CXFS^{TM} MultiOS for CXFS Client-Only Nodes: Installation and Configuration Guide

007–4507–006

CONTRIBUTORS

Written by Lori Johnson

Edited by Susan Wilkening

Illustrated by Chrystie Danzer

Production by Glen Traefald

Engineering contributions to the book by Vlad Apostolov, Neil Bannister, Dale Brantly, David Chatterton, Mark Cruciani, Tad Dolphay, Dave Ellis, Eric Eppe, Andrew Gildfind, Dennis Kender, Aaron Mantel, Troy McCorkell, Ken McDonell, Terry Merth, Daniel Moore, Max Matveev, Barry Naujok, Tim Sirianni, Michael Umansky, Madan Valluri, Geoffrey Wehrman

COPYRIGHT

LIMITED RIGHTS LEGEND

TRADEMARKS AND ATTRIBUTIONS

Cover design by Sarah Bolles, Sarah Bolles Design, and Dany Galgani, SGI Technical Publications.

New Features in this Guide

This guide contains the following new features:

- Support for Sun Microsystems Solaris 9 and specific Sun Fire systems. See "Requirements Specific to Solaris", page 24.
- Support for the JNI EZ Fibre release 2.2.1 or later. See "Installing and Running the EZ Fibre Configuration GUI", page 30. If you are running an earlier version of the EZ Fibre tool, you should upgrade to release 2.2.1 or later.
- A cluster of as many as 32 nodes, of which as many as 16 can be CXFS administration nodes; the rest will be client-only nodes. See "Requirements", page 6.
- When you define a node, you no longer need to specify the node weight. This has been replaced by the **Node Function** field. For Solaris and Windows nodes, **Client-Only** is automatically selected for you. Similar fields are provided for the `cmgr`(1M). For more information, see the *CXFS Version 2 Software Installation and Administration Guide*.
- Clarification that if the primary HBA path is at fault during the Windows boot up (for example, if the Fibre Channel cable is disconnected), no failover to the secondary HBA path will occur. This is a limitation of the QLogic driver. See "Requirements Specific to Windows", page 54.
- Reference to the availability of cluster information on Windows nodes. See "Windows Log Files and Cluster Status", page 55.
- Information about enabling Brocade Fibre Channel switch ports; see "Defining the Switch for I/O Fencing", page 101.
- Additional information about the following:
  - "Functional Limitations Specific to Windows", page 56
  - "Performance Considerations on a CXFS Windows Node", page 59
  - "Access Controls on a Windows Node", page 60

Record of Revision

Version	Description
001	March 2002 Original publication with the CXFS MultiOS Clients 2.0 release for IRIX 6.5.16f.
002	May 2002 Revised to support the CXFS MultiOS Clients 2.1 release for IRIX 6.5.16f. This release supports the Sun Microsystems Solaris and Microsoft Windows NT platforms.
003	June 2002 Revised to support the CXFS MultiOS Clients 2.1.1 release for IRIX 6.5.16f. This release supports the Sun Microsystems Solaris and Microsoft Windows NT platforms.
004	August 2002 Revised to support the CXFS MultiOS 2.2 Clients release for IRIX 6.5.17f. This release supports the Sun Microsystems Solaris, Microsoft Windows NT, and Microsoft Windows 2000 platforms.
005	November 2002 Revised to support the CXFS MultiOS Clients 2.3 release for IRIX 6.5.18f. This release supports the Sun Microsystems Solaris, Microsoft Windows NT, and Microsoft Windows 2000 platforms.
006	February 2003 Revised to support the CXFS MultiOS Clients 2.4 release for IRIX 6.5.19f. This release supports the Sun Microsystems Solaris, Microsoft Windows NT, and Microsoft Windows 2000 platforms.

Contents

Figures

About This Guide

This publication documents the CXFS MultiOS Clients 2.4 release for IRIX 6.5.19f. This release supports Sun Microsystems Solaris nodes, Microsoft Windows NT nodes, and Microsoft Windows 2000 nodes.

Prerequisites

This guide assumes the following:

- The reader is familiar with the information presented in the *CXFS Version 2 Software Installation and Administration Guide* and the operating system documentation.
- The IRIX CXFS cluster is installed and operational.
- The CXFS client-only nodes have the appropriate platform-specific operating system software installed.

Related Publications

The following documents contain additional information (if you are viewing this document online, you can click on `TPL Link` below to link to the book on the SGI TechPubs library):

- CXFS documentation:
  - Platform-specific release notes
  - *CXFS Version 2 Software Installation and Administration Guide* (`TPL link`)
- SGI TP9400 documentation :
  - *SGI TP9400 and SGI TP9500 Software Concepts Guide* (`TPL link`)
  - *SGI TP9400 and SGI TP9500 RAID Owner's Guide* (`TPL link`)
  - *SGI TP9400 and SGI TP9500 RAID Administration Guide* (`TPL link`)

  The SGI TP9400 documentation is available on the release CD in the following files:

  - `tp9400_sw_concepts_guide.pdf`

  - `tp9400_owners_guide.pdf`
  - `tp9400_admin_guide.pdf`
- SGI TP9100:
  - *TPM Installation Instructions and User's Guide for TP9100*
- JNI host bus adapter (HBA) card and driver documentation:
  - *Installation Guide, FCE-6460 and FCE2-6460 PCI-to-Fibre Channel Host Bus Adapters (Solaris, Windows NT/2000, Novell, AIX, HP-UX, MAC-OS, Linux) JNI FibreStar*
  - *Quick Installation Guide, Solaris, AIX and Windows JNI EZ Fibre*

  Also see the JNI website at:

  `http://www.jni.com`
- QLogic HBA card and driver documentation:
  - *Hardware Installation Guide for the QLA2xxx Board Family*
  - *Software Installation Guide for the QLA2xxx Board Family*

  See the QLogic website at:

  `http://www.qlogic.com`
- Solaris 8 documentation:
  - *Solaris 8 Installation Guide*
  - *Solaris 8 System Administration Collection*
  - *Solaris 8 Advanced Installation Guide*

  See the Sun Microsystems website at:

  `http://www.sun.com`
- Solaris 9 documentation:
  - *Solaris 9 Installation Guide*
  - *Solaris 9 System Administration Collection*

See the Sun Microsystems website at:

`http://www.sun.com`

- Sun Microsystems hardware documentation:
  - *Ultra Enterprise 450 Systems Manual*
  - *Ultra Enterprise 6000/5000/4000 Systems Manual*
  - *Ultra Enterprise 6000/5000/4000 Systems Installation Guide*
  - *Ultra Enterprise 10000 SSP 3.0 User's Guide*
- Windows software documentation:
  - *Start Here Microsoft Windows NT Workstation: Basics and Installation*
  - *Microsoft Windows 2000 Quick Start Guide*
- Hardware documentation for the Intel platform
- *Flexible License Manager End User Manual* from GLOBEtrotter Software, Inc.

Obtaining Publications

You can obtain SGI documentation in the following ways:

- See the SGI Technical Publications Library at: `http://docs.sgi.com`. Various formats are available. This library contains the most recent and most comprehensive set of online books, release notes, man pages, and other information.
- If it is installed on your SGI system, you can use InfoSearch, an online tool that provides a more limited set of online books, release notes, and man pages. With an IRIX system, select **Help** from the Toolchest, and then select **InfoSearch**. Or you can type `infosearch` on a command line.
- You can also view release notes by typing either `grelnotes` or `relnotes` on a command line.
- You can also view man pages by typing `man` *title* on a command line.

Conventions

Note: This guide uses *Windows* to refer to both Microsoft Windows NT and Microsoft Windows 2000 nodes when the information applies equally to both. Information that applies to only one of these types of nodes is identified.

The following conventions are used throughout this document:

Convention	**Meaning**
`command`	This fixed-space font denotes literal items such as commands, files, routines, path names, signals, messages, and programming language structures.
variable	Italic typeface denotes variable entries and words or concepts being defined.
`user input`	This bold, fixed-space font denotes literal items that the user enters in interactive sessions. (Output is shown in nonbold, fixed-space font.)
GUI	This font denotes the names of graphical user interface (GUI) elements such as windows, screens, dialog boxes, menus, toolbars, icons, buttons, boxes, fields, and lists.
[ ]	Brackets enclose optional portions of a command or directive line.
...	Ellipses indicate that a preceding element can be repeated.

Reader Comments

If you have comments about the technical accuracy, content, or organization of this document, contact SGI. Be sure to include the title and document number of the manual with your comments. (Online, the document number is located in the front matter of the manual. In printed manuals, the document number is located at the bottom of each page.)

You can contact SGI in any of the following ways:

- Send e-mail to the following address:

  techpubs@sgi.com

- Use the Feedback option on the Technical Publications Library Web page:

  `http://docs.sgi.com`

- Contact your customer service representative and ask that an incident be filed in the SGI incident tracking system.

- Send mail to the following address:

  Technical Publications
  SGI
  1600 Amphitheatre Parkway, M/S 535
  Mountain View, California 94043–1351

- Send a fax to the attention of "Technical Publications" at +1 650 932 0801.

SGI values your comments and will respond to them promptly.

Introduction

This guide provides an overview of the installation and configuration procedures for CXFS client-only nodes running SGI CXFS clustered filesystems. A *CXFS client-only node* runs a subset of CXFS processes and services.

This release supports client-only nodes running the following operating systems:

- Solaris 8
- Solaris 9
- Windows NT 4.0 Service Pack 6
- Windows 2000 Service Pack 2 or later

Note: This guide uses *Windows* to refer to both Microsoft Windows NT and Microsoft Windows 2000 nodes when the information applies equally to both. Information that applies to only one of these types of nodes is identified.

A cluster running multiple operating systems is known as a *multiOS cluster*.

Many of the procedures mentioned in this guide will be performed by SGI personnel or other qualified service personnel. Details for these procedures are provided in other documents.

⚠ **Caution:** CXFS is a complex product. To ensure that CXFS is installed and configured in an optimal manner, it is **mandatory** that you purchase SGI installation services developed for CXFS. Contact your local SGI sales representative for details.

For general information about CXFS terminology, concepts, and configuration, see the *CXFS Version 2 Software Installation and Administration Guide*.

This chapter discusses the following:

- "When to Use CXFS", page 2
- "CXFS on Client-Only Nodes", page 2
- "Overview of the Installation and Configuration Steps", page 10

When to Use CXFS

You should use CXFS when you have multiple hosts running applications that require high-bandwidth access to common filesystems.

CXFS performs best under the following conditions:

- Data I/O operations are greater than 16 KB.
- All processes that perform reads/writes for a given file reside on the same host.
- Multiple processes on multiple hosts read the same file.
- Direct-access I/O is used for reads/writes for multiple processes on multiple hosts.
- Large files and file accesses are being used.

For most filesystem loads, the preceding scenarios represent the bulk of the file accesses. Thus, CXFS delivers fast local file performance. CXFS is also useful when the amount of data I/O is larger than the amount of metadata I/O. (*Metadata* is information that describes a file, such as the file's name, size, location, and permissions.) CXFS is faster than NFS because the data does not go through the network.

CXFS on Client-Only Nodes

This section contains the following:

- "CXFS Processes", page 3
- "Licenses", page 3
- "Cluster Administration", page 3
- "User Administration for CXFS", page 4
- "Performance Considerations", page 5
- "Requirements", page 6
- "Recommendations", page 8

CXFS Processes

When CXFS is started on a client-only node, a user-space daemon/service is started that provides the required processes. This is a subset of the processes needed on an IRIX node.

