You can use the `passwd` command to make the password unique and more secure.

- **On the management controller, type.**

```
# passwd user
```

where *user* is name of the user who's password is to be changed. For example:

```
# passwd root
Changing password for user root.
New UNIX password: new_password
Retype new UNIX password: new_password
passwd: all authentication tokens updated successfully.
#
```

Related Information

- *Gateway Installation*, accessing the management controller
- "Restart the Management Controller" on page 31

Administering the I4 Switch Chip

These topics describe the administration of the I4 switch chip.

Related Information

- "Troubleshooting the Gateway" on page 1
- "Understanding Administrative Commands" on page 17
- "Administering the Chassis" on page 23
- "Administering Gateway Resources" on page 99
- "Administering the InfiniBand Fabric" on page 55
- "Administering the Subnet Manager" on page 43

Monitoring the I4 Switch Chip

These topics enable you to display and check the operation and status of the I4 switch chip.

Note – To use all the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell target of the Oracle ILOM CLI.

Related Information

- "Monitoring the Chassis" on page 23
- "Monitoring Gateway Resources" on page 128
- "Monitoring the InfiniBand Fabric" on page 55
- "Monitoring the Subnet Manager" on page 43
- "Controlling the I4 Switch Chip" on page 40

▼ Display the Switch Chip Port to QSFP Connector Mapping

You can use the `dcsport` command to display the mapping of a switch chip port to its respective QSFP connector and to display the mapping of a QSFP connector to its switch chip port.

- **On the management controller, type.**

```
# dcsport -port port
```

where *port* is the number of the port (`1–36`).

Alternatively, type.

```
# dcsport -connector connector
```

where *connector* is the number of the QSFP connector (`0A–15B`).

For example:

```
# dcsport -port 7
DCS-GW Switch port 7 maps to connector 14A
# dcsport -connector 14A
Connector 14A maps to Switch port 7
#
```

Note – For a complete mapping of all ports to all connectors, use the `dcsport -printconnnectors` command.

Related Information

- *Gateway Reference*, `dcsport` command
- "Understanding Routing Through the Gateway" on page 12

▼ Locate a Switch Chip or Connector From the GUID

The output of some InfiniBand commands identify a node by its GUID. You can determine the switch chip or InfiniBand cable connection associated with that GUID using this procedure.

1. **In the output of the InfiniBand command, identify the GUID and port in question.**

   For example, given this output from the `ibdiagnet` command:

```
-W- lid=0x000f guid=0x002128312345c0a0 dev=48438 Port=21
```

   In this example, the GUID is `0x002128312345c0a0` and the port is 21.

2. **Truncate all digits of the GUID except the four digits on the right.**

   For example, GUID `0x002128312345c0a0` is truncated to `c0a0`.

   The `c0` means that the node is inside of a Sun Network QDR InfiniBand Gateway Switch. The `a` means that the node is located in I4 switch chip A, the only switch chip. The port is 21.

3. **Use the `ibswitches` command to verify the information.**

```
# ibswitches
Switch  : 0x03ba01234567a0a0 ports 36 "SUN DCS 36P QDR mnm-36p-2 10.172.144.66"
enhanced port 0 lid 1 lmc 0
Switch  : 0x002128312345c0a0 ports 36 "SUN IB QDR GW switch mnm-gw-1
10.172.144.69" enhanced port 0 lid 15 lmc 0
#
```

4. **Use the tables in "Understanding Routing Through the Gateway" on page 12 to cross-reference the routing between ports and to connections.**

   For example, using "Switch Chip Port to QSFP Connectors and Link LED Routes" on page 12, I4 switch chip, port 21 routes to connector 1B.

5. **Alternatively, you can use the `dcsport` command to display routing between ports and connections.**

   See "Display the Switch Chip Port to QSFP Connector Mapping" on page 34.

Related Information

- "Switch GUIDs Overview" on page 14
- *Gateway Reference*, `dcsport` command

▼ Display Switch Chip Boot Status

If you believe there is a problem with the switch chip, the `checkboot` command provides a quick pass-fail check of the switch chip. Should the switch chip fail the test, you can use the `ibdevreset` command in an attempt to reboot the switch chip.

- **On the management controller, type.**

```
# checkboot
Switch OK
Bridge-0 OK
Bridge-1 OK
#
```

Related Information

- *Gateway Reference*, `checkboot` command
- "Display Gateway Environmental and Operational Data" on page 27
- *Gateway Reference*, `ibdevreset` command

▼ Display Link Status

In some situations, you might need to know the status of each route through the gateway. Additionally, the `listlinkup` command displays where InfiniBand cables are connected to the gateway.

- **On the management controller, type.**

```
# listlinkup connected
Connector  2A Present <-> Switch Port 24 up (Enabled)
Connector  3A Present <-> Switch Port 26 up (Enabled)
Connector  6A Present <-> Switch Port 35 up (Enabled)
Connector  8A Present <-> Switch Port 31 up (Enabled)
Connector 11A Present <-> Switch Port 12 down (Enabled)
Connector 13A Present <-> Switch Port 9 down (Enabled)
Connector 14A Present <-> Switch Port 7 down (Enabled)
Connector 0A-ETH Present
 Bridge-0 Port 0A-ETH-1 (Bridge-0-2) up (Enabled)
 Bridge-0 Port 0A-ETH-2 (Bridge-0-2) up (Enabled)
 Bridge-0 Port 0A-ETH-3 (Bridge-0-1) up (Enabled)
 Bridge-0 Port 0A-ETH-4 (Bridge-0-1) up (Enabled)
Connector 1A-ETH Present
 Bridge-1 Port 1A-ETH-1 (Bridge-1-2) up (Enabled)
 Bridge-1 Port 1A-ETH-2 (Bridge-1-2) up (Enabled)
 Bridge-1 Port 1A-ETH-3 (Bridge-1-1) up (Enabled)
```

```
 Bridge-1 Port 1A-ETH-4 (Bridge-1-1) up (Enabled)
Connector  9B Present <-> Switch Port 13 up (Enabled)
#
```

Related Information

- *Gateway Reference*, `listlinkup` command
- "Display the Link Status of a Node" on page 60

▼ Display Switch Chip Port Status

If by using other procedures, it seems that a particular port of the switch chip is problematic, you can use the `getportstatus` command to provide the port state, width, and speed.

- **On the management controller, type.**

```
# getportstatus connector | ibdev port
```

where:

- *connector* is the name of the connector (`0A – 15A`, `0A-ETH`, `1A-ETH`, `0B – 15B`).

Note – The subordinate names of the 10GbE connectors, `0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`, are also valid *connector* names.

- *ibdev* is the InfiniBand device name (`Switch`, `Bridge-0-0`, `Bridge-0-1`, `Bridge-1-0`, `Bridge-1-1`)
- *port* is the number of the port (`1–36`).

For example:

```
# getportstatus Switch 7
Port status for connector 14A Switch Port 7
Adminstate:.....................Enabled
LinkWidthEnabled:...............1X or 4X
LinkWidthSupported:.............1X or 4X
LinkWidthActive:................4X
LinkSpeedSupported:.............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkState:......................Active
PhysLinkState:..................LinkUp
LinkSpeedActive:................10.0 Gbps
LinkSpeedEnabled:...............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
#
```

Note – The parameters with the string `Active` indicate the current conditions.

Related Information

- *Gateway Reference*, `getportstatus` command
- "Display the Link Status of a Node" on page 60

▼ Display Switch Chip Port Counters

You can check the errors and throughput of a switch chip port using the `getportcounters` command.

- **On the management controller, type.**

```
# getportcounters port|connector
```

where:

- *port* is the number of the port (`1–36`).
- *connector* is the name of the connector (`0A – 15A`, `0A-ETH`, `1A-ETH`, `0B – 15B`).

Note – The subordinate names of the 10GbE connectors, `0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4` are also valid *connector* names.

For example, for port 1:

```
# getportcounters 1
# Port counters: Lid 6 port 1
PortSelect:......................1
CounterSelect:...................0x1b01
SymbolErrors:....................65535
LinkRecovers:....................0
LinkDowned:......................1
RcvErrors:.......................0
RcvRemotePhysErrors:.............0
RcvSwRelayErrors:................0
XmtDiscards:.....................1
XmtConstraintErrors:.............0
RcvConstraintErrors:.............0
LinkIntegrityErrors:.............0
ExcBufOverrunErrors:.............0
VL15Dropped:.....................0
XmtData:.........................277670000
RcvData:.........................321584043
XmtPkts:.........................4395402
RcvPkts:.........................5137415
#
```

For example, for connector 1A-ETH-1:

```
# getportcounters 1A-ETH-1
ETH Port 1A-ETH-1
---------------------------------
RX bytes:........................0x6a5559
RX packets:......................0xd650
RX Jumbo packets:................0x0
RX unicast packets:..............0x0
RX multicast packets:............0xbcfc
RX broadcast packets:............0x1954
RX no buffer:....................0x0
RX CRC:..........................0x0
RX runt:.........................0x0
RX errors:.......................0x0
TX bytes:........................0x0
TX packets:......................0x0
TX Jumbo packets:................0x0
TX unicast packets:..............0x0
```

```
TX multicast packets:.....0x0
TX broadcast packets:.....0x0
TX errors:.....0x0
#
```

Related Information

- *Gateway Reference*, `getportcounters` command
- "Display Counters for a Node" on page 61

Controlling the I4 Switch Chip

These topics describe how you can manage the I4 switch chip.

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell target of the Oracle ILOM CLI.

- "Disable a Switch Chip Port" on page 40
- "Enable a Switch Chip Port" on page 41

Related Information

- "Controlling the Chassis" on page 30
- "Controlling Gateway Ports and Parameters" on page 151
- "Controlling the InfiniBand Fabric" on page 66
- "Controlling the Subnet Manager" on page 47
- "Monitoring the I4 Switch Chip" on page 33

▼ Disable a Switch Chip Port

On occasion, you might need to turn off a port. For example, a cable might become damaged and cause symbol errors that affect the switch chip's performance. Use the `disableswitchport` command to disable a switch chip port and its respective cable connection.

- **On the management controller, type.**

```
# disableswitchport [--reason=reason] Switch port
```

where:

- *reason* is the reason for disabling the port, `Blacklist` or `Partition`.
- *port* is the number of the port (`1–36`).

Alternatively, type.

```
# disableswitchport [--reason=reason] connector
```

where *connector* is the number of the QSFP connector (`0A–15B`).

For example:

```
# disableswitchport Switch 14
Disable Switch port 14
Initial PortInfo:
# Port info: DR path slid 65535; dlid 65535; 0 port 14
LinkState:.....................Active
PhysLinkState:.................LinkUp
LinkWidthSupported:............1X or 4X
LinkWidthEnabled:..............1X or 4X
LinkWidthActive:...............4X
LinkSpeedSupported:............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:..............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:...............10.0 Gbps
After PortInfo set:
# Port info: DR path slid 65535; dlid 65535; 0 port 14
LinkState:.....................Down
PhysLinkState:.................Disabled
#
```

Related Information

- *Gateway Reference*, `disableswitchport` command
- "Disable a Port" on page 72
- "Enable a Switch Chip Port" on page 41

▼ Enable a Switch Chip Port

You can enable a disabled switch chip port with the `enableswitchport` command.

- **On the management controller, type.**

```
# enableswitchport [--reason=reason] Switch port
```

where:

- *reason* is the reason for disabling the port, `Blacklist` or `Partition`.
- *port* is the number of the port (`1–36`).

Note – If the port was disabled with the `--reason` option, (as seen with the `listlinkup` command) it can only be enabled with the same reason option.

Alternatively, type.

```
# enableswitchport [--reason=reason] connector
```

where *connector* is the number of the QSFP connector (`0A–15B`).

For example:

```
# enableswitchport Switch 14
Enable Switch port 14
Initial PortInfo:
# Port info: DR path slid 65535; dlid 65535; 0 port 14
LinkState:.....Down
PhysLinkState:.....Disabled
LinkWidthSupported:.....1X or 4X
LinkWidthEnabled:.....1X or 4X
LinkWidthActive:.....4X
LinkSpeedSupported:.....2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:.....2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:.....2.5 Gbps
After PortInfo set:
# Port info: DR path slid 65535; dlid 65535; 0 port 14
LinkState:.....Down
PhysLinkState:.....PortConfigurationTraining
#
```

Related Information

- *Gateway Reference*, `enableswitchport` command
- "Enable a Port" on page 73
- "Disable a Switch Chip Port" on page 40

Administering the Subnet Manager

These topics describe the administration of the Subnet Manager.

Related Information

- "Troubleshooting the Gateway" on page 1
- "Understanding Administrative Commands" on page 17
- "Administering the Chassis" on page 23
- "Administering the I4 Switch Chip" on page 33
- "Administering Gateway Resources" on page 99
- "Administering the InfiniBand Fabric" on page 55

Monitoring the Subnet Manager

These topics describe how to monitor the Subnet Manager.

Note – To use all the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell target of the Oracle ILOM CLI.

Related Information

- "Monitoring the Chassis" on page 23
- "Monitoring the I4 Switch Chip" on page 33
- "Monitoring Gateway Resources" on page 128
- "Monitoring the InfiniBand Fabric" on page 55
- "Controlling the Subnet Manager" on page 47

▼ Display Subnet Manager Status

If you want to quickly determine your Subnet Manager's priority and state, the `getmaster` command can also provide the LID and GUID of the hosting HCA.

- **On the management controller, type.**

```
# getmaster
Local SM enabled and running
20111020 08:55:02 Master SubnetManager on sm lid 1 sm guid 0x3ba01234567a0a0 :
SUN DCS 36P QDR o4nm2-36p-2 10.172.144.66
#
```

Related Information

- *Gateway Reference*, `getmaster` command
- "Display Recent Subnet Manager Activity" on page 44
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Verify the Subnet Manager Configuration Integrity" on page 45
- "Display the Subnet Manager Log" on page 46

▼ Display Recent Subnet Manager Activity

- **On the management controller, type.**

```
# getmaster -l
Local SM enabled and running
Last ring buffer history listed:
20100927 19:16:12 whereismaster started
20100927 19:16:12 Master SubnetManager on sm lid 0 sm guid 0x212856cfe2c0a0 :
20100927 19:16:29 Master SubnetManager on sm lid 6 sm guid 0x212856cfe2c0a0 :
SUN IB QDR GW switch mnm34-98
#
```

Related Information

- *Gateway Reference*, `getmaster` command
- "Display Subnet Manager Status" on page 44
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Verify the Subnet Manager Configuration Integrity" on page 45
- "Display the Subnet Manager Log" on page 46

▼ Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key

If you want to know your Subnet Manager's priority, controlled handover state, prefix, and management key, use the list option of the `setsmpriority` command.

- **On the management controller, type.**

```
# setsmpriority list
Current SM settings:
smpriority 8
controlled_handover TRUE
subnet_prefix 0xfe80000000000000
M_Key None
#
```

Related Information

- *Gateway Reference*, `setsmpriority` command
- "Controlling the Subnet Manager" on page 47
- "Display Subnet Manager Status" on page 44
- "Display Recent Subnet Manager Activity" on page 44
- "Verify the Subnet Manager Configuration Integrity" on page 45
- "Display the Subnet Manager Log" on page 46

▼ Verify the Subnet Manager Configuration Integrity

The `smconfigtest` command performs a check of the Subnet Manager's configuration and verifies the integrity.

Note – For the `spine` option of the `smconfigtest` command, the Subnet Manager is expected to have a priority of 5 or 8. For the `leaf` option, the expected priority is 5.

- **Verify the Subnet Manager configuration.**

```
# smconfigtest [spine|leaf]
```

where:

- `spine` – for spine switches.
- `leaf` – for leaf switches.

For example:

```
# smconfigtest leaf
#
```

Note – Output is only displayed if errors are found.

Related Information

- *Gateway Reference*, `smconfigtest` command
- "Display Subnet Manager Status" on page 44
- "Display Recent Subnet Manager Activity" on page 44
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Display the Subnet Manager Log" on page 46

▼ Display the Subnet Manager Log

1. **On the management controller, type.**

```
# showsmlog
Aug 13 15:58:22 679655 [B75F1B90] 0x01 -> __osm_mcmr_rcv_join_mgrp: ERR 1B11:
method = Subn
AdmSet, scope_state = 0x1, component mask = 0x0000000000010003, expected comp
mask = 0x0000
0000000130c7, MGID: ff12:e01b:2::22:2000 from port 0x0003ba000100e371 (mnm34-60
HCA-1)
```

```
Aug 13 15:58:22 692653 [B75F1B90] 0x01 -> __osm_mcmr_rcv_join_mgrp: ERR 1B11:
method = Subn
AdmSet, scope_state = 0x1, component mask = 0x0000000000010003, expected comp
mask = 0x0000
0000000130c7, MGID: ff12:e01b:3::22:b000 from port 0x00212800013ece9f (mnm34-55
HCA-1)
.
.
.
```

2. **Tap the space bar to display the next screen of the log.**
3. **Press the Q key to quit.**

Related Information

- *Gateway Reference*, `showsmlog` command
- "Display Subnet Manager Status" on page 44
- "Display Recent Subnet Manager Activity" on page 44
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Verify the Subnet Manager Configuration Integrity" on page 45

Controlling the Subnet Manager

You can enable the Subnet Manager with the `enablesm` command. When the Subnet Manager starts, it reads the configuration file for configuration information. You can disable the Subnet Manager with the `disablesm` command.

These topics describe how to control the Subnet Manager:

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell targets of the Oracle ILOM CLI.

- "Set the Subnet Manager Priority" on page 48
- "Set the Subnet Manager Prefix" on page 49
- "Enable Subnet Manager Controlled Handover" on page 50
- "Set the Subnet Manager Management Key" on page 50
- "Clear the Subnet Manager Management Key" on page 52

- "Enable the Subnet Manager" on page 52
- "Disable the Subnet Manager" on page 53

Related Information

- "Controlling the Chassis" on page 30
- "Controlling the I4 Switch Chip" on page 40
- "Controlling Gateway Ports and Parameters" on page 151
- "Controlling the InfiniBand Fabric" on page 66
- "Monitoring the Subnet Manager" on page 43

▼ Set the Subnet Manager Priority

By default, the Subnet Manager within the management controller is set to 0 priority. If there is more than one Subnet Manager in your InfiniBand fabric, you must set the priority of each Subnet Manager appropriately. The Subnet Manager with the highest priority is the primary (or Master) Subnet Manager.

1. **On the managment controller, disable the Subnet Manager.**

   See "Disable the Subnet Manager" on page 53.

