Sun™ Dual Port 4x IB Host Channel Adapter PCI-E Installation and User's Guide

Sun Microsystems, Inc.
www.sun.com

Part No. 819-4190-10
April 2006, Revision A

Submit comments about this document at: http://www.sun.com/hwdocs/feedback

Regulatory Compliance Statements

Your Sun product is marked to indicate its compliance class:

- Federal Communications Commission (FCC) — USA
- Industry Canada Equipment Standard for Digital Equipment (ICES-003) — Canada
- Voluntary Control Council for Interference (VCCI) — Japan
- Bureau of Standards Metrology and Inspection (BSMI) — Taiwan

Please read the appropriate section that corresponds to the marking on your Sun product before attempting to install the product.

FCC Class A Notice

This device complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions:

1. This device may not cause harmful interference.
2. This device must accept any interference received, including interference that may cause undesired operation.

Note: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy, and if it is not installed and used in accordance with the instruction manual, it may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference, in which case the user will be required to correct the interference at his own expense.

Modifications: Any modifications made to this device that are not approved by Sun Microsystems, Inc. may void the authority granted to the user by the FCC to operate this equipment.

FCC Class B Notice

This device complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions:

1. This device may not cause harmful interference.
2. This device must accept any interference received, including interference that may cause undesired operation.

Note: This equipment has been tested and found to comply with the limits for a Class B digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference in a residential installation. This equipment generates, uses and can radiate radio frequency energy and, if not installed and used in accordance with the instructions, may cause harmful interference to radio communications. However, there is no guarantee that interference will not occur in a particular installation. If this equipment does cause harmful interference to radio or television reception, which can be determined by turning the equipment off and on, the user is encouraged to try to correct the interference by one or more of the following measures:

- Reorient or relocate the receiving antenna.
- Increase the separation between the equipment and receiver.
- Connect the equipment into an outlet on a circuit different from that to which the receiver is connected.
- Consult the dealer or an experienced radio/television technician for help.

Modifications: Any modifications made to this device that are not approved by Sun Microsystems, Inc. may void the authority granted to the user by the FCC to operate this equipment.

ICES-003 Class A Notice - Avis NMB-003, Classe A

This Class A digital apparatus complies with Canadian ICES-003.

Cet appareil numérique de la classe A est conforme à la norme NMB-003 du Canada.

ICES-003 Class B Notice - Avis NMB-003, Classe B

This Class B digital apparatus complies with Canadian ICES-003.

Cet appareil numérique de la classe B est conforme à la norme NMB-003 du Canada.

VCCI 基準について

クラス A VCCI 基準について

クラス A VCCI の表示があるワークステーションおよびオプション製品は、クラス A 情報技術装置です。これらの製品には、下記の項目が該当します。

この装置は、情報処理装置等電波障害自主規制協議会 (VCCI) の基準に基づくクラス A 情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

クラス B VCCI 基準について

クラス B VCCI の表示  があるワークステーションおよびオプション製品は、クラス B 情報技術装置です。これらの製品には、下記の項目が該当します。

この装置は、情報処理装置等電波障害自主規制協議会 (VCCI) の基準に基づくクラス B 情報技術装置です。この装置は、家庭環境で使用することを目的としていますが、この装置がラジオやテレビジョン受信機に近接して使用されると、受信障害を引き起こすことがあります。取扱説明書に従って正しい取り扱いをしてください。

BSMI Class A Notice

The following statement is applicable to products shipped to Taiwan and marked as Class A on the product compliance label.

警告使用者：
這是甲類的資訊產品，在居住的環境中使用時，可能會造成射頻干擾，在這種情況下，使用者會被要求採取某些適當的對策。

CCC Class A Notice

The following statement is applicable to products shipped to China and marked with “Class A” on the product’s compliance label.

以下声明适用于运往中国且其认证标志上注有 "Class A" 字样的产品。

声明
此为A级产品，在生活环境中，该产品可能会造成无线电干扰。
在这种情况下，可能需要用户对其干扰采取切实可行的措施。

GOST-R Certification Mark

Declaration of Conformity

Compliance Model Number: **MHEL-CF256-T**
Product Family Name: **Sun Dual Port 4x IB Host Channel Adapter (HCA) PCI-E**

EMC

USA—FCC Class B

This equipment complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions:

1. This equipment may not cause harmful interference.
2. This equipment must accept any interference that may cause undesired operation.

European Union

This equipment complies with the following requirements of the EMC Directive 89/336/EEC:

As Information Technology Equipment (ITE) Class B per (as applicable):

EN 55022:1994 +A1:1995 +A2:1997	Class B
EN 61000-3-2:2000	Pass
EN 61000-3-3:1995 +A1:2001	Pass

EN 55024:1998 +A1:2001 +A2:2003 Required Limits:

IEC 61000-4-2	4 kV (Direct), 8kV (Air)
IEC 61000-4-3	3 V/m
IEC 61000-4-4	1 kV AC Power Lines, 0.5 kV Signal and DC Power Lines
IEC 61000-4-5	1 kV AC Line-Line and Outdoor Signal Lines, 2 kV AC Line-Gnd, 0.5 kV DC Power Lines
IEC 61000-4-6	3 V
IEC 61000-4-8	1 A/m
IEC 61000-4-11	Pass

Safety

This equipment complies with the following requirements of the Low Voltage Directive 73/23/EEC:

EC Type Examination Certificates:

IEC/EN 60950-1 Test Report Numbwe: MELSAF_EN.16591_rev1

Supplementary Information

This equipment was tested and complies with all the requirements for the CE Mark. This equipment complies with the Restriction of Hazardous Substances (RoHS) directive 2002/95/EC.