Licenses

You must have the following licenses:

- Brocade license. See "Required Brocade Fibre Channel Switch Firmware and License", page 17.
- CXFS FLEXlm license installed on every node in the cluster; see Chapter 4, "Obtaining CXFS and XVM FLEXlm Licenses", page 21.

> **Note:** XVM provides a mirroring feature. If you want to access a mirrored volume from a given node in the cluster, you must install the "XVM Volume Plexing" FLEXlm license on that node. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your SGI sales representative.

Cluster Administration

There must be at least one IRIX node in the cluster that is responsible for updating that filesystem's metadata. This node is referred to as the *CXFS metadata server*. Only IRIX nodes can be metadata servers; client-only nodes cannot be metadata servers. The CXFS cluster database is not stored on client-only nodes; only IRIX nodes contain the cluster database.

An IRIX node is required to perform administrative tasks, using either the `cmgr`(1M) command or the CXFS graphical user interface (GUI). For more information about using these tools, see the *CXFS Version 2 Software Installation and Administration Guide*.

Note: The NFS export scripts are supported on IRIX and Solaris nodes; they are not supported on Windows nodes. The scripts behave the same on IRIX and Solaris nodes, but the pathnames on Solaris are as follows:

```
/var/cluster/cxfs_client-scripts/cxfs-pre-mount
/var/cluster/cxfs_client-scripts/cxfs-post-mount
/var/cluster/cxfs_client-scripts/cxfs-pre-umount
/var/cluster/cxfs_client-scripts/cxfs-post-umount
```

For information about using these scripts on IRIX nodes or Solaris nodes, see the *CXFS Version 2 Software Installation and Administration Guide*.

User Administration for CXFS

A CXFS cluster requires a consistent user identification scheme across all hosts in the cluster so that one person using different cluster nodes has the same access to the files on the cluster.

The following must be observed to achieve this consistency:

- Users must have the same usernames on all nodes in the cluster. An individual user identifier (UID) should not be used by two different people anywhere in the cluster. Ideally, group names should also be the same on all nodes in the cluster.
- The `/etc/passwd` and `/etc/group` files from the CXFS metadata server must be installed on the client-only node. These files are used to determine the equivalent UNIX UID and group identifiers (GIDs) of each user.

  If these files are modified on the CXFS metadata server, the files should be reinstalled on the client-only node.

  Note: Under Windows, the CXFS software will detect that these files have changed and will apply the updated contents when mapping Windows users to UNIX UIDs. The `Administrator` user on the Windows node will be mapped to the `root` user on the UNIX nodes for the purpose of file access controls. Therefore, access to the `Administrator` user account on Windows nodes should be treated with the same caution as the `root` user on UNIX nodes.

Performance Considerations

CXFS may not give optimal performance under the following circumstances:

- When you are using NFS to export a CXFS filesystem from a CXFS client. Performance will be much better when the export is performed from a CXFS metadata server than when it is performed from a CXFS client-only node.
- When access would be as slow with CXFS as with network filesystems, such as with the following:
  - Small files
  - Low bandwidth
  - Lots of metadata transfer

    Metadata operations can take longer to complete through CXFS than on local filesystems. Metadata transaction examples include the following:

    - Opening and closing a file
    - Changing file size (usually extending a file)
    - Creating and deleting files
    - Searching a directory

    In addition, multiple processes on multiple hosts that are reading and writing the same file using buffered I/O can be slower when using CXFS than when using a local filesystem. This performance difference comes from maintaining coherency among the distributed file buffers; a write into a shared, buffered file will invalidate data (pertaining to that file) that is buffered in other hosts.
- When distributed applications write to shared files that are memory mapped.

Also see "Functional Limitations Specific to Windows", page 56.

Requirements

Using a client-only node in a CXFS cluster requires the following:

- A supported SAN hardware configuration.

  Note: For details about supported hardware, see the Entitlement Sheet that accompanies the base CXFS release materials. Using unsupported hardware constitutes a breach of the CXFS license. CXFS does **not** support the Silicon Graphics O2 workstation as a CXFS node nor does it support JBOD.

- At least one IRIX node to act as the metadata server and from which to perform cluster administration tasks. CXFS should be installed on the IRIX node before CXFS is installed on the client-only node.

- A private 100baseT TCP/IP network connected to each node, to be dedicated to the CXFS private heartbeat and control network. All nodes must be configured to use the same subnet.

- IRIX 6.5.19f or later, plus any required patches. For more details, see the platform-specific CXFS MultiOS Clients release notes.

- A FLEXlm license key for CXFS and optionally XVM. The CXFS license is required for all nodes in the pool; a license is required for each node from which you want to access a mirrored XVM volume.

- A Brocade Fibre Channel 2400, 2800, or 3800 switch that is sold and supported by SGI. The switch is required to protect data integrity.

  IRIX nodes use serial reset lines or I/O fencing to protect the integrity of the data stored in the cluster. (One of these methods is mandatory for the IRIX nodes in a cluster with only two server-capable nodes. Larger clusters should have an odd number of server-capable nodes.)

  The *I/O fencing* feature isolates a problem node so that it cannot access I/O devices and therefore cannot corrupt data in the shared CXFS filesystem. This feature can only be used with a Brocade Fibre Channel switch; therefore, the Brocade switch is a required piece of hardware in a multiOS cluster.

  I/O fencing differs from zoning:

  – *Fencing* is a generic cluster term that means to erect a barrier between a host and shared cluster resources.

- *Zoning* is the ability to define logical subsets of the switch (zones), with the ability to include or exclude hosts and media from a given zone. A host can only access media that are included in its zone. Zoning is one possible implementation of fencing.

Zoning implementation is complex and does not have uniform availability across switches. Instead, SGI chose to implement a simpler form of fencing, enabling/disabling a host's Brocade ports.

If there are problems with a node, the I/O fencing software sends a message via the `telnet` protocol to the appropriate Fibre Channel switch. The switch only allows one `telnet` session at a time; therefore, if you are using I/O fencing, you must keep the `telnet` port on the Fibre Channel switch free at all times.

⚠ **Caution:** Do not perform a `telnet` to the switch and leave the session connected.

- A cluster of as many as 32 nodes, of which as many as 16 can be CXFS administration nodes; the rest will be client-only nodes. At least one IRIX node must be a server-capable administration node in order to be a potential metadata server; other nodes can be CXFS clients. All Solaris nodes and Windows nodes are CXFS client-only nodes.

  A cluster in which both CXFS and IRIS FailSafe 2.1 or later are run (known as *coexecution*) is supported with a maximum of 32 nodes, as many as 8 of which can run FailSafe. However, FailSafe cannot run on Solaris nodes or Windows nodes.

- No nodes within the cluster running Trusted IRIX. A multiOS cluster cannot contain Trusted IRIX nodes.

- Ensure that the appropriate IRIX software is installed on the potential metadata server nodes. For example, if you want to use quotas and access control lists (ACLs) on any cluster node, the `eoe.sw.quotas`, `nfs.sw.acl_nfs`, and `eoe.sw.acl` subsystems must be installed on the IRIX nodes listed as potential metadata servers. For more information, see *IRIX Admin: Disks and Filesystems*, *IRIX Admin: Backup, Security and Accounting*, and your site's IRIX system administrator.

Also see "Requirements Specific to Solaris", page 24, and "Requirements Specific to Windows", page 54.

Recommendations

SGI recommends the following when running CXFS on a client-only node:

- Fix any network issues on the private network before trying to use CXFS.
- Use an Ethernet network switch rather than a hub for performance and control.
- A production cluster should be configured with a minimum of three server-capable administration nodes.
- For large clusters, SGI recommends that you define only the first three server-capable administration nodes and then continue on with the steps to define the cluster. After you have a successful small cluster, go back and add the remaining nodes.
- Any task initiated using `cron` on a CXFS filesystem should be launched from a single node in the cluster, preferably from the metadata server.

  The `cron`(1) daemon can cause severe stress on a CXFS filesystem if multiple nodes in a cluster start the same filesystem-intensive task simultaneously. An example of such a task is one that uses the `find`(1) command to search files in a filesystem.
- Do not run any defragmentation software on CXFS filesystems. This includes the IRIX `fsr`(1M) command and any similar commands on Solaris or Windows.
- Be very careful when running IRIX `xfs_repair`(1M) on CXFS filesystems. Only use `xfs_repair` on metadata servers and only when you have verified that all other cluster nodes have unmounted the filesystem. SGI recommends that you contact SGI technical support before using `xfs_repair`. For more details, see the *CXFS Version 2 Software Installation and Administration Guide*.
- Only those nodes that you want to be potential metadata servers should be CXFS administration nodes (installed with the `cluster_admin` software product). CXFS client administration nodes should only be used when necessary for coexecution with IRIS FailSafe. All other nodes should be client-only nodes (installed with `cxfs_client`). Use an odd number of server-capable administration nodes.
- Shut down cluster services before maintenance.
- In this release, relocation is disabled by default and recovery is supported only when using standby nodes.

A *standby node* is a metadata server-capable administration node that is configured as a potential metadata server for a given filesystem, but does not currently run any applications that will use that filesystem. To use recovery, you must not run any applications on any of the potential metadata servers for a given filesystem; after the active metadata server has been chosen by the system, you can then run applications that use the filesystem on the active metadata server and client-only nodes.

Relocation and recovery are fully implemented, but the number of associated problems prevents full support of these features in the current release. Although data integrity is not compromised, cluster node panics or hangs are likely to occur. Relocation and recovery will be fully supported in a future release when these issues are resolved.

- Use the following good practices:
  - Unmount the filesystems from the metadata server, shut down the node, and remount the filesystem when possible.
  - Do the following before shutting down a node:
    - Unmount filesystems.
    - Shut down cluster services.
- Do not run power management software, which may interfere with the CXFS cluster.
- Enable the *forced unmount* feature for CXFS filesystems, which is turned off by default. Normally, an unmount operation will fail if any process has an open file on the filesystem. However, a forced unmount allows the unmount to proceed regardless of whether the filesystem is still in use.

  Many sites have found that enabling this feature improves the stability of their CXFS cluster, particularly in situations where the filesystem must be unmounted.

  The method used to implement this feature is platform-specific:

  - On IRIX nodes, this feature uses the `umount -k` option. The `-k` option attempts to kill processes that have open files or current directories in the appropriate filesystems and then unmount them. That is, it attempts to terminate any I/O going to the filesystem, so that it can unmount it promptly, rather than having to wait for the I/O to finish on its own, causing the unmount to possibly fail.

- On Solaris nodes, a similar function is performed with the `fuser -k` command and a `umount -f` command.
- On Windows nodes, a forced unmount does not terminate user processes by default. For details, see "Forced Unmount on a Windows Node", page 58.

This feature is available on an IRIX node by the following CXFS GUI menu:

Tasks
 > Filesystems
  > Unmount a Filesystem

In the CXFS GUI, click the **Force** toggle in the **Unmount Filesystem** task.

You can also specify this feature using the `cmgr`(1M) commands to define the filesystem. For more information, see "Forced Unmount of CXFS Filesystems", page 106.

You must use `cmgr` from an IRIX node, and the GUI must be connected to an IRIX node.

For more information, see the *CXFS Version 2 Software Installation and Administration Guide*, the Solaris `fuser`(1M) man page, and the Solaris and IRIX `umount`(1M) man pages.