2. **Set the Subnet Manager priority.**

```
# setsmpriority priority
```

   where *priority* is `0` (lowest) to `13` (highest). For example:

```
# setsmpriority 5
Current SM settings:
smpriority 5
controlled_handover FALSE
subnet_prefix 0xfe80000000000000
M_Key None
#
```

3. **Enable the Subnet Manager.**

   See "Enable the Subnet Manager" on page 52.

Related Information

- *Gateway Reference*, `setsmpriority` command

- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Set the Subnet Manager Prefix" on page 49
- "Enable Subnet Manager Controlled Handover" on page 50
- "Set the Subnet Manager Management Key" on page 50
- "Clear the Subnet Manager Management Key" on page 52
- "Enable the Subnet Manager" on page 52
- "Disable the Subnet Manager" on page 53

▼ Set the Subnet Manager Prefix

The `setsubnetprefix` command writes a prefix value to the `subnet_prefix` parameter of the configuration file.

1. **On the managment controller, disable the Subnet Manager.**

   See "Disable the Subnet Manager" on page 53.

2. **Set the Subnet Manager prefix..**

```
# setsubnetprefix 0xabbababe
Current SM settings:
smpriority 5
controlled_handover FALSE
subnet_prefix 0xabbababe
M_Key None
#
```

3. **Enable the Subnet Manager.**

   See "Enable the Subnet Manager" on page 52.

Related Information

- *Gateway Reference*, `setsubnetprefix` command
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Set the Subnet Manager Priority" on page 48
- "Enable Subnet Manager Controlled Handover" on page 50
- "Set the Subnet Manager Management Key" on page 50
- "Clear the Subnet Manager Management Key" on page 52
- "Enable the Subnet Manager" on page 52
- "Disable the Subnet Manager" on page 53

▼ Enable Subnet Manager Controlled Handover

If your InfiniBand fabric has two or more Subnet Managers, you can force a constrained fallback protocol should the master Subnet Manager fail. See *Gateway Reference*, `setcontrolledhandover` command for more information.

1. **On the managment controller, disable the Subnet Manager.**

   See "Disable the Subnet Manager" on page 53.

2. **Enable controlled handover.**

```
# setcontrolledhandover TRUE
Current SM settings:
smpriority 5
controlled_handover TRUE
subnet_prefix 0xabbababe
M_Key None
#
```

3. **Enable the Subnet Manager.**

   See "Enable the Subnet Manager" on page 52.

Related Information

- *Gateway Reference*, `setcontrolledhandover` command
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Set the Subnet Manager Priority" on page 48
- "Set the Subnet Manager Prefix" on page 49
- "Set the Subnet Manager Management Key" on page 50
- "Clear the Subnet Manager Management Key" on page 52
- "Enable the Subnet Manager" on page 52
- "Disable the Subnet Manager" on page 53

▼ Set the Subnet Manager Management Key

For added security, a management key, or M_Key is used by the Subnet Manager Agent to initialize and configure network nodes. See *Gateway Reference*, `setsmmkey` command for more information.

Note – All Subnet Managers in the InfiniBand fabric must be configured with the same M_Key value.

1. **On the managment controller, disable the Subnet Manager.**

   See "Disable the Subnet Manager" on page 53.

2. **Set the management key.**

```
# setsmmkey m_key
```

   where *m_key* is a 12-digit hexadecimal number

   For example:

```
# setsmmkey 0xabbababe0001
Current SM settings:
smpriority 5
controlled_handover TRUE
subnet_prefix 0xabbababe
M_Key 0xabbababe0001
#
```

3. **Enable the Subnet Manager.**

   See "Enable the Subnet Manager" on page 52.

4. **Repeat Step 1 to Step 3 for all Subnet Managers of the InfiniBand fabric.**

Related Information

- *Gateway Reference*, `setsmmkey` command
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Set the Subnet Manager Priority" on page 48
- "Set the Subnet Manager Prefix" on page 49
- "Enable Subnet Manager Controlled Handover" on page 50
- "Clear the Subnet Manager Management Key" on page 52
- "Enable the Subnet Manager" on page 52
- "Disable the Subnet Manager" on page 53

▼ Clear the Subnet Manager Management Key

It might be necessary to remove the restrictions of the Subnet Manager management key. This procedure describes how to do so.

1. **On the managment controller, disable the Subnet Manager.**

   See "Disable the Subnet Manager" on page 53.

2. **Clear the management key.**

```
# setsmmkey none
Current SM settings:
smpriority 5
controlled_handover TRUE
subnet_prefix 0xabbababe
M_Key None
#
```

3. **Enable the Subnet Manager.**

   See "Enable the Subnet Manager" on page 52.

4. **Repeat Step 1 to Step 3 for all Subnet Managers requiring the change.**

Related Information

- *Gateway Reference*, `setsmmkey` command
- "Set the Subnet Manager Management Key" on page 50
- "Display Subnet Manager Priority, Controlled Handover State, Prefix, and Management Key" on page 45
- "Set the Subnet Manager Priority" on page 48
- "Set the Subnet Manager Prefix" on page 49
- "Enable Subnet Manager Controlled Handover" on page 50
- "Enable the Subnet Manager" on page 52
- "Disable the Subnet Manager" on page 53

▼ Enable the Subnet Manager

The Subnet Manager within the management controller is not enabled by default. Use the `enablesm` command to enable and start the Subnet Manager.

- **On the management controller, type.**

```
# enablesm
Starting IB Subnet Manager.                                [  OK  ]
Starting partitiond daemon.                                [  OK  ]
#
```

Related Information

- *Gateway Reference*, `enablesm` command
- "Disable the Subnet Manager" on page 53
- "Set the Subnet Manager Priority" on page 48
- "Set the Subnet Manager Prefix" on page 49
- "Enable Subnet Manager Controlled Handover" on page 50
- "Set the Subnet Manager Management Key" on page 50
- "Clear the Subnet Manager Management Key" on page 52

▼ Disable the Subnet Manager

If your InfiniBand fabric has too many Subnet Managers, you can disable and stop the Subnet Manager within the management controller by using the `disablesm` command.

- **On the management controller, type.**

```
# disablesm
Stopping partitiond daemon.                                [  OK  ]
Stopping IB Subnet Manager..                               [  OK  ]
#
```

Related Information

- *Gateway Reference*, `disablesm` command
- "Enable the Subnet Manager" on page 52
- "Set the Subnet Manager Priority" on page 48
- "Set the Subnet Manager Prefix" on page 49
- "Enable Subnet Manager Controlled Handover" on page 50
- "Set the Subnet Manager Management Key" on page 50
- "Clear the Subnet Manager Management Key" on page 52

Administering the InfiniBand Fabric

These topics describe the administration of the InfiniBand fabric.

Related Information

- "Troubleshooting the Gateway" on page 1
- "Understanding Administrative Commands" on page 17
- "Administering the Chassis" on page 23
- "Administering the I4 Switch Chip" on page 33
- "Administering Gateway Resources" on page 99
- "Administering the Subnet Manager" on page 43

Monitoring the InfiniBand Fabric

These topics enable you to display and check the operation and status of the InfiniBand fabric and components.

Note – To use all the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell target of the Oracle ILOM CLI.

- "Display a Route Through the Fabric" on page 59
- "Display the Link Status of a Node" on page 60
- "Display Counters for a Node" on page 61
- "Display Low-Level Detailed Information About a Node" on page 62
- "Display Low-Level Detailed Information About a Port" on page 63
- "Display the InfiniBand Fabric Partition Configuration" on page 65
- "Display the InfiniBand Fabric Configuration" on page 65

Related Information

- "Monitoring the Chassis" on page 23
- "Monitoring the I4 Switch Chip" on page 33
- "Monitoring Gateway Resources" on page 128
- "Monitoring the Subnet Manager" on page 43
- "Controlling the InfiniBand Fabric" on page 66
- "Configuring the Fabric Director Node List" on page 74
- "Partitioning the InfiniBand Fabric" on page 85

▼ Identify All Switches in the Fabric

If you need to know identity information about the switches in the InfiniBand fabric, you can use the `ibswitches` command. This command displays the GUID, name, LID, and LMC for each switch. The output of the command is a mapping of GUID to LID for switches in the fabric.

Note – The output for your InfiniBand fabric will differ greatly from the output in the example.

- **On the management controller, type.**

```
# ibswitches
Switch  : 0x00212856cfe2c0a0 ports 36 "SUN IB QDR GW switch mnm34-98" enhanced
port 0 lid 6 lmc 0
#
```

Related Information

- *Gateway Reference*, `ibswitches` command
- *Gateway Reference*, `ibhosts` command

■ "Identify All HCAs in the Fabric" on page 57

▼ Identify All HCAs in the Fabric

Similar to the `ibswitches` command, the `ibhosts` command displays identity information about the HCAs in the InfiniBand fabric. The output contains the GUID and name for each HCA.

Note – Unlike the `ibswitches` command, the `ibhosts` command does not display the LIDs.

---

1. **On the management controller, type.**

Note – The output in the example is just a portion of the full output and varies for each InfiniBand topology.

```
# ibhosts
Ca      : 0x00212856cd22c000 ports 2 "SUN IB QDR GW switch mnm34-97 Bridge 0"
Ca      : 0x00212856cd22c040 ports 2 "SUN IB QDR GW switch mnm34-97 Bridge 1"
Ca      : 0x0002c903000891aa ports 2 "mnm34-54 HCA-1"
Ca      : 0x00212800013ece9e ports 2 "mnm34-55 HCA-1"
Ca      : 0x0003ba000100e370 ports 2 "mnm34-60 HCA-1"
.
.
.
#
```

2. **(Optional) Save the output of `ibhosts` command as a text file named `hostGUIDs.txt`.**

Related Information

- *Gateway Reference*, `ibhosts` command
- *Gateway Reference*, `ibswitches` command
- "Identify All Switches in the Fabric" on page 56

▼ Display the InfiniBand Fabric Topology (Simple)

To understand the routing that happens within your InfiniBand fabric, the `showtopology` command displays the node-to-node connectivity. The output of the command depends on the size of your fabric.

- **On the management controller, type.**

```
# showtopology
SUNIBQDRGW mnm34-98 Bridge 0
  BX1/P1 -4x-10G-> I4_GENERIC mnm34-98 P4
      -4x-10G-> I4_GENERIC mnm34-98 P3
  BX3/P1 -4x-10G-> I4_GENERIC mnm34-98 P2
      -4x-10G-> I4_GENERIC mnm34-98 P1

DEV26428_02P mnm34-55_HCA-1
  P1 -4x-10G-> SUNIBQDRGW mnm34-97 C-7A
  P2 -4x-10G-> I4_GENERIC mnm34-98 P33

SUNIBQDRGW mnm34-97
  A-SW/P2 -4x-10G-> SUNIBQDRGW mnm34-97 BX3/P1
  A-SW/P4 -4x-10G-> SUNIBQDRGW mnm34-97 BX1/P1
  C-0B -4x-10G-> I4_GENERIC mnm34-98 P21
.
.
.
#
```

Related Information

- *Gateway Reference*, `showtopology` command
- "Determine Changes to the InfiniBand Fabric Topology" on page 67
- "Perform Comprehensive Diagnostics for the Entire Fabric" on page 67

▼ Display the InfiniBand Fabric Topology (Detailed)

To understand the routing that happens within your InfiniBand fabric, the `ibnetdiscover` command displays the node-to-node connectivity. The output of the command depends on the size of your fabric.

Note – You must be the `root` user to run the `ibnetdiscover` command.

Note – You can use the `ibnetdiscover` command to determine the LIDs of the HCAs.

- **On the management controller, type.**

```
# ibnetdiscover
#
# Topology file: generated on Sat Apr 13 22:28:55 2002
#
# Max of 1 hops discovered
# Initiated from node 0021283a8389a0a0 port 0021283a8389a0a0
vendid=0x2c9
devid=0xbd36
sysimgguid=0x21283a8389a0a3
switchguid=0x21283a8389a0a0(21283a8389a0a0)
Switch  36 "S-0021283a8389a0a0" # "Sun DCS 36 QDR switch localhost" enhanced port
0 lid 15 lmc 0
[23]    "H-0003ba000100e388"[2](3ba000100e38a) # "mnm33-43 HCA-1" lid 14 4xQDR
.
.
.
#
```

Related Information

- *Gateway Reference*, `ibnetdiscover` command
- "Perform Comprehensive Diagnostics for the Entire Fabric" on page 67

▼ Display a Route Through the Fabric

Often you need to know the route between two nodes. The `ibtracert` command can provide that information by displaying the GUIDs, ports, and LIDs of the nodes along the route.

- **On the management controller, type.**

```
# ibtracert slid dlid
```

where:

- *slid* is the LID of the source node.

- *dlid* is the LID of the destination node.

For example:

```
# ibtracert 15 14
From switch {0x00212856cd22c0a0} portnum 0 lid 15-15 "SUN IB QDR GW switch
mnm34-97"
[1] -> ca port {0x00212856cd22c042}[2] lid 14-14 "SUN IB QDR GW switch mnm34-97
Bridge 1"
To ca {0x00212856cd22c040} portnum 2 lid 14-14 "SUN IB QDR GW switch mnm34-97
Bridge 1"
#
```

Note – The output for your InfiniBand fabric will differ from that in the example.

For this example:

- The route starts at the switch with GUID `0x00212856cd22c0a0` and is using port `0`. The switch is LID `15` and in the description, the switch host's name is mnm34-97.
- The route enters at port 1 of the CA with GUID `0x00212856cd22c042` and exits at port `2`. The CA is LID `14` and is within bridge chip 1.

Related Information

- *Gateway Reference*, `ibtracert` command

▼ Display the Link Status of a Node

If you want to know the link status of a node in the InfiniBand fabric, the `ibportstate` command can tell you the state, width, and speed of that node.

- **On the management controller, type.**

```
# ibportstate lid port
```

where:

- *lid* is the LID of the node.

- *port* is the port of the node.

For example:

```
# ibportstate 15 23
PortInfo:
# Port info: Lid 15 port 23
LinkState:.....................Active
PhysLinkState:.................LinkUp
LinkWidthSupported:............1X or 4X
LinkWidthEnabled:..............1X or 4X
LinkWidthActive:...............4X
LinkSpeedSupported:............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:..............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:...............10.0 Gbps
Peer PortInfo:
# Port info: Lid 15 DR path slid 15; dlid 65535; 0,23
LinkState:.....................Active
PhysLinkState:.................LinkUp
LinkWidthSupported:............1X or 4X
LinkWidthEnabled:..............1X or 4X
LinkWidthActive:...............4X
LinkSpeedSupported:............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:..............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:...............10.0 Gbps
#
```

In the output, the `Active` parameters are the current state of the port.

Related Information

- *Gateway Reference*, `ibportstate` command
- "Display Switch Chip Port Status" on page 37
- "Display Link Status" on page 36
- "Find 1x, SDR, or DDR Links in the Fabric" on page 68
- "Set Port Speed" on page 71

▼ Display Counters for a Node

To help ascertain the health of a node, the `perfquery` command displays the performance, error, and data counters for that node.

- **On the management controller, type.**

```
# perfquery lid port
```

where:

- *lid* is the LID of the node.
- *port* is the port of the node.

Note – If a *port* value of 255 is specified for a switch node, the counters are the total for all switch ports.

For example:

Note – The output in the example is just a portion of the full output.

```
# perfquery 15 23
# Port counters: Lid 15 port 23
PortSelect:......................23
CounterSelect:...................0x1b01
SymbolErrors:....................0
.
.
.
VL15Dropped:.....................0
XmtData:.........................20232
RcvData:.........................20232
XmtPkts:.........................281
RcvPkts:.........................281
#
```

Related Information

- *Gateway Reference*, `perfquery` command
- "Display Switch Chip Port Counters" on page 38
- "Clear Data and Error Counters" on page 70

▼ Display Low-Level Detailed Information About a Node

If intensive troubleshooting is necessary to resolve a problem, the `smpquery` command can provide very detailed information about a node.

- **On the management controller, type.**

```
# smpquery switchinfo lid
```

where *lid* is the LID of the node.

For example, to see detailed information about a switch with LID 15, type.

```
# smpquery switchinfo 15
# Switch info: Lid 15
LinearFdbCap:.....................49152
RandomFdbCap:.....................0
McastFdbCap:......................4096
LinearFdbTop:.....................16
DefPort:..........................0
DefMcastPrimPort:.................255
DefMcastNotPrimPort:..............255
LifeTime:.........................18
StateChange:......................0
LidsPerPort:......................0
PartEnforceCap:...................32
InboundPartEnf:...................1
OutboundPartEnf:..................1
FilterRawInbound:.................1
FilterRawOutbound:................1
EnhancedPort0:....................1
#
```

Related Information

- *Gateway Reference*, `smpquery` command
- "Display Low-Level Detailed Information About a Port" on page 63

▼ Display Low-Level Detailed Information About a Port

The `smpquery` command can provide very detailed information about a port.

- **On the management controller, type.**

```
# smpquery portinfo lid port
```

where:

- *lid* is the LID of the node.

- *port* is the port of the node.

For example, to see detailed information about port 23 on the switch with LID 15, type.

Note – The output in the example is just a portion of the full output.

```
# smpquery portinfo 15 23
# Port info: Lid 15 port 23
Mkey:.............................0x0000000000000000
GidPrefix:........................0x0000000000000000
Lid:..............................0x0000
SMLid:............................0x0000
CapMask:..........................0x0
DiagCode:.........................0x0000
MkeyLeasePeriod:..................0
LocalPort:........................0
LinkWidthEnabled:.................1X or 4X
LinkWidthSupported:...............1X or 4X
LinkWidthActive:..................4X
LinkSpeedSupported:...............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkState:........................Active
PhysLinkState:....................LinkUp
LinkDownDefState:.................Polling
ProtectBits:......................0
LMC:..............................0
.
.
.
SubnetTimeout:....................0
RespTimeVal:......................0
LocalPhysErr:.....................8
OverrunErr:.......................8
MaxCreditHint:....................85
RoundTrip:........................16777215
#
```

Related Information

- *Gateway Reference*, `smpquery` command
- "Display Low-Level Detailed Information About a Node" on page 62

▼ Display the InfiniBand Fabric Partition Configuration

If you have partitioned the InfiniBand fabric, you can display the active (that which is in use) partition configuration or the modified (that which is in standby) partition configuration with the `smpartition` command.

- **On the management controller, type.**

```
# smpartition list partition
```

where *partition* is the partition configuration name (`active` or `modified`).