/S/
Dennis P. Symanski DATE
Worldwide Compliance Engineering
Sun Microsystems, Inc.
4150 Network Circle, MPK15-102
Santa Clara, CA 95054 U.S.A.
Tel: 650-786-3255
Fax: 650-786-3723

/S/
Donald Cameron DATE
Program Manager/Customer Quality
Sun Microsystems Scotland, Limited
Blackness Road, Phase I, Main Bldg.
Springfield, EH49 7LR
Scotland, United Kingdom
Tel: +44 1 506 672 539 Fax: +44 1 506 670 011

Contents

Figures

Tables

Preface

The *Sun Dual Port 4x IB Host Channel Adapter PCI-E Installation and User's Guide* provides installation instructions for the Sun™ Dual Port 4x IB Host Channel Adapter PCI-E card. This manual includes an overview of the adapter, supported platforms, cable information, and related documentation.

These instructions are designed for system administrators with experience installing network hardware and software.

Note – The Sun Dual Port 4x IB Host Channel Adapter PCI-E card is an InfiniBand (IB) host channel adapter (HCA) card. The adapter is referred to as *IB-HCA* in this manual.

How This Book Is Organized

The manual is organized as follows:

Chapter 1 provides a description of the adapter.

Chapter 2 details how to install the adapter card and cables, and how to verify the installation.

Chapter 3 contains information for configuring Internet Protocol over InfiniBand (IPoIB).

Chapter 4 gives an overview of the InfiniBand software.

Appendix A contains information on the InfiniBand cable and connector pin outs.

Appendix B provides an overview of the SunVTS™ diagnostic software and the `ibhcatest` test. The appendix also includes some troubleshooting tasks.

Using UNIX Commands

This document might not contain information about basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices. Refer to the following for this information:

- Software documentation that you received with your system
- Solaris™ Operating System documentation, which is at:

http://docs.sun.com

Shell Prompts

Shell	Prompt
C shell	*machine_name*%
C shell superuser	*machine_name*#
Bourne shell and Korn shell	$
Bourne shell and Korn shell superuser	#

Typographic Conventions

Typeface [1]	Meaning	Examples
`AaBbCc123`	The names of commands, files, and directories; on-screen computer output	Edit your `.login` file. Use `ls -a` to list all files. `% You have mail.`
`AaBbCc123`	What you type, when contrasted with on-screen computer output	`% su` `Password:`
AaBbCc123	Book titles, new words or terms, words to be emphasized	Read Chapter 6 in the *User's Guide*. These are called *class* options. You *must* be superuser to do this.
	Command-line variable; replace with a real name or value	To delete a file, type `rm` *filename*.

1 The settings on your browser might differ from these settings.

Related Documentation

The following documents provide additional information regarding the IB-HCA and related products. You can view or print Sun documentation at the following web site:

http://www.sun.com/documentation

Sun Dual Port 4x IB Host Channel Adapter PCI-E Release Notes (819-4191)

System Administration Guide: Devices and Files System (817-5093)

System Administration Guide: Network Services (816-4555)

Sun IB Switch 9P Hardware Installation Guide (816-0504)

Sun IB Switch 9P Administration Guide (819-0502)

Sun IB Switch 9P Release Notes (819-1278)

Documentation, Support, and Training

Sun Function	URL
Documentation	http://www.sun.com/documentation/
Support	http://www.sun.com/support/
Training	http://www.sun.com/training/

Third-Party Web Sites

Sun is not responsible for the availability of third-party web sites mentioned in this document. Sun does not endorse and is not responsible or liable for any content, advertising, products, or other materials that are available on or through such sites or resources. Sun will not be responsible or liable for any actual or alleged damage or loss caused by or in connection with the use of or reliance on any such content, goods, or services that are available on or through such sites or resources.

PCI-Express specifications can be found at this web site:

http://www.pcisig.com

Contacting Sun Technical Support

If you have technical questions about this product that are not answered in this document, go to:

http://www.sun.com/service/contacting

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. You can submit your comments by going to:

http://www.sun.com/hwdocs/feedback

Please include the title and part number of your document with your feedback:

Sun Dual Port 4x IB Host Channel Adapter PCI-E Installation and User's Guide, part number 819-4190-10

CHAPTER 1

Product Overview

This chapter provides an overview of the Sun Dual Port 4x IB Host Channel Adapter PCI-E card and includes the following sections:

Product Features

Sun Dual Port 4x IB Host Channel Adapter PCI-E card (X1236A-Z) features include:

- Dual 10 Gbyte/sec (4x) InfiniBand ports
- Integrated serializer/deserializer (SerDes)
- PCI-Express Revision 1.0a-compatible
- 8x (20 Gbyte/sec) PCI-Express (8 lanes)
- InfiniBand Trade Association (IBTA) version 1.1 interoperability
- 256 MByte of fixed local DDR memory
- Dual copper connectors (MicroGigaCN) with media detect circuit

FIGURE 1-1 shows the IB-HCA card.

