



Sun HPC 3.0 SCI Guide

901 San Antonio Road
Palo Alto, , CA 94303-4900
USA 650 960-1300 Fax 650 969-9131

Part No: 805-6263-10
June 1999, Revision A

Copyright Copyright 1999 Sun Microsystems, Inc. 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. All rights reserved.

This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, SunStore, AnswerBook2, docs.sun.com, and Solaris are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun[™] Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

RESTRICTED RIGHTS: Use, duplication, or disclosure by the U.S. Government is subject to restrictions of FAR 52.227-14(g)(2)(6/87) and FAR 52.227-19(6/87), or DFAR 252.227-7015(b)(6/95) and DFAR 227.7202-3(a).

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Copyright 1999 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, Californie 94303-4900 U.S.A. Tous droits réservés.

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées des systèmes Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, SunStore, AnswerBook2, docs.sun.com, et Solaris sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun[™] a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REPENDRE A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



Contents

	Preface	vii
1.	Preparing for SCI Installation	1
	Other Necessary Documentation	1
	SCI Adapter Cards	1
	Supported SCI Network Topologies	3
	Two-Node Networks	3
	Three-Node Networks	4
	Four-Node Networks	5
	SCI Adapter Card Scrubber Jumpers	7
2.	Network Connection Procedure	11
	Install SCI Adapter Cards	11
	Notes for Scrubber Jumper Settings	11
	Notes for Switched Two-Node Network	12
	Notes Regarding SBus Slots	12
	Connecting SCI Adapter Cards in a Two-Node Network	12
	Connecting SCI Adapter Cards in a Three- or Four-Node Network	14
3.	Configuring the SCI Network Interface	17
	Preconditions	17
	SCI Network Hardware Must Be Installed	17

	ClusterTools 3.0 Software Must Be Installed	18
	Caution for NIS+ Users	18
	Create a Temporary Network Map for Later Reference	19
	Create <code>sci_config.hpc</code>	20
	SCI Configuration Templates	20
	Editing a Configuration Template	20
	Propagate the SCI Configuration	28
	Run <code>sm_config</code>	28
	Compare <code>sm_config</code> Output With Contents of <code>sci_config.hpc</code>	28
	Verify the Rank of the SCI Interface	29
	Add Variable to <code>sci.conf</code> File	29
	Reboot Nodes	29
4.	Verify That the Network Is Functional	31
	Run <code>get_ci_status</code>	31
	Run <code>ifconfig --a</code>	31
	Ping the SCI Adapter Cards	31
	Do All-to-All Ping	32
	Check for RSM Daemon	32
	Verify Basic Communication Functionality	32
5.	Adding or Replacing SCI Adapter Cards	33
	Install/Replace SCI Adapter Card	33
	Connect New Adapter Card to Network	34
	Create a Temporary Network Map	34
	Run <code>sciconf</code>	34
	Update <code>sci_config.hpc</code>	35
	Run <code>sm_config</code>	35
	Confirm <code>sci_config.hpc</code> Contents	36
	Reboot the Node	36

	Verify the New Network	36
6.	SCI Interface Troubleshooting	37
	SCI Switch	37
	General Hardware Inspection	37
	SCI Switch Status LED Locations	37
	Port Status LEDs	38
	General Switch Status LED	39
	The <code>get_ci_status</code> Command	39
	Client Net Failure	40
	Incorrect Software Configuration	40
	Incorrect Firmware	41
A.	Man Pages	43
	<code>sm_config(1)</code>	43
	<code>get_sci_status(1m)</code>	47
B.	SMA Messages	51
	Message Directory	51
	SMA Messages	52

Preface

The *SunTM HPC 3.0 SCI Guide* is intended for experienced system administrators.

How This Book Is Organized

Chapter 1, provides an overview of the Sun HPC SCI subsystem, including descriptions of the principal hardware and software components.

Chapter 2, outlines the procedure for connecting the cluster nodes to an SCI network in the various supported topologies.

Chapter 3, explains how to configure the network interfaces on the cluster nodes.

Chapter 4, provides a set of procedures that can be used to check the basic functionality of the network.

Chapter 5, describes the procedure for configuring a newly installed SCI adapter card.

Chapter 6, describes techniques for solving failures of the SCI interface.

Appendix A, contains man pages related to the SCI interface.

Appendix B, lists system messages related to the SCI interface.

Using UNIX Commands

This document may not contain information on basic UNIX[®] commands and procedures such as shutting down the system, booting the system, and configuring devices.

See one or more of the following for this information:

- *Solaris 2.x Handbook for SMCC Peripherals*
- AnswerBook[™] online documentation for the Solaris[™] 2.x or Solaris 7 software environment.
- Other software documentation that you received with your system

Typographic Conventions

TABLE P-1 Typographic Conventions

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

Shell Prompts

TABLE P-2 Shell Prompts

Shell	Prompt
C shell	<i>machine_name%</i>
C shell superuser	<i>machine_name#</i>
Bourne shell and Korn shell	<i>\$</i>
Bourne shell and Korn shell superuser	<i>#</i>

Related Documentation

TABLE P-3 Related Documentation

Application	Title	Part Number
All	<i>Sun HPC ClusterToolsTM 3.0 Product Notes</i>	805-6262-10
Sun MPI Programming	<i>Sun MPI 4.0 User's Guide: With LSF</i>	805-7230-10
Sun MPI Programming	<i>Sun MPI 4.0 User's Guide With CRE</i>	806-0296-10
Sun MPI Programming	<i>Sun MPI 4.0 Programming and Reference Guide</i>	805-6269-10
Prism	<i>Prism 6.0 User's Guide</i>	805-6277-10
Prism	<i>Prism 6.0 Reference Manual</i>	805-6278-10

TABLE P-3 Related Documentation *(continued)*

Application	Title	Part Number
LSF	<i>LSF Batch User's Guide</i>	805-6258-10
LSF	<i>LSF Batch Programmer's Guide</i>	805-6260-10

Preparing for SCI Installation

This chapter describes the SCI network configurations that are supported by Sun HPC ClusterTools 3.0 software suite.

Other Necessary Documentation

The hardware installation and service manuals provided with your Sun server are likely to contain important information about installing SCI adapter cards and network cables. Please read the applicable sections of the Sun server hardware manuals before installing and configuring an SCI network on your Sun HPC cluster.

SCI Adapter Cards

Sun HPC 3.0 cluster nodes connect to the SCI network through SCI adapter cards installed in the node's SBus slots. In two-node SCI networks, the SCI adapter cards are ordinarily connected to each other directly, without going through a switch. One possible exception to this switchless approach is discussed in "Connecting SCI Adapter Cards in a Two-Node Network" on page 12.

In three- and four-node networks, each adapter card is connected to a port in a four-port SCI switch. These two basic SCI network connections schemes are shown in Figure 1-1.

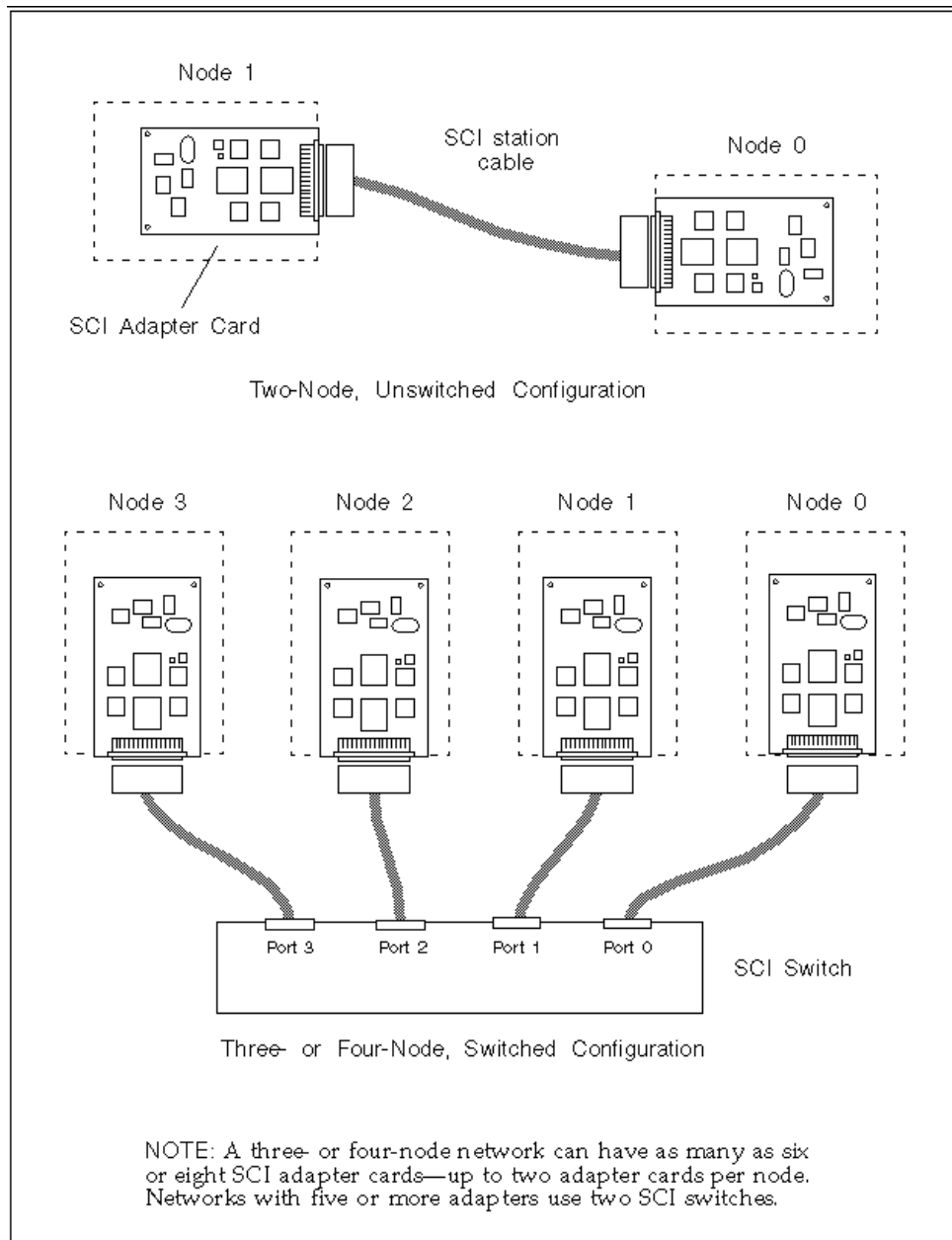


Figure 1-1 Basic SCI Network Connection Schemes for Sun HPC 3.0 Clusters

Supported SCI Network Topologies

A Sun HPC 3.0 cluster can have up to four nodes connected to an SCI-based private subnet. The nodes may connect to the SCI network through one or two SCI adapter cards. When each node in the network has two SCI adapter cards, communication bandwidth can be increased by striping messages across both network interfaces.

Note - See the *Sun HPC ClusterTools 3.0 Administrator's Guide: With LSF* or *Sun HPC ClusterTools 3.0 Administrator's Guide: With CRE* for additional information about configuring a Sun HPC cluster to support message striping.

Chapter 2 describes the procedure for connecting the nodes in the various network topologies described below. Chapter 3 describes the procedure for configuring the SCI drivers.

Two-Node Networks

Figure 1–2 shows how two nodes in a Sun HPC 3.0 cluster can be connected via an SCI network. The SCI adapter card in one node is connected directly to an SCI adapter card in the other node. There is no intervening SCI switch, which is the usual connection scheme for two-node networks.

If each node has two SCI adapter cards, messages can be striped across the two network interfaces. This is illustrated in lower schematic in “Connecting SCI Adapter Cards in a Two-Node Network” on page 12.