Overview of the Installation and Configuration Steps

This section provides an overview of the installation, verification, and configuration steps for each platform type:

- "Solaris Overview"
- "Windows Overview", page 12

Solaris Overview

Note: For additional details, see the CXFS MultiOS Clients release notes for Solaris. If you run into problems, see Chapter 8, "Troubleshooting", page 109.

Following is the order of installation and configuration steps for a CXFS Solaris node:

1. Install the Solaris 8 or Solaris 9 operating system according to the directions in the Solaris documentation (if not already done).
2. Install and verify the SGI TP9400 or SGI TP9100 RAID. See Chapter 2, "SGI RAID Firmware", page 15.
3. Install and verify the Brocade Fibre Channel switch. See Chapter 3, "Brocade Fibre Channel Switch Verification", page 17.
4. Obtain and install the CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must purchase the “XFS Volume Plexing” software option and obtain and install a FLEXlm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See Chapter 4, "Obtaining CXFS and XVM FLEXlm Licenses", page 21.
5. Install and verify the JNI host bus adapter (HBA). You will install the `JNIC146x` package, which provides software for the Fibre Channel card and driver. See "JNI Fibre Channel Host Bus Adapter Installation and Configuration", page 28.
6. Prepare the Solaris node, including adding a private network. See "Preinstallation Steps for Solaris", page 41.
7. Use the Solaris `pkgadd`(1M), `pkgrm`(1M), and `pkginfo`(1) commands as directed to install the `SGIcxfs` package, which provides the following:

   - The `/etc/init.d/cxfs_cluster` command and associated scripts in the `/etc/rc` directory for automatic startup and shutdown
   - Drivers required for CXFS (`xvm` and `cell`)
   - The `CXFS` module
   - The `cxfs_client` command

   See "Client Software Installation Steps for Solaris", page 47.

8. Create the I/O fencing file. See "Postinstallation Steps for Solaris: Creating the I/O Fencing File", page 49.

9. Configure the cluster to define the new Solaris node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration", page 99.

Windows Overview

This information applies to both Windows NT and Windows 2000 nodes unless otherwise noted.

Note: For additional details, see the CXFS MultiOS release notes for Windows. If you run into problems, see Chapter 8, "Troubleshooting", page 109.

Following is the order of installation and configuration steps for a CXFS Windows node:

1. Install the Windows operating system according to the directions in the Windows documentation (if not already done).

2. Install Windows NT Service Pack 6 or Windows 2000 Service Pack 2 according to the directions in the Windows documentation (if not already done).

3. Install and verify the SGI TP9400 or SGI TP9100 RAID. See Chapter 2, "SGI RAID Firmware", page 15.

4. Install and verify the Brocade Fibre Channel switch. See Chapter 3, "Brocade Fibre Channel Switch Verification", page 17.

5. Obtain the CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must purchase the “XFS Volume Plexing” software option and obtain and install a FLEXlm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See Chapter 4, "Obtaining CXFS and XVM FLEXlm Licenses", page 21.

6. Install and verify the QLogic host bus adapter (HBA) and driver. See "QLogic Fibre Channel Host Bus Adapter Installation for Windows", page 71.

7. Prepare the Windows node, including adding a private network. See "Preinstallation Steps for Windows", page 75.

8. Install the CXFS software. See "Client Software Installation Steps for Windows", page 81.

9. Perform post-installation configuration steps:

   - "Configuring the FLEXlm License for Windows", page 86
   - "Creating the Windows I/O Fencing File", page 87
   - "Performing User Configuration", page 89
   - "Creating a New Hardware Profile", page 90

10. Configure the cluster to define the new Solaris node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration", page 99.

11. Start CXFS services on the Windows node to see the mounted filesystems under the configured drive letter.

Chapter 2

SGI RAID Firmware

The SGI TP9400 or SGI TP9100 RAID will be initially installed and configured by SGI personnel.

Required SGI RAID Firmware

This section describes the required RAID firmware for the SGI TP9400 and SGI TP9100.

Required SGI TP9400 RAID Firmware

The TP9400 4.0 CD contains the required controller firmware and NVSRAM files for the 4774 or 4884 units:

- If you have a 4774 unit, the Mojave code must be installed according to FCO 1056.
- If you have a 4884 unit, the Mojave code is installed by default.

Note: By default, the TP9400 supports 32 logical units (LUNs). If additional LUNs are required, you must obtain a separate software enabling key; this key will support a maximum of 128 LUNs. Contact your SGI sales representative for the SGI software partitioning key.

Required SGI TP9100 RAID Firmware

The TP9100 4.0 CD contains the required version 7.75 controller firmware and NVSRAM files for the 1-Gbit TP9100. The TP9100 5.0 CD contains the required version 8.29 firmware and NVSRAM files for the 2-Gbit TP9100.

RAID Firmware Verification

To verify that the SGI RAID is properly installed and ready for use with CXFS, you can dump the RAID’s profile and verify the controller software revisions.

For More Information

The following documentation is used to install and verify the RAID:

- SGI TP9400:
  - *SGI TP9400 and SGI TP9500 Software Concepts Guide*
  - *SGI TP9400 and SGI TP9500 RAID Owner's Guide*
  - *SGI TP9400 and SGI TP9500 RAID Administration Guide*
- SGI TP9100:
  - *TPM Installation Instructions and User's Guide for TP9100*

Chapter 3

Brocade Fibre Channel Switch Verification

In order to protect data integrity, Solaris nodes and Windows nodes must use the *I/O fencing* feature, which isolates a problem node so that it cannot access I/O devices and therefore cannot corrupt data in the shared CXFS filesystem. This feature can only be used with a Brocade Fibre Channel switch sold and supported by SGI; therefore, the Brocade switch is a required piece of hardware in a multiOS cluster.

The Brocade Fibre Channel switches will be initially installed and configured by SGI personnel. You can use the information in this chapter to verify the installation.

Required Brocade Fibre Channel Switch Firmware and License

This release supports Brocade Silkworm 2400 (8-port), 2800 (16-port), 3200 (8–port, 2–Gbit), 3800 (16–port, 2-Gbit) Fibre Channel switches that are sold and supported by SGI.

All Brocade switches contained within the SAN fabric must have the appropriate Brocade license key installed. The following firmware is required:

- 2400 and 2800 switches: 2.6.0d or later
- 3200 and 3800 switches: 3.0.2c or later

If the current firmware level of the switches must be upgraded, please contact your local SGI service representative or customer support center.

The Brocade switch must be configured so that its Ethernet interface is accessible from all IRIX cluster nodes using `telnet`. The fencing network connected to the Brocade switch must be physically separate from the private heartbeat network.

⚠ **Caution:** The `telnet` port must be kept free in order for I/O fencing to succeed.

Verifying the Brocade License

To verify the Brocade license, log into the switch as user `admin` and use the `licenseshow` command, as shown in the following example:

```
brocade:admin> licenseshow
dcRyzyScSedSz0p:
    Web license
    Zoning license
    SES license
    Fabric license
SQQQSyddQ9TRRdUP:
    Release v2.2 license
```

Verifying the Brocade Switch Firmware Version

To verify the firmware version, log into the switch as user `admin` and use the `version` command, as shown in the following example:

```
brocade:admin> version
Kernel:     5.3.1
Fabric OS:  v2.2.1c                 <== Firmware Revision
Made on:    Mon Dec 18 11:39:26 PST 2000
Flash:      Mon Dec 18 11:40:01 PST 2000
BootProm:   Thu Jun 17 15:20:39 PDT 1999
```

Changing to the Brocade FC Cable Connections

To change Brocade Fibre Channel cable connections used by nodes in the CXFS cluster, do the following:

1. Cleanly shut down CXFS services on the nodes affected by the cable change using either the CXFS GUI or the `cmgr`(1M) command.
2. Rearrange the cables as required.
3. Restart CXFS services.
4. Reconfigure I/O fencing if required. You must perform this step if I/O fencing is enabled on the cluster and if you added/removed any Brocade switches. You

must use the CXFS GUI or the `cmgr`(1M) command to add/remove switches from the CXFS configuration as required.

5. If any CXFS client nodes are connected to a new (or different) Brocade switch, restart CXFS services on those nodes. This will ensure that the IRIX servers can correctly identify the Brocade ports used by all clients.

6. If connected to a different RAID device, restart the Windows NT node.

Consult the *CXFS Version 2 Software Installation and Administration Guide* for instructions to configure I/O fencing.

Obtaining CXFS and XVM FLEXlm Licenses

The software licensing used by CXFS is based on the FLEXlm product from GLOBEtrotter Software, Inc. For all supported platforms, a FLEXlm license is required to use CXFS. Perform the procedures in this chapter to satisfy this requirement.

XVM provides a mirroring feature. If you want to access a mirrored volume from a given node in the cluster, you must install the "XVM Volume Plexing" FLEXlm license on that node. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your SGI sales representative.

Obtain the Host Information Required for the License

When you order CXFS, you will receive an entitlement ID. You must submit the system host ID, host name, and entitlement ID when requesting your permanent CXFS license. The method used to obtain this information is platform-specific.

Solaris Host Information

To obtain the host identifier and hostname of the system on which you will run CXFS, execute the following Solaris commands:

```
/usr/bin/hostid
/usr/bin/hostname
```

For example:

```
# /usr/bin/hostid
830dad77
# /usr/bin/hostname
cxfssun2
```

When you are asked for the license manager host identifier, provide this information. You must have a separate license for each host on which CXFS is installed.

Windows Host Information

FLEXlm requires that you supply the Ethernet (MAC) address in order to generate the FLEXlm license. This address is known as the *Physical Address* in Windows. You can obtain this information in one of the following ways:

- View the network adapter properties in the **Windows Control Panel**
- Open a command prompt window and run the following command:

```
C:\> ipconfig /all
```

If the machine has more than one network interface, you should use the Physical Address of the private network interface.

Note: Windows NT licenses cannot be used under Windows 2000 and vice versa. If you are upgrading a Windows node to Windows 2000, you must obtain a new license.

Obtaining and Install the Licenses

Along with your entitlement number, you will receive a URL to a key generation page. To obtain your permanent CXFS and XVM licenses, follow the instructions on the key generation page. After the required information is provided, a key will be generated and displayed on the webpage along with installation instructions.

See also "FLEXlm License Verification for Solaris", page 28, and "Configuring the FLEXlm License for Windows", page 86.

For More Information

For more information about licensing, see the following webpage:

`http://www.sgi.com/support/licensing`

If you cannot use the web key generation page, you can contact the SGI order desk at 800 800 4SGI (800 800 4744).

For more information on FLEXlm, you may order the *Flexible License Manager End User Manual* from GLOBEtrotter Software, Inc.

Chapter 5

Solaris Platform

This chapter contains the following:

- "CXFS on Solaris"
- "FLEXlm License Verification for Solaris", page 28
- "JNI Fibre Channel Host Bus Adapter Installation and Configuration", page 28
- "Preinstallation Steps for Solaris", page 41
- "Client Software Installation Steps for Solaris", page 47
- "Postinstallation Steps for Solaris: Creating the I/O Fencing File", page 49
- "Manual CXFS Startup/Shutdown for Solaris", page 51
- "Software Maintenance for Solaris", page 52

CXFS on Solaris

This section contains the following:

- "Requirements Specific to Solaris", page 24
- "CXFS Commands Installed on Solaris", page 25
- "Solaris Log Files", page 25
- "Solaris Limitations and Considerations", page 25
- "Access Control Lists and Solaris", page 27

Requirements Specific to Solaris

In addition to the items listed in"Requirements", page 6, using a Solaris node to support CXFS requires the following:

- Solaris 8 or Solaris 9 operating system.
- One to four JNI FibreStar FCE-6460-N (PCI) 2-Gbit Fibre Channel host bus adapters (HBAs).