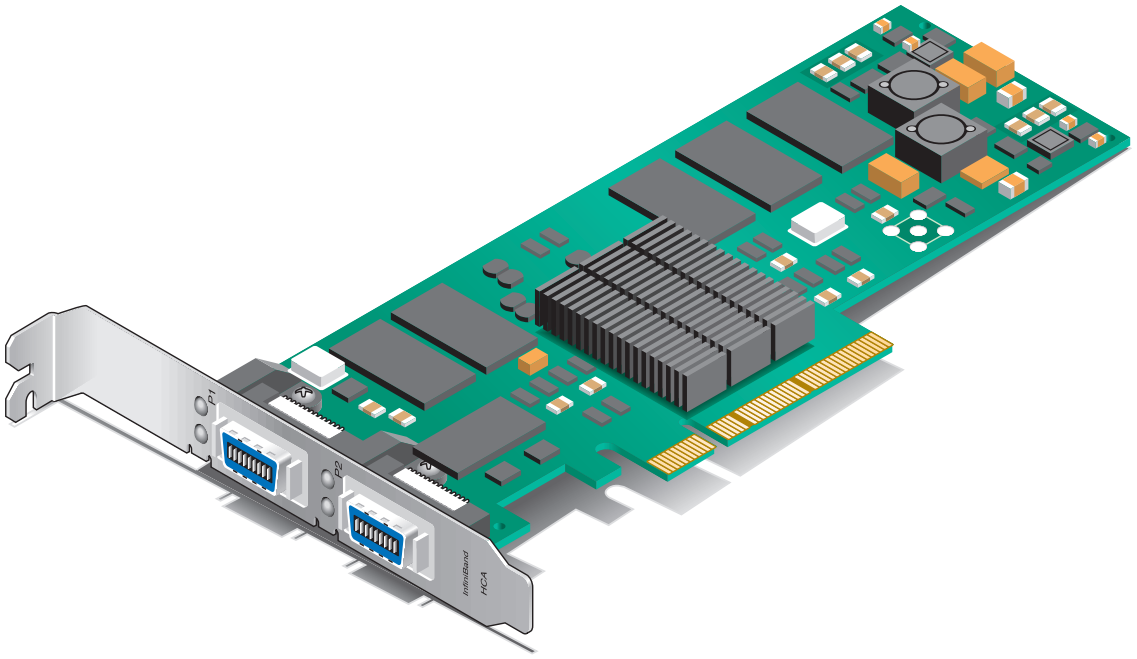

FIGURE 1-1 Sun Dual Port 4x IB Host Channel Adapter PCI-E Card

Hardware Description

The Sun Dual Port 4x IB Host Channel Adapter PCI-E card is an InfiniBand (IB) host channel adapter (HCA) card and is referred to as the IB-HCA in this manual. This adapter is a PCI-Express (Revision 1.0a), x8 compatible card with dual 4x MicroGigaCN InfiniBand-compliant connectors for copper cables. TABLE 1-1 lists the specifications for the IB-HCA card.

The board characteristics are:

- Power – approx. 10.3 watts total
- Low profile PCI form-factor card (approx. 2.5 in. by approx. 6.6 in.)
- Requires 4x InfiniBand cables (not included)
- 12V, 3.3V PCI-E x8 slot
- Link status indication LED
- Two 4x InfiniBand (IB) ports
- Two faceplate bracket sizes – standard and low profile (included)

The standard PCI-E interface is used for data traffic, management traffic, and to generate and receive inbound configuration packets.

TABLE 1-1 Board Specifications

Specification	Description	
Physical	Size	2.5 in X 6.6 in
	Air Flow	200LFM @ 55C
	10Gb/s Connector	InfiniBand Dual Copper MicroGigaCN
Power and environmental	Voltage	12V, 3.3V
	Max Power	10.3W
	Temperature	0 to 55 Celsius
Protocol support	InfiniBand	Autonegotiation 10Gb/s, 2.5 Gb/s
	QoS	8 InfiniBand virtual lanes for all ports
	RDMA Support	Yes, all ports

I/O Interfaces

The IB-HCA includes the following interfaces:

- Two 4x InfiniBand copper connectors
- PCI-Express 8-lane edge connector

InfiniBand Interface

The IB-HCA provides two 4x InfiniBand v1.1 connector (or ports) for external copper cables and is IBTA Specification 1.1-compliant.

PCI-E Interface

The PCI-E bus is a PCI-Express x8 version 1.0a-compliant interface. The adapter can be either a master initiating the PCI-Express bus operations or a slave responding to PCI bus operations. The PCI-E bus can connect to a host CPU in an HCA application.

LED Displays

Two InfiniBand ports connectors are located on the front panel of the adapter. Each port has two LEDs, as shown in FIGURE 1-2. An explanation of the LEDs is given in TABLE 1-2.

TABLE 1-2 IB-HCA LEDs

Color	Meaning
Green	Physical Link – The Physical Link LED illuminates when the physical connection is made.
Amber	Logical Link – The Logical Link LED illuminates once the driver is successfully attached to the port. This LED also blinks when there is traffic on the port.

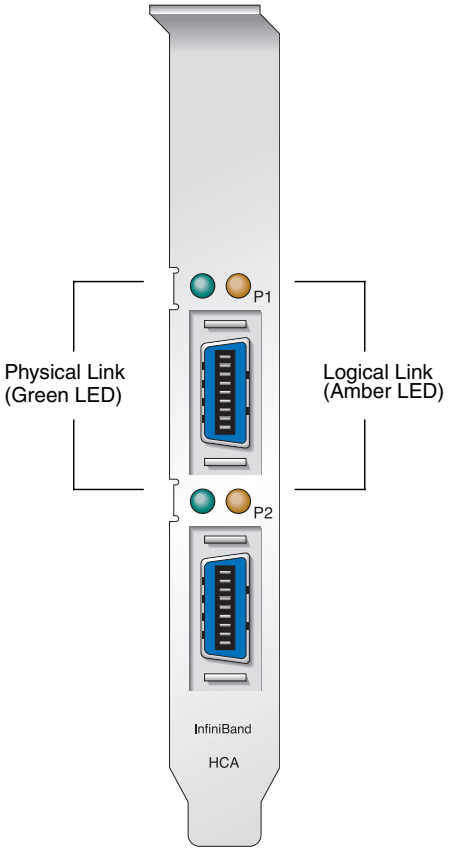

FIGURE 1-2 LED Layout

Hardware, Software, and Cable Requirements

The hardware, software, and cable requirements for the Sun Dual Port 4x IB Host Channel Adapter PCI-E card are listed in TABLE 1-3.

TABLE 1-3 Hardware, Software, and Cable Requirements

Operating System	**Servers**
Solaris 10 1/06 and newer	Sun Fire™ T1000, Sun Fire T2000, Sun Fire X2100
Diagnostics Version	
	SunVTS Version 6.0 Patch Set 1 and subsequent compatible releases
InfiniBand Switch [1]	
X3152A	Sun IB Switch 9P
Supported Cables	
X9280A	Sun 4x-to-12x InfiniBand cable, 2 meter length
X9281A	Sun 4x-to-12x InfiniBand cable, 5 meter length
	Note – Use only supported cables to ensure reliable InfiniBand interface connections. These cables can be ordered from Sun Microsystems, Inc.

[1] Check with your Sun representative for compatibility with other manufacturer's switches.