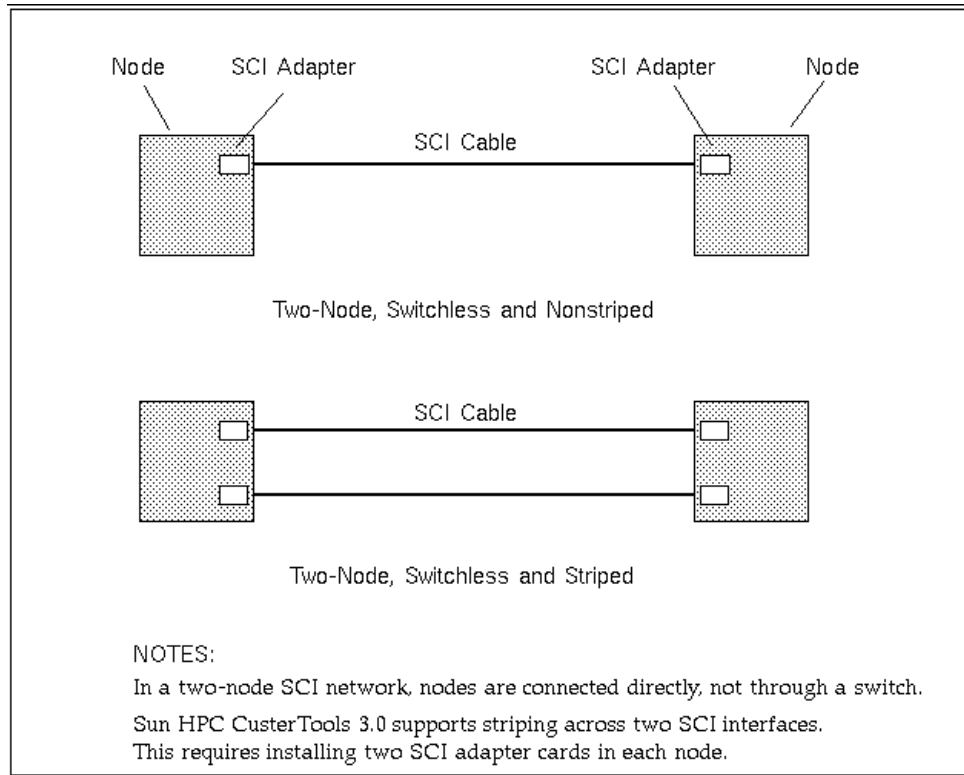


Figure 1-2 Two-Node SCI Interconnection, Without Striping

If you expect to add a node to a two-node network at a later time, you may want to connect the two nodes through a switch now, even though the switch is not needed. This would simplify the process of adding a third node later on. The chief disadvantage to using a switch in a two-node network is the latency it adds to the communication path between the nodes.

This alternate connection scheme is discussed further in “Connecting SCI Adapter Cards in a Two-Node Network” on page 12.

Three-Node Networks

Figure 1-3 shows examples of how three Sun HPC nodes can be connected to an SCI network, in both unstriped and striped modes.

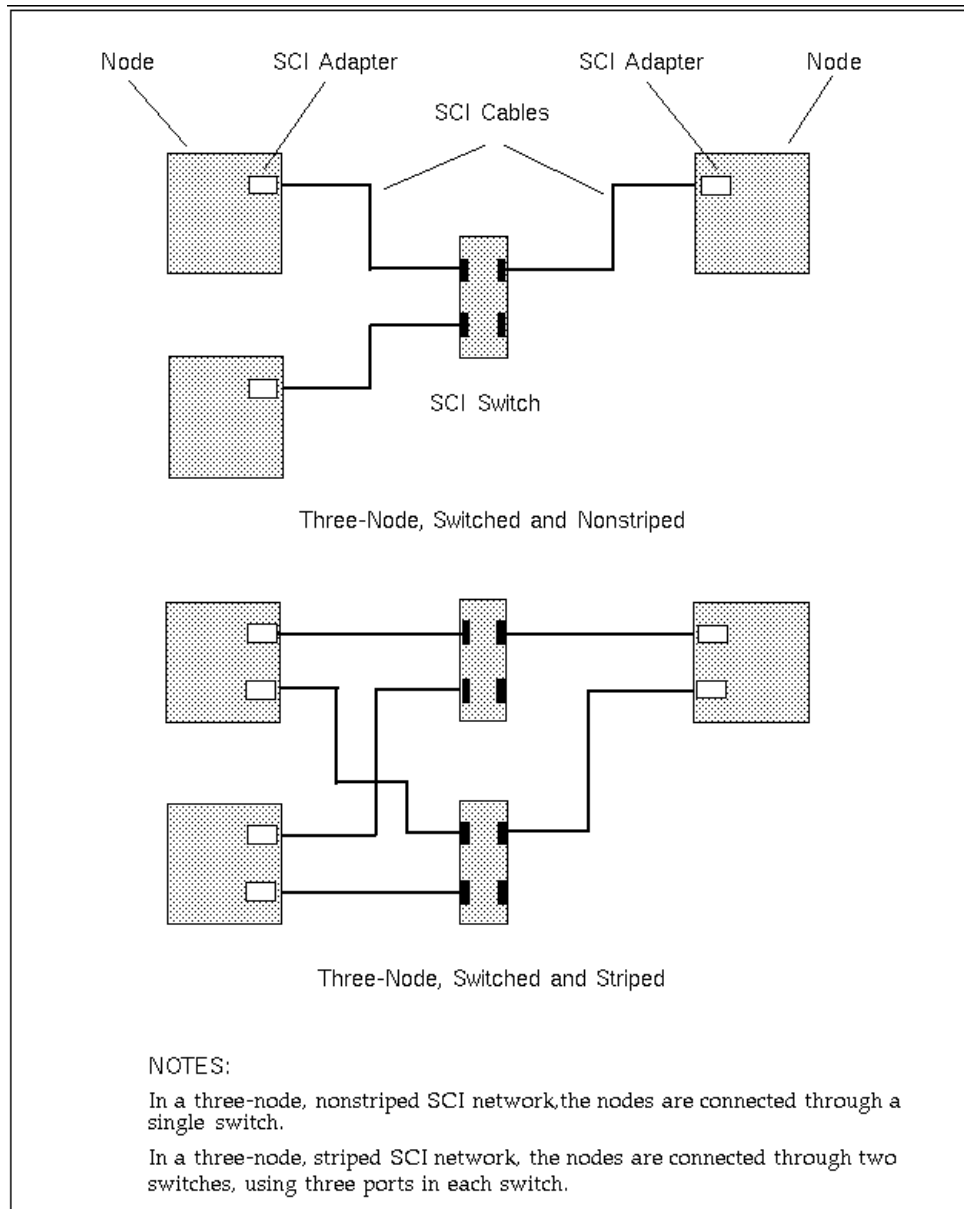


Figure 1-3 Supported Three-Node SCI Interconnections

Four-Node Networks

Figure 1-4 shows examples of how four Sun HPC nodes can be connected to an SCI network, in both unstriped and striped modes.

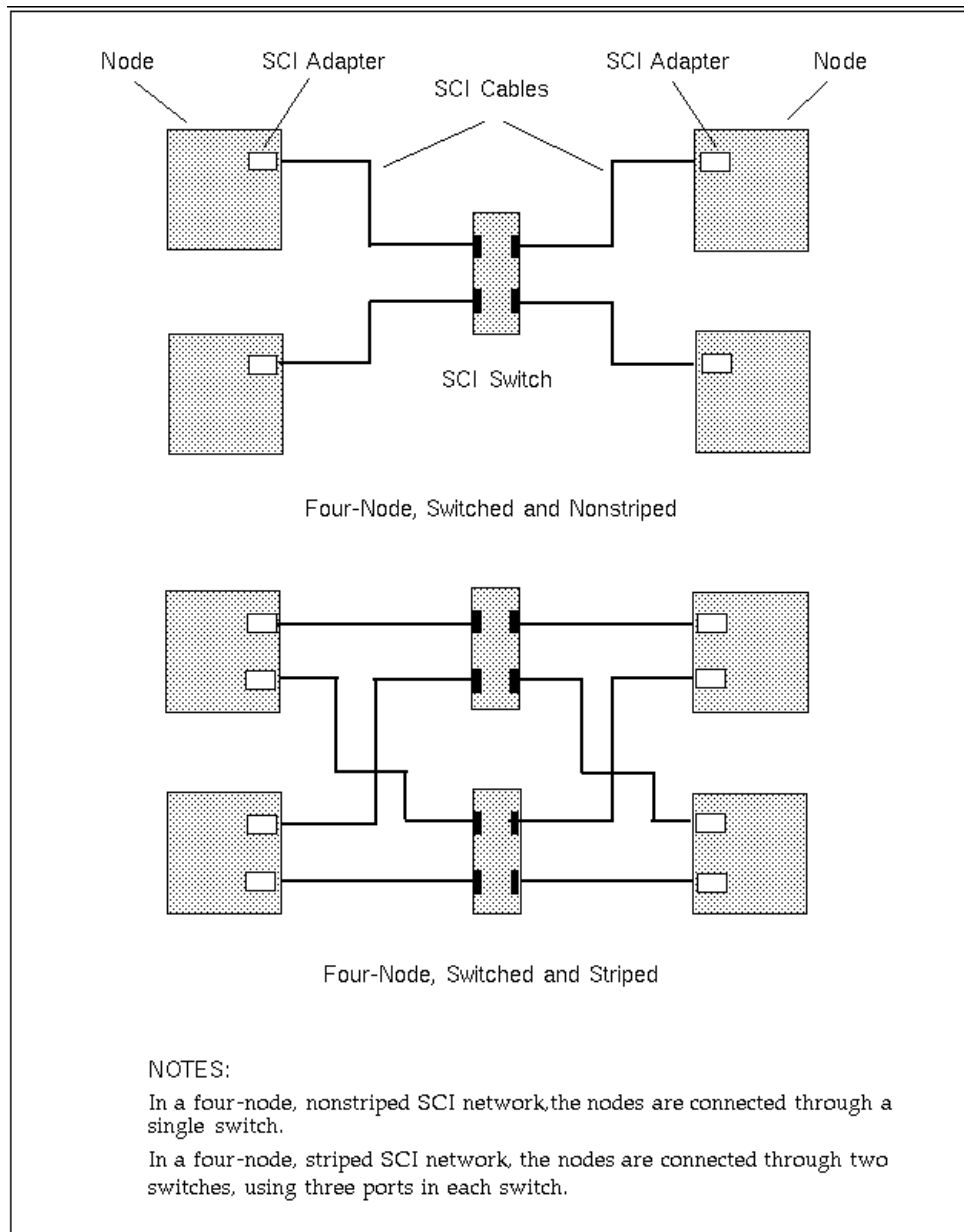


Figure 1-4 Supported Four-Node SCI Interconnections.

SCI Adapter Card Scrubber Jumpers

Each SCI adapter card has a jumper, called the scrubber jumper. This jumper configures the scrubber circuit, which controls link maintenance functions. Figure 1-5 shows its location on the SCI adapter card.

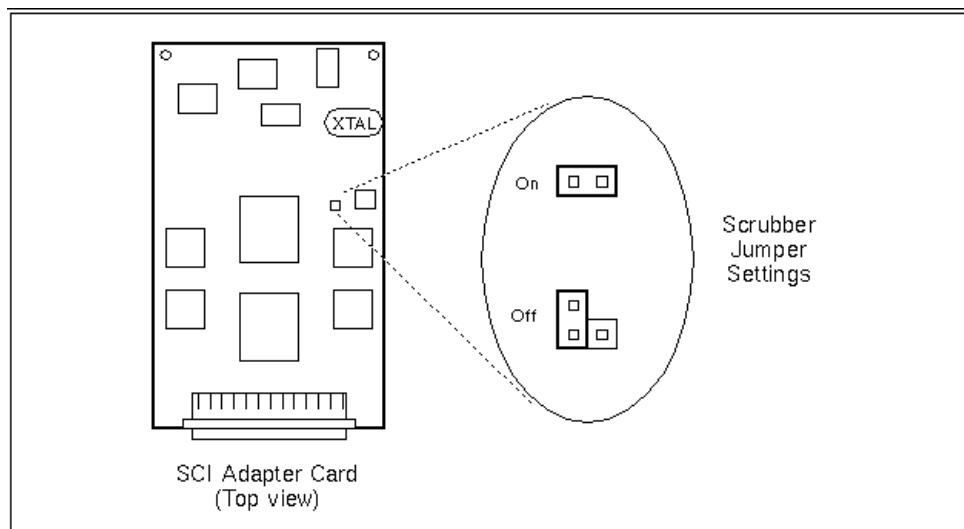


Figure 1-5 Location of the Scrubber Jumper

Table 1-1 specifies the appropriate scrubber jumper settings for unswitched and switched SCI networks.

TABLE 1-1 Scrubber Jumper Settings

Topology	SCI SBus Card Jumper Settings
Two nodes (unswitched)	Set one jumper ON and the other jumper OFF in each connected pair of SCI adapter cards. See Figure 1-6 for examples of two-node jumper settings.
Three or four nodes (switched)	Set all jumpers ON.