For example:

```
# smpartition list active
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 12
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
#
```

In this example, this is the 12th configuration file to be created. The partition key is set to a default of `7fff`. IP protocol over InfiniBand is configured. All hosts (`CAS`), all switches, and the management controller itself are granted `full` membership priviledges.

Related Information

- *Gateway Reference*, `smpartition` command
- "Partitioning the InfiniBand Fabric" on page 85
- "Display the InfiniBand Fabric Topology (Detailed)" on page 58

▼ Display the InfiniBand Fabric Configuration

If you have configured the InfiniBand fabric for SNMP operations, you can display the elements configured with the `fdconfig` command.

- **On the management controller, type.**

```
# fdconfig list-current-fabric-config
Name                              IP Addr              Type    Role(s)
isMaster        Fabricname
-----------------------------------         -----------------        ------
--------------------    --------         ----------
primary                           123.45.67.89         sw36gw  gateway
yes             test
secondary                         123.45.67.90         sw36    leafSwitch
no              test
#
```

In this example, the element named `primary` has IP address 123.45.67.89 and is a Sun Network QDR InfiniBand Gateway Switch. It is hosting the master Fabric Director daemon. The other element named `secondary` has an IP address of 123.45.67.90 and is a Sun Datacenter InfiniBand Switch 36. It is being used as a leaf switch.

Related Information

- *Gateway Reference*, `fdconfig` command
- "Configuring the Fabric Director Node List" on page 74

Controlling the InfiniBand Fabric

You can perform these tasks to manage the InfiniBand fabric and its components.

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Fabric_Mgmt` Linux shell target of the Oracle ILOM CLI.

- "Enable a Port" on page 73

Related Information

- "Controlling the Chassis" on page 30
- "Controlling the I4 Switch Chip" on page 40
- "Controlling Gateway Ports and Parameters" on page 151
- "Controlling the Subnet Manager" on page 47
- "Monitoring the InfiniBand Fabric" on page 55
- "Configuring the Fabric Director Node List" on page 74
- "Partitioning the InfiniBand Fabric" on page 85

▼ Perform Comprehensive Diagnostics for the Entire Fabric

If you require a full testing of your InfiniBand fabric, the `ibdiagnet` command can perform many tests with verbose results. The command is a useful tool to determine the general overall health of the InfiniBand fabric.

- **On the management controller, type.**

```
# ibdiagnet -v -r
```

Related Information

- *Gateway Reference*, `ibdiagnet` command
- *Gateway Reference*, `ibdiagpath` command
- "Display the InfiniBand Fabric Topology (Detailed)" on page 58

▼ Determine Changes to the InfiniBand Fabric Topology

If your fabric has a number of nodes that are suspect, the `generatetopology` command enables you to take a snapshot of your fabric. At a later time, use the `matchtopology` command to compare the topology file to the present conditions.

1. **Take a snapshot of the fabric topology.**

```
# generatetopology date.topo
```

where *date* is the date of the snapshot. For example:

```
# generatetopology Sept8.topo
It will take some time to generate a topology file. Please wait!
topo.conf exist! will move it to topo.conf.old
will create new topo.conf
Topo.conf file is created. Will now start generating the topo file
Wrote Topology file:Sept8.topo
#
```

2. **After an event, compare the present topology to that saved in the topology file.**

```
# matchtopology date.topo
```

where *date* is the date of the snapshot. For example:

```
# matchtopology Sept8.topo
Topology matching will take some time. Please wait!
-I---------------------------------------------------
-I- Topology matching results
-I---------------------------------------------------
-I- The topology defined in ib_topology.topo perfectly matches the
discovered fabric.
----------------------------------------------------------------
#
```

Related Information

- *Gateway Reference*, `generatetopology` command
- *Gateway Reference*, `matchtopology` command
- "Display the InfiniBand Fabric Topology (Simple)" on page 58

▼ Find 1x, SDR, or DDR Links in the Fabric

You can use the `ibdiagnet` command to determine which links are at 1x bandwidth, 2.5 Gbps, or 5 Gbps data rate.

- **On the management controller, type.**

```
# ibdiagnet -lw 4x -ls 10 -pc -pm -skip all
```

In this instance of the `ibdiagnet` command, there is a check for all links to be at 4x QDR (`-lw 4x -ls 10`), and if not, to report those links that are not 4x QDR.

Related Information

- *Gateway Reference*, `ibdiagnet` command
- "Display the Link Status of a Node" on page 60
- "Determine Which Links Are Experiencing Significant Errors" on page 69

▼ Determine Which Links Are Experiencing Significant Errors

You can use the `ibdiagnet` command to determine which links are experiencing symbol errors and recovery errors by injecting packets.

1. **On the management controller, type.**

```
# ibdiagnet -c 1000 -P all=1
```

In this instance of the `ibdiagnet` command, 1000 test packets are injected into each link. The `-P all=1` option returns all Performance Monitor counters that increment during the test, respective to the GUID and port of the InfiniBand device.

2. **In the output of the `ibdiagnet` command, search for the `symbol_error_counter` string.**

That line contains the symbol error count in hexadecimal. The preceding lines identify the node and port with the errors. Symbol errors are minor errors. If there are relatively few minor errors during the diagnostic, they can be monitored.

Note – According to InfiniBand specification 10E-12 BER, the maximum allowable symbol error rate is 120 errors per hour.

---

3. **Also in the output of the `ibdiagnet` command, search for the `link_error_recovery_counter` string.**

   That line contains the recovery error count in hexadecimal. The preceding lines identify the node and port with the errors. Recovery errors are major errors. The respective links must be investigated for the cause of the rapid symbol error propagation.

Related Information

- *Gateway Reference*, `ibdiagnet` command
- "Find 1x, SDR, or DDR Links in the Fabric" on page 68
- "Display Counters for a Node" on page 61

▼ Clear Data and Error Counters

When you are optimizing the InfiniBand fabric for performance, you might want to know how the throughput increases or decreases according to changes you are making. Alternatively, if you are troubleshooting a port, the `perfquery` command provides counters of errors occurring at that port. To determine the throughput as a function of time, or if the problem at the port has been resolved, the `ibdiagnet -pc` command enables you to reset the data counters and error counters to 0.

- **On the management controller, type.**

```
# ibdiagnet -pc
```

Related Information

- *Gateway Reference*, `ibdiagnet` command
- *Gateway Reference*, `perfquery` command
- "Display Counters for a Node" on page 61

▼ Reset a Port

You might need to reset a port to determine its functionality.

- **On the management controller, type.**

```
# ibportstate lid port reset
```

where:

- *lid* is the LID of the node.

- *port* is the port of the node.

For example:

```
# ibportstate 15 23 reset
Initial PortInfo:
# Port info: Lid 15 port 23
LinkState:.....................Down
PhysLinkState:.................Disabled
LinkWidthSupported:............1X or 4X
LinkWidthEnabled:..............1X or 4X
LinkWidthActive:...............4X
LinkSpeedSupported:............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:..............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:...............2.5 Gbps

After PortInfo set:
# Port info: Lid 15 port 23
LinkState:.....................Down
PhysLinkState:.................Disabled

After PortInfo set:
# Port info: Lid 15 port 23
LinkState:.....................Down
PhysLinkState:.................PortConfigurationTraining
#
```

Related Information

- *Gateway Reference*, `ibportstate` command
- "Disable a Port" on page 72
- "Enable a Port" on page 73

▼ Set Port Speed

You can manually set the speed of a single port to help determine symbol error generation. The `ibportstate` command can set the speed to 2.5, 5.0, or 10.0 Gbyte/sec.

- **On the management controller, type.**

```
# ibportstate lid port speed speed
```

where:

- *lid* is the LID of the node.
- *port* is the port of the node.

- *speed* is the speed of the port: `1` for 2.5 Gbyte/sec, `2` for 5.0 Gbyte/sec, and `4` for 10.0 Gbyte/sec.

Note – Adding speed values enables either speed. For example, speed 7 is 2.5, 5.0, and 10.0 Gbyte/sec.

For example:

```
# ibportstate 15 23 speed 1
Initial PortInfo:
# Port info: Lid 15 port 23
LinkSpeedEnabled:.................2.5 Gbps or 5.0 Gbps or 10.0 Gbps

After PortInfo set:
# Port info: Lid 15 port 23
LinkSpeedEnabled:.................2.5 Gbps
# ibportstate 15 23 speed 7
Initial PortInfo:
# Port info: Lid 15 port 23
LinkSpeedEnabled:.................2.5 Gbps
After PortInfo set:
# Port info: Lid 15 port 23
LinkSpeedEnabled:.................2.5 Gbps or 5.0 Gbps or 10.0 Gbps
#
```

Related Information

- *Gateway Reference*, `ibportstate` command
- "Display the Link Status of a Node" on page 60

▼ Disable a Port

If a port is found to be problematic, you can disable it.

- **On the management controller, type.**

```
# ibportstate lid port disable
```

where:

- *lid* is the LID of the node.

- *port* is the port of the node.

For example:

```
# ibportstate 15 23 disable
Initial PortInfo:
# Port info: Lid 15 port 23
LinkState:.....................Active
PhysLinkState:.................LinkUp
LinkWidthSupported:............1X or 4X
LinkWidthEnabled:..............1X or 4X
LinkWidthActive:...............4X
LinkSpeedSupported:............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:..............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:...............10.0 Gbps

After PortInfo set:
# Port info: Lid 15 port 23
LinkState:.....................Down
PhysLinkState:.................Disabled
#
```

Related Information

- *Gateway Reference*, `ibportstate` command
- "Disable a Switch Chip Port" on page 40
- "Enable a Port" on page 73
- "Reset a Port" on page 70

▼ Enable a Port

After disabling a port, you can enable the port with the `ibportstate` command.

- **On the management controller, type.**

```
# ibportstate lid port enable
```

where:

- *lid* is the LID of the node.

- *port* is the port of the node.

For example:

```
# ibportstate 15 23 enable
Initial PortInfo:
# Port info: Lid 15 port 23
LinkState:.....................Down
PhysLinkState:.................Disabled
LinkWidthSupported:............1X or 4X
LinkWidthEnabled:..............1X or 4X
LinkWidthActive:...............4X
LinkSpeedSupported:............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedEnabled:..............2.5 Gbps or 5.0 Gbps or 10.0 Gbps
LinkSpeedActive:...............2.5 Gbps

After PortInfo set:
# Port info: Lid 15 port 23
LinkState:.....................Down
PhysLinkState:.................PortConfigurationTraining
#
```

Related Information

- *Gateway Reference*, `ibportstate` command
- "Enable a Switch Chip Port" on page 41
- "Disable a Port" on page 72
- "Reset a Port" on page 70

Configuring the Fabric Director Node List

These topics enable you to configure the Fabric Director node list.

- "Fabric Director and Fabric Elements" on page 75
- "`fdconfig` Command" on page 76
- "Create a Fabric Configuration" on page 77
- "Add or Remove an Element From the Fabric Configuration" on page 80
- "Modify an Element of the Fabric Configuration" on page 82
- "Exchange Master Fabric Director Status" on page 84

Related Information

- "Monitoring the InfiniBand Fabric" on page 55
- "Controlling the InfiniBand Fabric" on page 66
- "Partitioning the InfiniBand Fabric" on page 85

Fabric Director and Fabric Elements

The Fabric Director monitors the InfiniBand fabric by polling the nodes or *elements* of the fabric. Each fabric element has a Fabric Director instance, and the two communicate with eachother in a one-to-one relationship. The Fabric Director gathers information about the fabric element through the LDA and ENVD daemons running on the fabric element. From the information gathered, the Fabric Director creates a local SNMP `fabricMIB`.

One Fabric Director is assigned the role of master. The master Fabric Director gathers information about all of the fabric elements by communicating out-of-band (management network) with the Fabric Directors instances on those fabric elements. This master Fabric Director creates a model of the fabric which is also in the form of an SNMP MIB, as defined by the `SUN-FABRIC-MIB` specification.

By default, Fabric Directors run in non-master mode. Only when a Fabric Director reads the fabric node list and discovers that its respective fabric element is to host the master Fabric Director, does it become the master Fabric Director.

Fabric Directors generate `fabricMIB` tables such as:

- `fabricElemPortTable`
- `fabricElemConnectorTable`

and also generate `fabricMIB` scalars such as:

- `fabricElemMgrIpAddress`
- `fabricElemName`
- `fabricElemNumPorts`
- `fabricElemOperStatus`
- `fabricElemNumConnectors`

Master Fabric Directors additionally generate `fabricMIB` tables such as:

- `fabricMgmtElemTable`
- `fabricMgmtExtElemTable`
- `fabricMgmtGatewayTable`
- `fabricMgmtLinkTable`

and generate `fabricMIB` scalars such as:

- fabricMgmtFabricDescr
- fabricMgmtFabricType
- fabricMgmtFabricTopology
- fabricMgmtFabricOperStatus
- fabricMgmtFabricName

Related Information

- *Gateway Reference*, SUN-FABRIC-MIB mib
- "fdconfig Command" on page 76
- "Create a Fabric Configuration" on page 77
- "Add or Remove an Element From the Fabric Configuration" on page 80
- "Modify an Element of the Fabric Configuration" on page 82
- "Exchange Master Fabric Director Status" on page 84

fdconfig Command

The fdconfig command is used to configure a list of fabric elements expected to be in the fabric, the fabric node list. The list is used by the Fabric Directors, and provides these attributes for each fabric element:

- Name – The name of the fabric element. Can be the host name if desired.
- IP address – The out-of-band management network IP address of the management controller hosting the fabric element and Fabric Director.
- Fabric element type – The name of the type of fabric element.
  - sw36 – Sun Datacenter InfiniBand Switch 36
  - sw36gw – Sun Network QDR InfiniBand Gateway Switch
  - unknown – The element's type is unknown.
- Role type – The purpose of the fabric element:
  - gateway – The element acts as a gateway.
  - spineSwitch – The element acts as a spine switch.
  - leafSwitch – The element acts as a leaf switch.
  - unknown – The element's role is unknown.
- Master – Whether the Fabric Director instance on the fabric element is the master.

Note – At this time, the fdconfig command is only available to the root user.

Note – See *Gateway Reference*, `fdconfig` command, for more information.

Related Information

- *Gateway Reference*, `fdconfig` command
- "Fabric Director and Fabric Elements" on page 75
- "Create a Fabric Configuration" on page 77
- "Add or Remove an Element From the Fabric Configuration" on page 80
- "Modify an Element of the Fabric Configuration" on page 82
- "Exchange Master Fabric Director Status" on page 84

▼ Create a Fabric Configuration

Before you can use the fabric MIBs, you must first create a fabric configuration. You must create identical configurations on all elements of the fabric.

Note – At this time, the `fdconfig` command is only available to the `root` user.

1. On the management controller, initiate a fabric configuration seesion.

```
# fdconfig start-fabric-config empty
#
```

2. Define the fabric's name.

```
# fdconfig define-fabric-name name
```

where *name* is the identifier of the fabric. For example:

```
# fdconfig define-fabric-name test
Fabric name updated
#
```

3. **Define an element of the fabric.**

```
# fdconfig define-element -name name -ip IP_address -type type -role
role1 [-role role2]
```

where:

- *name* is the identifier of the element (can be host name).
- *IP_address* is the IP address of an element.
- *type* is the identifer of the type of element:
  - `sw36` – Sun Datacenter InfiniBand Switch 36
  - `sw36gw` – Sun Network QDR InfiniBand Gateway Switch
  - `unknown` – The element's type is unknown.
- *role* is the purpose of the element:
  - `gateway` – The element acts as a gateway.
  - `spineSwitch` – The element acts as a spine switch.
  - `leafSwitch` – The element acts as a leaf switch.
  - `unknown` – The element's role is unknown.

For example:

```
# fdconfig define-element -name primary -ip 123.45.67.89 -type sw36gw -role
leafSwitch -role gateway
Element added
#
```

4. **Verify the fabric configuration.**

```
# fdconfig list-in-progress-fabric-config
Name                              IP Addr             Type    Role(s)
isMaster        Fabricname
---------------------------------         ----------------        ------
--------------------    --------        ----------
primary                          123.45.67.89         sw36gw  gateway
yes             test
#
```

5. **Consider your next step.**
   - If you want to add elements, perform Step 3 to Step 4 for each new element.

- If you want to abort the configuration altogether, type.

```
# fdconfig abort
In progress config aborted
#
```

- Otherwise, go to Step 6.

6. **End the configuration session and commit to the new configuration.**

```
# fdconfig complete-fabric-config
#
```

7. **Perform Step 1 through Step 6 on the management controllers of all elements of the fabric.**

8. **Reinitiate the fabric configuration on the element which is to host the master Fabric Director.**

```
# fdconfig start-fabric-config from-current
#
```

9. **Configure the element to host the master Fabric Director.**

Note – There must be only one element hosting the master Fabric Director within the fabric.

```
# fdconfig set-master name
```

where *name* is the identifier assigned to the element to host the master Fabric Director. For example:

```
# fdconfig set-master primary
Element primary updated
#
```

10. **End the configuration session and commit to the new configuration.**

```
# fdconfig complete-fabric-config
#
```

11. **Perform** Step 8 **through** Step 10 **on the management controllers of all elements of the fabric.**

    The configuration files and fabric MIBs are created.

Related Information

- *Gateway Reference*, `fdconfig` command
- "Fabric Director and Fabric Elements" on page 75
- "`fdconfig` Command" on page 76
- "Add or Remove an Element From the Fabric Configuration" on page 80
- "Modify an Element of the Fabric Configuration" on page 82
- "Exchange Master Fabric Director Status" on page 84

▼ Add or Remove an Element From the Fabric Configuration

Note – At this time, the `fdconfig` command is only available to the `root` user.

1. **Initiate a fabric configuration seesion.**

```
# fdconfig start-fabric-config from-current
#
```

2. **Add an element to the fabric configuration.**

```
# fdconfig define-element -name name -ip IP_address -type type -role
role1 [-role role2]
```

where:

- *name* is the identifier of the element (can be host name).
- *IP_address* is the IP address of an element.
- *type* is the identifer of the type of element:
  - `sw36` – Sun Datacenter InfiniBand Switch 36
  - `sw36gw` – Sun Network QDR InfiniBand Gateway Switch
  - `unknown` – The element's type is unknown.
- *role* is the purpose of the element:
  - `gateway` – The element acts as a gateway.

- `spineSwitch` – The element acts as a spine switch.
- `leafSwitch` – The element acts as a leaf switch.
- `unknown` – The element's role is unknown.