  Note: 1-Gbit HBAs and Sbus HBAs are not supported.

- One or more of the following Sun Microsystems hardware platform series:
  - Sun Fire 280R
  - Sun Fire V480
  - Sun Fire V880
  - Sun Fire 4800/4810 (PCI slots only, cPCI is not supported)
  - Sun Fire 6800 (PCI slots only, cPCI is not supported)
  - Sun Fire 12K
  - Sun Fire 15K
  - Ultra Enterprise 250
  - Ultra Enterprise 450
  - Ultra Enterprise 4000
  - Ultra Enterprise 3000
  - Ultra Enterprise 5000
  - Ultra Enterprise 6000
  - Ultra Enterprise 10000

IRIX nodes do not permit nested mount points on CXFS filesystems; that is, you cannot mount an IRIX XFS or CXFS filesystem on top of an existing CXFS filesystem. Although it is possible to mount a UFS or NFS filesystem on top of a Solaris CXFS filesystem, this is not recommended.

CXFS Commands Installed on Solaris

The following commands are shipped as part of the CXFS Solaris package:

- `/usr/cxfs_cluster/bin/cxfs_client` (the CXFS client service)
- `/usr/cxfs_cluster/bin/cxfslicense`
- `/usr/cxfs_cluster/bin/xvmprobe`

These commands provide all of the services needed to include a Solaris node in a CXFS cluster. The `pkgadd`(1M) output lists all software added; see "Installation Overview", page 48.

For more information, see the `cxfs_client`(1M) and `xvmprobe`(1M) man pages.

Solaris Log Files

The `cxfs_client` command creates a `/var/log/cxfs_client` log file. (There is no `/var/cluster` log on Solaris nodes.) This log file is not rotated or truncated.

For information about the log files created on IRIX nodes, see the *CXFS Version 2 Software Installation and Administration Guide*.

Solaris Limitations and Considerations

CXFS for Solaris has the following limitations and considerations:

- For optimal performance, you should set the value of the Solaris system tunable parameter `maxphys` in the `/etc/system` file. Do the following:

  1. Make a backup copy of the `/etc/system` file.

     Note: Exercise extreme caution in changing `/etc/system` and always make a backup copy.

  2. Change the value of `maxphys` to `0x800000` (hexadecimal) in the `/etc/system` file.

  3. Reboot the Solaris node. This causes the change to take effect.

4. Verify that the new value for maxphys is in effect by running the following command:

   ```
   # echo "maxphys/X" | adb -k
            physmem 1f03f
            maxphys:
            maxphys:        800000
   ```

- CXFS Solaris nodes cannot view or edit user and group quotas because CXFS administration must be performed from an IRIX node. However, user and group quotas are enforced correctly by the metadata server.

  To view or edit your quota information, you must log in to an IRIX cluster node and make any necessary changes. If you would like to provide a viewing command such as repquota, you could construct a Solaris shell script similar to the following:

  ```
  #! /bin/sh
  #

  # Where repquota lives on IRIX
  repquota=/usr/etc/repquota

  # The name of an IRIX node in the cluster
  irixnode=cain

  rsh $irixnode "$repquota $*"
  exit
  ```

- The minimum block size supported is 2 KB, determined by a maximum of 4 extents per 8-KB page. (XFS uses a default block size of 4 KB unless overridden by an administrator to a different blocksize value, for example 2 KB or 8 KB.)

- All disk devices attached to JNI controllers must be for use only by CXFS disks; do not attach non-disk devices to any JNI controllers that are configured for CXFS use. This restriction is required because all disk devices on JNI controllers configured for CXFS make use of the whole disk volume, which must be conveyed to Solaris via modification in the JNI driver to the value returned by the READ_CAPACITY SCSI command.

Access Control Lists and Solaris

All CXFS files have UNIX mode bits (read, write, and execute) and optionally an access control list (ACL). For more information, see the `chmod`(1) and `setfacl`(1) man pages.

If you restore a CXFS file that had an ACL containing only owner-ACL entries (that is, owner/group/other/mask) from a Solaris node, upon restoration one of the following will happen:

- **When using `tar`(1), `cpio`(1), and Legato Networker:** The ACL will be lost because these tools behave "intelligently" by not calling `acl`(2) to set an ACL if the file has only owner/group/other/mask entries. These tools will only set the file mode. However, this does not present a change in functionality because an access permissions check on the mode and the ACL containing only owner entries will give the same result.
- **When using other backup/restore utilities:** A mask will be added to the ACL if the application calls `acl`(2) for every file.

  A backup/restore utility that calls `acl`(2) to set an ACL for every file will result in a file being restored with four ACL entries (that is, owner/group/other/mask), even though it may have originally had only three (that is, owner/group/other). This is due to a requirement in `getfacl`(1) that it receive four ACL entries for the `GETACL` command to `acl`(2). (If fewer than four entries are returned, `getfacl` will report an error).

Note: Normally, Solaris filesystem ACLs can have up to 1024 entries for a file and a directory can have 1024 entries as well as an additional 1024 entries for the default ACL. However, CXFS filesystems on Solaris nodes in a multiOS cluster must maintain compatibility with the IRIX metadata server. The CXFS filesystems on a Solaris node are limited to a maximum of 25 ACL entries for a file and a maximum total of 50 for a directory (that is, the directory ACL plus the default ACL).

When using the `ls`(1) command to display access permissions for a file with an ACL, the mode reported for a CXFS file follows IRIX semantics instead of Solaris/UFS semantics.

On Solaris, a UFS file mode reports the group permission as the intersection of the `GROUP` and `MASK` entries in the ACL. If the `GROUP` entry is `r-x` and the `MASK` entry is `rw-`, the group permission ill be reported as `r--`.

The IRIX model calls for reporting the ACL MASK for the group permission in the mode. Therefore, using the example above, the group permission will be reported as rw-. Although, it appears that the group has write permission, it does not and an attempt to write to the file will be rejected. You can obtain the real (that is, effective) group permission by using the Solaris getfacl(1) command.

FLEXlm License Verification for Solaris

Use the cxfslicense command with the -d option to verify that the FLEXlm licenses have been installed properly.

If the CXFS license is properly installed, you will see the following:

```
# /usr/cxfs_cluster/bin/cxfslicense -d
CXFS license granted.
```

If you do not have the CXFS license properly installed, you will see the following error on the console when trying to run CXFS:

```
Cluster services:CXFS not properly licensed for this host.  Run
        '/usr/cxfs_cluster/bin/cxfslicense -d'
for detailed failure information.  After fixing the
license, please run '/etc/init.d/cxfs_cluster restart'.
```

An error such as the following example will appear in the SYSLOG file:

```
Mar  4 12:58:05 6X:typhoon-q32 crsd[533]: <<CI> N crs 0> Crsd restarted.
Mar  4 12:58:05 6X:typhoon-q32 clconfd[537]: <<CI> N clconf 0>
Mar  4 12:58:05 5B:typhoon-q32 CLCONFD failed the CXFS license check.Use the
Mar  4 12:58:05 5B:typhoon-q32    '/usr/cxfs_cluster/bin/cxfslicense -d'
Mar  4 12:58:05 5B:typhoon-q32 command to diagnose the license problem.
```

JNI Fibre Channel Host Bus Adapter Installation and Configuration

This section provides an overview of the JNI HBA installation and verification for Solaris nodes:

- "Installing the JNI HBA"
- "Installing and Running the EZ Fibre Configuration GUI", page 30
- "Verifying the JNI HBA Installation", page 39

These procedures may be performed by you or by a qualified Sun service representative. You must be logged in as `root` to perform the steps listed in this section.

Installing the JNI HBA

You can use one to four JNI HBAs for CXFS per Sun machine. (Other HBAs may be present that are not shared with the CXFS cluster.)

To install the JNI HBA, perform the following steps. Additional details are provided in various chapters/sections of the *Installation Guide, FCE-6460 and FCE2-6460 PCI-to-Fibre Channel Host Bus Adapters (Solaris, Windows NT/2000, Novell, AIX, HP-UX, Mac OS, Linux) JNI FibreStar*, as noted.

1. Install the JNI host bus adapter (HBA) into the Solaris system. Perform the steps in the following chapter:

   - "Hardware Installation"

   Note: The JNI card **must** be installed in a 66MHz slot; if it is installed in another type of slot, CXFS will not work properly.

2. Bring the system back up using the steps listed in the following "Verifying" sections (the following represents the location of these sections in the manual):

   - "Unix Server DriverSuite"
     - "Solaris Driver"
       - "Verifying Hardware in OpenBoot PROM"
       - "Verifying Hardware in Solaris"

   You will be required to perform a Solaris `boot -r` after installing hardware.

   Caution: If you do not see the expected results, do not attempt to continue. Instead, diagnose why the card is not being seen by the hardware.

3. Install the latest JNI HBA driver software (`JNIC146x`) and Storage Networking Industry Association (SNIA) application programming interface package (`JNIsnia`), according to the instructions in the following "Installing" section:

- “Unix Server DriverSuite”
  - ”Solaris Driver”
    - “Installing the Software”

You can retrieve the driver and SNIA package from the following JNI website:

`http://www.jni.com/Drivers`

a. Under **Locate Driver by Product**, click on **FCE-6460**.

b. Under the **Solaris** section, left click **JNIC146x.pkg** and save as the following pathname:

```
/var/tmp/JNIC146x.pkg
```

Verify that the driver attached correctly to the HBA and that the package installed correctly by following the verification steps at the end of the section. Do not proceed until the verification succeeds.

4. Set the HBA to fabric mode:

a. In the `/kernel/drv/jnic146x.conf` file, change the following lines:

```
# FcLoopEnabled=1;
# FcFabricEnabled=0;
```

Delete the # character at the beginning of each line to uncomment it and change the values so that loop is disabled (0) and fabric is enabled (1). When you are done, the lines will appear as follows:

```
FcLoopEnabled=0;
FcFabricEnabled=1;
```

b. Reboot the Solaris node by entering the following command:

```
# init 6
```

Installing and Running the EZ Fibre Configuration GUI

After you have verified the installation of the HBA and the driver’s attachment to it, you are ready to install and run the JNI EZ Fibre program. This graphical user interface (GUI) will modify the driver’s configuration file, `/kernel/drv/jnic146x.conf`, so that it lists the worldwide node name (WWNN) and worldwide port name (WWPN) for the devices on your Fibre Channel.

For general installation information, see *Quick Installation Guide, Solaris, AIX and Windows JNI EZ Fibre*.

Do the following:

1. Install the GUI and change to the appropriate directory:

   a. Download the latest EZ Fibre GUI from the JNI website (EZ Fibre 2.2.1 or later):

      i. Go to the following website:

         `http://www.jni.com/Drivers`

      ii. Under **Locate Driver by Product**, click on **FCE-6460**

      iii. Under the **Solaris** section, left click **EZF_221.tar** or later and save as the following pathname (for example):

         `/var/tmp/EZF_221.tar`

   b. Extract the saved file using the `tar`(1) command. For example:

      ```
      # tar xvf /var/tmp/EZF_221.tar
      ```

   c. Change to the directory where the extracted GUI command is located:

      ```
      # cd /var/tmp/EZF_22
      ```

   d. Run the `install.sh` script:

```
# ./install.sh
Checking for required and recommended patches...
checkpatches.sh: Note - The following OS vendor recommended patches are
not installed or have been superseded -- please consult the EZ-Fibre
read me:
 108434-01 108435-01
<press enter to continue>

InstallAnywhere is preparing to install...
Installer using temporary disk space in ’/tmp’ ($TMPDIR not set).
```

You should install the GUI into the default location. When you see **Congratulations!**, click on **Done**.