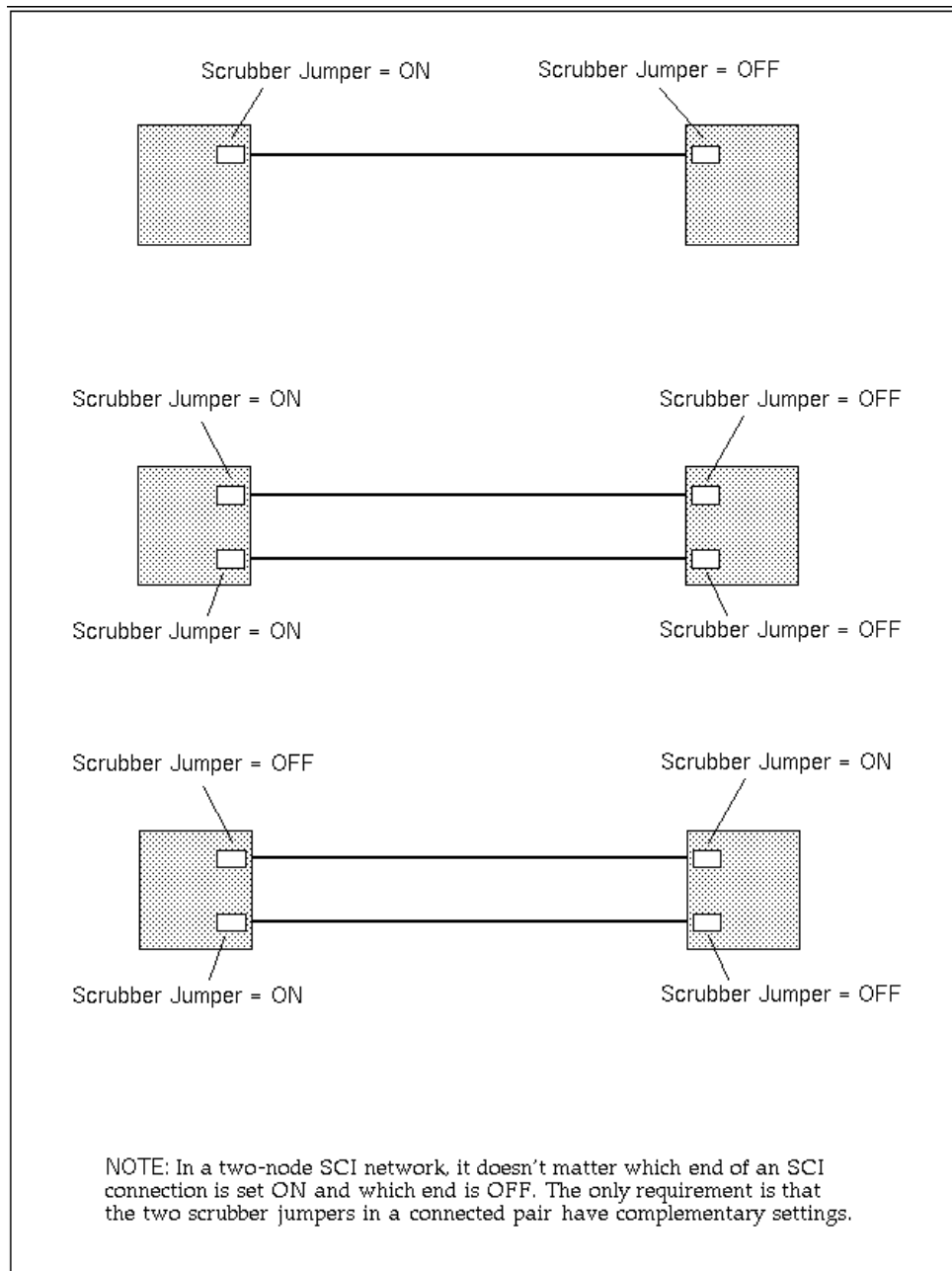


Figure 1-6 Examples of Scrubber Jumper Settings in Two-Node Networks

Note - SCI adapter cards are shipped without a default setting. Therefore, examine the setting on each SCI adapter card and adjust it if necessary. If scrubber jumpers are not set correctly when installed, communication between nodes may experience intermittent faults.

Network Connection Procedure

The Sun HPC ClusterTools 3.0 software may—but need not be—already installed on the nodes before you perform the procedures described in this chapter. The ClusterTools 3.0 software *must* be installed before you configure the SCI drivers, as described in Chapter 3.

Install SCI Adapter Cards

Notes for Scrubber Jumper Settings

If the nodes are not already in a power-off state, turn their power switches off and disconnect their power cables from the power outlet.

Verify that each SCI adapter card's scrubber jumper has the correct setting for the network configuration in which it will be used.

- Switchless two-node network – If you are creating a two-node network in which the nodes will connect directly, not through a switch, set the scrubber jumper at one end of a connection to ON and the scrubber jumper at the other end of the connection to OFF.
- Switched two-node network – If you plan to add one or two nodes to a two-node network at a later time, you may want to connect the nodes through an SCI switch when you first set up the two-node network. See the section “Notes for Switched Two-Node Network” on page 12 for an explanation.
- Three- and four-node networks – The scrubber jumper on every SCI adapter card must be set to ON.

Notes for Switched Two-Node Network

In a switchless two-node network, one scrubber jumper must be ON and the other must be OFF. In a three- or four-node network, all scrubber jumpers must be ON. This means, if you add a node to a two-node network that was originally set up without a switch, you must:

- add a switch to the network
- change the OFF scrubber jumper to ON and reboot that node

If, however, you configure the two-node network with a switch from the beginning, you will initially set all scrubber jumpers to ON. As a result, the adapter cards will be initialized to communicate through a switch from the beginning and it will not be necessary to change any scrubber jumpers when you add nodes to the network.

If you plan to do this, read the instructions for creating three- and four-node networks instead of the section on two-node networks.

Note - The chief disadvantage to using a switch when it is not required is that the switch adds some latency that would otherwise not be in the network.

Notes Regarding SBus Slots

If the SCI adapter cards are not already installed in the nodes, install them in available SBus slots. Refer to the hardware installation manual that was supplied with your server for installation instructions.

If possible, apply the following guidelines when choosing which SBus slots to use:

- Install the cards in the lowest-numbered available SBus slots.
- If you are installing two adapter cards per node, locate them in adjacent SBus slots.

Adhering to these guidelines will make it easier for you to edit the SCI configuration file, `sci_config.hpc`, later in the installation process. Chapter 3 explains how to edit this file.

Connecting SCI Adapter Cards in a Two-Node Network

This section explains how to create a two-node SCI network that does not use an SCI switch.

1. **Position the nodes in the desired locations.**

2. Use an 80-line SCI station cable to connect the SCI adapter card of one node to the SCI adapter card of the other node.
If each of your nodes has two SCI adapter cards, use two SCI cables, one for each pair of adapter cards. See Figure 2-1 for examples.
3. Connect the node power cords to the appropriate power outlets.
4. Turn the node power switches on and boot the nodes.

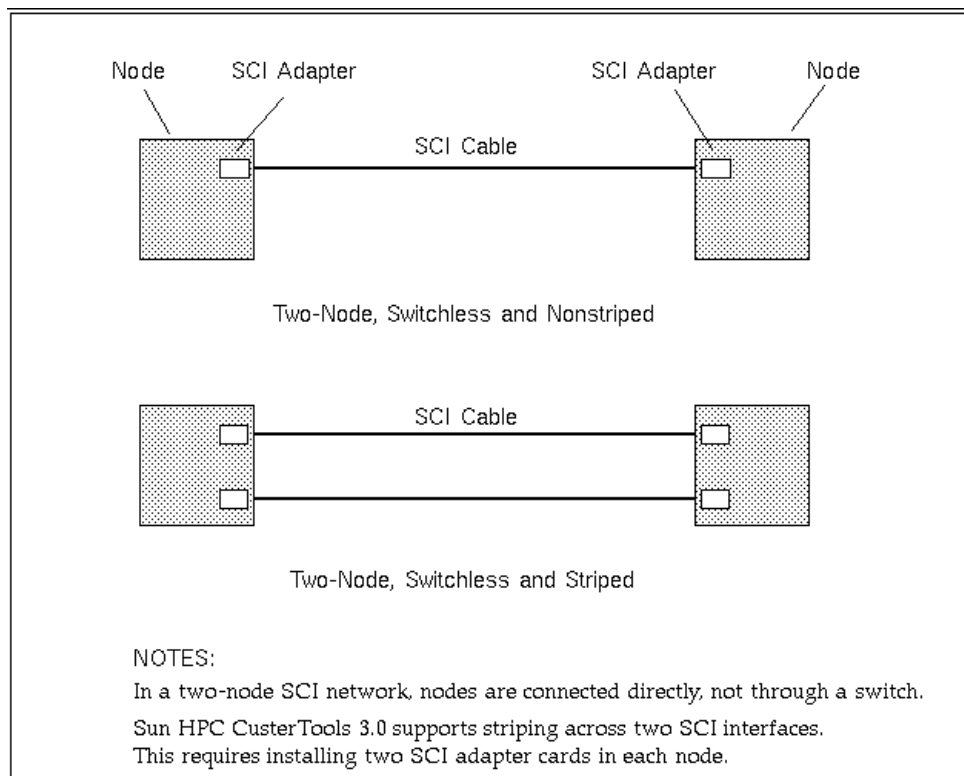


Figure 2-1 Examples of Two-Node Unswitched SCI Connections

Connecting SCI Adapter Cards in a Three- or Four-Node Network

Perform the following steps to create a three- or four-node SCI network.

1. **Position the nodes in the desired locations.**
2. **Use a separate 80-line SCI station cable to connect each SCI adapter card to a port on an SCI switch.**
You can connect any SCI adapter card to any switch port, but you should follow a logical order in making your connections. This will simplify network configuration and administration tasks. See Figure 2-2 and Figure 2-3 for connection examples.
3. **Connect the node power cords to the appropriate power outlets.**
4. **Turn the node power switches on and boot the nodes.**