For example:

```
# fdconfig define-element -name secondary -ip 123.45.67.90 -type
sw36 -role leafSwitch
Element added
#
```

3. Or, remove an element from the fabric configuration.

```
# fdconfig remove-element -name name
```

where *name* is the identifier of the element. For example.

```
# fdconfig remove-element -name secondary
Element deleted
#
```

4. Verify the fabric configuration.

```
# fdconfig list-in-progress-fabric-config
Name                              IP Addr              Type    Role(s)
isMaster        Fabricname
---------------------------------          -----------------         ------
--------------------    --------         ----------
primary                           123.45.67.89          sw36gw  gateway
yes             test
secondary                         123.45.67.90          sw36    leafSwitch
no              test
#
```

5. End the configuration session and commit to the new configuration.

```
# fdconfig complete-fabric-config
#
```

6. Perform Step 1 through Step 5 on the management controllers of all elements of the fabric.

7. **If you added a new element, create a fabric configuration on the management controller of that element using the same configuration information.**

   See "Create a Fabric Configuration" on page 77.

Related Information

- *Gateway Reference*, `fdconfig` command
- "Fabric Director and Fabric Elements" on page 75
- "`fdconfig` Command" on page 76
- "Create a Fabric Configuration" on page 77
- "Modify an Element of the Fabric Configuration" on page 82
- "Exchange Master Fabric Director Status" on page 84

▼ Modify an Element of the Fabric Configuration

You can modify the parameters of an element.

Note – To change the master Fabric Director status, see "Exchange Master Fabric Director Status" on page 84.

Note – At this time, the `fdconfig` command is only available to the `root` user.

---

1. **Initiate a fabric configuration seesion.**

```
# fdconfig start-fabric-config from-current
#
```

2. **Modify an element of the fabric configuration.**

```
# fdconfig redefine-element -name name [-ip IP_address|-type
type|-role role1 [-role role2]]
```

   where:

   - *name* is the identifier of the element (can be host name).
   - *IP_address* is the IP address of an element.
   - *type* is the identifer of the type of element:
     - `sw36` – Sun Datacenter InfiniBand Switch 36
     - `sw36gw` – Sun Network QDR InfiniBand Gateway Switch

- `unknown` – The element's type is unknown.

- *role* is the purpose of the element:
  - `gateway` – The element acts as a gateway.
  - `spineSwitch` – The element acts as a spine switch.
  - `leafSwitch` – The element acts as a leaf switch.
  - `unknown` – The element's role is unknown.

For example:

```
# fdconfig redefine-element -name secondary -role spineSwitch
Element modified
#
```

3. **Verify the fabric configuration.**

```
# fdconfig list-in-progress-fabric-config
Name                              IP Addr              Type    Role(s)
isMaster        Fabricname
---------------------------------         -----------------        ------
--------------------    --------        ----------
primary                           123.45.67.89         sw36gw  gateway
yes             test
secondary                         123.45.67.90         sw36    spineSwitch
no              test
#
```

4. **End the configuration session and commit to the new configuration.**

```
# fdconfig complete-fabric-config
#
```

5. **Perform Step 1 through Step 4 on the management controllers of all elements of the fabric.**

Related Information

- *Gateway Reference*, `fdconfig` command
- "Fabric Director and Fabric Elements" on page 75
- "`fdconfig` Command" on page 76
- "Create a Fabric Configuration" on page 77
- "Add or Remove an Element From the Fabric Configuration" on page 80
- "Exchange Master Fabric Director Status" on page 84

▼ Exchange Master Fabric Director Status

Should you need to move the master Fabric Director status from one element to another, follow this procedure.

Note – There must be one element in the fabric hosting the master Fabric Director.

Note – At this time, the `fdconfig` command is only available to the `root` user.

---

1. Initiate a fabric configuration seesion.

```
# fdconfig start-fabric-config from-current
#
```

2. Remove master Fabric Director status from the first element.

```
# fdconfig reset-master name
```

where *name* is the identifier of the element. For example:

```
# fdconfig reset-master primary
Element primary updated
#
```

3. Assign master Fabric Director status to the new element.

```
# fdconfig set-master name
```

where *name* is the identifier of the element. For example:

```
# fdconfig set-master secondary
Element secondary updated
#
```

4. Verify the fabric configuration.

```
# fdconfig list-in-progress-fabric-config
Name                            IP Addr             Type    Role(s)
isMaster        Fabricname
--------------------------------          -----------------         ------
--------------------    --------         ----------
```

```
primary                         123.45.67.89         sw36gw  gateway
no              test
secondary                       123.45.67.90         sw36    spineSwitch
yes             test
#
```

5. **End the configuration session and commit to the new configuration.**

```
# fdconfig complete-fabric-config
#
```

6. **Perform Step 1 through Step 5 on the management controllers of all elements of the fabric.**

Related Information

- *Gateway Reference*, `fdconfig` command
- "Fabric Director and Fabric Elements" on page 75
- "`fdconfig` Command" on page 76
- "Create a Fabric Configuration" on page 77
- "Add or Remove an Element From the Fabric Configuration" on page 80
- "Modify an Element of the Fabric Configuration" on page 82

Partitioning the InfiniBand Fabric

These topics enable you to partition the InfiniBand fabric and modify the partition's configuration.

- "`smpartition` Command" on page 86
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Use the `smnodes` Command" on page 88
- "Determine Partitions and P_Keys" on page 89
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Modify a Partition or Port" on page 95
- "Delete a Partition" on page 97

Related Information

- "Monitoring the InfiniBand Fabric" on page 55
- "Controlling the InfiniBand Fabric" on page 66
- "Configuring the Fabric Director Node List" on page 74

`smpartition` Command

You can partition your InfiniBand fabric using the `smpartition` command and its subcommands. The `smpartition` command is issued from the gateway that has the master Subnet Manager and acts upon the partition configuration. There are two configurations, the *active* configuration is the one currently in use, and the *modified* configuration is the one which you can affect.

You begin a configuration session with the `smpartition start` command. You then make changes to the modified configuration with the `smpartition create`, `smpartition add`, `smpartition remove`, `smpartition modify`, and `smpartition delete` commands. You finally make the modified configuration into the active configuration with the `smpartition commit` command.

See *Gateway Reference*, `smpartition` command, for more information.

Note – There is a hardware limitation of 127 user-defined and 1 default partition for each port of the ConnectX2 chips (in HCAs) and for each InfiniBand port of the BridgeX chips (in gateways).

Related Information

- *Gateway Reference*, `smpartition` command
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Use the `smnodes` Command" on page 88
- "Determine Partitions and P_Keys" on page 89
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Modify a Partition or Port" on page 95
- "Delete a Partition" on page 97

▼ Remove User Partitions for Firmware Downgrade

You must perform these steps before downgrading the firmware to a version before 2.0.

Note – Only perform this procedure if you are downgrading the firmware.

Note – By removing user partitions, you might lose connectivity that was available only through those partitions. Additionally, you might gain undesired connectivity because all hosts become full members of the default partition.

---

1. **On the management controller of the gateways where VNICs using user partitions were created, remove all VNICs that are using user partitions.**

   See "Delete VNICs" on page 150.

2. **Remove all VLANs mapped to user partitions.**

   See "Delete VLANs" on page 144.

3. **Remove all user defined partitions.**

   See "Delete a Partition" on page 97.

4. **Restore the default partition to default settings.**

   See "Modify a Partition or Port" on page 95, however use this command line to modify the partition.

```
# smpartition modify -pkey 0x7fff -port ALL_CAS -flag ipoib -m full
```

5. **Consider your next step:**
   - If your InfiniBand fabric has just one Subnet Manager, downgrade the firmware.

     See *Gateway Remote Administration*, upgrading the firmware.
   - If your InfiniBand fabric has multiple Subnet Managers, go to Step 6.

6. **Remove all entries from the Subnet Manager node list.**

```
# smnodes delete IP_address [IP_address ...]
```

where *IP_address* are the IP addresses of the Subnet Manager nodes. For example:

```
# smnodes delete 123.45.67.89 123.45.67.90
#
```

7. **Repeat Step 6 on the management controllers of all Subnet Managers in the InfiniBand fabric.**

8. **Downgrade the firmware.**

   See *Gateway Remote Administration*, upgrading the firmware.

Related Information

- *Gateway Reference*, `smpartition` command
- *Gateway Reference*, `smnodes` command
- "`smpartition` Command" on page 86
- "Use the `smnodes` Command" on page 88
- "Determine Partitions and P_Keys" on page 89
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Modify a Partition or Port" on page 95
- "Delete a Partition" on page 97
- "Delete VNICs" on page 150
- "Delete VLANs" on page 144

▼ Use the `smnodes` Command

If you are partitioning your InfiniBand fabric, a list of valid Subnet Manager nodes must exist in the file system of every management controller running a Subnet Manager. You create this file with the `smodes` command. The file contains a list of IP addresses of all active management controllers running a Subnet Manager in your fabric. The file should have an entry for every Sun Datacenter InfiniBand Switch 36 and Sun Network QDR InfiniBand Gateway Switch that runs a Subnet Manager in your InfiniBand fabric.

Note – If the Subnet Manager nodes of your InfiniBand fabric ever change (disabled, added, and so on), you must update all copies of the Subnet Manager nodes file.

1. **Determine the IP addresses of all management controllers in your InfiniBand fabric.**
2. **On management controller, propagate the Subnet Manager nodes file with the IP addresses of all Subnet Manager nodes.**

```
# smnodes add IP_address IP_address ...
```

where *IP_address* is the IP address of each management controller hosting a Subnet Manager. For example:

```
# smnodes add 123.45.67.89 123.45.67.90
#
```

3. **Repeat** Step 2 **for all management controllers in the InfiniBand fabric.**
4. **Determine how you will partition your InfiniBand fabric.**

   See "Determine Partitions and P_Keys" on page 89.

Related Information

- *Gateway Reference*, `smodes` command
- "`smpartition` Command" on page 86
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Determine Partitions and P_Keys" on page 89
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Modify a Partition or Port" on page 95
- "Delete a Partition" on page 97

▼ Determine Partitions and P_Keys

The gateway supports several partitions in the InfiniBand fabric. You assign a P_Key to each partition as a simplified means of identifying the partition to the Subnet Manager. P_Keys are 15-bit integers and have a value of `0x1` to `0x7fff`. A P_Key value of `0x7fff` represents the default partition.

An additional bit, the membership bit, can identify the membership of the partition:

- Full – The membership bit is `1`. Full membership permits communication to all members within a partition.
- Limited – The membership bit is `0`. Limited membership permits communication only with a full member.

Combined together, P_Key and the membership bit comprise a 16-bit integer, and the most significant bit is the membership bit. In a full membership, the membership bit is set high. When this happens, the P_Key value is effectively increased by `0x8000`. Similarly, if you were to define a P_Key with a value greater than `0x8000`, the membership bit is automatically set to `1`, and consequently is given full membership.

1. **Use the output of the `ibswitches`, `ibhosts`, and `showgwports` commands to identify your switch, HCA, and gateway node GUIDs.**

   See:

   - "Identify All Switches in the Fabric" on page 56
   - "Identify All HCAs in the Fabric" on page 57
   - "Display Gateway Port Information" on page 133

2. **Determine the partitions you will have, their names, and their respective P_Keys.**

3. **Collate the GUIDs, partition names, and P_Keys into a partition information text file.**

4. **(Optional) If you will also configure VLANs, assign a unique VLAN identifier to similar P_Keys in the text file.**

Related Information

- *Gateway Reference*, `ibswitches` command
- *Gateway Reference*, `ibhosts` command
- *Gateway Reference*, `showgwports` command
- "`smpartition` Command" on page 86
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Use the `smnodes` Command" on page 88
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Modify a Partition or Port" on page 95
- "Delete a Partition" on page 97

▼ Create a User Partition

Note – After creating user partitions, consider modifying the default partition to limited membership for CAs and disabling IPoIB. The reason being that nodes with full membership in different user partitions can communicate with each other through the default partition. See "Modify a Partition or Port" on page 95 for an example.

1. **If you have not already done so, create the Subnet Manager nodes file.**

   See "Use the `smnodes` Command" on page 88.

2. **Initiate a partition configuration session on the management controller.**

```
# smpartition start
#
```

3. **Create the user partition.**

```
# smpartition create -n partition_name -pkey p_key[-flag [ipoib, mtu
mtu, rate rate, sl sl, scope scope]][-m defmember]
```

   where:

   - *partition_name* is the identifier of the InfiniBand partition.
   - *p_key* is the partition key (`1` – `7fff` or `default`).
   - *mtu* is the value of the MTU.
   - *rate* is the throughput of a link (link width + link speed) in Gbps from `2.5`, `5`, `10`, `20`, `30` to a maximum of `40`.
   - *sl* is the service level.
   - *scope* is the multicast address scope value (`0` – `F`).

Note – The `mtu`, `rate`, `sl`, and `scope` parameters are for the multicast group created when `ipoib` (IP over InfiniBand) is configured for the partition.

   - *defmember* is the default membership type (`full`, `limited`, or `both`) for the partition.

Note – If ports are added to the partition without specifying the membership type, the default membership type is applied to the port

For example:

```
# smpartition create -n testpartition -pkey 5 -m full
#
```

4. **Verify the partition configuration.**

```
# smpartition list modified
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 13
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=
full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
testpartition = 0x0005,defmember=full:
#
```

5. **Consider your next steps.**
   - If you want to add GUIDs, see "Add or Remove a Port From a Partition" on page 93.
   - If you want to modify the configuration, see "Modify a Partition or Port" on page 95.
   - If you want to delete the configuration altogether, see "Delete a Partition" on page 97.
   - Otherwise, go to Step 6.

6. **End the configuration session and commit the configuration to the active partition configuration.**

```
# smpartition commit
#
```

Related Information

- *Gateway Reference*, `smpartition` command
- "`smpartition` Command" on page 86
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Use the `smnodes` Command" on page 88
- "Determine Partitions and P_Keys" on page 89

▼ Add or Remove a Port From a Partition

You can add or remove one or more ports from the partition at one time.

1. **Initiate a partition configuration session on the management controller.**

```
# smpartition start
#
```

2. **Add ports to the partition.**

```
# smpartition add -n partition_name|-pkey p_key -port
port|ALL|ALL_CAS|ALL_SWITCHES|ALL_ROUTERS|SELF [-m member]
```

where:

- *partition_name* is the identifier of the InfiniBand partition.
- *p_key* is the partition key (`1` – `7fff`).
- *port* is the GUID of the port, or the special parameter, to add:
  - `ALL` – All of the CAs, switches, and routers in the InfiniBand fabric.
  - `ALL_CAS` – All CAs in the InfiniBand fabric.
  - `ALL_SWITCHES` – All switches.
  - `ALL_ROUTERS` – All routers.
  - `SELF` – The Master Subnet Manager.
- *member* is the membership type (`full`, `limited`, or `both`) for the port.

For example:

```
# smpartition add -n testpartition -port 00212800013e9313 00212800013e9314
00212800013e93f7
#
```

3. **Or, remove ports from the partition.**

```
# smpartition remove -n partition_name|-pkey p_key -port
port|ALL|ALL_CAS|ALL_SWITCHES|ALL_ROUTERS|SELF
```

where:

- *partition_name* is the identifier of the InfiniBand partition.
- *p_key* is the partition key (`1` – `7fff`).
- *port* is the GUID of the port, or the special parameter, to remove:
  - `ALL` – All of the CAs, switches, and routers in the InfiniBand fabric.
  - `ALL_CAS` – All CAs in the InfiniBand fabric.
  - `ALL_SWITCHES` – All switches.
  - `ALL_ROUTERS` – All routers.
  - `SELF` – The Master Subnet Manager.

For example:

```
# smpartition remove -n testpartition -port 00212800013e9314
#
```

4. **Verify the partition configuration.**

```
# smpartition list modified
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 13
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=
full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
testpartition = 0x0005,defmember=full:
0x00212800013e9313,
0x00212800013e93f7;
#
```

5. **End the configuration session and commit the configuration to the active partition configuration.**

```
# smpartition commit
#
```

Related Information

- *Gateway Reference*, `smpartition` command

▼ Modify a Partition or Port

You can modify the partition's configuration or the port(s) membership with the `smpartition modify` command.

1. Initiate a partition configuration session on the management controller.

```
# smpartition start
#
```

2. Modify the partition or port(s).

```
# smpartition modify -n partition_name|-pkey p_key[-flag [ipoib, mtu mtu, rate rate,
sl sl, scope scope]]|[-port port|ALL|ALL_CAS|ALL_SWITCHES|ALL_ROUTERS|SELF [-m
member]
```

where:

- *partition_name* is the identifier of the InfiniBand partition.
- *p_key* is the partition key (`1` – `7fff`).
- *mtu* is the value of the MTU.
- *rate* is the throughput of a link (link width + link speed) in Gbps from `2.5`, `5`, `10`, `20`, `30` to a maximum of `40`.
- *sl* is the service level.
- *scope* is the multicast address scope value (`0` – `F`).

Note – The `mtu`, `rate`, `sl`, and `scope` parameters are for the multicast group created when `ipoib` (IP over InfiniBand) is configured for the partition.

- *port* is the GUID of the port, or the special parameter, to modify:
  - `ALL` – All of the CAs, switches, and routers in the InfiniBand fabric.

- `ALL_CAS` – All CAs in the InfiniBand fabric.
- `ALL_SWITCHES` – All switches.
- `ALL_ROUTERS` – All routers.
- `SELF` – The Master Subnet Manager.

- *member* is the membership type (`full`, `limited`, or `both`) for the port.

For example, to configure the default partition for limited membership for CAs and no IP over InfiniBand support:

```
# smpartition modify -pkey 0x7fff -port ALL_CAS -flag -m limited
#
```

3. Verify the partition configuration.

```
# smpartition list modified
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 13
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=
full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
testpartition = 0x0005,defmember=full,ipoib:
0x00212800013e9313,
0x00212800013e93f7;
#
```

4. End the configuration session and commit the configuration to the active partition configuration.

```
# smpartition commit
#
```

Related Information

- *Gateway Reference*, `smpartition` command
- "`smpartition` Command" on page 86
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Use the `smnodes` Command" on page 88
- "Determine Partitions and P_Keys" on page 89
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Delete a Partition" on page 97

▼ Delete a Partition

When you delete a partition, you effectively commit a blank default configuration.

1. Initiate a partition configuration session on the management controller.

```
# smpartition start
#
```

2. Delete the partition.

```
# smpartition delete -n partition_name|-pkey p_key
```

where:

- *partition_name* is the identifier of the InfiniBand partition.
- *p_key* is the partition key (`1` – `7fff`).