2. Change to the following directory and read any `README` files you find there:

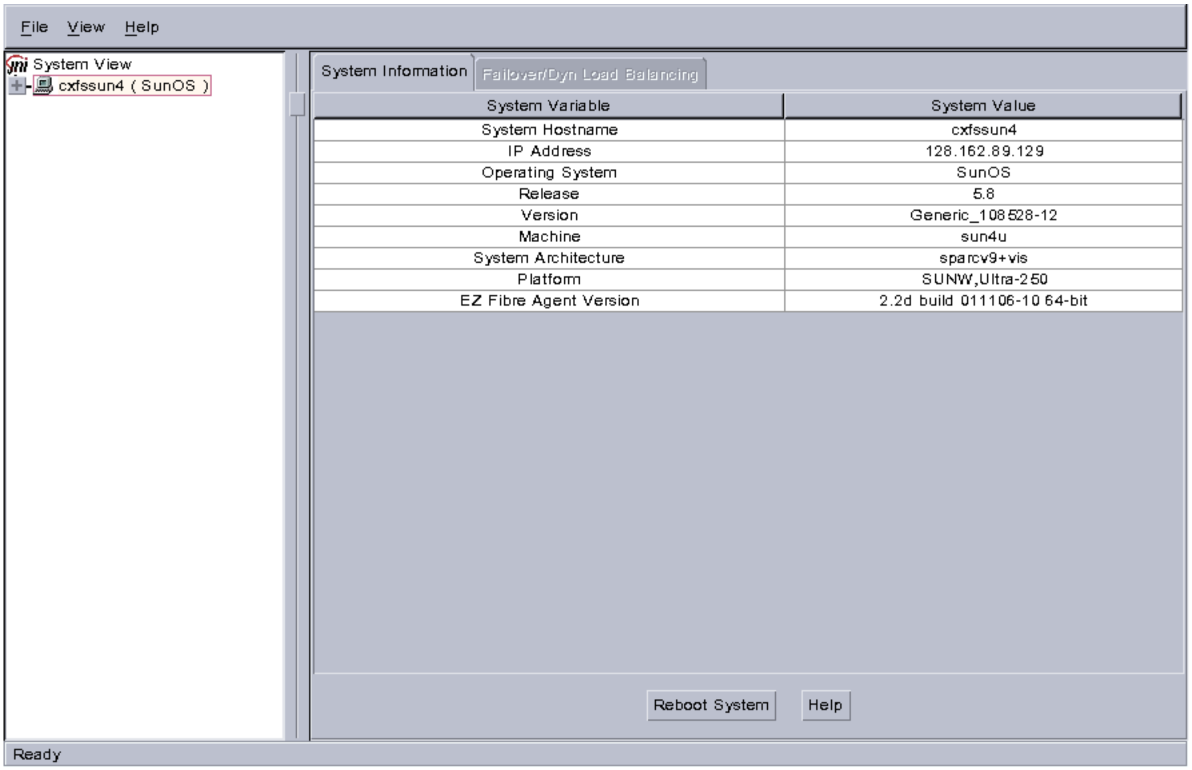
   ```
   # cd /opt/jni/ezfibre/standalone
   ```

3. Invoke the EZ Fibre GUI from within the same directory by entering the following command:

   ```
   # ./ezf
   ```

   Two windows will appear. The first, titled **Refresh SAN information**, will say **Discovering LUNs for HBA#0**.

   After a short while, this window will be replaced by a larger window, as shown in the example in Figure 5-1. (The example screen snaps in this document may not exactly match the display you see.)



System Variable	System Value
System Hostname	cxfssun4
IP Address	128.162.89.129
Operating System	SunOS
Release	5.8
Version	Generic_108528-12
Machine	sun4u
System Architecture	sparcv9+vis
Platform	SUNW,Ultra-250
EZ Fibre Agent Version	2.2d build 011106-10 64-bit

Figure 5-1 Example: Second Window: EZ Fibre Configuration Utility - Standalone

The left-hand pane of this window displays a listing of systems. Find the system you are configuring and click on the + sign next to it; this action expands the display so that it shows the installed JNI HBA on the system. Figure 5-2 highlights the + sign.

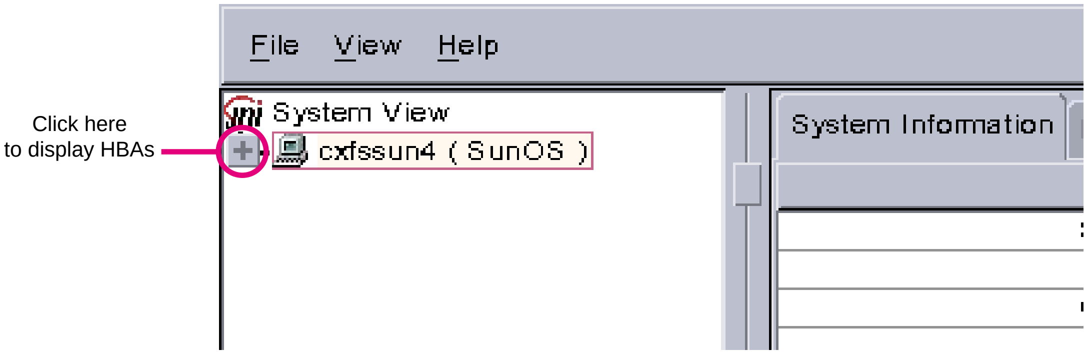
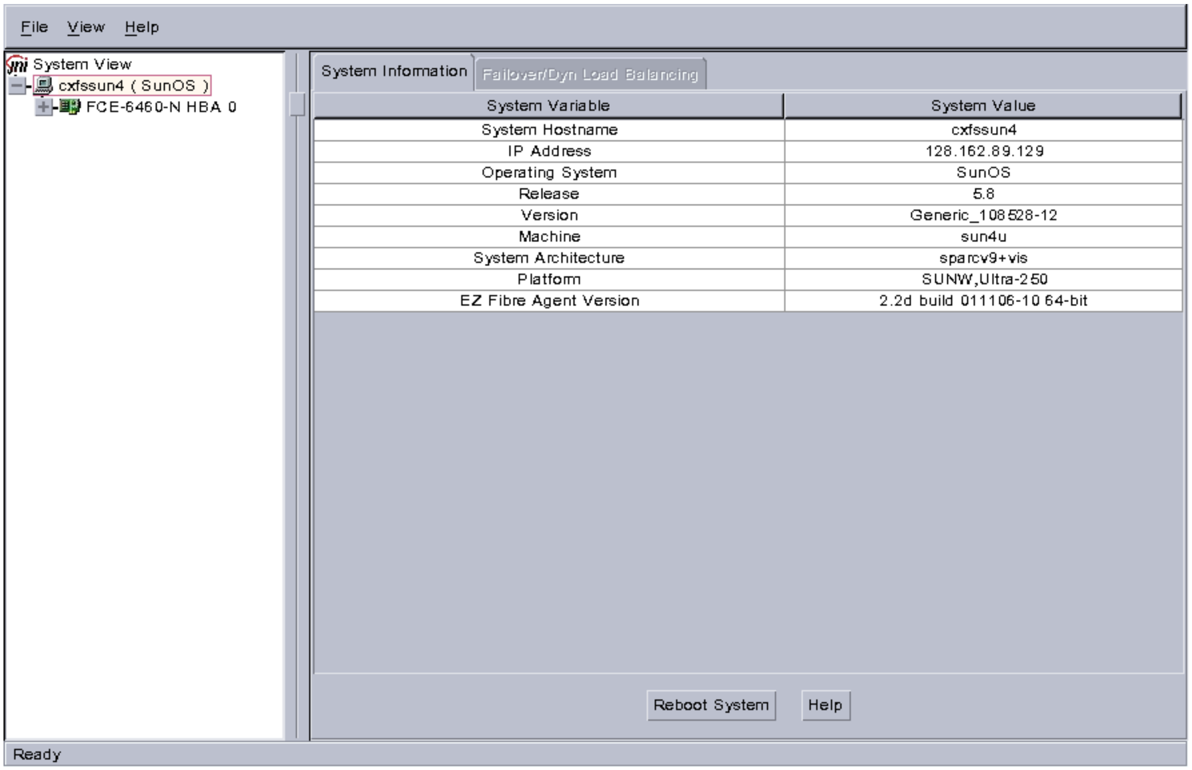


Figure 5-2 Location of icon (+) to Display the HBA

Figure 5-3 shows an example of the display after clicking on the + sign for `cxfssun4`, which shows the JNI HBA.



System Variable	System Value
System Hostname	cxfssun4
IP Address	128.162.89.129
Operating System	SunOS
Release	5.8
Version	Generic_108528-12
Machine	sun4u
System Architecture	sparcv9+vis
Platform	SUNW,Ultra-250
EZ Fibre Agent Version	2.2d build 011106-10 64-bit

Figure 5-3 Example: After Clicking + to Display the HBA

4. Click on the icon to the right (not the + sign to the left). Figure 5-4 shows the icon.

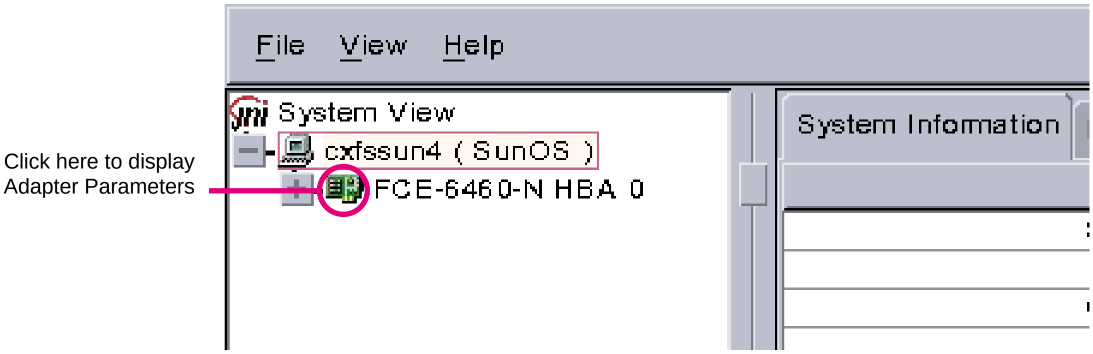
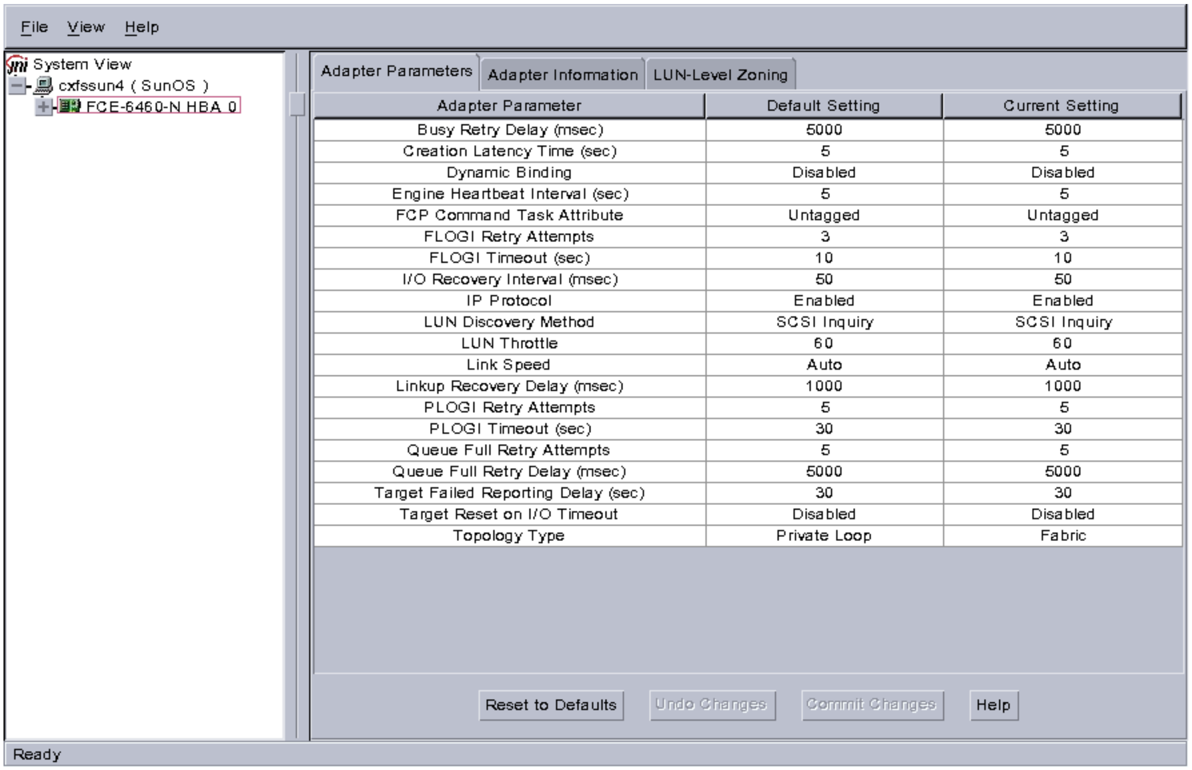