CHAPTER 2

Adapter Installation

This chapter describes how to change the faceplate, install the adapter in your system and verify that it has been installed correctly.

▼ Changing the Faceplate

The IB-HCA is shipped with two faceplates, a standard faceplate and a low-profile faceplate. Use the following procedure to change the faceplate:

1. **Remove the retaining clips from around each connector on the front of the faceplate (**FIGURE 2-1**).**

   Save the clips.

2. **Remove the two faceplate mounting screws and washers.**

3. **Remove the faceplate and replace with the new faceplate.**

4. **Attach the new faceplate by installing the washers, screws, and retaining clips.**

   Ensure that the washers and retaining clips are properly installed.

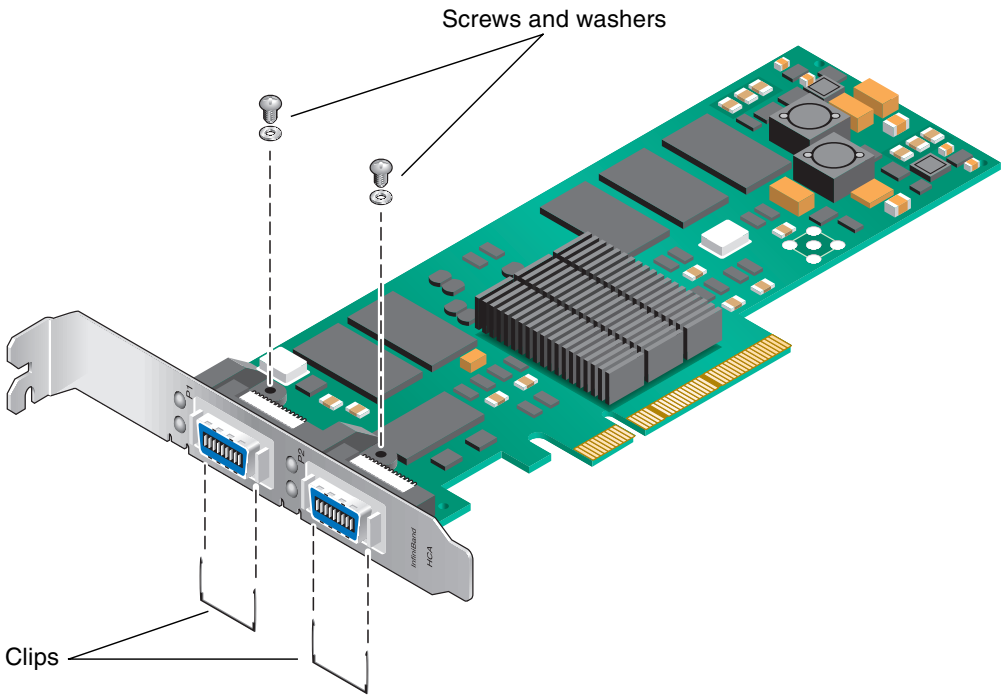

FIGURE 2-1 Changing the Adapter's Faceplate

▼ Installing the Adapter

Caution – Electronic components on printed circuit boards are extremely sensitive to static electricity. Ordinary amounts of static electricity generated by your clothing or work environment can damage the electronic equipment. When installing the IB-HCA in a system, use antistatic grounding straps and antistatic mats to help prevent damage due to electrostatic discharge.

Note – Refer to your system installation or service manual for detailed instructions for the following steps.

1. **Power off your server, using the standard shutdown procedures described in the *Solaris Handbook for Sun Peripherals* or your system service manual.**

   The *Solaris Handbook for Sun Peripherals* is shipped with the Solaris Operating System software and is available on the `http://docs.sun.com` web site.

2. **Remove the cover from the unit to access the card slots and connectors.**

3. **Select an available PCI-E x8 slot and remove the blank front panel.**

   If you are replacing an existing card, remove the card.

4. **Install the IB-HCA card into the slot, pushing the card's edge connector into the connector on the chassis.**

   Ensure that the front plate on the IB-HCA card mounts flush with the chassis panel opening.

5. **If applicable, install the screw in front plate to secure the IB-HCA card into the chassis.**

6. **Connect the 4x end of the InfiniBand I/O cable(s) to the IB-HCA port connector(s).**

   Ensure that the connectors are properly engaged.

Caution – Avoid putting unnecessary stress on the connection. Do not bend or twist the cable near the connectors and avoid cable bends of more than 90 degrees.

7. **Replace the cover on the unit.**

8. **If not already connected, connect the 12x end of the InfiniBand I/O cables to the appropriate ports on the switch or switches.**

   The IB-HCA ports can be connected to different ports on the same switch or to a port on different switches. See FIGURE 2-2 for an example of cabling an adapter to two switches.

9. **Turn power back on to the system and allow the server to reboot.**

   This completes the hardware installation. Proceed to the verification instructions in "Verifying the Installation" on page 10.