For example:

```
# smpartition delete -n testpartition
#
```

3. End the configuration session and commit the configuration to the active partition configuration.

```
# smpartition commit
#
```

Related Information

- *Gateway Reference*, `smpartition` command
- "`smpartition` Command" on page 86
- "Remove User Partitions for Firmware Downgrade" on page 87
- "Use the `smnodes` Command" on page 88
- "Determine Partitions and P_Keys" on page 89
- "Create a User Partition" on page 91
- "Add or Remove a Port From a Partition" on page 93
- "Modify a Partition or Port" on page 95

Administering Gateway Resources

These topics describe the host and BridgeX functionality of the gateway and how to configure, monitor, and control that functionality.

Related Information

- "Troubleshooting the Gateway" on page 1
- "Understanding Administrative Commands" on page 17
- "Administering the Chassis" on page 23
- "Administering the I4 Switch Chip" on page 33
- "Administering the InfiniBand Fabric" on page 55
- "Administering the Subnet Manager" on page 43

Installing Gateway Supportive Software (Linux)

Install the BXOFED software on each host of the InfiniBand fabric that utilizes the VNIC and VLAN features of the gateway. Within the BXOFED software package are drivers and applications you use to configure VNICs and VLANs on the hosts.

- "Install the BXOFED Software (Linux)" on page 101

Related Information

- "Creating VNICs Under Gateway Manual Mode (Linux)" on page 103
- "Creating VNICs Under Host Manual Mode (Linux)" on page 107
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118
- "Monitoring Gateway Resources" on page 128
- "Controlling LAGs" on page 137
- "Controlling VLANs and VNICs" on page 142
- "Controlling Gateway Ports and Parameters" on page 151

▼ Acquire the BXOFED Software (Linux)

1. **Open a web browser on a host that will receive the BXOFED software.**
2. **Go to this URL:**

   `http://support.oracle.com`

   Oracle's My Oracle Support page is displayed.
3. **Sign in if you already have an account.**

   The dashboard page is displayed.

Note – If you do not have an account, you must register.

---

4. **Click the Patches & Updates tab.**

   The Patches & Updates page is displayed.
5. **In the Patch Search window, click the click Product or Family (Advanced Search).**

   The Patch Search window updates.
6. **In the Product is field, type BridgeX.**

   Possible products are suggested.
7. **Click on the most appropriate link.**

   The Release is field might autopropogate with the most current version.
8. **In the Release drop-down menu, select the most current version of the BridgeX OFED software.**

   For example, BridgeX OFED 1.5.1.

9. **In the Platform is drop-down menu, select Linux x86 or Linux x86-64.**

   For example, Linux x86-64.

10. **Click close.**

11. **Click Search.**

    The Patch Search window expands with the search results.

12. **In the Patch Name column, click the respective patch number link.**

    For example, `12621910`. The Patch Search window reformats.

13. **Click Download.**

    The File Download window opens.

14. **Click the *filename*`.zip` link to initiate the download.**

    For example, `p12621910_151_Linux-x86-64.zip`.

15. **Indicate where the file should be saved.**

    The file is downloaded and saved.

16. **In your receiving directory, decompress the *filename*`.zip` file.**

    The BXOFED software is in the `BXOFED-1.5.1-`*version*`_for Oracle.tgz` file. There are also readme, release notes, installation guide and user manual files in the *filename*`.zip` file.

17. **Read the readme, release notes, and installation guide files for information how to install the BXOFED software.**

Related Information

- *Gateway Remote Administration*, acquiring the gateway firmware package
- "Install the BXOFED Software (Linux)" on page 101

▼ Install the BXOFED Software (Linux)

When you install the BXOFED software, any previous installations of OFED or BXOFED software are removed. Configuration files are not removed.

Note – If you are installing the BXOFED software on a cluster, install the software onto one of the cluster nodes, then install the `.rpm` files in the `OFED-1.5.1/RPMS` on all remaining cluster nodes using cluster-aware tools.

---

1. **Become superuser of the host that received the BXOFED software.**

2. **Change to the directory where you extracted the `.tgz` file.**

3. **Run the installation script.**

```
# ./BXOFED-1.5.1-1.6.3/install.pl
```

The script begins. Interactive menus direct you through the installation process. During the installation, two configuration files are created:

- `ofed.conf` – contains the names of the software modules installed and the configuration settings chosen during the installation.
- `ofed_net.conf` – contains the IPoIB configuration settings chosen during the installation.

The script finishes. This information is found in the respective locations:

- Man pages are installed in `/usr/share/man`.
- Documentation is installed under the `/usr/share/doc` directory.
- IPoIB configuration information is installed under the `/etc/sysconfig/network*` directory.
- The `openibd` daemon is installed under the `/etc/init.d` directory.
- BXOFED commands are located in the `/usr/bin` and `/usr/sbin` directories.
- BXOFED software installation information is displayed with the `/etc/infiniband/info` script.

4. **(Optional) If the `.tgz` file was extracted to a NFS shared directory for a cluster, then to install the BXOFED software onto any remaining nodes in that cluster.**

   a. **Login as superuser of a node to receive the BXOFED software.**

   b. **Change to the directory where the `.tgz` file was extracted.**

   c. **Install the BXOFED software automatically.**

```
# ./BXOFED-1.5.1-1.6.3/install.pl -c path/ofed.conf -n path/ofed_net.conf
```

where *path* is the directory path to the `ofed.conf` and `ofed_net.conf` files.

   d. **Repeat from Step a for all nodes to receive the BXOFED software.**

5. **Reboot the Linux InfiniBand host(s).**

Related Information

- *Gateway Remote Administration*, upgrading the gateway firmware
- "Acquire the BXOFED Software (Linux)" on page 100

Creating VNICs Under Gateway Manual Mode (Linux)

These topics describe how to create VNICs under gateway manual mode:

Related Information

- "Installing Gateway Supportive Software (Linux)" on page 99
- "Creating VNICs Under Host Manual Mode (Linux)" on page 107
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118
- "Monitoring Gateway Resources" on page 128
- "Controlling LAGs" on page 137
- "Controlling VLANs and VNICs" on page 142
- "Controlling Gateway Ports and Parameters" on page 151

Gateway Manual Mode Overview (Linux)

The default and suggested means of creating and managing VNICs on the gateway is through use of BridgeX gateway administrative commands. To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Gateway_Mgmt` or `/SYS/Fabric_Mgmt` Linux shell targets of the Oracle ILOM CLI. The VNICs you create through the commands are based on the InfiniBand hosts.

When creating VNICs, consider this:

- You must specify the MAC for the VNIC. If you do not, the default MAC for the VNIC is 00:00:00:00:00:00, which renders the VNIC unusable.
- You must ensure that each MAC is unique.
- If you do not have a list of unique global MACs, use locally administrated MACs.
- VNICs will not go to an up state without an active Subnet Manager present.
- Manually created VNICs are persistent, and survive reboots and power cycles.

Related Information

- "Host Manual Mode Overview (Linux)" on page 108
- "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104
- "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105
- "Configure and Create VNICs for Gateway Manual Mode (Linux)" on page 106

▼ Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)

MAC addresses, GUIDs, and connector names or LAGs are assigned to VNICs.

Note – This procedure creates a *gateway* MAC address file, used to create VNICs in gateway manual mode.

---

1. **Determine if you will use unique global MAC addresses or locally administrated MAC addresses.**
2. **Open a text editor on the management controller and create a list of MAC addresses, one per row.**

Note – Each MAC address must be unique and not 00:00:00:00:00:00.

Note – Only even numbers are supported for the most significant byte of the MAC address (unicast).

---

3. **To the left of each MAC address, provide the HCA port GUID to receive that MAC address.**
4. **To the left of each GUID and MAC address pair, provide the name of the connector or LAG where the host will be physically connected to the 10GbE network.**

   There are eight connector names, `0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`.

5. **Save the file for when you create the VNICs.**

   An example entry in the gateway MAC address file might look like this:

```
0A-ETH-1 0003ba000100c70b 00:30:48:7d:de:e4
```

   For this example:

   - `0A-ETH-1` is the gateway connector.
   - `0003ba000100c70b` is the HCA port GUID.
   - `00:30:48:7d:de:e4` is the MAC address.

Related Information

- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "Gateway Manual Mode Overview (Linux)" on page 103
- "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105
- "Configure and Create VNICs for Gateway Manual Mode (Linux)" on page 106

▼ Determine VLAN Associations for Gateway Manual Mode (Linux)

If you plan to associate a VNIC with a VLAN, you also must partition the InfiniBand fabric and provide a partition key. If the VLAN is associated with the default P_Key (`7fff`), you need not create the default partition, for it already exists.

1. **If you have not done already, partition the InfiniBand fabric.**

   See "Partitioning the InfiniBand Fabric" on page 85.

2. **Use the `smpartition` command on the management controller to list the GUIDs associated with the partition and partition key.**

   See "Display the InfiniBand Fabric Partition Configuration" on page 65

3. **Open the gateway MAC address file on the management controller, and determine which addresses will be associated with a VLAN.**

4. **For each MAC address, append the VLAN identifier (`NO`, `0`, or `2` - `4094`) and partition key number to the right of the MAC address.**

Note – Do not use VLAN identifiers `1` or `4095`.

Note – Due to hardware limitations for MultiCast groups, there is a maximum of 1000 VLANs.

If MAC address will not be associated with a VLAN, the partition key will be `default`.

Note – When you associate a MAC with a VLAN, the GUID respective to the MAC must be a member of the partition associated with the partition key.

5. **Save the file for when you create the VLANs and VNICs.**

   An example entry in the gateway MAC address file might look like this:

```
0A-ETH-1 0003ba000100c70b 00:30:48:7d:de:e4 3 default
```

   For this example:

   - `0A-ETH-1` is the gateway connector.
   - `0003ba000100c70b` is the HCA port GUID.
   - `00:30:48:7d:de:e4` is the MAC address.
   - `3` is the VLAN identifier.
   - `default` is the partition key.

Related Information

- "Gateway Manual Mode Overview (Linux)" on page 103
- "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104
- "Configure and Create VNICs for Gateway Manual Mode (Linux)" on page 106

▼ Configure and Create VNICs for Gateway Manual Mode (Linux)

1. **Determine the VNIC configuration parameters.**

   See "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104.

2. **Determine VLAN associations.**

   See "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105.

3. **Create VLANs.**

   See "Create VLANs" on page 143.

4. **Create VNICs.**

   See "Create VNICs" on page 145.

Related Information

- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117
- "Gateway Manual Mode Overview (Linux)" on page 103
- "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104
- "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105

Creating VNICs Under Host Manual Mode (Linux)

These topics describe how to create VNICs under host manual mode:

- "Host Manual Mode Overview (Linux)" on page 108
- "Central Configuration File (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110
- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "Create the Central Confguration File (Linux)" on page 113
- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

Related Information

- "Installing Gateway Supportive Software (Linux)" on page 99
- "Creating VNICs Under Gateway Manual Mode (Linux)" on page 103
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118
- "Monitoring Gateway Resources" on page 128
- "Controlling LAGs" on page 137
- "Controlling VLANs and VNICs" on page 142

- "Controlling Gateway Ports and Parameters" on page 151

Host Manual Mode Overview (Linux)

Before creating VNICs in host manual mode, you must install the BXOFED software onto the hosts. Afterwards, you create VNICs using static configuration files located on the hosts. These configuration files define the number of VNICs and the VNICs' properties. The `mlx4_vnic_confd` daemon in the BXOFED software reads the configuration files and passes the relevant data to the `mlx4_vnic` module, which creates the VNICs.

The two types of configuration files for VNICs provide the same functionality:

- A central configuration file (`mlx4_vnic.conf`)
- VNIC-specific configuration files (`ifcfg-eth`*XX* where *XX* is the `eth` number of the VNIC)

If both forms of configuration files exist, the central configuration file has precedence and is the only file used.

Related Information

- "Gateway Manual Mode Overview (Linux)" on page 103
- "Central Configuration File (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110
- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "Create the Central Confguration File (Linux)" on page 113
- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

Central Configuration File (Linux)

The central configuration file is the `/etc/infiniband/mlx4_vnic.conf` file and consists of multiple single-line entries, each describing a VNIC. Each VNIC entry has this format and parameters:

```
name=name mac=mac ib_port=device:port [vid=vlan_id] vnic_id=number bx=string eport=
connector
```

where:

- *name* is the VNIC device name or `eth` number.
- *mac* is the MAC assigned to the VNIC.
- *device* is the device name retrieved from the `hca_id` field in the output of the `ibv_devinfo` command.
- *port* is the port number, either `1` or `2`.
- *vlan_id* is the VLAN identifier to assign to the VNIC **(**`NO`**,** `0`**, or** `2` **-** `4094`**)**.

Note – Do not use VLAN identifiers `1` and `4095`.

Note – The `vid` parameter is optional.

- *number* is a unique number to assign to the VNIC. The value is `1–32767`.
- *string* is either the HCA port GUID or the system name.
- *connector* is the gateway connector assigned to the VNIC (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).

Related Information

- "Host Manual Mode Overview (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110
- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "Create the Central Confguration File (Linux)" on page 113
- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

VNIC-Specific Configuration File (Linux)

The `ifcfg-eth`*XX* file already exists as a means for the network service to derive information about a node. The *XX* represents the `eth` number of the VNIC. To configure VNICs, additional parameters are added to the base file format. For Red Hat Linux, the `ifcfg-eth`*XX* file has this format:

```
DEVICE=name
HWADDR=mac
BOOTPROTO=dhcp
ONBOOT=yes
BXADDR=string
BXEPORT=connector
VNICVLAN=vlan_id
VNICIBPORT=device:port
```

where:

- *name* is the VNIC device name or `eth` number.

Note – The `DEVICE` parameter is optional. If the parameter is missing, the suffix of the configuration file name (`eth`*XX*) is used.

- *mac* is the MAC assigned to the VNIC.
- *string* is either the HCA port GUID or the system name.
- *connector* is the gateway connector assigned to the VNIC (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).
- *vlan_id* is the VLAN identifier to assign to the VNIC **(**`NO`**,** `0`**, or** `2` **-** `4094`**)**.

Note – Do not use VLAN identifiers `1` and `4095`.

Note – The `VNICVLAN` parameter is optional.

- *device* is the device name retrieved from the `hca_id` field in the output of the `ibv_devinfo` command.
- *port* is the port number, either `1` or `2`.

Note – Additional parameters used for regular `eth` interfaces can be appended to the `ifcfg-eth`*XX* file.

Related Information

- "Host Manual Mode Overview (Linux)" on page 108
- "Central Configuration File (Linux)" on page 108
- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "Create the Central Confguration File (Linux)" on page 113
- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

▼ Determine VNIC Configuration Parameters for Host Manual Mode (Linux)

MAC addresses, GUIDs, and other parameters are assigned to VNICs.

Note – This procedure creates a *host* MAC address file, used to create VNICs in host manual mode.

---

1. **Determine if you will use unique global MAC addresses or locally administrated MAC addresses.**
2. **As superuser of a host in the InfiniBand fabric, open a text editor and create a numbered list, one number per row.**
3. **For each number, provide a MAC address to the right.**

Note – Each MAC address must be unique and not 00:00:00:00:00:00.

Note – Only even numbers are supported for the most significant byte of the MAC address (unicast).

---

4. **For each MAC address, open a terminal window and log in to the host to receive that MAC address.**

5. **Use the** `ibv_devinfo` **command to determine the** `hca_id` **value and HCA port GUID.**

   For example:

```
# ibv_devinfo
hca_id:mlx4_0
  fw_ver:          2.5.9266
  node_guid:       0003:ba00:0100:c708
  sys_image_guid:  0003:ba00:0100:c70b
.
.
.
#
```

   In the example, the `hca_id` value is `mlx4_0`, and the HCA port GUID (sys_image_guid) is `003ba000100c70b`.

6. **For each MAC address, append the host name, the respective** `hca_id` **value, the port (1 or 2), and the HCA port GUID to the right of the MAC address.**

7. **For each MAC address that will be associated with a VLAN, append the VLAN identifier (**`NO`**,** `0`**, or** `2` **-** `4094`**) to the right of the GUID.**

Note – Do not use VLAN identifiers `1` and `4095`.

Note – Due to hardware limitations for MultiCast groups, there is a maximum of 1000 VLANs.

---

8. **For each MAC address, append the name of the gateway connector where the host will physically connect to the 10GbE network to the right of the GUID or VLAN identifier (if present).**

9. **Save the file for future use.**

   An example entry in the host MAC address file might look like this:

```
1 00:30:48:7d:de:e4 cupcake mlx4_0 1 0003ba000100c70b 0 0A-ETH-1
```

   For this example:

   - `1` is the entry number. This will become the VNIC number and `eth` number.
   - `00:30:48:7d:de:e4` is the MAC address.
   - `cupcake` is the host name and is used to identify the location of the VNIC.
   - `mlx4_0` is the value of the `hca_id` field and is called the device name.

- `1` is the port of the HCA.
- `0003ba000100c70b` is the HCA port GUID.
- `0` is the VLAN identifier.
- `0A-ETH-1` is the connector.

Related Information

- "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104
- "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105
- "Host Manual Mode Overview (Linux)" on page 108
- "Central Configuration File (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110
- "Create the Central Confguration File (Linux)" on page 113
- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

▼ Create the Central Confguration File (Linux)

The `etc/infiniband/mlx4_vnic.conf` file is shared by all hosts in the InfiniBand fabric. You create the configuration file from the host MAC address file.

1. **As superuser of a host in the InfiniBand fabric, open the host MAC address file that you created previously.**

   See "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111.

2. **Open a text editor to create the configuration file.**

3. **Create an entry in the configuration file with this format:**

```
name=name mac=mac ib_port=device:port [vid=vlan_id] vnic_id=number bx=string eport=
connector
```

where:

- *name* is `eth`*XX* and *XX* is column one of the host MAC address file.
- *mac* is column two of the host MAC address file.
- *device* is column four of the host MAC address file.
- *port* is column five of the host MAC address file.

- *vlan_id* is column seven of the host MAC address file.

Note – The `vid` parameter is optional.

- *number* is column one of the host MAC address file.
- *string* is column six of the host MAC address file.
- *connector* is column eight of the host MAC address file.

For example:

```
name=eth1 mac=00:30:48:7d:de:e4 ib_port=mlx4_0:1 vid=0 vnic_id=1 bx=
0003ba000100c70b eport=0A-ETH-1
```

4. **Repeat** Step 3 **for all rows of the host MAC address file.**
5. **Save the file with the name** `mlx4_vnic.conf`**.**
6. **Copy the** `mlx4_vnic.conf` **file to the** `/etc/infiniband` **directory of all hosts in the InfiniBand fabric.**

Related Information

- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "Host Manual Mode Overview (Linux)" on page 108
- "Central Configuration File (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110
- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

▼ Create the VNIC-Specific Confguration Files (Linux)

Each host will have a VNIC-specific configuration file for the VNICs it will host. Each line of the MAC address file becomes a VNIC-specific configuration file. Repeat this procedure ofr each host in the InfiniBand fabric.