Figure 5-4 Location of the Icon to Display the Adapter Parameters

The right-hand pane will change to show **Adapter Parameters** for the selected HBA, as shown in Figure 5-5.

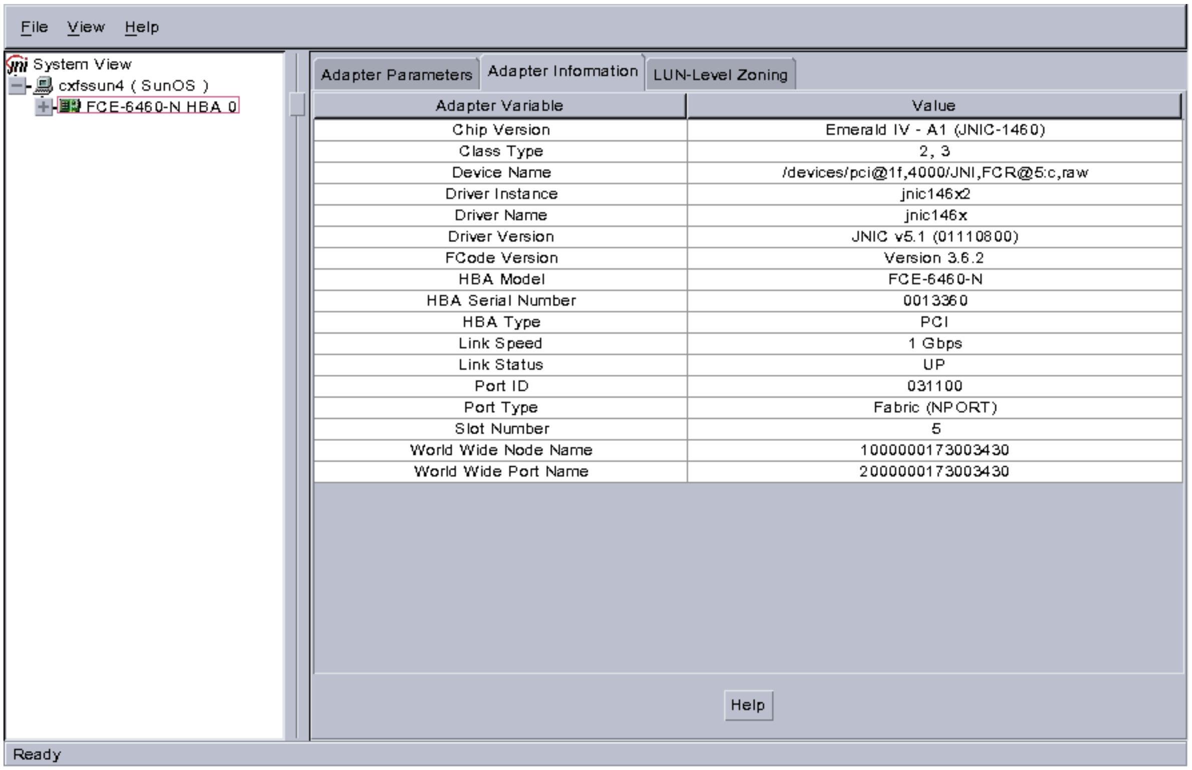


Adapter Parameter	Default Setting	Current Setting
Busy Retry Delay (msec)	5000	5000
Creation Latency Time (sec)	5	5
Dynamic Binding	Disabled	Disabled
Engine Heartbeat Interval (sec)	5	5
FCP Command Task Attribute	Untagged	Untagged
FLOGI Retry Attempts	3	3
FLOGI Timeout (sec)	10	10
I/O Recovery Interval (msec)	50	50
IP Protocol	Enabled	Enabled
LUN Discovery Method	SCSI Inquiry	SCSI Inquiry
LUN Throttle	60	60
Link Speed	Auto	Auto
Linkup Recovery Delay (msec)	1000	1000
PLOGI Retry Attempts	5	5
PLOGI Timeout (sec)	30	30
Queue Full Retry Attempts	5	5
Queue Full Retry Delay (msec)	5000	5000
Target Failed Reporting Delay (sec)	30	30
Target Reset on I/O Timeout	Disabled	Disabled
Topology Type	Private Loop	Fabric

Figure 5-5 Example: After Clicking the HBA Icon to Show the Adapter Parameters

a. Click on the **Adapter Information** tab to see the information in Figure 5-6.

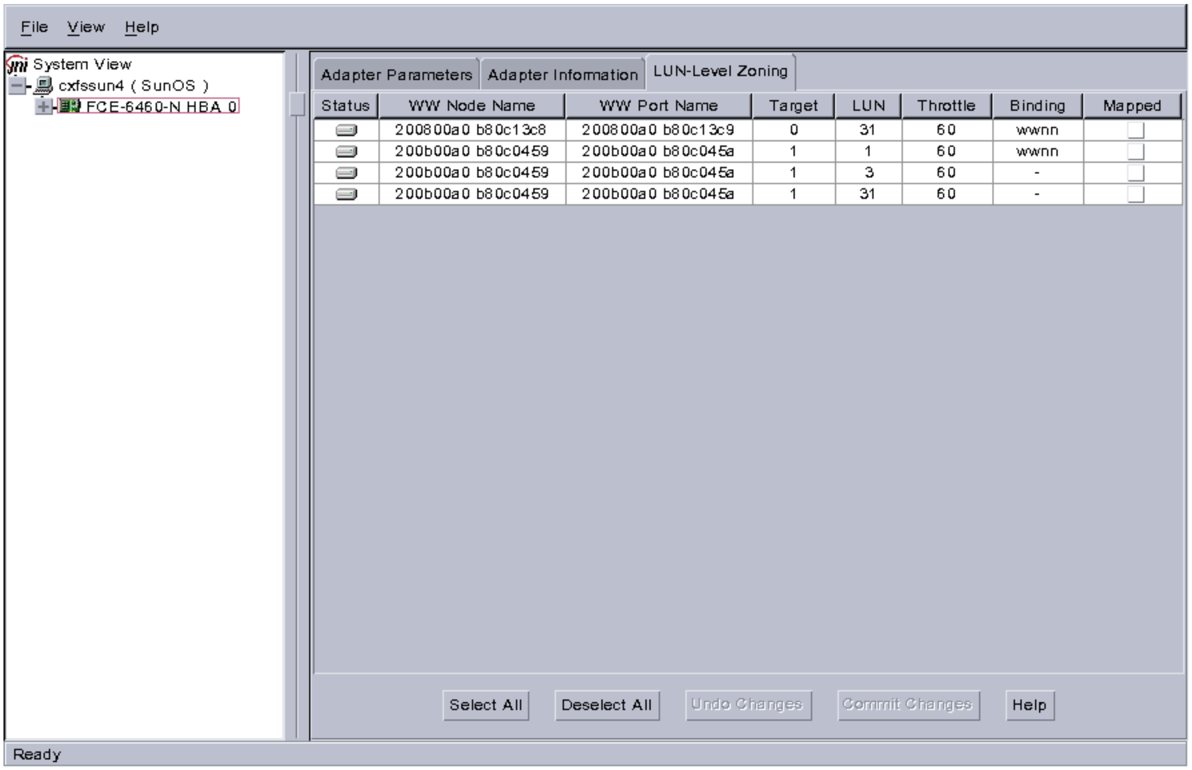
The last two lines show the WWNN and WWPN of the JNI HBA. You will need the WWPN numbers when you configure the `/etc/fencing.conf` file, so take note of them now; for more information about the `/etc/fencing.conf` file, see "Postinstallation Steps for Solaris: Creating the I/O Fencing File", page 49.



Adapter Variable	Value
Chip Version	Emerald IV - A1 (JNIC-1460)
Class Type	2, 3
Device Name	/devices/pci@1f,4000/JNI,FCR@5:c,raw
Driver Instance	jnic146x2
Driver Name	jnic146x
Driver Version	JNIC v5.1 (01110800)
FCode Version	Version 3.6.2
HBA Model	FCE-6460-N
HBA Serial Number	0013360
HBA Type	PCI
Link Speed	1 Gbps
Link Status	UP
Port ID	031100
Port Type	Fabric (NPORT)
Slot Number	5
World Wide Node Name	1000000173003430
World Wide Port Name	2000000173003430

Figure 5-6 After Clicking the Adapter Information Tab

b. Click on the **LUN-Level Zoning** tab in the left-hand pane to display a list of all the known devices on the selected HBA, as shown in Figure 5-7.