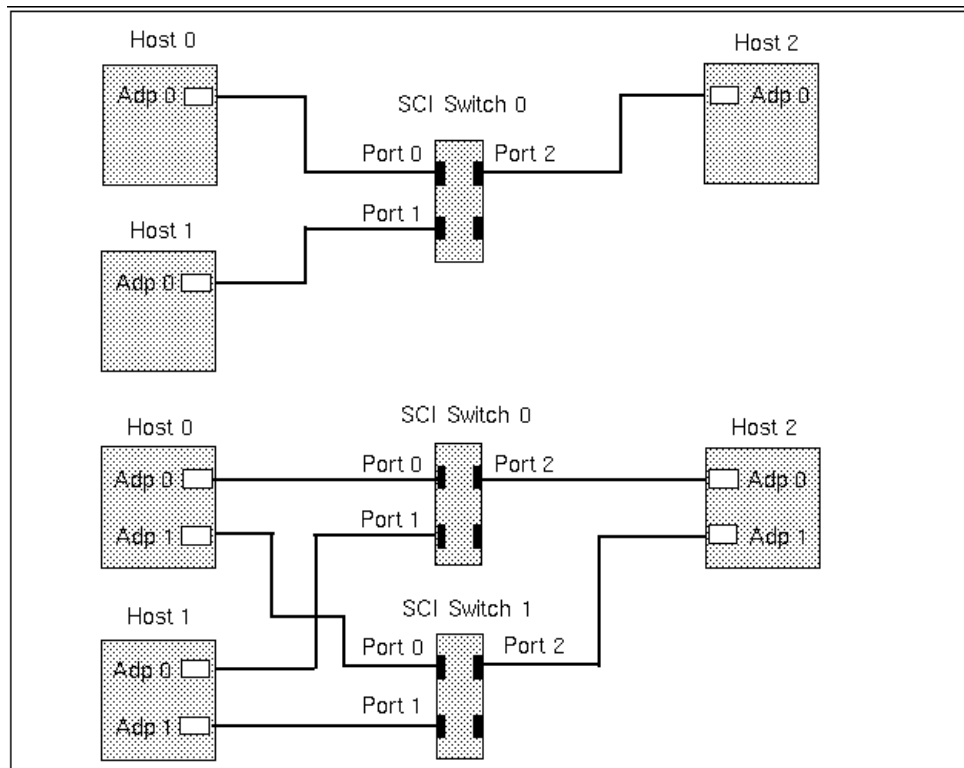


Figure 2-2 Examples of Three-Node Switched SCI Connections

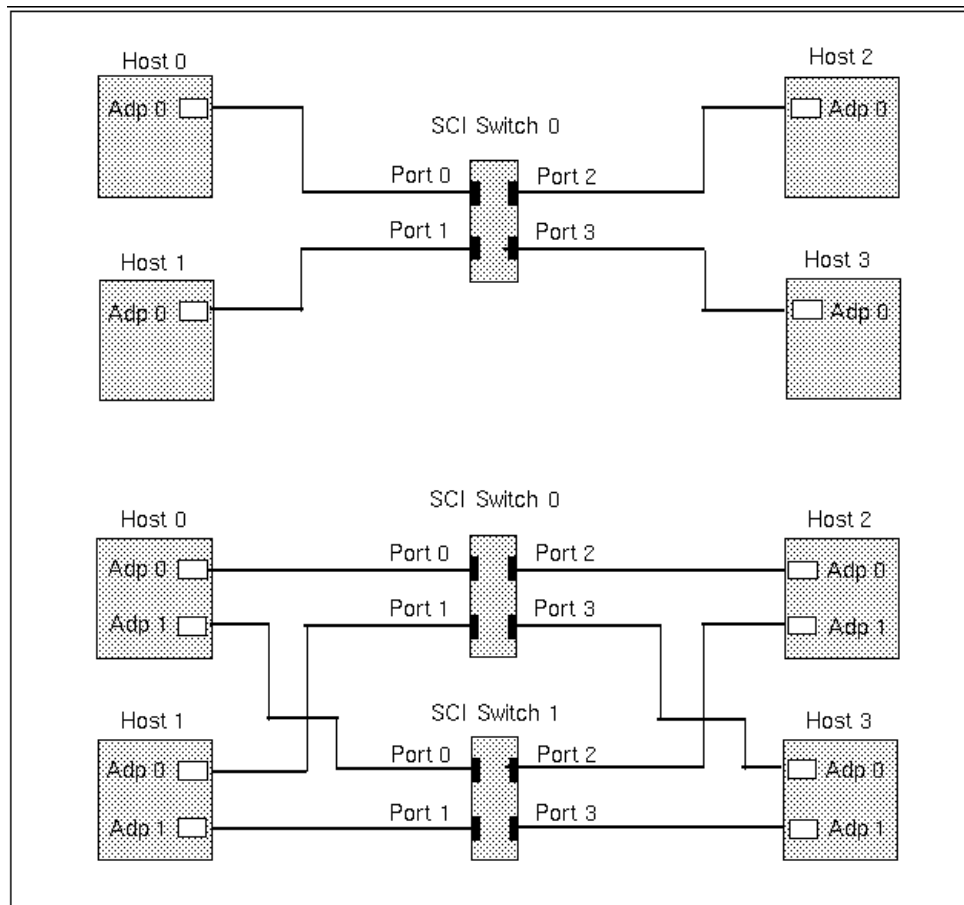


Figure 2-3 Examples of Four-Node SCI Switched Connections

Configuring the SCI Network Interface

This chapter explains how to configure the SCI network interface in a Sun HPC 3.0 cluster. It covers the following procedures:

- Creating a temporary network map for later reference – See “Create a Temporary Network Map for Later Reference” on page 19
- Creating the SCI configuration file, `sci_config.hpc` – See “Create `sci_config.hpc`” on page 20
- Running `sm_config` to initialize the SCI network interface – See “Propagate the SCI Configuration” on page 28
- Verifying the rank of the SCI interfaces – “Verify the Rank of the SCI Interface” on page 29
- Rebooting the cluster nodes – “Reboot Nodes” on page 29

The first section, “Preconditions” on page 17, describes a few conditions that must be in effect before the SCI interface can be configured.

Preconditions

Ensure that the conditions described below are in effect before proceeding with the SCI setup procedure.

SCI Network Hardware Must Be Installed

The SCI network hardware must already have been installed.

ClusterTools 3.0 Software Must Be Installed

The Sun HPC ClusterTools 3.0 software must be installed on the cluster—either on an NFS server or locally on each node. If this has not yet been done, refer to the Sun HPC ClusterTools 3.0 Installation Guide for instructions.

If you will be using the LSF suite as the cluster's workload manager, the LSF components must be installed before the ClusterTools software.

Verify that the following Sun HPC ClusterTools 3.0 packages are present.

- SUNWsci
- SUNWscid
- SUNWscidx – required only on clusters with 64-bit Solaris 7
- SUNWsma
- SUNWsmmax – required only on clusters with 64-bit Solaris 7
- SUNWrsmop
- SUNWrsm

These SCI and RSM packages may not be present if, during installation of the ClusterTools 3.0 software, the installer responded “No” when asked if SCI packages should be installed. If they are not present, they must be installed.

The installation GUI (graphical user interface) includes an option for installing the SCI-related packages exclusively. See the Sun HPC ClusterTools 3.0 Installation Guide for more information.

If any of these packages were installed from another source, you must remove them and install the packages provided on the Sun HPC ClusterTools 3.0 software release.

Caution for NIS+ Users

`sm_config` may have trouble contacting other node in the cluster in an NIS+ environment. By default, the NIS+ version of `/etc/nsswitch.conf` specifies the services entry as: `nisplus [NOTFOUND=return] files`. Since the `/etc/services` file is modified and used by `SUNWsma` and other packages, the `/etc/nsswitch.conf` entry should be as follows:

```
services:  files nisplus
```

Place the term `files` first before other entries.

Create a Temporary Network Map for Later Reference

Before you create the SCI configuration file, `sci_config.hpc`, sketch out a rough map of the physical network connections and identify each SCI adapter by its serial number. Look on the connector panel of the node. An adapter's serial number will be printed on a white label in the upper left corner of the adapter's connector. It is usually a four- or five-digit number.

Figure 3-1 shows an example of a temporary map of a four-node configuration without striping—that is, with one SCI network adapter connection per node.

This information will make it easier to verify that the adapter ID values specified in the `sci_config.hpc` file match the actual values assigned by the device driver. Instructions for using this map are provided in the section “SCI Configuration Templates” on page 20.

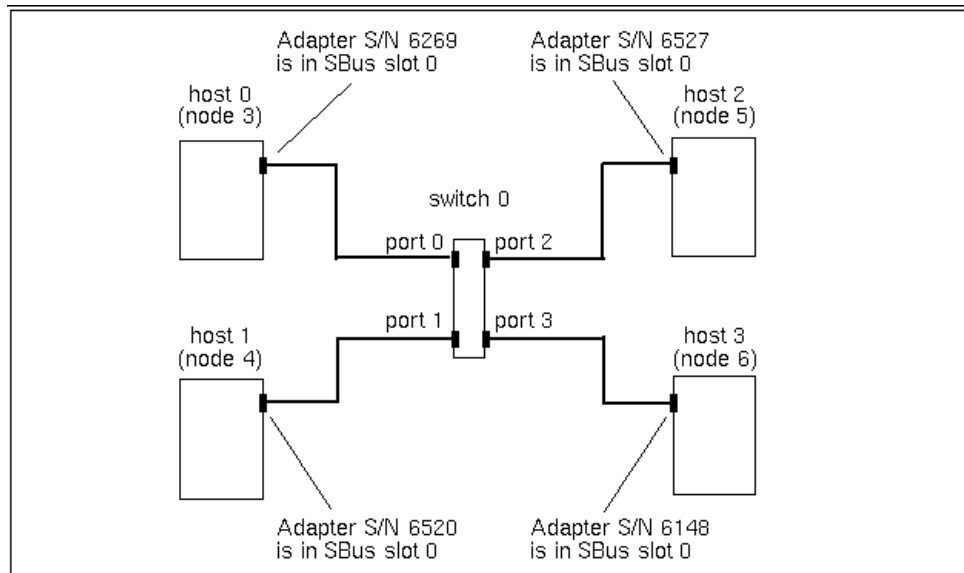


Figure 3-1 Sample Temporary Map of a Four-Node, Nonstriped Configuration

Create `sci_config.hpc`

The SCI configuration procedure reads network mapping information from the configuration file `/opt/SUNWsmas/sci_config.hpc`, which you create.

SCI Configuration Templates

The Sun HPC ClusterTools 3.0 software release includes the following SCI configuration templates to simplify the creation of `sci_config.hpc`. Each template represents a supported SCI network topology.

- `sma2.hpc` – Two nodes connected directly by a single SCI station cable. There is no intervening SCI switch. Because this configuration uses only one network interface per node, it does not support message striping.
- `sma3.hpc` – Three nodes connected through an SCI switch. Each node is connected to the switch by a single station cable. Consequently, message striping is not supported.
- `sma4.hpc` – Four nodes connected through a single SCI switch. Each node is connected to the switch by a single station cable. Again, message striping is not supported.
- `sma2-2stripes.hpc` – Two nodes connected directly by two SCI station cables. There is no intervening SCI switch. Because both nodes have two network interfaces, messages can be striped across both cables.
- `sma4-2stripes.hpc` – Four nodes connected through two SCI switches via two station cables per node. Because each node has two network interfaces, striping is supported. Two SCI switches are needed because each switch has only four ports.

These templates are in `/opt/SUNWhpc/bin/Install_Uutilities/config_dir`.

Editing a Configuration Template

Copy the applicable template to `/opt/SUNWsmas/sci_config.hpc`. For example, to create a configuration file for the two-node striped topology,

```
# cd /opt/SUNWsmas # cp /opt/SUNWhpc/bin/Install_Uutilities/config_dir/sma2-2stripes.hpc sci_config.hpc
```

Use the `sma4-2stripes.hpc` template for creating a three-node, striped configuration.

Next, edit `sci_config.hpc`. Every template type is organized into eight sections. Instructions for editing each section are provided below.

Section 1 – Cluster Configuration Type

Section 1 asks you to specify the type of cluster you have; you are given the options: SC (Sun Cluster) or HPC. Enter HPC, as follows:

```
Cluster is configured as = HPC
```

SC is not a valid entry for clusters running Sun HPC ClusterTools software.

Section 2 – Host Names

List all of the nodes in the cluster by replacing `<host_namen>` placeholders with the host names of the cluster's nodes. For example, if your cluster contains the nodes: node3, node4, node5, and node6, Section 2 should look like this:

```
HOST 0 = node3
HOST 1 = node4
HOST 2 = node5
HOST 3 = node6
```

The nodes can be listed in any order.

Section 3 – Number of Switches

Specify the number of SCI switches in the cluster. This will be determined by which network topology you implement, as follows:

- Two-node cluster – either nonstriped or striped, set

```
Number of Switches in cluster = 0
```

- Three- or four-node cluster – nonstriped, set

Number of Switches in cluster = 1

- Three- or four-node cluster – striped, set

Number of Switches in cluster = 2

Section 4 – Number of Direct Links

Specify the number of unswitched node-to-node connections your cluster has. Again, this will depend on which topology you implement, as follows:

- Two-node cluster – nonstriped, set

Number of Direct Links in cluster =
1

- Two-node cluster – striped, set

Number of Direct Links in cluster =
2

- Three- or four-node cluster – either nonstriped or striped, set

Number of Direct Links in cluster =
0

Section 5 – Number of Ring Connections

Ring connections are not supported by Sun HPC ClusterTools 3.0 software. Therefore, always specify

Number of Rings in cluster = 0

Section 6– Adapter Information

List all SCI adapters in the cluster and describe the connection details for each.

Use a separate line for each adapter description. The format for describing unswitched connections is

```
host n :: adp n is connected to = link  
n :: endpt n
```

When no switch is used, an adapter (adp) is connected to a particular endpoint (endpt n) on a particular channel (link n). See Figure 3–2 and Figure 3–3.

Each adapter has its own endpoint. That is why two different endpoints are shown on one link.

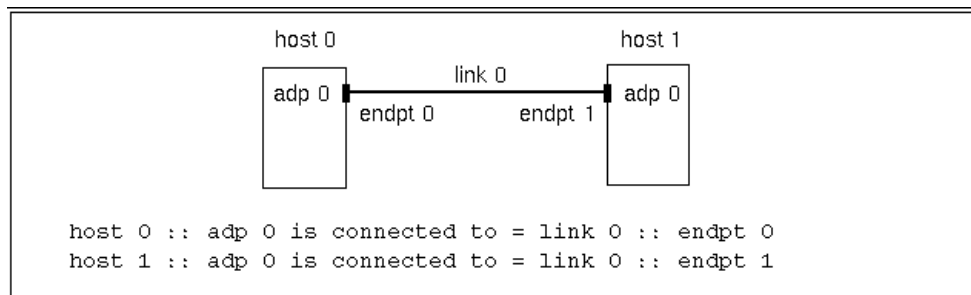


Figure 3–2 Two-Node Nonstriped Configuration

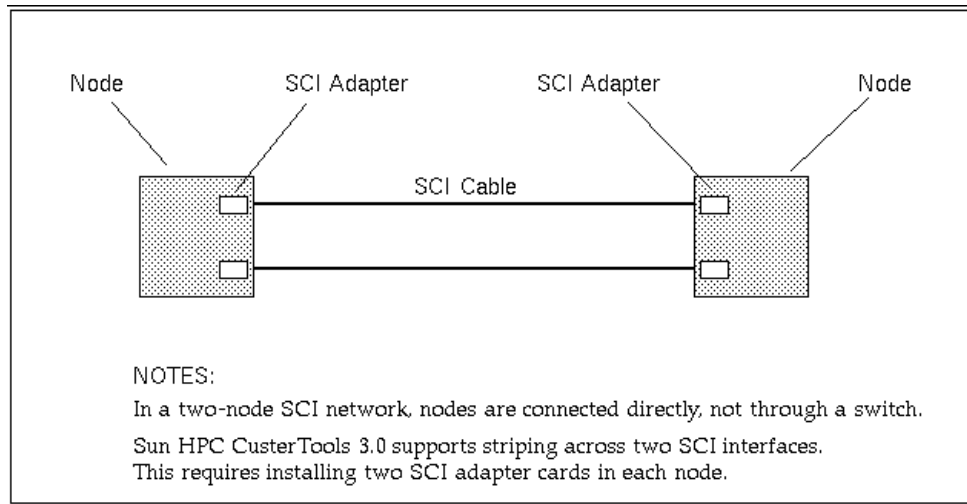


Figure 3-3 Two-Node, Striped Configuration

The format for describing switched connections is slightly different.

```
host n :: adp n is connected to = switch
n :: port n
```

Here, an adapter is connected to port n of switch n. Figure 3-4 through Figure 3-7 show examples of this format.