1. **As superuser of a host in the InfiniBand fabric, open the host MAC address file that you created previously.**

   See "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111.

2. **Look in column three of the host MAC address file to find entries for the respective host.**

   Note the row numbers in column one.

3. **Open a text editor to create a configuration file.**

4. **Enter text from the entry appropriate for the respective host into the file with this format.**

```
DEVICE=name
HWADDR=mac
BOOTPROTO=dhcp
ONBOOT=yes
BXADDR=string
BXEPORT=connector
VNICVLAN=vlan_id
VNICIBPORT=device:port
```

   where:

   - *name* is `eth`*XX* and *XX* is column one of the host MAC address file.
   - *mac* is column two of the host MAC address file.
   - *string* is column six of the host MAC address file.
   - *connector* is column eight of the host MAC address file.
   - *vlan_id* is column seven of the host MAC address file.

Note – The `VNICVLAN` parameter is optional.

   - *device* is column four of the host MAC address file.
   - *port* is column five of the host MAC address file.

   For example:

```
DEVICE=eth1
HWADDR=00:30:48:7d:de:e4
BOOTPROTO=dhcp
ONBOOT=yes
BXADDR=0003ba000100c70b
BXEPORT=0A-ETH-1
VNICVLAN=0
VNICIBPORT=mlx4_0:1
```

5. **Save the file with the name of `ifcfg-eth`*XX* where *XX* is the value of column one.**

   For example, `ifcfg-eth1`.

6. **Repeat** Step 3 **to** Step 5 **for each entry appropriate for the host.**
7. **Repeat** Step 2 **to** Step 6 **for each host to have a VNIC configuration file.**

Related Information

- "Create the Central Confguration File (Linux)" on page 113
- "Host Manual Mode Overview (Linux)" on page 108
- "Central Configuration File (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110
- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "`mlx4_vnic_confd` Daemon (Linux)" on page 116
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

`mlx4_vnic_confd` Daemon (Linux)

The `mlx4_vnic_confd` daemon is used to manage host administrated VNICs. The daemon is located in the `/etc/init.d` directory of each InfiniBand host. The syntax for the command line is:

`mlx4_vnic_confd` *operation*

where *operation* sets the daemon's interaction with VNICs.

This table describes the *operations* supported by the `mlx4_vnic_confd` daemon.

Operation	Description
`start`	Starts, loads, or creates new host administrated VNICs.
`stop`	Stops all host administrated VNICs.
`restart`	Closes and then reopens all host administrated VNICs.
`reload`	Updates the system with the most current configuration files. This operation does not modify host administrated VNICs with unchanged configuration files.

Related Information

- "Host Manual Mode Overview (Linux)" on page 108
- "Central Configuration File (Linux)" on page 108
- "VNIC-Specific Configuration File (Linux)" on page 110

- "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111
- "Create the Central Confguration File (Linux)" on page 113
- "Create the VNIC-Specific Confguration Files (Linux)" on page 114
- "Configure and Create VNICs for Host Manual Mode (Linux)" on page 117

▼ Configure and Create VNICs for Host Manual Mode (Linux)

1. **Determine if you will use a central configuration file or VNIC-specific configuration files.**

   See:

   - "Central Configuration File (Linux)" on page 108
   - "VNIC-Specific Configuration File (Linux)" on page 110

2. **Determine the VNIC configuration parameters.**

   See "Determine VNIC Configuration Parameters for Host Manual Mode (Linux)" on page 111.

3. **Create the appropriate configuration files.**

   See:

   - "Create the Central Confguration File (Linux)" on page 113
   - "Create the VNIC-Specific Confguration Files (Linux)" on page 114

4. **On the management controller, enable host manual mode.**

```
# allowhostconfig
Stopping Bridge Manager..                                  [  OK  ]
Starting Bridge Manager.                                   [  OK  ]
#
```

5. **Login as superuser and start the `mlx4_vnic_confd` daemon.**

```
# /etc/init.d/mlx4_vnic_confd start
```

   The VNICs are created.

6. **Repeat Step 5 for all hosts of the InfiniBand fabric.**

Related Information

- "Configure and Create VNICs for Gateway Manual Mode (Linux)" on page 106

Creating Virtual IO Adapters (Oracle Solaris)

These topics discuss steps to enable gateway functionality for Oracle Solaris 11 hosts.

Step	Description	Links
1.	Review background information.	"Ethernet Over InfiniBand Overview (Oracle Solaris)" on page 119
2.	Verify the driver is installed.	"Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)" on page 120
3.	(Optional) Install the driver.	"Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121
4.	Create the VIOAs.	"Set Up VIOAs (Oracle Solaris)" on page 121
5.	Create IP interfaces and assign IP addresses to the VIOAs.	"Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125
6,	Create Oracle Solaris VNICs.	"Create Oracle Solaris VNICs (Oracle Solaris)" on page 127

Related Information

- "Controlling LAGs" on page 137
- "Controlling VLANs and VNICs" on page 142
- "Controlling Gateway Ports and Parameters" on page 151

Ethernet Over InfiniBand Overview (Oracle Solaris)

The Ethernet over InfiniBand driver is available in current releases of the Oracle Solaris 11 image. This driver supports the Data Link Provider Interface over all InfiniBand ports of an Oracle Solaris 11 host connected to the gateway. The driver uses the IBA unreliable datagram mode to enable initialization, gateway handshake, heartbeat management, frame transmit and receive functions, multicast support, and statistical reporting. The driver is delivered through an IPS-based package with the file name of `ethernet-over-ib`.

Network interfaces in the Oracle Solaris 11 operating system correspond to instances of network hardware devices, and the interfaces are configured over data links. The Ethernet over InfiniBand driver supports virtual network hardware on HCAs and enables interfaces to be assigned to that virtual hardware.

Administrators create IP interfaces on top of data links. A data link represents a link object in the second layer of the OSI model. The respective physical link is directly associated with a device (physical or virtual) and device instance name. The device instance name is comprised of the device driver name augmented with the instance number, which has a value of `0` to *n* instances of network devices (physical or virtual) using that driver. For each virtual device and interface created on the gateway, there is a corresponding data link named `eoib`*X* created on the Oracle Solaris 11 host.

Associating the InfiniBand port of an HCA to a gateway Ethernet port, and assigning one or more MAC addresses to the pair creates a virtual IO adapter (VIOA). The `createvnic` command of the gateway is used to fulfill this objective. Oracle Solaris discovers the VIOA, binds an `eoib`*X* datalink instance to the VIOA, and then manages the VIOA as if it were a physical network interface card (NIC). Like a physical NIC, the VIOA represents the access path to an Ethernet port. Because more than one MAC address can be assigned to a VIOA, the VNIC management commands of the gateway are actually managing the MAC addresses within the gateway itself.

Related Information

- "Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)" on page 120
- "Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121

- "Set Up VIOAs (Oracle Solaris)" on page 121
- "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125
- "Create Oracle Solaris VNICs (Oracle Solaris)" on page 127

▼ Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)

1. **Become superuser of the Oracle Solaris 11 host.**
2. **Type.**

```
# pkg info ethernet-over-ib
 Name: system/io/infiniband/ethernet-over-ib
 Summary: Ethernet over InfiniBand (EoIB) Drivers
 Description: InfiniBand device driver implementing Ethernet over InfiniBand
 Category: System/Hardware
 State: Installed
 Publisher: solaris
 Version: 0.5.11
 Build Release: 5.11
 Branch: 0.175.1.0.0.2.17991
Packaging Date: Mon Oct 24 10:50:29 2011
 Size: 304.74 kB
 FMRI:
pkg://solaris/system/io/infiniband/ethernet-over-ib@0.5.11,5.11-0.175.1.0.0.2.
17991:20111024T105029Z
#
```

3. **If there is no information returned, or an error message is displayed, install the Ethernet over InfiniBand driver.**

   See "Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121.

Related Information

- pkg command man page
- "Ethernet Over InfiniBand Overview (Oracle Solaris)" on page 119
- "Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121
- "Set Up VIOAs (Oracle Solaris)" on page 121
- "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125
- "Create Oracle Solaris VNICs (Oracle Solaris)" on page 127

▼ Install the Ethernet Over InfiniBand Driver (Oracle Solaris)

1. **Consider your first steps:**
   - If you know for certain that your Oracle Solaris 11 image contains the `ethernet-over-ib` package, go to Step 4.
   - If you are not certain, go to Step 2.
2. **Download the Oracle Solaris 11 operating system.**

   See this URL for instructions.

   `http://www.oracle.com/technetwork/server-storage/solaris11/downloads/index.html`
3. **Install the Oracle Solaris 11 operating system according to the documentation included with the software.**
4. **As superuser of the Oracle Solaris 11 host, install the Ethernet over InfiniBand package.**

```
# pkg install ethernet-over-ib
```

   Follow the prompts.

Related Information

- pkg command man page
- "Ethernet Over InfiniBand Overview (Oracle Solaris)" on page 119
- "Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)" on page 120
- "Set Up VIOAs (Oracle Solaris)" on page 121
- "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125
- "Create Oracle Solaris VNICs (Oracle Solaris)" on page 127

▼ Set Up VIOAs (Oracle Solaris)

1. **Become superuser of the Oracle Solaris 11 host.**

   For this procedure, the Oracle Solaris 11 host name is `solaris01`.

2. **Display the HCA GUIDs and port GUIDs configured on the host.**

```
# dladm show-ib
LINK         HCAGUID          PORTGUID         PORT STATE  PKEYS
ibp0         21280001A0A590   21280001A0A591   1    up     FFFF
ibp1         21280001A0A590   21280001A0A592   2    up     FFFF
#
```

3. **Record the host name, HCA GUIDs, and port GUIDs.**

   In this example, the HCA GUID is `21280001A0A590` and the port GUIDs are `21280001A0A591` and `21280001A0A592`.

4. **From the management controller of the gateway, display the HCAs recognized by the gateway.**

```
# ibhosts
Ca      : 0x0021280001A0A590 ports 2 "solaris01"
Ca      : 0x00212856cd22c040 ports 2 "SUN IB QDR GW switch mnm34-97 Bridge 1"
Ca      : 0x0002c903000891aa ports 2 "mnm34-54 HCA-1"
Ca      : 0x00212800013ece9e ports 2 "mnm34-55 HCA-1"
Ca      : 0x0003ba000100e370 ports 2 "mnm34-60 HCA-1"
.
.
.
#
```

5. **Compare the output of the `ibhosts` command with the recorded host name, HCA GUIDs, and port GUIDs.**

   The port GUIDs are the HCA GUID +1 (port 1) and HCA GUID +2 (port 2), respectively.

6. **Verify that the gateway recognizes the Oracle Solaris 11 host correctly.**

   If not, check the cabling between the gateway and the Oracle Solaris 11 host.

7. **Determine if a VIOA has already been assigned to the Oracle Solaris 11 host.**

```
# showvnics
ID  STATE    FLG IOA_GUID             NODE      IID  MAC             VLN PKEY   GW
--- --------  --- ---------------------- ---------- ---- ----------------- ---
----   --------
 4 UP        N 00:03:BA:00:01:00:E3:71 mnm34-60 0000 02:02:02:02:02:03 NO  ffff
0A-ETH-1
 5 UP        N 00:03:BA:00:01:00:E3:71 mnm34-60 0002 02:02:02:02:02:04 NO  ffff
0A-ETH-1
```

```
 3 DISABLED   N 00:03:BA:00:01:00:E3:71 mnm34-60 0000 02:02:02:02:02:02 NO  ffff
0A-ETH-1
 6 UP         N 00:03:BA:00:01:00:E3:72 mnm34-60 0000 02:02:02:02:02:05 NO  ffff
0A-ETH-1
#
```

8. **Examine the output by column:**
   - `STATE` – The state of the VIOA.
   - `IOA_GUID` – The port GUID of the VIOA. The GUID might be listed several times.
   - `NODE` – The host name of the node hosting the VIOA.
   - `MAC` – The MAC address assigned to the VIOA. There might be several MAC addresses.
   - `PKEY` – The partition key assigned.
   - `GW` – The physical connector cabled to the node or host.
9. **Determine your next steps.**
   - If no port GUID of the Oracle Solaris 11 host is listed, or there are no MAC addressess assigned to the port GUID, go to Step 10.
   - Otherwise, go to "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125.
10. **Determine your MAC addresses and whether they are to be global or locally administrated.**

Note – Each MAC address must be unique and not `00:00:00:00:00:00`.

Note – Only even numbers are supported for the most significant byte of the MAC address (unicast).

---

11. **Activate the VIOA by assigning a MAC address to it.**

```
# createvnic connector -guid guid -mac mac_address -pkey default
```

where:

- *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).
- *guid* is the global unique identifier of the target port on the host associated with the VIOA.
- *mac_address* is the MAC address to be assigned to the VIOA.

Note – You can use the information from Step 8 as a guide.

For example, to assign the MAC address of `a0:a5:91:95:30:9a` to the VIOA for port GUID `21280001A0A591` through connector `0A-ETH-2`:

```
# createvnic 0A-ETH-2 -guid 0021280001A0A591 -mac a0:a5:91:95:30:9a -pkey
default
VNIC created
#
```

The VIOA is created and an Ethernet over InfiniBand data link is bound to the VIOA.

12. **Verify that the VIOA is active.**

```
# showvnics
ID  STATE    FLG IOA_GUID             NODE      IID  MAC             VLN PKEY   GW
--- --------  --- ---------------------- ---------- ---- ----------------- ---
----   --------
 7 UP          N 00:21:28:00:01:A0:A5:91 solaris01 0000 a0:a5:91:95:30:9a NO
ffff 0A-ETH-2
 4 UP        N 00:03:BA:00:01:00:E3:71 mnm34-60 0000 02:02:02:02:02:03 NO  ffff
0A-ETH-1
 5 UP        N 00:03:BA:00:01:00:E3:71 mnm34-60 0002 02:02:02:02:02:04 NO  ffff
0A-ETH-1
 3 DISABLED   N 00:03:BA:00:01:00:E3:71 mnm34-60 0000 02:02:02:02:02:02 NO  ffff
0A-ETH-1
 6 UP        N 00:03:BA:00:01:00:E3:72 mnm34-60 0000 02:02:02:02:02:05 NO  ffff
0A-ETH-1
#
```

13. **Configure an IP address on the data link.**

    See "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125.

Related Information

- `dladm` command man page
- *Gateway Reference*, `ibhosts` command
- *Gateway Reference*, `showvnics` command
- *Gateway Reference*, `createvnics` command
- "Ethernet Over InfiniBand Overview (Oracle Solaris)" on page 119
- "Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)" on page 120
- "Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121

- "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125
- "Create Oracle Solaris VNICs (Oracle Solaris)" on page 127

▼ Configure IP Addresses on the Data Link (Oracle Solaris)

1. **Become superuser of the Oracle Solaris 11 host.**
2. **Verify that the data link corresponding to the activated VIOA has been created.**

```
# dladm show-phys -m
LINK                SLOT     ADDRESS             INUSE CLIENT
net1                primary  0:21:28:80:f2:e5    no    --
net2                primary  a0:a5:91:95:30:9a   no    --
#
```

In the `ADDRESS` column, the previously assigned MAC address is displayed. It corresponds to the link, `net2`.

Without the `-m` option, the `dladm show-phys` command displays all data links corresponding to all hardware devices. For example:

```
# dladm show-phys
LINK              MEDIA                STATE      SPEED  DUPLEX    DEVICE
net1              Ethernet             unknown    0      unknown   igb1
net2              Ethernet             up         10000  full      eoib0
#
```

Similarly, the `dladm show-ether` command displays all data links supporting the Ethernet protocol. For example:

```
# dladm show-ether
LINK              PTYPE    STATE    AUTO  SPEED-DUPLEX                    PAUSE
net1              current  unknown  yes   0M                               bi
net2              current  up       no    10G-f                            none
#
```

3. **Create an IP interface on the data link:**

```
# ipadm create-ip eoibX
```

where *X* is the instance of the IP interface. For example:

```
# ipadm create-ip eoib0
#
```

Note – You must create the IP interface before an IP address is assigned to it.

4. **Verify that the interface was created.**

```
# ipadm show-if
IFNAME CLASS STATE ACTIVE OVER
eoib0 ip down no --
#
```

5. **Assign an IP address to the IP interface.**

```
# ipadm create-addr -T static -a IP_address/24 IP_interface/v4
```

where:

- *IP_address* is the IP adress to be assigned to the IP interface.
- *IP_interface* is the interface corresponding to the data link and VIOA.

For example, to assign the IP address of `192.168.17.24` to IP interface `eoib0`:

```
# ipadm create-addr -T static -a 192.168.17.24/24 eoib0/v4
#
```

6. **Verify that the IP address was properly assigned.**

```
# ipadm show-addr IP_interface/v4
```

where *IP_interface* is the interface corresponding to the data link and VIOA. For example:

```
# ipadm show-addr eoib0/v4
ADDROBJ TYPE STATE ADDR
eoib0/v4 static ok 192.168.17.24/24
#
```

Related Information

- `dladm` command man page
- `ipadm` command man page
- "Ethernet Over InfiniBand Overview (Oracle Solaris)" on page 119
- "Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)" on page 120
- "Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121
- "Set Up VIOAs (Oracle Solaris)" on page 121
- "Create Oracle Solaris VNICs (Oracle Solaris)" on page 127

▼ Create Oracle Solaris VNICs (Oracle Solaris)

Note – This procedure does not create VNICs on Linux hosts or the gateway. This procedure is independent of and unrelated to Host Manual Mode or Gateway Manual Mode.

1. Create a Oracle Solaris VNIC on the data link.

```
# dladm create-vnic -l IP_interface  name
```

where:

- *IP_interface* is the interface corresponding to the data link and VIOA.
- *name* is the name of the VNIC. For ease of identification, the VNIC name can be formatted as `e`*X*`_vnic`*Y*, where:
  - *X* is the instance number of the Ethernet over InfiniBand IP interface.
  - *Y* is the instance of the VNIC, independent of the interface instance.