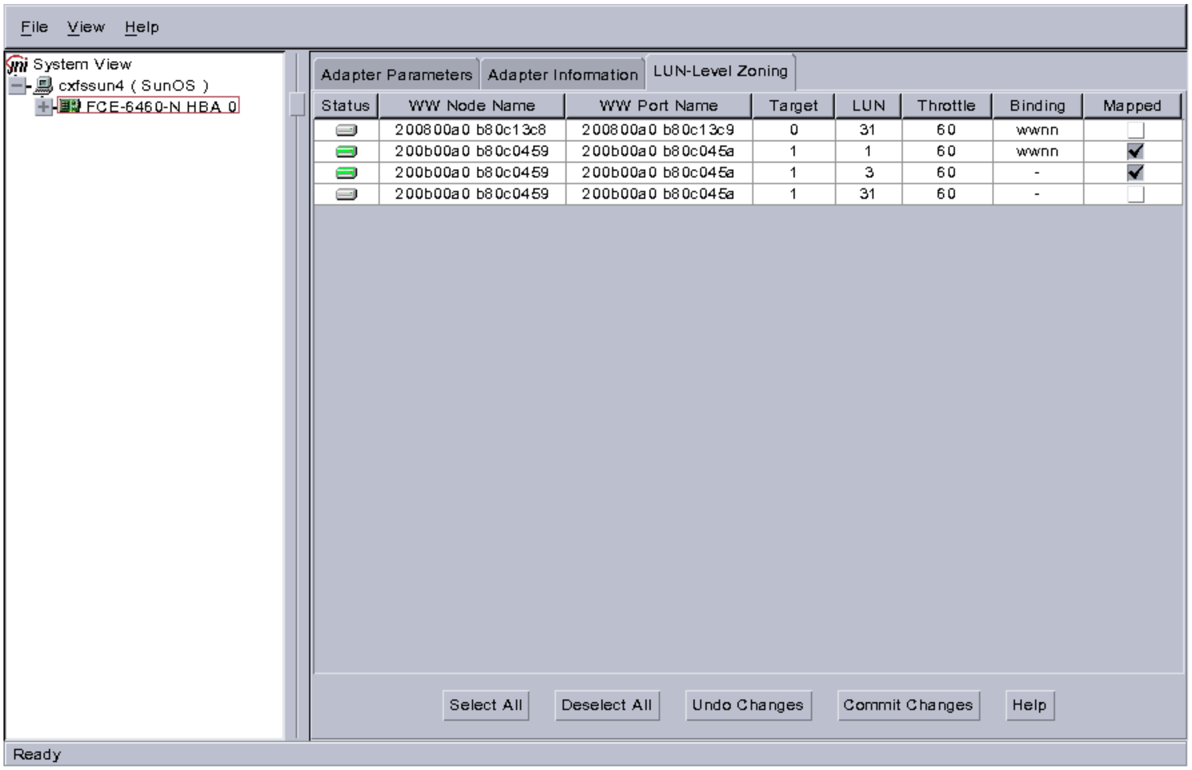
Status	WW Node Name	WW Port Name	Target	LUN	Throttle	Binding	Mapped
▭	200800a0 b80c13c8	200800a0 b80c13c9	0	31	60	wwnn	□
▭	200b00a0 b80c0459	200b00a0 b80c045a	1	1	60	wwnn	□
▭	200b00a0 b80c0459	200b00a0 b80c045a	1	3	60	-	□
▭	200b00a0 b80c0459	200b00a0 b80c045a	1	31	60	-	□

Figure 5-7 After Clicking on LUN-Level Zoning

5. Select the devices that should be accessed through the HBA.

   For each device you want to access, click on the corresponding box in the **Mapped** column to make a check mark appear, as shown in Figure 5-8. After you have selected all the desired devices for the HBA, click on **Commit Changes**. The LUNs you map will depend upon your own site's needs.

⚠ **Caution:** In this example, LUN 31 is used for administration by the TP9400. This LUN must not be used for other purposes; do not map it or use it for XVM volumes.



Status	WW Node Name	WW Port Name	Target	LUN	Throttle	Binding	Mapped
⬜	200800a0 b80c13c8	200800a0 b80c13c9	0	31	60	wwnn	□
🟩	200b00a0 b80c0459	200b00a0 b80c045a	1	1	60	wwnn	☑
🟩	200b00a0 b80c0459	200b00a0 b80c045a	1	3	60	-	☑
⬜	200b00a0 b80c0459	200b00a0 b80c045a	1	31	60	-	□

Figure 5-8 Example: After Mapping the LUNs and Committing the Changes

6. Reboot the system to make the changes take effect:

```
# init 6
```

Verifying the JNI HBA Installation

After the system reboots, you should verify that the devices were correctly configured by running the Solaris `format`(1M) command. You should see a list of each device you selected.

For example:

```
# format
Searching for disks...  done

c4t1d1: configured with capacity of 133.99GB
c4t1d3: configured with capacity of 133.99GB


AVAILABLE DISK SELECTIONS:
       0. c0t0d0 <SUN9.0G cyl 4924 alt 2 hd 27 sec 133>
          /pci@1f,4000/scsi@3/sd@0,0
       1. c4t1d1 <SGI-TP9400-0401 cyl 65533 alt 2 hd 64 sec 67>
          /pci@1f,4000/JNI,FCR@5/sd@1,1
       2. c4t1d3 <GI-TP9400-0401 cyl 65533 alt 2 hd 64 sec 67>
          /pci@1f,4000/JNI,FCR@5/sd@1,3
Specify disk (enter its number):
```

In this example, disks 1 and 2 are being addressed by the JNI driver, as indicated by the presence of JNI,FCR in the pathname.

The system log and console display may display warning messages because the disks have IRIX labels on them. For example:

```
Mar  5 14:17:33 cxfssun4 scsi: WARNING: /pci@1f,4000/JNI,FCR@5/sd@1,1 (sd154):
Mar  5 14:17:33 cxfssun4        corrupt label - wrong magic number
Mar  5 14:17:33 cxfssun4 scsi:     Vendor 'SGI', product 'TP9400', 284203008 512 byte blocks
Mar  5 14:17:33 cxfssun4 scsi:  WARNING: /pci@1f,4000/JNI,FCR@5/sd@1,3 (sd155):
Mar  5 14:17:33 cxfssun4        corrupt label - wrong magic number
Mar  5 14:17:33 cxfssun4 scsi:     Vendor 'SGI', product 'TP9400', 284203008 512 byte blocks
```

This situation will be corrected automatically by CXFS after it is installed.

Note: You should not be alarmed by the preceding messages, nor should you try to relabel the disks with the format command. At this point, you are only trying to achieve connectivity to the devices, and the content is not important.

If you are having trouble with the verification steps, see "Common Solaris Problems: JNI Problems", page 115.

Preinstallation Steps for Solaris

When you install the CXFS software on the client-only node, you must modify certain system files. **The network configuration is critical.** Each node in the cluster must be able to communicate with every other node in the cluster by both logical name and IP address without going through any other network routing; proper name resolution is key. SGI recommends static routing.

This section provides an overview of the steps that you or a qualified Sun service representative will perform on your Solaris nodes prior to installing the CXFS software. It contains the following sections:

- "Hostname Resolution and Network Configuration Rules for Solaris"
- "Adding a Private Network for Solaris Nodes", page 42
- "Verifying the Private and Public Networks for Solaris", page 47

Hostname Resolution and Network Configuration Rules for Solaris

⚠ **Caution:** It is critical that you understand these rules before attempting to configure a CXFS cluster.

The following hostname resolution rules and recommendations apply to client-only nodes:

- Hostnames cannot begin with an underscore (_) or include any whitespace characters.
- The private network IP addresses on a running node in the cluster cannot be changed while cluster services are active.
- You must be able to communicate directly between every node in the cluster (including client-only nodes) using IP addresses and logical names, without routing.
- A private network must be dedicated to be the heartbeat and control network. No other load is supported on this network.
- The heartbeat and control network must be connected to all nodes, and all nodes must be configured to use the same subnet for that network.

If you change hostname resolution settings in the `/etc/nsswitch.conf` file after you have defined the first IRIX node (which creates the cluster database), you must recreate the cluster database.

Adding a Private Network for Solaris Nodes

The following procedure provides an overview of the steps required to add a private network to the Solaris system.

Note: A private network is **required** for use with CXFS. Only the private network is used by CXFS for heartbeat/control messages.

You may skip some steps, depending upon the starting conditions at your site. For details about any of these steps, see the Solaris documentation.

1. If your system is already operational and on the network, skip to step 2.

   If your Solaris system has **never** been set up, bring the system to single-user mode. For example, go to the PROM prompt and boot the Solaris node into single-user mode:

   ```
   > boot -s
   ```

   As a last resort, you can reach the PROM prompt by pressing the `L1-A` (or `Stop-A`) key sequence.

2. Edit the `/etc/inet/hosts` (or `/etc/hosts`) file so that it contains entries for every node in the cluster and their private interfaces as well.

   The `/etc/inet/hosts` (or `/etc/hosts`) file has the following format, where *primary_hostname* can be the simple hostname or the fully qualified domain name:

   ```
   IP_address    primary_hostname        aliases
   ```

   You should be consistent when using fully qualified domain names in the `/etc/inet/hosts` (or `/etc/hosts`) file. If you use fully qualified domain names on a particular node, then all of the nodes in the cluster should use the fully qualified name of that node when defining the IP/hostname information for that node in their `/etc/inet/hosts` (or `/etc/hosts`) file.

The decision to use fully qualified domain names is usually a matter of how the clients (such as NFS) are going to resolve names for their client server programs, how their default resolution is done, and so on.

Even if you are using the domain name service (DNS) or the network information service (NIS), you must add every IP address and hostname for the nodes to /etc/inet/hosts (or /etc/hosts) on all nodes. For example:

```
190.0.2.1 server1-company.com server1
190.0.2.3 stocks
190.0.3.1 priv-server1
190.0.2.2 server2-company.com server2
190.0.2.4 bonds
190.0.3.2 priv-server2
```

You should then add all of these IP addresses to /etc/inet/hosts (or /etc/hosts) on the other nodes in the cluster.

For more information, see the hosts(4), named(1M), and nis(1) man pages.

Note: Exclusive use of NIS or DNS for IP address lookup for the nodes will reduce availability in situations where the NIS or DNS service becomes unreliable.

For more information, see "Hostname Resolution and Network Configuration Rules for Solaris", page 41.

3. Edit the /etc/nsswitch.conf file so that local files are accessed before either NIS or DNS. That is, the hosts line in /etc/nsswitch.conf must list files first.

   For example:

   ```
   hosts:      files nis dns
   ```

   (The order of nis and dns is not significant to CXFS, but files must be first.)

4. Determine the name of the private interface by using the ifconfig(1M) command as follows:

   ```
   # ifconfig -a
   ```

   If the second network does not appear, it may be that a network interface card must be installed in order to provide a second network, or it may be that the network is not yet initialized.

For example, on an Ultra Enterprise 250, the integrated Ethernet is hme0; this is the public network. The following ifconfig output shows that only the public interface exists:

```
# ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
        inet 127.0.0.1 netmask ff000000
hme0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
        inet 128.162.2.91 netmask ffffff00 broadcast 128.162.2.255
        ether 8:0:20:d2:29:c5
```

If the second network does not appear, do the following:

a. If you do not have the PCI card installed, install it. Refer to your PCI documentation for instructions.