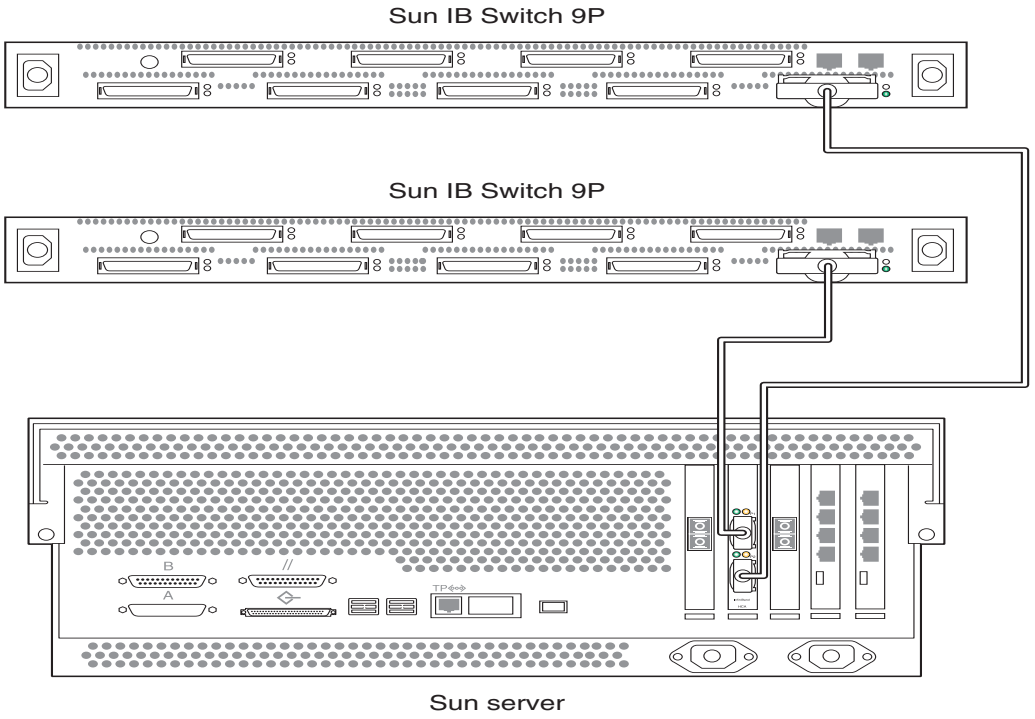

FIGURE 2-2 Adapter Connected to Two Sun IB Switches

▼ Verifying the Installation

Before you can verify the installation, you must install the adapter in the chassis, power on the server, and cable it to an operational InfiniBand switch. The InfiniBand switch should automatically recognize the IB-HCA when it is connected to the fabric if the IB Subnet Manager is running on the switch.

1. **Ensure that the cables are connected to the adapter and switches.**
2. **Verify that the IB Subnet Manager is running on the IB switch.**

   Refer to the *Sun IB Switch 9P Administration Guide* (819-0502) manual for information on the IB Subnet Manager.
3. **Check that the green LED is illuminated for each port that is connected to the switch.**

   If the green LED is not on, check the cable connections at the adapter and at the switch.

4. **Check that the amber LED is illuminated for each port that is connected to the switch.**

   If the amber LED is not on, refer to Appendix B, "Diagnostic Software and Troubleshooting" on page 23 for more information.

5. **Verify that the IB-HCA ports are up and the driver is attached.**

   To do this, type the following command:

```
# dmesg | grep tavor
```

   The output shows recently printed system diagnostic messages that have the string `tavor` in the message. Included in the output is a message that tells whether the port is up or down. For example:

```
Feb 17 20:30:42 hostname px_pci: [ID 370704 kern.info] PCI-device:
pciex15b3,6278@0, tavor0
Feb 17 20:30:42 hostname genunix: [ID 936769 kern.info] tavor0 is
/pci@7c0/pci@0/pci@8/pciex15b3,6278@0
Feb 17 20:30:42 hostname tavor: [ID 548581 kern.info] tavor0: FW ver:
0004.0007.0190, HW rev: a0
Feb 17 20:30:42 hostname tavor: [ID 827127 kern.info] tavor0: MT23108
Sun Microsystems InfiniBand HCA (0x0003ba0001001c18)
Feb 17 20:30:45 hostname genunix: [ID 979957 kern.info] NOTICE: tavor0:
port 1 up
Feb 17 20:30:45 hostname genunix: [ID 611667 kern.info] NOTICE: tavor0:
port 2 up
Feb 17 20:30:58 hostname genunix: [ID 408114 kern.info]
/pci@7c0/pci@0/pci@8/pciex15b3,6278@0 (tavor0) online
```

CHAPTER 3

Configuring IP Over InfiniBand

This chapter contains the following sections:

IPoIB Configuration

The following procedure can be used when making changes to your configuration. See the *System Administration Guide: Basic Administration* (817-1985) for Solaris setup information.

▼ Configuring IPoIB

1. **Confirm that the** IB-HCA **ports are up.**

   To do this, type the following command:

```
# dmesg | grep tavor
```

   The output shows recently printed system diagnostic messages that have the string `tavor` in the message. Included in the output is a message that tells whether the port is up or down.

For example:

```
Feb 17 20:30:42 hostname px_pci: [ID 370704 kern.info] PCI-device:
pciex15b3,6278@0, tavor0
Feb 17 20:30:42 hostname genunix: [ID 936769 kern.info] tavor0 is
/pci@7c0/pci@0/pci@8/pciex15b3,6278@0
Feb 17 20:30:42 hostname tavor: [ID 548581 kern.info] tavor0: FW ver:
0004.0007.0190, HW rev: a0
Feb 17 20:30:42 hostname tavor: [ID 827127 kern.info] tavor0: MT23108
Sun Microsystems InfiniBand HCA (0x0003ba0001001c18)
Feb 17 20:30:45 hostname genunix: [ID 979957 kern.info] NOTICE: tavor0:
port 1 up
Feb 17 20:30:45 hostname genunix: [ID 611667 kern.info] NOTICE: tavor0:
port 2 up
Feb 17 20:30:58 hostname genunix: [ID 408114 kern.info]
/pci@7c0/pci@0/pci@8/pciex15b3,6278@0 (tavor0) online
```

Refer to dmesg(1M) for more information.

2. **Create the InfiniBand devices using the following command:**

```
# devfsadm -C
```

A /dev/ibd*n* device is created for each IB-HCA port in the system where *n* is the port number. Refer to devfsadm(1M) for more information.

3. **Verify InfiniBand devices.**

Check that the InfiniBand devices were created. To do this, type the following command:

```
# ls -l /dev/ibd*
```

A list of InfiniBand devices (/dev/ibd*) is displayed. For example:

```
# ls -l /dev/ibd*
lrwxrwxrwx   1 root     other         29 Apr  4 15:07 /dev/ibd ->
../devices/pseudo/clone@0:ibd
lrwxrwxrwx   1 root     root          88 Apr  4 15:26 /dev/ibd1 ->
../devices/pci@1b,0/pci1022,7450@3/pci15b3,5a46@1/pci15b3,5a44@0/ibport@1,8001
,ipib:ibd1
lrwxrwxrwx   1 root     root          88 Apr  4 15:26 /dev/ibd3 ->
../devices/pci@1b,0/pci1022,7450@3/pci15b3,5a46@1/pci15b3,5a44@0/ibport@2,8001
,ipib:ibd3
```

4. **Plumb each of the IB devices.**

   To do this, type the following command for each IB-HCA port in the system:

   ```
   # ifconfig ibdn plumb host-ipv4-address up
   ```

   where *n* is the instance number of the interface and *host-ipv4-address* is the designated IPv4 address for the device.