For example:

```
# dladm create-vnic -l eoib0 e0_vnic0
#
```

2. Verify the creation of the VNIC.

```
# dladm show-vnic
LINK OVER SPEED MACADDRESS MACADDRTYPE VID
e0_vnic0 eoib0 10000 2:8:20:7c:5d:d3 random 0
#
```

3. **Display a list of all data links.**

```
# dladm show-link
LINK CLASS MTU STATE OVER
eoib0 phys 1500 up --
e0_vnic0 vnic 1500 up eoib0
```

Note – For more information about network virtualization, see *Oracle Solaris Administration: Network Interfaces and Network Virtualization*, network virtualization and resource control.

Related Information

- `dladm` command man page
- "Ethernet Over InfiniBand Overview (Oracle Solaris)" on page 119
- "Verify That the Ethernet Over InfiniBand Driver Is Installed (Oracle Solaris)" on page 120
- "Install the Ethernet Over InfiniBand Driver (Oracle Solaris)" on page 121
- "Set Up VIOAs (Oracle Solaris)" on page 121
- "Configure IP Addresses on the Data Link (Oracle Solaris)" on page 125

Monitoring Gateway Resources

These topics describe how to monitor the gateway interfaces.

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Gateway_Mgmt` or `/SYS/Fabric_Mgmt` Linux shell targets of the Oracle ILOM CLI.

- "Display LAG Information" on page 129
- "Display the VLANs" on page 130
- "Display the VNICs" on page 131
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133
- "Display the Gateway Instance Number" on page 135
- "Display Information About the Gateway" on page 136

Related Information

- "Monitoring the Chassis" on page 23
- "Monitoring the I4 Switch Chip" on page 33
- "Monitoring the InfiniBand Fabric" on page 55
- "Monitoring the Subnet Manager" on page 43
- "Installing Gateway Supportive Software (Linux)" on page 99
- "Creating VNICs Under Gateway Manual Mode (Linux)" on page 103
- "Creating VNICs Under Host Manual Mode (Linux)" on page 107
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118
- "Controlling LAGs" on page 137
- "Controlling VLANs and VNICs" on page 142
- "Controlling Gateway Ports and Parameters" on page 151

▼ Display LAG Information

The `showlag` command displays general infromation about all LAGs or detailed information for a specific LAG.

1. On the management console, display general information about all LAGs.

```
# showlag
Name       Iport      Members Vnics Gw Port ID BANDWIDTH Distribution LACP  Mode
---------- ---------- ------- ----- ---------- --------- ------------ ----------
test       Bridge-0-1 2       0     112        20        Default      Off
#
```

2. Display detailed information about a specific LAG.

```
showlag lagname
```

where *lagname* is the identifier of the LAG

For example

```
# showlag test
Name       enabled active RC
---------- ------- ------ ---
0A-ETH-2   1       1      0x0
0A-ETH-3   1       1      0x0
#
```

Related Information

- *Gateway Reference*, `showlag` command
- "Controlling LAGs" on page 137
- "Display the VLANs" on page 130
- "Display the VNICs" on page 131
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133
- "Display the Gateway Instance Number" on page 135
- "Display Information About the Gateway" on page 136

▼ Display the VLANs

You can use the `showvlan` command to identify which VLANs are associated with the gateway 10GbE connectors.

- **On the management controller, type.**

```
# showvlan
  Connector/LAG  VLN  PKEY
  -------------  ---  ----
  0A-ETH-1         3   ffff
  0A-ETH-1         0   ffff
  0A-ETH-2         3   ffff
  0A-ETH-2         0   ffff
  0A-ETH-3         3   ffff
  0A-ETH-3         0   ffff
  0A-ETH-4         3   ffff
  0A-ETH-4         0   ffff
  1A-ETH-1         3   ffff
  1A-ETH-1         0   ffff
  1A-ETH-2         3   ffff
  1A-ETH-2         0   ffff
  1A-ETH-3         3   ffff
  1A-ETH-3         0   ffff
  1A-ETH-4         3   ffff
  1A-ETH-4         0   ffff
#
```

Note – A VLAN identifier (`VLN`) of `0`, `-1`, or `NO` indicates that there is no VLAN assigned.

Note – Your output will differ from that in the example.

The `showvlan` command lists the connectors on the left, and the associated VLANs and respective partition keys on the right.

Related Information

- *Gateway Reference*, `showvlan` command
- "Controlling VLANs and VNICs" on page 142
- "Display LAG Information" on page 129
- "Display the VNICs" on page 131
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133
- "Display the Gateway Instance Number" on page 135
- "Display Information About the Gateway" on page 136

▼ Display the VNICs

The `showvnics` command displays information about and status of the VNICs.

- **On the management controller, type.**

```
# showvnics
ID  STATE    FLG IOA_GUID             NODE      IID  MAC             VLN PKEY   GW
--- --------  --- ----------------------- ---------- ---- ----------------- ---
----   --------
 16 UP          N 00:21:28:00:01:3E:CE:A0 mnm34-55   0000 02:02:02:02:02:04 3
ffff 1A-ETH-4
 13 UP          N 00:03:BA:00:01:00:E3:71 mnm34-60   0000 02:02:02:02:02:01 3
ffff 1A-ETH-3
 15 UP          N 00:21:28:00:01:3E:CE:9F mnm34-55   0000 02:02:02:02:02:03 3
ffff 1A-ETH-1
 14 UP          N 00:03:BA:00:01:00:E3:72 mnm34-60   0000 02:02:02:02:02:02 3
ffff 1A-ETH-2
#
```

Note – Your output will differ from that in the example.

For each VNIC, the `showvnics` command displays the VNIC ID, its state, any flags, its associated host GUID and node description, its MAC address, its associated VLAN ID, its partition key, and connector.

Related Information

- *Gateway Reference*, `showvnics` command
- "Controlling VLANs and VNICs" on page 142
- "Display LAG Information" on page 129
- "Display the VLANs" on page 130
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133
- "Display the Gateway Instance Number" on page 135
- "Display Information About the Gateway" on page 136

▼ Display Gateway Ethernet Port Information

You can display Ethernet port information with the `setgwethport` command. The output is similar to that of the `getportstatus` command.

- **On the management controller, type.**

```
# setgwethport connector
```

where *connector* is either `0A-ETH` or `1A-ETH`. For example:

Note – Your output will differ from that of the example.

```
# setgwethport 0A-ETH
Port status for connector 0A-ETH-1:
Adminstate.....................Enabled
State..........................Up
Link state.....................Up
Protocol.......................Ethernet
Link Mode......................XFI
Speed..........................10Gb/s
MTU............................9600
Tx pause.......................Global
Rx pause.......................Global
Port status for connector 0A-ETH-2:
Adminstate.....................Enabled
.
.
.
MTU............................9600
```

```
Tx pause.....................Global
Rx pause.....................Global
#
```

Related Information

- *Gateway Reference*, `setgwethport` command
- *Gateway Reference*, `getportstatus` command
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Display LAG Information" on page 129
- "Display the VLANs" on page 130
- "Display the VNICs" on page 131
- "Display Gateway Port Information" on page 133
- "Display the Gateway Instance Number" on page 135
- "Display Information About the Gateway" on page 136

▼ Display Gateway Port Information

Use the `showgwports` command to display detailed information about the gateway ports.

- **On the management controller, type.**

```
# showgwports -v

BRIDGE DEVICES:
---------------
NodeGUID           NodeDescrition
--------------------------------
0x0021283bad45c000 SUN IB QDR GW switch gw-2 Bridge 0
0x0021283bad45c040 SUN IB QDR GW switch gw-2 Bridge 1

INTERNAL PORTS:
---------------
Device   Port Portname  PeerPort PortGUID          LID    IBState  GWState
Speed  VLs MTU
-------------------------------------------------------------------------------
-----------------
Bridge-0 1  Bridge-0-1   4   0x0021283bad45c001 0x0007 Active  Up       40Gbs
2 4096
Bridge-0 2  Bridge-0-2   3   0x0021283bad45c002 0x0008 Active  Up       40Gbs
2 4096
Bridge-1 1  Bridge-1-1   2   0x0021283bad45c041 0x0009 Active  Up       40Gbs
2 4096
Bridge-1 2  Bridge-1-2   1   0x0021283bad45c042 0x000a Active  Up       40Gbs
2 4096

CONNECTOR 0A-ETH:
-----------------
Port      Bridge      Adminstate Link  State       Linkmode       Speed   MTU
TxPause  RxPause
-------------------------------------------------------------------------------
--------------------
0A-ETH-1  Bridge-0-2  Enabled    Up    Up          XFI            10Gb/s  9600
Global   Global
0A-ETH-2  Bridge-0-2  Enabled    Up    Up          XFI            10Gb/s  9600
Global   Global
0A-ETH-3  Bridge-0-1  Enabled    Up    Up          XFI            10Gb/s  9600
Global   Global
0A-ETH-4  Bridge-0-1  Enabled    Up    Up          XFI            10Gb/s  9600
Global   Global

CONNECTOR 1A-ETH:
-----------------
Port      Bridge      Adminstate Link  State       Linkmode       Speed   MTU
TxPause  RxPause
-------------------------------------------------------------------------------
--------------------
```

```
1A-ETH-1  Bridge-1-2  Enabled    Up    Up          XFI           10Gb/s  9600
Global    Global
1A-ETH-2  Bridge-1-2  Enabled    Up    Up          XFI           10Gb/s  9600
Global    Global
1A-ETH-3  Bridge-1-1  Enabled    Up    Up          XFI           10Gb/s  9600
Global    Global
1A-ETH-4  Bridge-1-1  Enabled    Up    Up          XFI           10Gb/s  9600
Global    Global
#
```

When the `-v` option is used, this information is provided:

- BridgeX devices – The node GUID and node description are provided.
- Internal ports – The device, port, port name, peer port, port GUID, LID, InfiniBand and gateway state, speed and MTU are displayed.
- Ethernet connectors – For each connector, the port, associated bridge, the administrative, link, and physical state, link mode, speed, MTU, and transmit and receive pauses are provided.

Note – Your output will differ from that of the example.

Related Information

- *Gateway Reference*, `showgwports` command
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Display LAG Information" on page 129
- "Display the VLANs" on page 130
- "Display the VNICs" on page 131
- "Display Gateway Ethernet Port Information" on page 132
- "Display the Gateway Instance Number" on page 135
- "Display Information About the Gateway" on page 136

▼ Display the Gateway Instance Number

Gateways use instance numbers to identify themselves to each other.

- **On the management controller, type.**

```
# setgwinstance list
Current GW instance: 34
#
```

Related Information

- *Gateway Reference*, `setgwinstance` command
- "Set the Gateway Instance Number" on page 155
- "Display LAG Information" on page 129
- "Display the VLANs" on page 130
- "Display the VNICs" on page 131
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133
- "Display Information About the Gateway" on page 136

▼ Display Information About the Gateway

You can display version, status, and operational information about the gateway with the `showgwconfig` command.

- **On the management controller, type.**

```
# showgwconfig
BXM (pid 17030) is running
BXM versions: bxm_user 1.3.6-0, BXM-API 1.6.0, bxm_libs 1.3.6-0, bxm_main 1.30
mlx_bx_core 1.30
Parameter        Configured Value     Running Value
-------------------------------------------------
GWInstance       None                 34
SystemName       None                 h2onm2-gw-2
EoIB Data SL     1                    1
EoIB Control SL  1                    1
#
```

Related Information

- *Gateway Reference*, `showgwconfig` command
- "Display LAG Information" on page 129
- "Display the VLANs" on page 130
- "Display the VNICs" on page 131
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133
- "Display the Gateway Instance Number" on page 135

Controlling LAGs

LAGs are a means of effectively increasing the bandwidth of an Ethernet network by grouping ports.

Related Information

- "Display LAG Information" on page 129
- "Controlling VLANs and VNICs" on page 142
- "Controlling Gateway Ports and Parameters" on page 151
- "Controlling the Chassis" on page 30
- "Controlling the I4 Switch Chip" on page 40
- "Controlling the InfiniBand Fabric" on page 66
- "Controlling the Subnet Manager" on page 47
- "Installing Gateway Supportive Software (Linux)" on page 99
- "Creating VNICs Under Gateway Manual Mode (Linux)" on page 103
- "Creating VNICs Under Host Manual Mode (Linux)" on page 107
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118
- "Monitoring Gateway Resources" on page 128

▼ Enable LAG Mode

Note – LAG mode must be enabled on all InfiniBand gateways in the fabric before LAGs are created.

1. **On the management controller, type.**

```
# enablelagmode
VNICs for host nodes with EoIB driver version that do not support LAG will go
down if LAG mode is enabled.
Do you still want to enable LAG(y/n)?y
Stopping Bridge Manager..                                    [  OK  ]
Starting Bridge Manager.                                     [  OK  ]
#
```

2. **Repeat** Step 1 **for all InfiniBand gateways in the fabric.**

Related Information

- *Gateway Reference*, `enablelagmode` command
- "Create LAGs" on page 138
- "Add or Delete Connectors From a LAG" on page 139
- "Delete a LAG" on page 140
- "Disable LAG Mode" on page 141

▼ Create LAGs

Once LAG mode is enabled, you can create LAGs.

1. **Determine which connectors will be grouped into a LAG and what the LAG names will be.**

Note – A connector is assigned to only one LAG at a time.

Note – If a VNIC has been assigned to a connector, that connector cannot be added to a LAG until the respective VNIC has been deleted.

---

2. **On the management controller, create the first LAG.**

```
# createlag lagname connector1 [connector2 [...connectorN]][-n][-lacpmode
active|passive|off]
```

where:

- *lagname* is the identifier of the LAG

- *connector1* is the name of the first connector to be added to the LAG (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).
- *connectorN* is the name of the last connector to be added to the LAG (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).

For example:

```
# createlag 0A12 0A-ETH-1 0A-ETH-2
lag created
#
```

3. Repeat Step 2 for any other LAGs to be created.

Related Information

- *Gateway Reference*, `createlag` command
- "Enable LAG Mode" on page 137
- "Add or Delete Connectors From a LAG" on page 139
- "Delete a LAG" on page 140
- "Disable LAG Mode" on page 141

▼ Add or Delete Connectors From a LAG

You can add more connectors to a LAG to increase its bandwidth. Or, you can remove connectors from a LAG to free them up for other purposes.

Note – A connector is assigned to only one LAG at a time.

Note – If a VNIC has been assigned to a connector, that connector cannot be added to a LAG until the respective VNIC has been deleted.

---

1. To add a connector to a LAG, on the management controller, type.

```
# addlagport lagname connector1 [connector2 [...connectorN]]
```

where:

- *lagname* is the identifier of the LAG
- *connector1* is the name of the first connector to be added to the LAG (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).

- *connectorN* is the name of the last connector to be added to the LAG (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).

For example:

```
# addlagport 0A12 0A-ETH-3
lag updated
#
```

2. Or, to delete a connector from a LAG, type.

```
# dellagport lagname connector1 [connector2 [...connectorN]]
```

where:

- *lagname* is the identifier of the LAG
- *connector1* is the name of the first connector to be removed from the LAG (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).
- *connectorN* is the name of the last connector to be removed from the LAG (`0A-ETH-1` to `0A-ETH-4` and `1A-ETH-1` to `1A-ETH-4`).

For example:

```
# dellagport 0A12 0A-ETH-3
lag updated
#
```

Related Information

- *Gateway Reference*, `dellagport` command
- "Enable LAG Mode" on page 137
- "Create LAGs" on page 138
- "Delete a LAG" on page 140
- "Disable LAG Mode" on page 141

▼ Delete a LAG

When a LAG is no longer needed, or its connectors must be re-purposed, you can delete the LAG.

- **On the management controller, type.**

```
# deletelag lagname
```

where *lagname* is the identifier of the LAG

For example:

```
# deletelag 0A12
LAG 0A12 deleted
#
```

Related Information

- *Gateway Reference*, `deletelag` command
- "Enable LAG Mode" on page 137
- "Create LAGs" on page 138
- "Add or Delete Connectors From a LAG" on page 139
- "Disable LAG Mode" on page 141

▼ Disable LAG Mode

When you disable LAG mode, the LAG configurations remain. Upon re-enabling LAG mode, the LAGs become active again.

Note – LAG mode must be disabled for all InfiniBand gateways in the fabric.

---

1. **On the management controller, type.**

```
# disablelagmode
Stopping Bridge Manager..                                  [  OK  ]
Starting Bridge Manager.                                   [  OK  ]
#
```

2. **Repeat** Step 1 **for all InfiniBand gateways in the fabric.**

Related Information

- *Gateway Reference*, `disablelagmode` command
- "Enable LAG Mode" on page 137
- "Create LAGs" on page 138
- "Add or Delete Connectors From a LAG" on page 139

- "Delete a LAG" on page 140

Controlling VLANs and VNICs

These topics describe how you can control the gateway ports and parameters:

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Gateway_Mgmt` or `/SYS/Fabric_Mgmt` Linux shell targets of the Oracle ILOM CLI..

- "Create VLANs" on page 143
- "Delete VLANs" on page 144
- "Create VNICs" on page 145
- "Enable VNICs" on page 148
- "Disable VNICs" on page 149
- "Delete VNICs" on page 150

Related Information

- "Controlling LAGs" on page 137
- "Controlling Gateway Ports and Parameters" on page 151
- "Controlling the Chassis" on page 30
- "Controlling the I4 Switch Chip" on page 40
- "Controlling the InfiniBand Fabric" on page 66
- "Controlling the Subnet Manager" on page 47
- "Installing Gateway Supportive Software (Linux)" on page 99
- "Creating VNICs Under Gateway Manual Mode (Linux)" on page 103
- "Creating VNICs Under Host Manual Mode (Linux)" on page 107
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118
- "Monitoring Gateway Resources" on page 128

▼ Create VLANs

When you create a VLAN, you create a mapping between a gateway connector, a VLAN identifier, and a partition key. You can create VLANs with the `createvlan` command. If a VNIC will be associated to a VLAN, you must create the VLAN before the VNIC.

Note – Due to hardware limitations for MultiCast groups, there is a maximum of 1000 VLANs.

---

1. **If unknown, determine the connectors, VNICs, and partition keys to be associated with the VLANs.**

   See "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105.

2. **(Optional) Open the gateway MAC address file that you created earlier.**

   See "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105.

3. **Create the VLAN on the management controller.**

```
# createvlan connector|lagname -vlan vlan_ID -pkey p_key
```

   where:

   - *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`). This value is column one of the gateway MAC address file.
   - *lagname* is the identifier of the LAG.
   - *vlan_ID* is the VLAN identifier **(**`NO`**, **`0`**, or** `2` **-** `4094`**)**. This value is column four of the gateway MAC address file.

Note – Do not use VLAN identifiers `1` or `4095`.

   - *p_key* is the partition identifier (`1` – `7fff` or `default`). This value is column five of the gateway MAC address file.