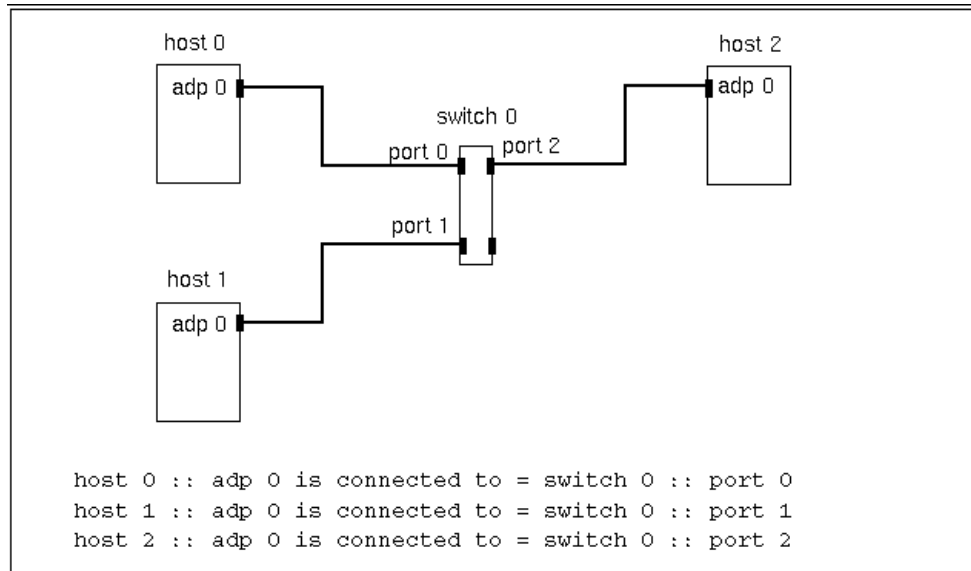


Figure 3-4 Three-Node, Nonstriped Configuration

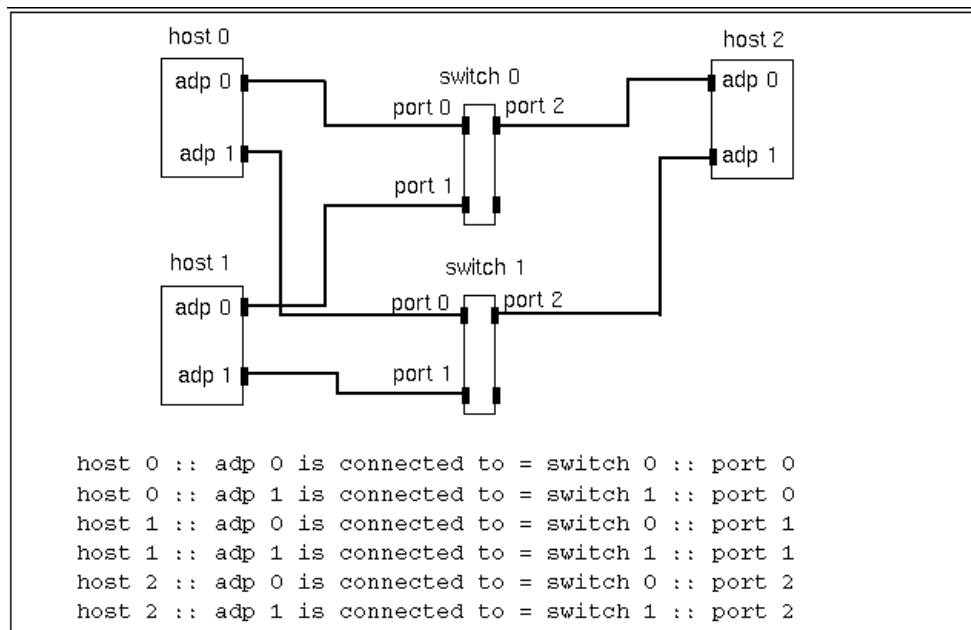


Figure 3-5 Three-Node, Striped Configuration

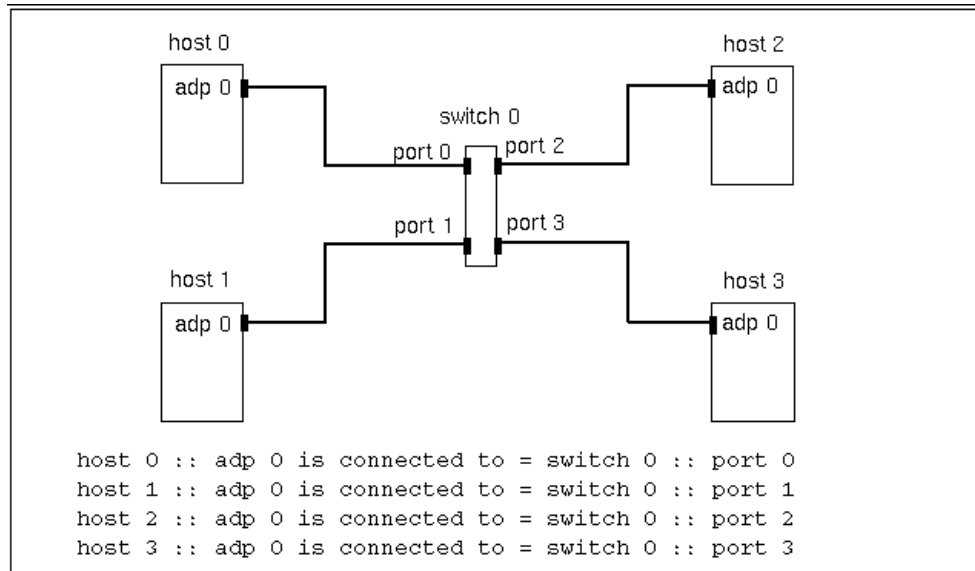


Figure 3-6 Four-Node, Nonstriped Configuration

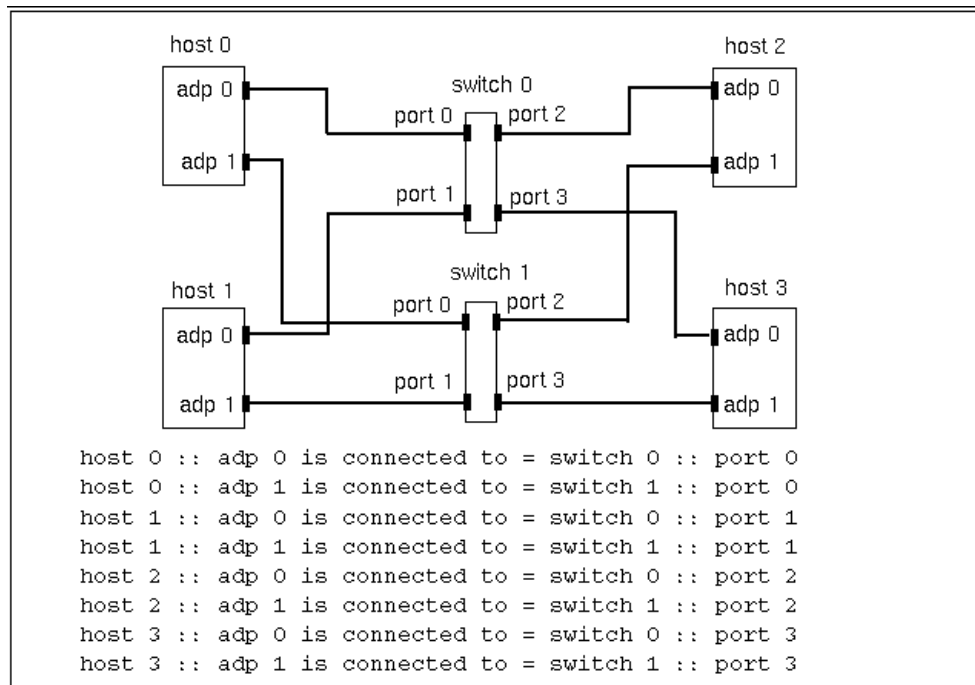


Figure 3-7 Four-Node, Striped Configuration

Adapter ID values are assigned automatically by the device driver. Initially, the device driver assigns ID 0 to the adapter installed in the lowest-numbered SBus slot,

ID 1 to the adapter in the next higher-numbered slot, and so forth. ID value assignments are always consecutive, even if the adapters are not installed in adjacent slots. Consequently, if adapter cards are not installed in adjacent slots, adapter ID values do not necessarily correspond to SBus slot numbers.

Once an ID value is assigned, it will not be reassigned, even if the adapter to which it was originally assigned is removed.

Because a ID values do not necessarily correspond to SBus slot numbers, the adapter values that you assign in the `sci_config.hpc` file may not match the actual ID assignments made by the device driver. For this reason, you may need to revise this adapter connection description (that is, the contents of Section 7) to match the actual adapter connections.

This is why you were advised to make the temporary map of the physical network layout. Instructions for ensuring that the `sci_config.hpc` file matches the actual adapter ID values are provided in “Compare `sm_config` Output With Contents of `sci_config.hpc`” on page 28

Section 7 – Network IP Addresses

Specify the first three octets of the IP address of each link or switch. For example, for a two-node, striped configuration,

```
Network IP address for Link 0 = 204.71.29
Network IP address for Link 1 = 204.71.15
```

An example of a four-node, nonstriped configuration might be

```
Network IP address for Switch 0 = 204.101.30
```

Section 8 – Netmask Value

Specify the netmask to be used for the private SCI subnet. For example,

```
Netmask = e0
```

This netmask value will support up to eight subnets with up to 30 hosts per subnet.

Propagate the SCI Configuration

Perform the following steps on the node that contains the `sci_config.hpc` file.

Run `sm_config`

Go to `/opt/SUNWsm/bin` and run the SCI setup program, `sm_config`.

If possible, do this from a console terminal so you can see the output generated by `sm_config`. If this is not possible, examine the output in `/var/adm/messages`. The following example shows the output that would be generated for the four-node, nonstriped sample configuration shown in Figure 3-1.

```
# cd /opt/SUNWsm/bin
# sm_config -f ../sci_config.hpc
For Host #0 (node3), adapter details :-
    Adp #0 :- serial no = 6269; bus slot = 0;
For Host #1 (node4), adapter details :-
    Adp #0 :- serial no = 6520; bus slot = 0;
For Host #2 (node5), adapter details :-
    Adp #0 :- serial no = 6527; bus slot = 0;
For Host #3 (node6), adapter details :-
    Adp #0 :- serial no = 6148; bus slot = 0;
Press Return to continue:
```

Do not press Return yet. Instead, go to the next section.

Compare `sm_config` Output With Contents of `sci_config.hpc`

Compare the list of serial numbers in the `sm_config` output with the serial numbers in the temporary map you made of the actual network configuration. Verify that the adapter IDs and connection details you entered in Section 6 of `sci_config.hpc` correspond to your temporary network map. If not, change the contents of `sci_config.hpc` to correspond to physical configuration described by the temporary map.

When examining the `sm_config` output, also look for any error messages reported by `sm_config` or `sm_configd`.

If the `sm_config` output conflicts with Section 6 of the `sci_config.hpc` file, stop execution of `sm_config` (press Control-C) and correct the configuration file. Then run `sm_config` again and compare its output with `sci_config.hpc` again.

When the contents of the `sci_config.hpc` file are confirmed by the `sm_config` output, press Return to allow `sm_config` to complete execution.