   If you are using IPv6, use the following command instead:

   ```
   # ifconfig ibdn inet6 plumb host-ipv6-address up
   ```

5. **Verify that the devices are plumbed.**

   To do this, type the command:

   ```
   # ifconfig -a
   ```

   The output includes information on the IB devices. For example:

```
# ifconfig -a
lo0: flags=2001000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4,VIRTUAL> mtu 8232
index 1
        inet 127.0.0.1 netmask ff000000
ibd1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 2044 index 2
        inet 192.128.91.198 netmask ffffff00 broadcast 192.128.91.255
        ipib 0:0:4:7:0:0:0:0:0:0:1:0:3:ba:0:1:0:1c:19
ibd6: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 2044 index 3
        inet 192.128.92.198 netmask ffffff00 broadcast 192.128.92.255
        ipib 0:0:4:7:0:0:0:0:0:0:1:0:2:c9:2:0:0:55:d9
#
```

6. **Check that data can be send over the InfiniBand link.**

   Use the `ping` command to send ICMP (ICMP6) ECHO_REQUEST packets and receive ICMP (ICMP6) ECHO_RESPONSE packets between InfiniBand hosts. For example:

   ```
   # ping host-ip-address
   ```

   where *host-ip-address* is the IP address of another IB host. See `ping`(1M) for details.

   Repeat this procedure on all hosts that are connected to the IB switch.

CHAPTER 4

InfiniBand Software Overview

This chapter includes the following topics:

InfiniBand Software

The InfiniBand software stack is included in the Solaris 10 Operating System. FIGURE 4-1 shows a block diagram of the InfiniBand software stack modules:

- User Direct Access Programming Library (uDAPL)
- IP over InfiniBand Driver (IPoIB)
- InfiniBand Transport Framework (IBTF)
- InfiniBand Management Framework (IBMF)
- Tavor HCA driver

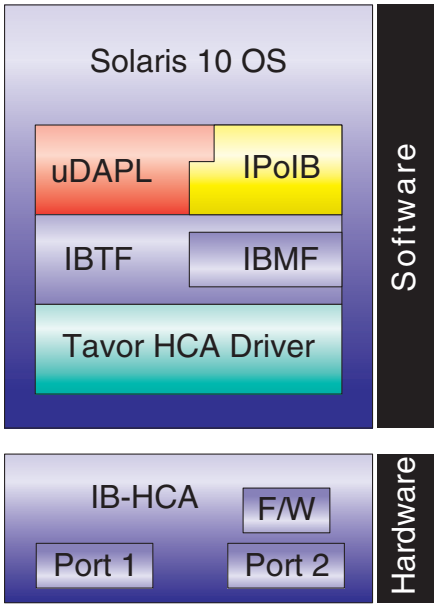

FIGURE 4-1 Solaris 10 OS InfiniBand Software Stack and IB-HCA

User Direct Access Programming Library

This library supports service provider modules written according to the uDAPL 1.2 specification. Sun provides default service provider modules to support the Tavor HCA driver.

IP over InfiniBand Driver

The IPoIB driver implements the IBTF IPoIB protocol and provides IPoIB services for all IB-HCA ports present in the system. This driver enables the standard Solaris TCP/IP stack to operate over InfiniBand. The link layer driver encapsulates IPv4/6 packets, adds IB multicast support, extends ARP address resolution, and changes the network stack to support link addresses greater than 6 bytes.

InfiniBand Transport Framework

IBTF is the Solaris implementation of the InfiniBand Transport Layer (IBTL). IBTF consists of the InfiniBand Device Manager (IBDM), the InfiniBand Communications Manager (IBCM), and an InfiniBand Nexus driver. The Nexus driver and the IBDM create and manage InfiniBand devices. The following devices are supported:

- IB I/O controller (IOC) devices
- IB pseudo devices
- IB service devices
- HCA, port, virtual physical point of attachment (VPPA) service devices

See `ib`(4) and `ib`(7d) for details.

IBCM provides a transport layer abstraction to IB clients to set up reliable connected channels along with service, multicast, and path lookup-related functionality. See `ibcm`(7d) for details.

InfiniBand Management Framework

IBMF is a complementary component to IBTF that enables IB clients to provide and access management services over the IB fabrics. It has two logical components. One component deals with sending and receiving management data to any type of management entity. The other component deals with sending and receiving management data to the subnet administrator (SA) only. Since every IB client needs to send and receive management data to the SA, IBMF provides functions that specifically handle the SA interaction.

Tavor HCA Driver

The Tavor HCA driver controls the Tavor hardware and enables the transfer of data over the IB wire. This driver interfaces with the IBTF and provides an implementation of channel interfaces that are defined by IBTF. The driver also enables management applications and agents to access the IB fabric.

Related Software Documentation

Manuals

- *System Administration Guide: Devices and Files System* (817-5093)

  For information on InfiniBand devices, see the *System Administration Guide: Devices and Files System* (817-5093).

- *System Administration Guide: Network Services* (816-4555)

  For general information on the InfiniBand network, see the *System Administration Guide: Network Services* (816-4555).

- *Sun Cluster 3.x Hardware Administration Manual for Solaris OS*

  For information on administering Sun Cluster hardware, see the *Sun Cluster 3.x Hardware Administration Manual for Solaris OS* (817-0168).