   For example, to associate connector `1A-ETH-3` with VLAN `3` and default partition key:

```
# createvlan 1A-ETH-3 -vlan 3 -pkey default
#
```

4. **Repeat** Step 3 **for all VLANs to be associated with connectors.**

5. **Verify the VLAN configuration.**

```
# showvlan
  Connector/LAG  VLN  PKEY
  -------------  ---  ----
  0A-ETH-1         0   ffff
  0A-ETH-2         0   ffff
  0A-ETH-3         0   ffff
  0A-ETH-4         0   ffff
  1A-ETH-1         3   ffff
  1A-ETH-1         0   ffff
  1A-ETH-2         3   ffff
  1A-ETH-2         0   ffff
  1A-ETH-3         3   ffff
  1A-ETH-3         0   ffff
  1A-ETH-4         3   ffff
  1A-ETH-4         0   ffff
#
```

Related Information

- *Gateway Reference*, `createvlan` command
- *Gateway Reference*, `showvlan` command
- "Delete VLANs" on page 144
- "Create VNICs" on page 145
- "Enable VNICs" on page 148
- "Disable VNICs" on page 149
- "Delete VNICs" on page 150
- "Display the VLANs" on page 130

▼ Delete VLANs

You can delete VLANs with the `deletevlan` command.

1. **Delete all VNICs associated with the connector for the VLAN you are deleting.**
   See "Delete VNICs" on page 150.
2. **On the management controller, type.**

```
# deletevlan lagname|connector -vlan vlan_ID
```

   where:

   - *lagname* is the identifier of the LAG.

- *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`).
- *vlan_ID* is the VLAN identifier.

For example, to delete the association of connector `1A-ETH-1` to VLAN `3`, type.

```
# deletevlan 1A-ETH-1 -vlan 3
#
```

3. **Repeat from** Step 1 **for all VLANs to be deleted.**
4. **Verify that the VLANs have been deleted.**

```
# showvlan
  Connector/LAG  VLN  PKEY
  -------------  ---  ----
  0A-ETH-1         0   ffff
  0A-ETH-2         0   ffff
  0A-ETH-3         0   ffff
  0A-ETH-4         0   ffff
  1A-ETH-1         0   ffff
  1A-ETH-2         0   ffff
  1A-ETH-3         0   ffff
  1A-ETH-4         0   ffff
#
```

Related Information

- *Gateway Reference*, `deletevlan` command
- *Gateway Reference*, `showvlan` command
- "Create VLANs" on page 143
- "Create VNICs" on page 145
- "Enable VNICs" on page 148
- "Disable VNICs" on page 149
- "Delete VNICs" on page 150
- "Display the VLANs" on page 130

▼ Create VNICs

If you create a VNIC that uses the `default` partition key (PKey), you do not need to set up a VLAN for the VNIC. However, if you want to create a VNIC that uses a PKey other than the `default`, you must create the VLAN first.

1. **If unknown, determine the HCA port GUIDs, MAC addresses, the partition keys, and the connectors that will be assigned to the VNICs.**

   See "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104.

2. **If a VLAN will be associated with the VNICs, create the VLANs before the VNICs.**

   See "Create VLANs" on page 143.

3. **(Optional) Open the gateway MAC address file you created earlier.**

   See "Determine VNIC Configuration Parameters For Gateway Manual Mode (Linux)" on page 104 or "Determine VLAN Associations for Gateway Manual Mode (Linux)" on page 105.

4. **Create the VNIC on the management controller.**

```
# createvnic connector|lagname <-guid guid>|<-host hostname -port
port>|<-node string -port port> -mac mac -vlan vlan_ID -pkey p_key
```

   where:

   - *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`). This value is column one of the gateway MAC address file.
   - *lagname* is the identifier of the LAG.
   - *guid* is the global unique identifier of the target port on the host associated with the VNIC. This value is column two of the gateway MAC address file.
   - *hostname* is the host name of the management controller or system name recognized by the BridgeX manager.
   - *port* is the port of the host.
   - *string* is the character string in the node description field.
   - *mac* is the machine address code. This value is column three of the gateway MAC address file.

Note – Only even numbers are supported for the most significant byte of the MAC address (unicast).

   - *vlan_ID* is the VLAN identifier. This value is column four of the gateway MAC address file.

- *p_key* is the partition identifier (`default`). This value is column five of the gateway MAC address file.

For example, to create a persistent VNIC for connector 1A-ETH-3 associated with VLAN 3 with partition key `default`, type.

```
# createvnic 1A-ETH-3 -guid 00:03:BA:00:01:00:E3:71 -mac 02:02:02:02:02:01 -vlan
3 -pkey default
#
```

Note – If the VNIC is to be non-persistent, append the `-n` option to the command line.

Note – If no VLAN will be associated with the VNIC, do not use the `-vlan` option.

5. **Repeat** Step 4 **for all VNICs to be created.**

6. **Verify the status and configuration of the VNICs.**

```
# showvnics
ID  STATE    FLG IOA_GUID             NODE      IID  MAC             VLN PKEY   GW
--- -------- --- ---------------------- ---------- ---- ----------------- ---
----   --------
 16 UP         N 00:21:28:00:01:3E:CE:A0 mnm34-55   0000 02:02:02:02:02:04 3
ffff 1A-ETH-4
 13 UP         N 00:03:BA:00:01:00:E3:71 mnm34-60   0000 02:02:02:02:02:01 3
ffff 1A-ETH-3
 15 UP         N 00:21:28:00:01:3E:CE:9F mnm34-55   0000 02:02:02:02:02:03 3
ffff 1A-ETH-1
 14 UP         N 00:03:BA:00:01:00:E3:72 mnm34-60   0000 02:02:02:02:02:02 3
ffff 1A-ETH-2
#
```

Related Information

- *Gateway Reference*, `createvnic` command
- *Gateway Reference*, `showvnics` command
- "Create VLANs" on page 143
- "Delete VLANs" on page 144
- "Enable VNICs" on page 148
- "Disable VNICs" on page 149
- "Delete VNICs" on page 150
- "Display the VNICs" on page 131

▼ Enable VNICs

VNICs are automatically enabled when they are created. You can use this procedure to enable a previously disabled VNIC.

1. **Identify the VNIC you want to enable.**

   See "Display the VNICs" on page 131.

2. **Enable the VNIC on the management controller.**

```
# enablevnic connector VNIC_ID
```

   where:

   - *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`).
   - *VNIC_ID* is the VNIC identifier.

   For example, to enable VNIC `15` on connector `1A-ETH-1`, type.

```
# enablevnic 1A-ETH-1 15
#
```

3. **Repeat from Step 1 for all VNICs to be enabled.**

4. **Verify that the VNICs have been enabled.**

```
# showvnics
ID  STATE    FLG IOA_GUID             NODE      IID  MAC             VLN PKEY   GW
--- --------  --- ---------------------- ---------- ---- ----------------- ---
----   --------
 16 UP          N 00:21:28:00:01:3E:CE:A0 mnm34-55   0000 02:02:02:02:02:04 3
ffff 1A-ETH-4
 13 UP          N 00:03:BA:00:01:00:E3:71 mnm34-60   0000 02:02:02:02:02:01 3
ffff 1A-ETH-3
 15 UP          N 00:21:28:00:01:3E:CE:9F mnm34-55   0000 02:02:02:02:02:03 3
ffff 1A-ETH-1
 14 UP          N 00:03:BA:00:01:00:E3:72 mnm34-60   0000 02:02:02:02:02:02 3
ffff 1A-ETH-2
#
```

Related Information

- *Gateway Reference*, `enablevnic` command
- *Gateway Reference*, `showvnics` command
- "Create VLANs" on page 143

- "Delete VLANs" on page 144
- "Create VNICs" on page 145
- "Disable VNICs" on page 149
- "Delete VNICs" on page 150
- "Display the VNICs" on page 131

▼ Disable VNICs

If you believe that there is a problem with a VNIC, you can disable it.

1. **Identify the VNIC you want to disable.**

   See "Display the VNICs" on page 131.

2. **Disable the VNIC on the management controller.**

```
# disablevnic connector VNIC_ID
```

   where:

   - *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`).
   - *VNIC_ID* is the VNIC identifier.

   For example, to disable VNIC `15` on connector `1A-ETH-1`, type.

```
# disablevnic 1A-ETH-1 15
#
```

3. **Repeat from Step 1 for all VNICs to be disabled.**

4. **Verify that the VNICs have been disabled.**

```
# showvnics
ID  STATE    FLG IOA_GUID             NODE      IID  MAC             VLN PKEY   GW
--- -------- --- -------------------- ---------- ---- ----------------- --- ----
----   --------
 16 UP          N 00:21:28:00:01:3E:CE:A0 mnm34-55   0000 02:02:02:02:02:04 3
ffff 1A-ETH-4
 13 UP          N 00:03:BA:00:01:00:E3:71 mnm34-60   0000 02:02:02:02:02:01 3
ffff 1A-ETH-3
```

```
 15 DISABLED    N 00:21:28:00:01:3E:CE:9F mnm34-55   0000 02:02:02:02:02:88 3
ffff 1A-ETH-1
 14 UP          N 00:03:BA:00:01:00:E3:72 mnm34-60   0000 02:02:02:02:02:02 3
ffff 1A-ETH-2
#
```

Related Information

- *Gateway Reference*, `disablevnic` command
- *Gateway Reference*, `showvnics` command
- "Create VLANs" on page 143
- "Delete VLANs" on page 144
- "Create VNICs" on page 145
- "Enable VNICs" on page 148
- "Delete VNICs" on page 150
- "Display the VNICs" on page 131

▼ Delete VNICs

Before you can delete a VLAN, you must delete the VNIC associated with that VLAN and connector.

1. **Identify the VNIC you want to delete.**

   See "Display the VNICs" on page 131.

2. **Delete the VNIC on the management controller.**

```
# deletevnic connector VNIC_ID
```

   where:

   - *connector* is the name of the connector (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`).
   - *VNIC_ID* is the VNIC identifier.

   For example, to delete VNIC `15` on connector `1A-ETH-1`, type.

```
# deletevnic 1A-ETH-1 15
#
```

3. **Repeat from Step 1 for all VNICs to be deleted.**

4. **Verify that the VNICs have been deleted.**

```
# showvnics
ID  STATE    FLG IOA_GUID              NODE      IID  MAC             VLN PKEY   GW
--- --------  --- ----------------------- ---------- ---- ----------------- ---
----   --------
 16 UP          N 00:21:28:00:01:3E:CE:A0 mnm34-55   0000 02:02:02:02:02:04 3
ffff 1A-ETH-4
 13 UP          N 00:03:BA:00:01:00:E3:71 mnm34-60   0000 02:02:02:02:02:01 3
ffff 1A-ETH-3
 14 UP          N 00:03:BA:00:01:00:E3:72 mnm34-60   0000 02:02:02:02:02:02 3
ffff 1A-ETH-2
#
```

5. **If you deleted the VNICs in order to delete VLANs, delete the VLANs.**

   See "Delete VLANs" on page 144.

Related Information

- *Gateway Reference*, `deletevnic` command
- *Gateway Reference*, `showvnics` command
- "Create VLANs" on page 143
- "Delete VLANs" on page 144
- "Create VNICs" on page 145
- "Enable VNICs" on page 148
- "Disable VNICs" on page 149
- "Display the VNICs" on page 131

Controlling Gateway Ports and Parameters

These topics describe how you can control the gateway ports and parameters:

Note – To use the commands described in these topics, you must either be the `root` user of the management controller or access them through the `/SYS/Gateway_Mgmt` or `/SYS/Fabric_Mgmt` Linux shell targets of the Oracle ILOM CLI.

- "Enable a Gateway Ethernet Port" on page 152

- "Disable a Gateway Ethernet Port" on page 153
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Set the Gateway Service Level" on page 155
- "Set the Gateway Instance Number" on page 155
- "Set the Gateway System Name" on page 156

Related Information

- "Monitoring Gateway Resources" on page 128
- "Controlling LAGs" on page 137
- "Controlling VLANs and VNICs" on page 142
- "Controlling the Chassis" on page 30
- "Controlling the I4 Switch Chip" on page 40
- "Controlling the InfiniBand Fabric" on page 66
- "Controlling the Subnet Manager" on page 47
- "Installing Gateway Supportive Software (Linux)" on page 99
- "Creating VNICs Under Gateway Manual Mode (Linux)" on page 103
- "Creating VNICs Under Host Manual Mode (Linux)" on page 107
- "Creating Virtual IO Adapters (Oracle Solaris)" on page 118

▼ Enable a Gateway Ethernet Port

- **On the management controller, type.**

```
# enablegwport connector
```

where *connector* is the name of the connector hosting the port (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`).

For example, to enable the port associated with connector 0A-ETH-3, type.

```
# enablegwport 0A-ETH-3
#
```

Related Information

- *Gateway Reference*, `enablegwport` command
- "Disable a Gateway Ethernet Port" on page 153
- "Set the Gateway Ethernet Port Parameters" on page 153

- "Set the Gateway Service Level" on page 155
- "Set the Gateway Instance Number" on page 155
- "Set the Gateway System Name" on page 156

▼ Disable a Gateway Ethernet Port

You can disable an enabled gateway Ethernet port with the `disablegwport` command.

- **On the management controller, type.**

```
# disablegwport connector
```

where *connector* is the name of the connector hosting the port (`0A-ETH-1` to `0A-ETH-4`, `0A-ETH`, `1A-ETH-1` to `1A-ETH-4`, and `1A-ETH`).

For example, to disable the port associated with connector 0A-ETH-3, type.

```
# disablegwport 0A-ETH-3
#
```

Related Information

- *Gateway Reference*, `disablegwport` command
- "Enable a Gateway Ethernet Port" on page 152
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Set the Gateway Service Level" on page 155
- "Set the Gateway Instance Number" on page 155
- "Set the Gateway System Name" on page 156

▼ Set the Gateway Ethernet Port Parameters

Use the `setgwethport` command to configure the gateway Ethernet ports.

- **On the management controller, type.**

```
# setgwethport connector [-linkmode linkmode][-mtu mtu][-txpause
pause][-rxpause pause]
```

where:

- *connector* is the name of the connector (`0A-ETH` or `1A-ETH`).

- *linkmode* is the is the link mode, `XFI` (10Gb/s).
- *mtu* is the value of the MTU.
- *pause* is the priority flow control pause in quanta (`0– 255`, `Global`, or `None`)

For example, to set the speed of connector `1A-ETH` to 10 GB/s, type.

```
# setgwethport 1A-ETH -linkmode XFI
Port status for connector 0A-ETH-1:
Adminstate.....................Enabled
State..........................Up
Link state.....................Up
Protocol.......................Ethernet
Link Mode......................XFI
Speed..........................10Gb/s
MTU............................9600
Tx pause.......................Global
Rx pause.......................Global
Port status for connector 1A-ETH-2:
.
.
.
Port status for connector 1A-ETH-4:
Adminstate.....................Enabled
State..........................Up
Link state.....................Up
Protocol.......................Ethernet
Link Mode......................XFI
Speed..........................10Gb/s
MTU............................9600
Tx pause.......................Global
Rx pause.......................Global
#
```

Related Information

- *Gateway Reference*, `setgwethport` command
- "Enable a Gateway Ethernet Port" on page 152
- "Disable a Gateway Ethernet Port" on page 153
- "Set the Gateway Service Level" on page 155
- "Set the Gateway Instance Number" on page 155
- "Set the Gateway System Name" on page 156
- "Display Gateway Ethernet Port Information" on page 132
- "Display Gateway Port Information" on page 133

▼ Set the Gateway Service Level

You can set the Ethernet data traffic and control traffic service levels with the `setgwsl` command.

- **On the management controller, type.**

```
# setgwsl eoib|ctrl level
```

where *level* is the service level. For example, to set the data traffic service level to 2, type.

```
# setgwsl eoib 2
Stopping Bridge Manager..                                  [  OK  ]
Starting Bridge Manager.                                   [  OK  ]
#
```

Related Information

- *Gateway Reference*, `setgwsl` command
- "Enable a Gateway Ethernet Port" on page 152
- "Disable a Gateway Ethernet Port" on page 153
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Set the Gateway Instance Number" on page 155
- "Set the Gateway System Name" on page 156
- "Display Information About the Gateway" on page 136

▼ Set the Gateway Instance Number

By default, gateway instance numbers are set according the 6 lowest bits of the gateway IP address, and must be unique for VNICs to be operational. When two or more gateways share the same InfiniBand fabric, they might have duplicate gateway instance numbers. In this situation, you must assign unique instance numbers to all of the gateways in the fabric.

1. **On each gateway's management controller, display the gateway's instance number.**

```
# setgwinstance list
```

2. **Determine which gateway will receive a new instance number and what that number will be.**

   The gateway instance is an number from 0 to 63.

3. **On that gateway's management controller, type.**

```
# setgwinstance instance
```

   where *instance* is an number from `0` to `63`. For example, to the set the instance to `3`:

```
# setgwinstance 3
Stopping Bridge Manager..                                  [  OK  ]
Starting Bridge Manager.                                   [  OK  ]
#
```

Related Information

- *Gateway Reference*, `setgwinstance` command
- "Enable a Gateway Ethernet Port" on page 152
- "Disable a Gateway Ethernet Port" on page 153
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Set the Gateway Service Level" on page 155
- "Set the Gateway System Name" on page 156
- "Display the Gateway Instance Number" on page 135

▼ Set the Gateway System Name

The BridgeX manager within the management controller recognizes a system name that is independent of the management controller's host name. If no system name is configured, the host name is used. The system name is a maximum of 19 characters.

1. **On the management controller, type.**

```
# setgwsystemname systemname
```

   where *systemname* is the BridgeX manager's system name. For example:

```
# setgwsystemname brooklyn
Stopping Bridge Manager..                                  [  OK  ]
Starting Bridge Manager.                                   [  OK  ]
```

```
# setgwsystemname --list
BXM system name set to --myname
#
```

2. **Use the `--list` option of the `setgwsystemname` command to verify the name configuration.**

```
# setgwsystemname --list
BXM system name set to brooklyn
#
```

Related Information

- *Gateway Reference*, `setgwsystemname` command
- "Enable a Gateway Ethernet Port" on page 152
- "Disable a Gateway Ethernet Port" on page 153
- "Set the Gateway Ethernet Port Parameters" on page 153
- "Set the Gateway Service Level" on page 155
- "Set the Gateway Instance Number" on page 155
- "Display Information About the Gateway" on page 136

Index

D

E

F

G

H

I

L

M

N

O

P

Q

R

S

T

V