Verify the Rank of the SCI Interface

Look in the file `hpc.conf` and change the default ranking of the SCI interface to give it the highest priority. That is, give it a lower number in the RANK column than any other interface listed in the file—for example, change its rank to 1.

If you don't know the `hpc.conf` file's location, do one of the following:

- LSF – If your cluster is running LSF, open the LSF file `/etc/lsf.conf`. The `LSF_CONFDIR` entry in `lsf.conf` identifies the directory containing `hpc.conf`.
- CRE – If your cluster is running the CRE, look in `/opt/SUNWhpc/conf/hpc.conf`.

Add Variable to `sci.conf` File

Add the following line to `/usr/kernel/drv/sci.conf`.

```
max-vc-number = 1024;
```

Reboot Nodes

Reboot all the nodes in the cluster.

Next, verify that the SCI network is correctly configured. Instructions for verifying the network are provided in Chapter 4.

Verify That the Network Is Functional

Perform the steps described below to verify that the SCI network functions correctly.

Run `get_ci_status`

Execute `get_ci_status` on all cluster nodes to verify interconnectivity.

Run `ifconfig --a`

Execute `ifconfig -a` to verify that all the nodes are up with the SCI daemons running.

Ping the SCI Adapter Cards

Ping each SCI network interface, using their IP addresses. This verifies that every network interface can be reached from a single point in the network.

Do All-to-All Ping

Cause each node to ping the SCI interface on every other node. This verifies that any node can reach any other node over the SCI network.

Check for RSM Daemon

Check to see that the RSM daemon, `hpc_rsm`, is running. This verifies that the RSM layer is functional.

Verify Basic Communication Functionality

Chapter 5 of the *Sun HPC ClusterTools 3.0 Installation Guide* describes procedures for verifying that the cluster can successfully perform a basic set of functions. These procedures include tests of internode communication. Execute the procedures provided in that chapter as a final verification of the SCI network.

Adding or Replacing SCI Adapter Cards

If you add or replace an SCI adapter card on a node that has already been configured by `sm_config`, perform the steps described in this chapter to initialize the new adapter card.

Note - Because the node has already been configured by `sm_config`, it should also already contain `SUNWsci` and the other SCI-related packages. If not—that is, if you will be installing these SUNW packages as part of the adapter card installation process—perform the steps described in Chapter 2 and Chapter 3 instead, but only on the node receiving the SCI-related packages and new adapter card.

Install/Replace SCI Adapter Card

Turn the node's power switch off and disconnect it from the power outlet. Note that this only needs to be done on the node receiving the new adapter card.

Check the scrubber jumper setting on the new card and change its setting if needed.

- If replacing an existing SCI adapter card without changing any other features of the network, simply give the new adapter card the same scrubber jumper setting as the one it is replacing.
- If installing a new adapter card that will be connected directly to another adapter card—not through a switch—set its scrubber jumper to the opposite of the other adapter card. That is, if the other card's scrubber jumper is ON, set this card's jumper to OFF. If the other is OFF, set this one to ON.
- If installing a new adapter card that will be connected to an SCI switch, set its scrubber jumper to ON.

If replacing an existing SCI adapter card, simply install the new card in the same SBus slot.

If adding a new SCI adapter card, install it in an available SBus slot. If possible, install the new card in the next higher slot. This will promote an orderly relationship between SBus slot numbers and network ID values, which can simplify network administration.

Connect New Adapter Card to Network

Connect the new card to the network, either directly to an SCI adapter in another node or to a port on an SCI switch. See the sample two-, three-, and four-node topologies illustrated in Figure 2-1, Figure 2-2, and Figure 2-3.

Create a Temporary Network Map

Either update the temporary network map that was made when the network was created or last changed or create a new temporary network map as described in Chapter 3. Include the new adapter card in the map, identifying its SBus location and serial number.

Run `sciconf`

Run the command `sciconf` as root on the node that contains the new SCI adapter card. This will initialize the card so it can be used by the local SCI driver software. In the following example, `sciconf` is run on a node that has been added to an existing network. This node contains a single SCI adapter whose serial number is 6269. The `sciconf` output indicates the network ID and SBus slot of this SCI adapter.

```
# /opt/SUNWsci/bin/sciconf
Dolphin SCI SBus2b card configuration script
== Doing drvconfig, please wait ...
== Doing devlinks, please wait ...
== drvconfig and devlinks completed ...
== 1 SCI card was detected in the system
== Programming SCI card #1 with nodeid, adapter #, and firmware ...
== this takes 20 seconds, please wait ...
```

(continued)

```

== Programming is done ... Please verify the following information:
sciadm $Revision: 2.30 $
DOLPHIN SBus card (SBus2b) found in SBus slot 0 on Board# 0, card slot 0.
Adapter number:          0
NodeId:                  220 (0xdc)
Slot Number:             0 (0x00)
System Board Number:     0 (0x00)
Card Slot Number:        0 (0x00)
Adapter Serial Number:   6269
Fcode Version:           Revision: 2.18
Fcode checksum:          0x8906bca8
Basic FPGA checksum:     0xaf0c0c7d
Alternative FPGA checksum: 0xaf0c0c7d
== Programming done ... remember to reboot the system when finished

```

Update sci_config.hpc

Update the /opt/SUNWsmi/sci_config.hpc file to incorporate the information provided by the sciconf output.

Run sm_config

Execute the sm_config command as root on the node that contains the sci_config.hpc file.

If possible, do this from a console terminal so you can see the output it generates. If this is not possible, examine the output in /var/adm/messages. The following example shows the output that would be generated for the four-node, nonstriped sample configuration shown in Figure 3-1.

Note - Do not press Return when the sm_config output pauses. Go to the next section first.

```

# /opt/SUNWsmi/bin/sm_config -f ../sci_config.hpc
For Host #0 (node3), adapter details :-
    Adp #0 :- serial no = 6148; bus slot = 0;
For Host #1 (node4), adapter details :
    Adp #0 :- serial no = 6520; bus slot = 0;

```

```
For Host #2 (node5), adapter details :  
-      Adp #0 :- serial no = 6527; bus slot = 0;  
For Host #3 (node6), adapter details :-  
      Adp #0 :- serial no = 6269; bus slot = 0;  
Press Return to continue:
```

Confirm `sci_config.hpc` Contents

Compare the list of serial numbers in the `sm_config` output with the serial numbers in the temporary network map.

If the `sm_config` output conflicts with Section 6 of the `sci_config.hpc` file, stop execution of `sm_config` (press Control-C) and correct the configuration file. Then run `sm_config` again and compare its output with `sci_config.hpc` again.

When the contents of the `sci_config.hpc` file are confirmed by the `sm_config` output, press Return to allow `sm_config` to complete execution.

Reboot the Node

The `sm_config` output will also identify which nodes need to be rebooted. Reboot those nodes.

Verify the New Network

Perform the network verification steps described in Chapter 4 to be certain that the network still functions correctly.

SCI Interface Troubleshooting

SCI Switch

General Hardware Inspection

Perform the following checks to determine the physical state of various SCI subsystem components. Verify that:

- All SCI scrubber jumpers are properly set, depending on the cluster topology.
- All SCI cables are properly seated.
- All SCI switches have power applied
- No SCI status LEDs are red—see Table 6-1 and Table 6-2

SCI Switch Status LED Locations

Clusters with three or four nodes can be connected through one or two SCI switches. The switch status LEDs provide information that can be used to troubleshoot SCI switch failures (Figure 6-1). Guidelines for interpreting these LEDs are provided in “Port Status LEDs” on page 38” and “General Switch Status LED” on page 39.

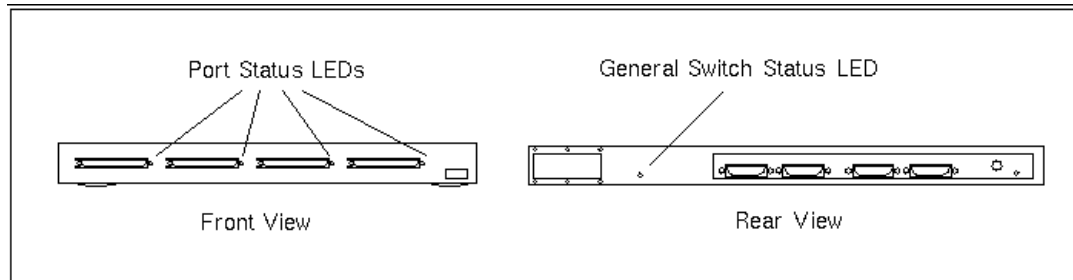


Figure 6-1 SCI Status LED Locations

Port Status LEDs

The four port status LEDs located on the switch front panel can be used to troubleshoot individual port failures (Table 6-1).

Note - A switch port sync error can result from a cable being removed.

TABLE 6-1 SCI Switch Port Status LEDs

Situation	Port LED Status
No power	All four LEDs not lit
Fatal switch errors: fatal hardware error, temperature too high, fan(s) not operative, power supply problem	All four LEDs red
Port errors: SCI cable out, sync error	Associated port LED is red
Port operative, no transactions	Associated port LED is green
Port operative, with transactions	Associated port LED is blinking green

General Switch Status LED

The switch status LED located on the rear panel indicates overall switch failures (Table 6-2).

TABLE 6-2 SCI Switch Rear Panel LED

Situation	LED Status
Fatal switch errors: fatal hardware error, temperature too high, fan(s) not operative, power supply problem	Red
Switch operational	Green

The get_ci_status Command

You can use the results of the `get_ci_status` command to troubleshoot clusters that have SCI switches. For example, for the configuration in Figure 6-2, if the `get_ci_status` command is used on `interconn1`, a typical output would be:

```
# /opt/SUNWsm/bin/get_ci_status
sma: sci #0: sbus_slot# 1; adapter_id 8 (0x08); ip_address 1; switch_id# 0; port_id# 0; Adapter Status - UP; Link Status - UP
sma: Switch_id# 0
sma: port_id# 1: host_name = interconn2; adapter_id = 72; active | operational
sma: port_id# 2: host_name = interconn3; adapter_id = 136; active | operational
sma: port_id# 3: host_name = interconn4; adapter_id = 200;inactive|inoperational
#
```

In this example, the line

```
sma: port_id# 3: host_name = interconn4; adapter_id = 200;inactive|inoperational
```

indicates that the path between SCI switch 0, port 3 and `interconn4` is inactive and not operational.

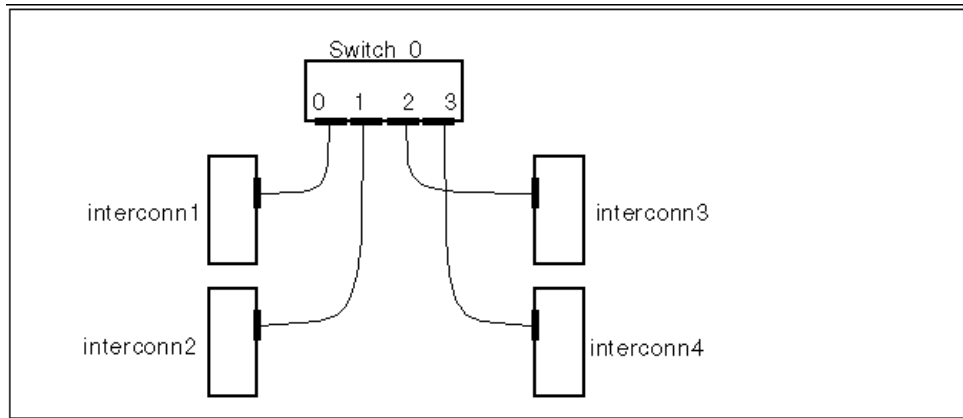


Figure 6-2 Typical Four-Node Configuration with an SCI Switch

In this instance, if the `get_ci_status` command were run on all four nodes, and if the same path was inactive and inoperative between SCI switch 0, port 3 and interconn4, it is more than likely that either the SCI switch 0, port 3, the cable, or the interconn4 host adapter is faulty.

However, if the `get_ci_status` command indicates that the same path is inactive and inoperative for one node only, such as in the instance of interconn1, then it is more than likely that either the interconn 1 host adapter, the cable, or SCI switch 0, port 0 is faulty.

Note that some aspects of the `get_ci_status` command output, such as host names, will vary according to your configuration.

Client Net Failure

System console messages will identify the specific port that has failed. Otherwise, for information on test commands as well as additional troubleshooting, refer to the documentation that came with your client network interface card.

Incorrect Software Configuration

Make sure that:

- The working copy of the `sm_config` template file correctly matches the hardware configuration and cluster topology.
- `sm_config` ran successfully on only one of the cluster nodes.
- All nodes were rebooted after `sm_config` was executed.