Note – You can view or print Sun documentation at the following web site:
`http://www.sun.com/documentation`

Man Pages

- `cfgadm_ib`(1M) – InfiniBand hardware-specific commands for `cfgadm`
- `datadm`(1M) – Maintain DAT static registry file
- `ifconfig`(1M) – Configure network interface parameters
- `libdat`(3LIB) – Direct access transport library
- `ib`(4) – InfiniBand device driver configuration files
- `ibmf`(7) – InfiniBand Management Transport Framework
- `daplt`(7D) – Tavor uDAPL service driver
- `ib`(7D) – InfiniBand Bus Nexus Driver
- `ibcm`(7D) – InfiniBand Communication Manager
- `ibd`(7D) – Infiniband IPoIB device driver
- `ibdm`(7D) – Solaris InfiniBand device manager
- `tavor`(7D) – InfiniBand (IB) Tavor driver

APPENDIX **A**

InfiniBand Cable Pin Assignments

FIGURE A-1 shows the InfiniBand cable connectors and pin assignments.

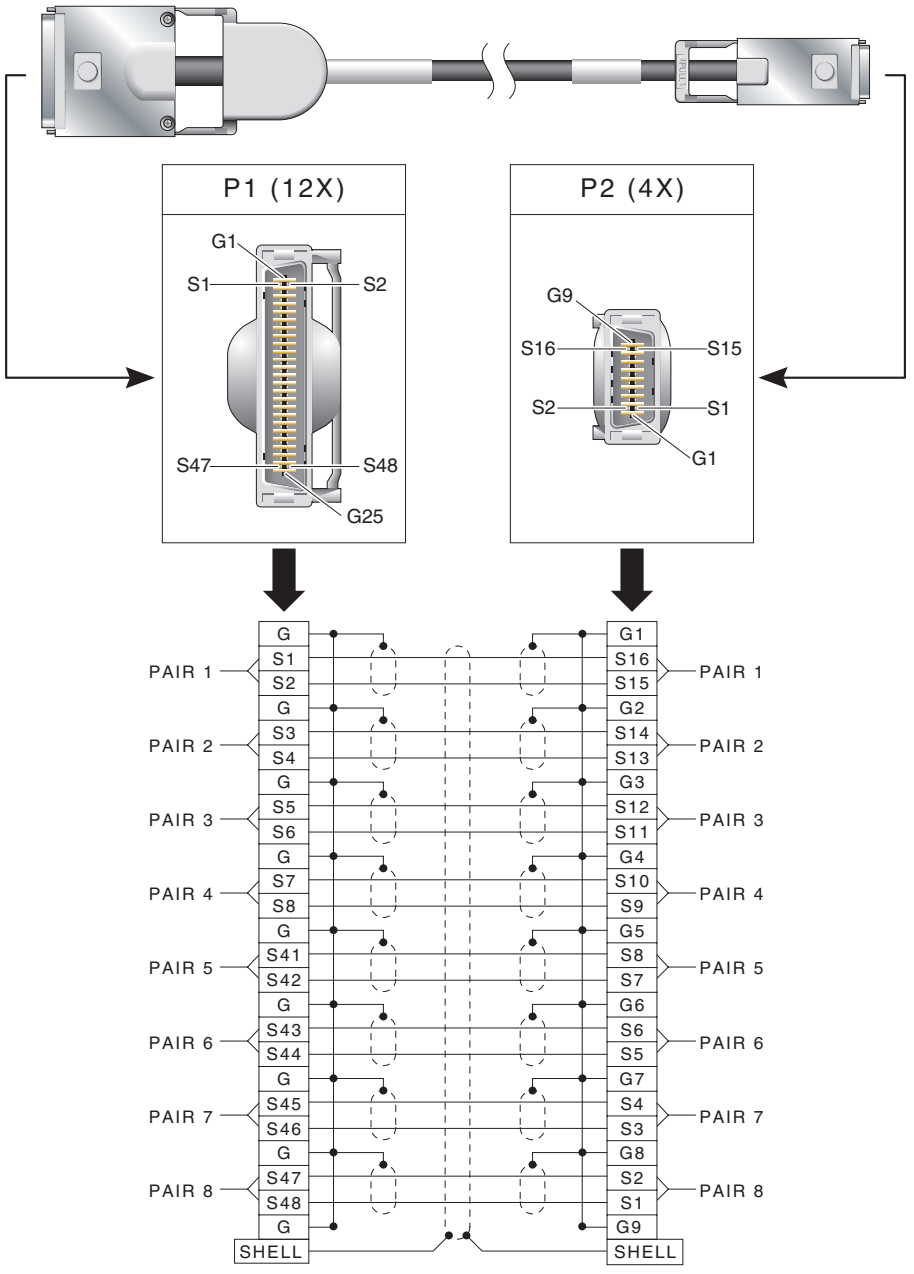

FIGURE A-1 Sun 4x-to-12x InfiniBand Cable Pin Assignments

APPENDIX **B**

Diagnostic Software and Troubleshooting

This appendix provides an overview of the SunVTS diagnostic application and troubleshooting tips for checking the adapter. There is also a section outlining some common troubleshooting issues. This appendix contains the following sections:

SunVTS Diagnostic Testing

The SunVTS software executes multiple diagnostic hardware tests from a single user interface, and is used to verify the configuration and functionality of most hardware controllers and devices. The SunVTS software operates primarily from a graphical user interface, enabling test parameters to be set quickly and easily while a diagnostic test operation is being performed.

Refer to the SunVTS documents (listed in TABLE B-1) for instructions on how to run and monitor the `nettest` diagnostic. These SunVTS documents are available online at the following URL:

`http://www.sun.com/documentation`

Search for the document for the Solaris release on your system.

TABLE B-1 SunVTS Documentation

Title	Descriptions
SunVTS 6.0 Patch Set 1 Documentation Supplement (819-1804)	Describes the SunVTS 6.0 Patch Set 1 features and tests, including the `ibhcatest`.
SunVTS 6.0 User's Guide (817-7664)	Describes the SunVTS diagnostic environment.
SunVTS 6.0 Test Reference Manual (817-7665)	Describes each SunVTS test and describes the various test options and command-line arguments.
SunVTS 6.0 Quick Reference Card (817-7686)	Provides an overview of the user interface.