Incorrect Firmware

If an SCI adapter card is loaded with the wrong firmware, the SCI cards will not be detected upon system power-on or reboot/reset.

Improper loading of the firmware can happen two ways:

- Old firmware programmed into new SBus2b cards
- New firmware programmed into old SBus2 cards

If proper firmware is loaded, a banner (containing the word `FCode`) will be printed from each SCI card *twice* during power-on or reboot or reset. No banner will be printed at all for a card loaded with improper firmware.

The following are sample console messages (which are not saved in the message file):

1. One SCI card is working in the node:

```
rebooting...
Resetting ...

DOLPHIN SBus-to-SCI (SBus2b) Adapter - 9029, Serial #5017
FCode 9029 $Revision: 2.3 $ - d9029_52 $Date: 1996/10/
30 07:47:53 $

Executing SCI adapter selftest.    Adapter OK.
screen not found.
Can't open input device.
Keyboard not present.  Using ttya for input and output.

DOLPHIN SBus-to-SCI (SBus2b) Adapter - 9029, Serial #5017
FCode 9029 $Revision: 2.3 $ - d9029_52 $Date: 1996/10/
30 07:47:53 $

Executing SCI adapter selftest.    Adapter OK.

Sun Ultra 1 SBus (UltraSPARC 167MHz), No Keyboard
```

2. No SCI cards are working in the node:

```
rebooting...
Resetting ...

screen not found.
Can't open input device.
Keyboard not present.  Using ttya for input and output.

Sun Ultra 1 SBus (UltraSPARC 167MHz), No Keyboard
```

(continued)

Note - If SCI cards do not show up during boot time, check the physical installation of the cards. If reseating the cards does not correct the problem, the SCI cards may be damaged and should be returned.

If you suspect that an SCI SBus interface card is loaded with the wrong firmware, perform the following steps to investigate:

1. **With the system powered off, note the serial numbers of the adapter cards that are physically installed.**
2. **Turn the system power on.**
3. **Run `/opt/SUNWsci/bin/sciadm` and enter the `identify` command.**
This command displays the firmware version, fcode version, and serial number of each adapter board found.
4. **Compare the number of cards found by `sciadm` against the number of adapters physically installed.**
Two cards should be displayed in the output. If not, there is at least one bad card in the system.
5. **Compare the adapter board serial numbers from the output of the `identify` command, to the serial number on each adapter card physically installed.**
Note which serial number(s) are displayed. Cards that do not have their serial numbers displayed are bad and need replacement.

Man Pages

This appendix contains man pages for:

- `sm_config`
- `get_ci_status`.

`sm_config(1)`

CODE EXAMPLE A-1 `sm_config(1)` man Page

```
sm_config(1M)           Maintenance Commands           sm_config(1M)

NAME
    sm_config - SCI adapter configuration utility for clusters

SYNOPSIS
    sm_config [-t] -f filename

AVAILABILITY
    SUNWsmc

INTERFACE CLASSIFICATION
    Sun Private

DESCRIPTION
    sm_config is the SCI adapter configuration utility. It acts
    as a client of sm_configd(1M) daemon. sm_config contacts
    the sm_configd(1M) daemon on all the hosts and works in a
    distributed fashion to retrieve the adapter inventory and
    configure the adapter cards on these hosts. The configura-
```

(continued)

tion process involves programming -
 (a) the adapter Node-Ids into the adapter's flash memory
 and
 (b) IP addresses into the cards.

Upon successful completion, a configuration file named /etc/sma.config is installed on all the hosts in the cluster. This file contains a snapshot view of the cluster members, switches, adapters etc. It also installs a file called /etc/sma.ip which contains the IP addresses of all the SCI interfaces in the cluster.

OPTIONS

-t starts sm_config in debug mode.

-f filename takes the filename as an input file. The input file template is available in /opt/SUNWsma/bin directory as (a) template.pdb (for PDB clusters) and (b) template.hpc (for HPC clusters). These template files provide detailed information about the type of information required by sm_config.

This input file template contains 8 sections -

1. Cluster configuration section - specifies the type of cluster being configured (PDB or HPC). A sample template for this section -

Cluster is configured as = PDB

2. Host names section - requires the names of all the hosts in the cluster. If the hosts in the cluster do not have full public-net connectivity then the name of the host without connectivity must be preceded by ``_%``. This indicates to sm_config not to contact this host via the public-net.

For example, consider a case where host2 in a cluster of host1, host2, host3 and host4 lacks public-net connectivity. When sm_config is started with the following template_1, it will contact host1, host3 and host4 over the net and configure their SCI interfaces. However, it is now the user's responsibility to run sm_config on host2 in the stand-alone mode using template_2 below.

template_1 - used on host1, host3 and host4 :-

```
HOST 0 = host1
HOST 1 = _%host2
HOST 2 = host3
HOST 3 = host4
```

template_2 - used on host2 :-

```
HOST 0 = _%host1
```

(continued)

```

HOST 1 = host2
HOST 2 = _%host3
HOST 3 = _%host4

```

A caveat to keep in mind when running `sm_config` in stand-alone mode is that, `sm_config` cannot guarantee the coherency of the `/etc/sma.config` generated during the different invocations (for eg. in the above case - `/etc/sma.config` on `host2` versus the ones on `host1`, `host3` and `host4`) if the user were to supply inconsistent input data for the two cases.

3. Number of Switches section - Accepts input for the total no. of switches in the cluster.

However, if the cluster being configured has some unused SCI adapters meant for use in the future, then the cluster should be configured as it would look in the future, when all the adapters are fully connected. For instance, a 1-switch cluster containing 4 hosts with 2 adapters on each (second set of adapters idle), which will later evolve into a 2-switch cluster should be configured as a 2-switch cluster.

This ensures that when the cluster evolves to its final form in future, new communication channels (SMA sessions) will be created on the new links (say, through a new switch) on the fly. This eliminates having to run `sm_config` later and rebooting the machine. A detailed example of this is given in the input template file.

A sample template for this section -

```

Number of Switches in cluster = 2

```

4. Number of Direct Links section - Accepts input for the total no. of direct SCI links in the cluster. A sample template for this section -

```

Number of Direct Links in cluster = 2

```

5. Allow Rings section - Whether the cluster supports configurations with multiple hosts connected to the same port of the same switch. A sample template for this section -

```

Allow Rings in cluster (Y/N)? = N

```

6. Adapter information section - Accepts detailed information for each adapter on each host. A sample template for this section is -

```

host 0 :: adp 0 is connected to = switch 0 :: port 0
host 0 :: adp 1 is connected to = link 1 :: endpt 0

```

7. Network IP address section - Accepts the first 3 octets (network) of the IP address for a particular switch or link.

A sample template for this section -

```

Network IP address for Switch 0 = 204.152.65

```

(continued)

8. Netmask section - Accepts the netmask for the private SCI sub-nets. For example, a cluster with less than 15 hosts per switch should select a netmask of 0xf0 while a cluster with 15 hosts or more but less than 31 hosts would choose 0xe0. A sample template for this section -
 Netmask = f0

USAGE

The root user can start `sm_config` from the command line using the `-f` option to provide an input file to it.

A cluster can have 3 topologies -

- (i) Switched - All hosts are connected to each other via SCI switches. Can have 2 or more hosts.
- (ii) Non-switched - Two hosts connected directly via SCI cables (direct links). Has exactly 2 hosts.
- (iii) Hybrid - Contains switches and direct links. Can have 2 or more hosts.

NOTE - At this point, PDB does not support more than two switches in a cluster (see (i) above), nor does it support case (iii) from above.

NOTE - `sm_config` can be run on any host in the cluster, but it should not be run on multiple hosts simultaneously (eg. via `cconsole`). If this occurs, the results are unpredictable - in the worst case, the adapter flash memory might get programmed with corrupt data.

NOTE - After running `sm_config`, the system should be rebooted.

FILES

`/opt/SUNWcluster/bin/sm_config`
`/etc/sma.config`
`/etc/sma.ip`

SEE ALSO

`sm_configd(1M)`

DIAGNOSTICS

`sm_config` prints error and warning messages to `stderr`. If a fatal error occurs on any host or locally where `sm_config` is running, then the process is aborted and no `/etc/sma.config` is generated till the error is rectified. Do not reboot the machine till a successful run of `sm_config` has been completed.

RELEASE NOTES -

If `nis+` is being used as the name service then the default behaviour is to look up the global `nis+` map but if that doesn't exist, the local `/etc/services` file is not searched. This behaviour is different from the default `nis` behaviour. In this scenario `inetd` will be unable to start the

(continued)


```
sm_configd daemon.
```

```
SunOS 5.5.1      Last change: 30 March 1997
```

get_sci_status(1m)

CODE EXAMPLE A-2 get_sci_status(1m) man Page

```
get_ci_status(1M)      Maintenance Commands      get_ci_status(1M)
```

NAME

get_ci_status - Displays the Cluster configuration, the SCI adapter status and the SMA session status.

SYNOPSIS

```
get_ci_status [ -l ]
```

AVAILABILITY

SUNWsma

INTERFACE CLASSIFICATION

Sun Private

DESCRIPTION

get_ci_status displays the cluster configuration, the SCI adapter status and the SMA session status. It queries the SCI driver for information about the local SCI adapters and tests the connectivity to SCI adapters on other hosts, either via a switch or a direct link.

For each adapter in the cluster, get_ci_status displays the host it is on, the port on a switch it is connected to, its adapter-id and whether the local adapters can communicate with the adapters on other hosts.

In addition, for each local adapter get_ci_status displays the SBus slot# it is attached to, the host part of its IP address and whether the adapter is functional.

OPTIONS

```
-l      Displays the local SCI adapter status only.
```

(continued)

no option

Displays the local SCI adapter status and global cluster status. When displaying the global status, get_ci_status reports whether the remote adapter can be reached at the hardware level (via SCI_PROBES) and/or at the software session level (via SMA sessions).

The SCI Probe reachability is indicated by active or inactive keywords following the status for the remote adapter in question. The software SMA session reachability is indicated by operational or inoperational keywords. For example, an output of the following form -

```
sma: sci #0: sbus_slot# 1; adapter_id 8 (0x08); ip_address
1; switch_id# 0; port_id# 0; Adapter Status - UP; Link
Status - UP
sma: sci #1: sbus_slot# 2; adapter_id 12 (0x0c); ip_address
17; switch_id# 1; port_id# 0; Adapter Status - UP; Link
Status - UP
sma: Switch_id# 0
sma: port_id# 1: host_name = interconn2; adapter_id = 72;
active | operational
sma: port_id# 2: host_name = interconn3; adapter_id = 136;
active | operational
sma: port_id# 3: host_name = interconn4; adapter_id = 200;
active | operational
sma: Switch_id# 1
sma: port_id# 1: host_name = interconn2; adapter_id = 76;
active | inoperational
sma: port_id# 2: host_name = interconn3; adapter_id = 140;
inactive | operational
sma: port_id# 3: host_name = interconn4; adapter_id = 204;
inactive | inoperational
```

indicates that there are 2 local adapters (adapter_id 8 and 12), both of which are functioning OK (keyword UP) with respect to SCI Probes to themselves. In case, a local adapter is unable to complete a successful SCI Probe to itself, the status of that local adapter is shown as DOWN.

The global status is shown in the set of lines associated with a switch. The status of the communication channel from the local adapter (adapter_id 8) to remote adapters on interconn2, interconn3 and interconn4 via the first switch (Switch_id# 0) is - SCI Probe status OK (keyword active) and SMA sessions functional (keyword operational).

However, the status of the communication channels from the local adapter (adapter_id 12) to remote adapters via the second switch (Switch_id #1) have the following problems -

1. Adapter_id 76 => SCI Probes - reachable (keyword active) and SMA session - not established (keyword inoperational)

(continued)

2. Adapter_id 140 => SCI Probes - unreachable (keyword inactive) and SMA session - established (keyword inoperational). This is a brief transitional stage.

3. Adapter_id 204 => SCI Probes - unreachable (keyword inactive) and SMA session - not established (keyword inoperational)

USAGE

get_ci_status can be run from the command line by any user. However, it can only be run after the adapter cards have been initialized using sm_config(1M). This ensures that all the adapter node-ids have been properly programmed and the configuration file /etc/sma.config exists.

FILES

/opt/SUNWsma/bin/get_ci_status
/etc/sma.config

SEE ALSO

sm_config(1M),

DIAGNOSTICS

get_ci_status prints error and warning messages to stderr.

SunOS 5.5.1

Last change: 30 March 1997

SMA Messages

If the HPC cluster uses an SCI switch, the SMA software can display the following messages.