Using the SunVTS `ibhcatest`

The `ibhcatest` diagnostic test checks the functionality of Sun Dual Port 4x IB Host Channel Adapter PCI-E card. This test can be run from the SunVTS user interface, or it can be run from the command line. See the *SunVTS 6.0 Test Reference Manual* (817-7665) for more information about the `ibhcatest` test.

The `ibhcatest` diagnostic test is included in the SunVTS 6.0 Patch Set 1 and subsequent SunVTS software releases. SunVTS 6.0 Patch Sets are available for downloading from the SunSolve℠ web site `http://sunsolve.sun.com` using the following patch numbers:

- SPARC – Download patch 118962-01, 119881-01, 120639-01, or higher
  - 118962-01 – SunVTS 6.0 Patch Set 1
  - 119881-01 – SunVTS 6.0 Patch Set 2
  - 120639-01 – SunVTS 6.0 Patch Set 3
- X86 – Download patch 118961-01, 119882-01, 120640-01, or higher
  - 118961-01 – SunVTS 6.0 Patch Set 1
  - 119882-01 – SunVTS 6.0 Patch Set 2
  - 120640-01 – SunVTS 6.0 Patch Set 3
- X64 – Download patch 120640-01 or higher

The IB-HCA card and Tavor device driver must be installed, and the IB port interface must be configured `offline` for the `ibhcatest` to run. A loopback cable is not needed because `ibhcatest` includes an internal loopback test. Use the following procedure when running the `ibhcatest`.

▼ Using the ibhcatest

1. **Ensure that the SunVTS software and the Tavor driver are installed on your system.**

   To do this, type the following command:

   ```
   # pkginfo SUNWvts SUNWvtsx SUNWtavor
   ```

   If a SunVTS software package is not installed, refer to the *SunVTS User's Guide* for installation instructions. If the `SUNWtavor` package is not installed, check your Solaris Operating System documentation for software package information.

2. **Unplumb the interface from the system, using the `ifconfig` command:**

   ```
   # ifconfig ibdn down unplumb
   ```

   where *n* is the instance number of the interface.

3. **Refer to** *SunVTS 6.0 PS1 Documentation Supplement* (819-1804) **for instructions on how to run `ibhcatest`.**

Troubleshooting Tasks

The following tasks can be useful when troubleshooting the IB-HCA and the link.

- **Verify that the InfiniBand software packages are installed on the host.**

  To do this, type the command:

  ```
  pkginfo | egrep "InfiniBand | Tavor | udapl"
  ```

  Check that the following packages are installed:

  - `SUNWib` – Sun InfiniBand Framework
  - `SUNWtavor` – Sun Tavor HCA Driver
  - `SUNWipoib` – Sun IP over InfiniBand
  - `SUNWudaplr` – Direct access transport (DAT) registry package (`root`)
  - `SUNWudaplu` – Direct access transport (DAT) registry packages (`usr`)
  - `SUNWudapltr` – Service provider for Tavor packages (`root`)
  - `SUNWudapltu` – Service provider for Tavor packages

If an InfiniBand software package is not installed, check your Solaris Operating System documentation for software package information.

- **Check the system log or console for Tavor driver error messages.**

  See `tavor`(7D) for error messages and descriptions. When the driver is attached to a port on the adapter, the following message is sent.

```
tavorn: port m up (link width 4x).
```

  Where *n* is the instance of the Tavor device number and *m* is the port number on the adapter.

  One way to check Tavor messages is by typing the following command:

```
# dmesg | grep tavor
```

Other Useful Utilities

These utilities can display status and other information about InfiniBand devices:

- `cfgadm`
- `snoop`
- `netstat`
- `kstat`

`cfgadm`

The `cfgadm` utility displays status and other information about the IB-HCA and IB fabric. See `cfgadm_ib`(1M) for details. For example:

```
# cfgadm -al
  Ap_Id                    Type       Receptacle   Occupant      Condition
  hca:21346543210a987        IB-HCA     connected    configured    ok
  ib                         IB-FABRIC  connected    configured    ok
  ib::80020123456789a        IB-IOC     connected    configured    ok
  ib::802abc9876543          IB-IOC     connected    unconfigured  unknown
  ib::80245678,ffff,ipib     IB-VPPA    connected    configured    ok
  ib::12245678,0,nfs         IB-PORT    connected    configured    ok
  ib::21346543,0,hnfs        IB-HCA_SVC connected    configured    ok
  ib::sdp,0                  IB-PSEUDO  connected    configured    ok
```

snoop

The snoop program captures and inspects network packets. See the snoop(1M) man page for details. For example:

```
# snoop -d ibd1
Using device /dev/ibd1 (promiscuous mode)
    ib-1-167 -> *            ARP C Who is 199.1.1.168, ib-1-168 ?
    ib-1-168 -> ib-1-167     ARP R 199.1.1.168, ib-1-168 is
0:2:4:7:0:0:0:0:a:4:7c:4f:0:2:c9:2:0:0:55:91
    ib-1-167 -> ib-1-168     ICMP Echo request (ID: 35608 Sequence number: 0)
    ib-1-168 -> ib-1-167     ICMP Echo reply (ID: 35608 Sequence number: 0)
```

netstat

netstat shows network status. See the netstat(1M) man page for details. For example:

```
# netstat -I ibd 4
    input   ibd1      output       input  (Total)    output
packets errs  packets errs  colls  packets errs  packets errs  colls
2458394 0     2458268 0     0      2467288 0     2465951 0     0
92233   0     92237   0     0      92247   0     92238   0     0
92703   0     92702   0     0      92709   0     92704   0     0
```

kstat

kstat displays kernel statistics. See the kstat(1M) man page for details. For example:

```
# kstat ibd:1
module: ibd                             instance: 1
name:   ibd1                            class:    net

                                                   0
        opackets                        27381595
        opackets64                      27381595
        promisc                         off        xmt_badinterp
              0
        xmtretry                        4
```

Index

S

T

U