Message Directory

The following pages list SMA messages in alphabetical order:

- SUNWcluster.sma.smactl.4007 - Cannot create logical adapter: None found
- SUNWcluster.sma.smactl.4008 - Cannot create logical adapter:no response
- SUNWcluster.sma.smad.1030 - \$clustername adapter \$adp selected
- SUNWcluster.sma.smad.1101 - smad(\$pid): entering stand-alone mode
- SUNWcluster.sma.smad.1102 - smad: Cluster
- SUNWcluster.sma.smad.1103 - smad: Cluster
- SUNWcluster.sma.smad.1104 - smad: Cluster
- SUNWcluster.sma.smad.1105 - smad: Cluster
- SUNWcluster.sma.smad.3104 - SCI Adapter \$adp (\$from_aid): Session to
- SUNWcluster.sma.smad.4004 - smad(\$pid): exiting by request
- SUNWcluster.sma.smad.4030 - \$clustername no adapter available to select
- SUNWcluster.sma.smad.5010 - \$clustername adapter \$adp de-selected
- SUNWcluster.sma.smak.1001 - SCI Adapter \$adp: Card operational
- SUNWcluster.sma.smak.1051 - SCI Adapter \$adp: Link operational
- SUNWcluster.sma.smak.3001 - SCI Adapter \$adp (\$from_aid): Session to

- SUNWcluster.sma.sma.3002 - SCI Adapter \$adp (\$from_aid): Session to
- SUNWcluster.sma.sma.3003 - SCI Adapter \$adp (\$from_aid): Session to
- SUNWcluster.sma.sma.4001 - SCI Adapter \$adp: Card not operational
- SUNWcluster.sma.sma.4051 - SCI Adapter \$adp: Link not operational
- SUNWcluster.sma.watchdog.1002 - smad(\$pid) started
- SUNWcluster.sma.watchdog.2001 - child (\$pid) \$exit died \$status
- SUNWcluster.sma.watchdog.4002 - smad (\$pid) exiting

SMA Messages

CODE EXAMPLE B-1 SMA Messagees

```
#####
# SUNWcluster.sma.watchdog.1002 - smad($pid) started
#####
msgid  ``SUNWcluster.sma.watchdog.1002.message``
msgstr  ``SMAD is starting.``
msgid  ``SUNWcluster.sma.watchdog.1002.error``
msgstr  ``The process-id is mentioned in brackets.``
msgid  ``SUNWcluster.sma.watchdog.1002.fix``
msgstr  ``Not Applicable. ``

#####
# SUNWcluster.sma.watchdog.2001 - child ($pid) $exit died $status
#####
msgid  ``SUNWcluster.sma.watchdog.2001.message``
msgstr  ``The SMAD child daemon is dead. If necessary, another SMAD child``
        ``daemon will be started up automatically.``
msgid  ``SUNWcluster.sma.watchdog.2001.error``
msgstr  ``The process-id exit code and exit status are mentioned in the``
        ``message.``
msgid  ``SUNWcluster.sma.watchdog.2001.fix``
msgstr  ``Not Applicable. ``

#####
# SUNWcluster.sma.watchdog.4002 - smad ($pid) exiting
#####
msgid  ``SUNWcluster.sma.watchdog.4002.message``
msgstr  ``The SMAD watchdog parent daemon is also exiting, because of a``
        ``request to exit, or a fatal error.``
msgid  ``SUNWcluster.sma.watchdog.4002.error``
msgstr  ``The process-id of the parent daemon is in brackets.``
msgid  ``SUNWcluster.sma.watchdog.4002.fix``
msgstr  ``Not Applicable. ``

#####
# SUNWcluster.sma.sma.4004 - smad($pid): exiting by request
```

(continued)

```

#####
msgid    ``SUNWcluster.sma.smad.4004.message``
msgstr   ``The SMAD child daemon is dying by request.``
          ``request to exit, or a fatal error.``
msgid    ``SUNWcluster.sma.smad.4004.error``
msgstr   ``Probably a due to a shutdown or a pkgm.``
msgid    ``SUNWcluster.sma.smad.4004.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smak.1001 - SCI Adapter $adp: Card operational
#####
msgid    ``SUNWcluster.sma.smak.1001.message``
msgstr   ``The adapter is working.``
msgid    ``SUNWcluster.sma.smak.1001.error``
msgstr   ``Not Applicable. ``
msgid    ``SUNWcluster.sma.smak.1001.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smak.1051 - SCI Adapter $adp: Link operational
#####
msgid    ``SUNWcluster.sma.smak.1051.message``
msgstr   ``The link from the mentioned adapter is working.``
msgid    ``SUNWcluster.sma.smak.1051.error``
msgstr   ``Not Applicable. ``
msgid    ``SUNWcluster.sma.smak.1051.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smak.4001 - SCI Adapter $adp: Card not operational
#####
msgid    ``SUNWcluster.sma.smak.4001.message``
msgstr   ``The adapter is not working.``
msgid    ``SUNWcluster.sma.smak.4001.error``
msgstr   ``Not Applicable. ``
msgid    ``SUNWcluster.sma.smak.4001.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smak.4051 - SCI Adapter $adp: Link not operational
#####
msgid    ``SUNWcluster.sma.smak.4051.message``
msgstr   ``The link from the mentioned adapter is not working.``
msgid    ``SUNWcluster.sma.smak.4051.error``
msgstr   ``Not Applicable. ``
msgid    ``SUNWcluster.sma.smak.4051.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smad.1030 - $clustername adapter $adp selected
#####
msgid    ``SUNWcluster.sma.smad.1030.message``
msgstr   ``The mentioned adapter has been selected to act as the``

```

(continued)

```

        ``logical adapter.``
msgid    ``SUNWcluster.sma.smad.1030.error``
msgstr   ``Not Applicable. ``
msgid    ``SUNWcluster.sma.smad.1030.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smad.4030 - $clustername no adapter available to select
#####
msgid    ``SUNWcluster.sma.smad.4030.message``
msgstr   ``There is no adapter available which can act as the logical adapter.``
msgid    ``SUNWcluster.sma.smad.4030.error``
msgstr   ``This indicates significantly broken connectivity.``
msgid    ``SUNWcluster.sma.smad.4030.fix``
msgstr   ``It is likely that the cluster will reconfigure itself. If the cluster``
        ``does not reconfigure, administrator needs to take some corrective``
        ``action, such as aborting a node or fixing the link problems.``

#####
# SUNWcluster.sma.smad.5010 - $clustername adapter $adp de-selected
#####
msgid    ``SUNWcluster.sma.smad.5010.message``
msgstr   ``The mentioned adapter has been de-selected and is no longer the``
        ``logical adapter.``
msgid    ``SUNWcluster.sma.smad.5010.error``
msgstr   ``This adapter was acting as the logical adapter and it is no longer``
        ``the logical adapter. If recovery happens, some other adapter will be``
        ``chosen to act as the logical adapter.``
msgid    ``SUNWcluster.sma.smad.5010.fix``
msgstr   ``Not Applicable. ``

#####
# SUNWcluster.sma.smad.1101 - smad($pid): entering stand-alone mode
#####
msgid    ``SUNWcluster.sma.smad.1101.message``
msgstr   ``SMAD is starting up and things seem to be fine. No PDB software is ``
        ``running.``
msgid    ``SUNWcluster.sma.smad.1101.error``
msgstr   ``Not Available.``
msgid    ``SUNWcluster.sma.smad.1101.fix``
msgstr   ``Not Applicable.``

#####
# SUNWcluster.sma.smad.1102 - smad: Cluster '$clustername' monitoring
#####
msgid    ``SUNWcluster.sma.smad.1102.message``
msgstr   ``CMM has asked SMAD to monitor the cluster status and it is now being ``
        ``monitored. It can also happen if the SMAD was already in 'monitor' ``
        ``mode and had died and restarted.``
msgid    ``SUNWcluster.sma.smad.1102.error``
msgstr   ``Not Available.``
msgid    ``SUNWcluster.sma.smad.1102.fix``
msgstr   ``Not Applicable.``

#####

```

(continued)


```

# SUNWcluster.sma.smad.1103 - smad: Cluster '$clustname' running
#####
msgid ``SUNWcluster.sma.smad.1103.message``
msgstr ``CMM has informed SMAD of the cluster membership. SMAD continues to ``
      ``monitor the entire cluster. It can also happen if the SMAD was ``
      ``already in 'cluster' mode and had died and restarted.``
msgid ``SUNWcluster.sma.smad.1103.error``
msgstr ``Not Available.``
msgid ``SUNWcluster.sma.smad.1103.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smad.1104 - smad: Cluster '$clustname' returning
#####
msgid ``SUNWcluster.sma.smad.1104.message``
msgstr ``The cluster is entering the 'return' transition. SMAD does not ``
      ``manage the logical adapter in this state, until 'step1' (running) ``
      ``occurs. SMAD will continue to monitor the cluster for physical ``
      ``failures. It can also happen if the SMAD was already in 'return' ``
      ``mode and had died and restarted.``
msgid ``SUNWcluster.sma.smad.1104.error``
msgstr ``Not Available.``
msgid ``SUNWcluster.sma.smad.1104.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smad.1105 - smad: Cluster '$clustname' no longer running
#####
msgid ``SUNWcluster.sma.smad.1105.message``
msgstr ``This node has left the cluster. CMM has informed SMAD of this. The ``
      ``cluster is no longer being monitored. It can also happen if the SMAD ``
      ``had died and restarted.``
msgid ``SUNWcluster.sma.smad.1105.error``
msgstr ``Not Available.``
msgid ``SUNWcluster.sma.smad.1105.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smad.3104 - SCI Adapter $adp ($from_aid): Session to
#                               $to_aid not open
#####
msgid ``SUNWcluster.sma.smad.3104.message``
msgstr ``The session from adapter ($adp), whose SCI id is $from_aid, to the ``
      ``SCI adapter with SCI id $to_aid could not be opened.``
msgid ``SUNWcluster.sma.smad.3104.error``
msgstr ``It will be opened when possible, later.
msgid ``SUNWcluster.sma.smad.3104.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smak.3001 - SCI Adapter $adp ($from_aid): Session to
#                               $to_aid active
#####
msgid ``SUNWcluster.sma.smak.3001.message``

```

(continued)

```

msgstr ``The session from adapter ($adp), whose SCI id is $from_aid, to the ``
msgstr ``SCI adapter with SCI id $to_aid is active.``
msgid ``SUNWcluster.sma.smak.3001.error``
msgstr ``Not Available.``
msgid ``SUNWcluster.sma.smak.3001.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smak.3002 - SCI Adapter $adp ($from_aid): Session to
# $to_aid closed
#####
msgid ``SUNWcluster.sma.smak.3002.message``
msgstr ``The session from adapter ($adp), whose SCI id is $from_aid, to the ``
msgstr ``SCI adapter with SCI id $to_aid has been closed.``
msgid ``SUNWcluster.sma.smak.3002.error``
msgstr ``This is probably because a node is leaving the cluster, or ``
msgstr ``rebooting.``
msgid ``SUNWcluster.sma.smak.3002.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smak.3003 - SCI Adapter $adp ($from_aid): Session to
# $to_aid failed
#####
msgid ``SUNWcluster.sma.smak.3003.message``
msgstr ``The session from adapter ($adp), whose SCI id is $from_aid, to the ``
msgstr ``SCI adapter with SCI id $to_aid has been closed.``
msgid ``SUNWcluster.sma.smak.3003.error``
msgstr ``This is probably because of a failure, or a shutdown of the remote ``
msgstr ``node.``
msgid ``SUNWcluster.sma.smak.3003.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smactl.4007 - Cannot create logical adapter: None found
#####
msgid ``SUNWcluster.sma.smactl.4007.message``
msgstr ``The unavailability of a logical adapter has been detected at step1 ``
msgstr ``of the reconfiguration process. This node will abort itself out of ``
msgstr ``the cluster.``
msgid ``SUNWcluster.sma.smactl.4007.error``
msgstr ``Not Available.``
msgid ``SUNWcluster.sma.smactl.4007.fix``
msgstr ``Not Applicable.``

#####
# SUNWcluster.sma.smactl.4008 - Cannot create logical adapter:no response
# from SMAD
#####
msgid ``SUNWcluster.sma.smactl.4008.message``
msgstr ``SMAD has not responded to a request to create a logical adapter.It ``
msgstr ``is likely that SMAD has aborted / died in a fatal manner.``
msgid ``SUNWcluster.sma.smactl.4008.error``

```

(continued)

```
msgstr ``Not Available.``  
msgid ``SUNWcluster.sma.smactl.4008.fix``  
msgstr ``Not Applicable.``
```