   If your card is already installed, skip to step b.

b. Use the output from the dmesg(1M) command to determine the interface name for the private network; look for the network interface that immediately follows the public network; you may wish to search for Found. For example:

```
# dmesg

Feb  6 09:38:36 ue250 last message repeated 42 times
Feb  6 11:38:40 ue250 pseudo: [ID 129642 kern.info] pseudo-device: devinfo0
Feb  6 11:38:40 ue250 genunix: [ID 936769 kern.info] devinfo0 is /pseudo/devinfo@0
Feb  6 11:38:41 ue250 hme: [ID 517527 kern.info] SUNW,hme0 : PCI IO 2.0 (Rev Id = c1) Found
Feb  6 11:38:41 ue250 genunix: [ID 936769 kern.info] hme0 is /pci@1f,4000/network@1,1
Feb  6 11:38:41 ue250 hme: [ID 517527 kern.info] SUNW,hme1 : PCI IO 2.0 (Rev Id = c1) Found
Feb  6 11:38:41 ue250 hme: [ID 517527 kern.info] SUNW,hme1 : Local Ethernet address = 8:0:20:cc:43:48
Feb  6 11:38:41 ue250 pcipsy: [ID 370704 kern.info] PCI-device: SUNW,hme@1,1, hme1
Feb  6 11:38:41 ue250 genunix: [ID 936769 kern.info] hme1 is /pci@1f,2000/SUNW,hme@1,1
```

   The second network is hme1; this is the private network, and is displayed after hme0 in the dmesg output. In this example, hme1 is the value needed in step c and in step 5 below.

c. Initialize the private network's interface by using the ifconfig(1M) command as follows, where *interface* is the value determined in step b:

   ifconfig *interface* plumb

For example:

```
# ifconfig hme1 plumb
```

After performing the `plumb`, the `hme1` interface will appear in the `ifconfig` output, although it will not contain the appropriate information (the correct information will be discovered after the system is rebooted later in step 8). For example, at this stage you would see the following:

```
ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
        inet 127.0.0.1 netmask ff000000
hme0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
        inet 128.162.2.91 netmask ffffff00 broadcast 128.162.2.255
        ether 8:0:20:d2:29:c5
hme1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
        inet 0.0.0.0 netmask ff000000 broadcast 255.0.0.0
        ether 8:0:20:d2:29:c5
```

5. Create a file named `/etc/hostname.`*interface*, where *interface* is the value determined in step 4. This file must contain the name of the **private** network. For example:

   ```
   # cat /etc/hostname.hme1
   cxfssun3-priv
   ```

   Note: In this scenario, `/etc/hostname.hme0` must contain the same value as the `/etc/nodename` file. For example:

   ```
   # cat /etc/hostname.hme0
   cxfssun3
   # cat /etc/nodename
   cxfssun3
   ```

   The Solaris `/etc/nodename` file is analogous to the IRIX `/etc/sys_id` file.

6. Edit the `/etc/netmasks` file to include the appropriate entries.

7. *(Optional)* Edit the `/.rhosts` file if you want to use remote access or if you want to use the connectivity diagnostics provided with CXFS.

   Ensure that the `/.rhosts` file on each Solaris node allows all of the nodes in the cluster to have access to each other. The connectivity tests execute a `ping`(1) command from the local node to all nodes and from all nodes to the local node. To execute `ping` on a remote node, CXFS uses `rsh`(1) as user `root`.

For example, suppose you have a cluster with three nodes: `irix0`, `sun1`, and `sun2`. The `/.rhosts` files could be as follows (the prompt denotes the node name):

```
irix0# cat /.rhosts
sun1 root
sun1-priv root
sun2 root
sun2-priv root

sun1# cat /.rhosts
irix0 root
irix0-priv root
sun2 root
sun2-priv root

sun2# cat /.rhosts
irix0 root
irix0-priv root
sun1 root
sun1-priv root
```

8. Reboot the Solaris system:

```
# init 6
```

At this point, `ifconfig` will show the correct information for the private network.

For example:

```
ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
        inet 127.0.0.1 netmask ff000000
hme0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
        inet 128.162.2.91 netmask ffffff00 broadcast 128.162.2.255
        ether 8:0:20:d2:29:c5
hme1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
        inet 10.1.1.36 netmask ffffff00 broadcast 10.1.1.255
        ether 8:0:20:d2:29:c5
```

Verifying the Private and Public Networks for Solaris

For each private network on each Solaris node in the pool, verify access with the Solaris `ping`(1) command. Enter the following, where *nodeIPaddress* is the IP address of the node:

```
# /usr/sbin/ping -s -c 3 nodeIPaddress
```

For example:

```
# /usr/sbin/ping -s -c 3 128.162.2.91
PING 128.162.2.91: 56 data bytes
64 bytes from cxfssun3.americas.sgi.com (128.162.2.91): icmp_seq=0. time=0. ms
64 bytes from cxfssun3.americas.sgi.com (128.162.2.91): icmp_seq=1. time=0. ms
64 bytes from cxfssun3.americas.sgi.com (128.162.2.91): icmp_seq=2. time=0. ms
64 bytes from cxfssun3.americas.sgi.com (128.162.2.91): icmp_seq=3. time=0. ms
```

Also execute a `ping` on the public networks. If `ping` fails, follow these steps:

1. Verify that the network interface was configured up using `ifconfig`; for example:

```
# /usr/sbin/ifconfig eri0
eri0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
        inet 128.162.2.127 netmask ffffff00 broadcast 128.162.2.255
        ether 0:3:ba:d:ad:77
```

   In the first output line above, `UP` indicates that the interface was configured up.

2. Verify that the cables are correctly seated.

Repeat this procedure on each node.

Client Software Installation Steps for Solaris

The CXFS software will be initially installed and configured by SGI personnel. This section provides an overview of those procedures. You can use the information in this section to verify the installation.

Installation Overview

Installing the CXFS client CD for Solaris requires approximately 20 MB of space.

To install the required software on a Solaris node, SGI personnel will do the following:

1. Verify that the node has been upgraded to Solaris 8 (also known as SunOS 5.8) or Solaris 9 (also known as SunOS 5.9) according to the Solaris installation guide. Use the following command to display the currently installed system:

   ```
   # uname -r
   ```

   This command should return a value of 5.8 or 5.9.

2. Do the following:

   a. Insert the CXFS MultiOS CD-ROM.

   b. Read the already inserted CD-ROM as follows:

      - Solaris 8:

        ```
        # pkgadd -d /cdrom/cdrom01/solaris/SGIcxfs-sol8.pkg
        ```

      - Solaris 9

        ```
        # pkgadd -d /cdrom/cdrom01/solaris/SGIcxfs-sol9.pkg
        ```

      For example, installing SGIcxfs-sol8.pkg under Solaris 8 will display at least the following output, although the exact version numbers may differ:

```
# pkgadd -d /cdrom/cdrom01/solaris/SGIcxfs-sol8.pkg
The following packages are available:
  1 SGIcxfs     SGI CXFS client software
                (sparc) release 2.4

Select package(s) you wish to process (or ²all² to process
all packages). (default: all) [?,??,q]:

Processing package instance <SGIcxfs> from </cdrom/solaris/SGIcxfs-sol8.pkg>
. . .
```

   c. Verify that the CXFS license key has been installed. See Chapter 4, "Obtaining CXFS and XVM FLEXlm Licenses", page 21.

For example:

```
# /usr/cxfs_cluster/bin/cxfslicense -d
CXFS license granted.
```

Verifying the Solaris Installation

To verify that the CXFS software has been installed properly, use the `pkginfo`(1M) command as follows:

```
pkginfo -l SGIcxfs
```

For example, the following output indicates that the CXFS package installed properly:

```
% pkginfo -l SGIcxfs
   PKGINST:  SGIcxfs
      NAME:  SGI CXFS MultiOS client software
  CATEGORY:  system
      ARCH:  sparc
   VERSION:  release 2.4
   BASEDIR:  /
    VENDOR:  Silicon Graphics Inc.
```

Postinstallation Steps for Solaris: Creating the I/O Fencing File

Note: Solaris nodes use the *I/O fencing* feature, which isolates a problem node so that it cannot access I/O devices and therefore cannot corrupt data in the shared CXFS filesystem. Solaris nodes do not have reset lines and therefore require I/O fencing to protect data integrity. I/O fencing can only be used with a Brocade Fibre Channel switch that is sold and supported by SGI; therefore, the Brocade switch is a required piece of hardware in a multiOS cluster.

To use I/O fencing, you must create the Solaris `/etc/fencing.conf` file, which enumerates the worldwide port name for all of the JNI host bus adapters (HBA) that will be used to mount a CXFS filesystem. There must be a line for the JNI HBA worldwide port name (WWPN) as a 64-bit hexadecimal number.

You must update the `/etc/fencing.conf` file whenever the JNI HBA configuration changes, including the replacement of a JNI HBA.

To determine the HBA WWPN, you must first set up the Brocade Fibre Channel switch and JNI HBA according to the directions in Chapter 3, "Brocade Fibre Channel Switch Verification", page 17, and "JNI Fibre Channel Host Bus Adapter Installation and Configuration", page 28.

At this point, you might be able to determine the HBA WWPN by running the EZ Fibre Configuration GUI: see "Installing and Running the EZ Fibre Configuration GUI", page 30, and Figure 5-6, page 37. If so, and you are **completely certain** that you can determine the correct WWPN of the HBA (and **not** that of any of the SAN targets), you can enter this value in the `/etc/fencing.conf` file.

Note: The WWPN is that of the JNI HBA itself, **not** any of the devices that are visible to that HBA in the fabric.

If you are **not** completely certain which number you should use, do the following:

1. Follow the Fibre Channel cable on the back of the Solaris host to determine the port to which it is connected in the Brocade Fibre Channel switch. Ports are numbered beginning with 0. (For example, if there are 8 ports, they will be numbered 0 through 7.)

2. Use the `telnet`(1) command to connect to the Brocade Fibre Channel switch and log in as user `admin` (the password is `password` by default).

3. Execute the `switchshow` command to display the switches and their WWPN numbers.

   For example:

   ```
   brocade04:admin> switchshow
   switchName:     brocade04
   switchType:     2.4
   switchState:    Online
   switchRole:     Principal
   switchDomain:   6
   switchId:       fffc06
   switchWwn:      10:00:00:60:69:12:11:9e
   switchBeacon:   OFF
   port  0: sw  Online        F-Port  20:00:00:01:73:00:2c:0b
   port  1: cu  Online        F-Port  21:00:00:e0:8b:02:36:49
   port  2: cu  Online        F-Port  21:00:00:e0:8b:02:12:49
   port  3: sw  Online        F-Port  20:00:00:01:73:00:2d:3e
   ```

```
port  4: cu  Online       F-Port  21:00:00:e0:8b:02:18:96
port  5: cu  Online       F-Port  21:00:00:e0:8b:00:90:8e
port  6: sw  Online       F-Port  20:00:00:01:73:00:3b:5f
port  7: sw  Online       F-Port  20:00:00:01:73:00:33:76
port  8: sw  Online       F-Port  21:00:00:e0:8b:01:d2:57
port  9: sw  Online       F-Port  21:00:00:e0:8b:01:0c:57
port 10: sw  Online       F-Port  20:08:00:a0:b8:0c:13:c9
port 11: sw  Online       F-Port  20:0a:00:a0:b8:0c:04:5a
port 12: sw  Online       F-Port  20:0c:00:a0:b8:0c:24:76
port 13: sw  Online       L-Port  1 public
port 14: sw  No_Light
port 15: cu  Online       F-Port  21:00:00:e0:8b:00:42:d8
```

The WWPN is the hexadecimal string to the right of the port number. For example, the WWPN for port 0 is `2000000173002c0b` (you must remove the colons from the WWPN reported in the `switchshow` output to produce the string to be used in the `/etc/fencing.conf` file).

4. Edit or create the `/etc/fencing.conf` file and add the WWPN for the port determined in step 1. (Comment lines begin with #.)

   For example, if you determined that port 0 is the port connected to the Brocade Fibre Channel switch, your `/etc/fencing.conf` file should appear as follows:

   ```
   # WWPN of the JNI HBA installed on this system
   #
   2000000173002c0b
   ```

5. After the Solaris node is added to the cluster (see Chapter 7, "Cluster Configuration", page 99), enable the fencing feature by using the CXFS GUI or `cmgr` command on an IRIX node; for more information, see the *CXFS Version 2 Software Installation and Administration Guide*.

Manual CXFS Startup/Shutdown for Solaris

The `/etc/init.d/cxfs_cluster` script will be invoked automatically during normal system startup and shutdown procedures. This script starts and stops the processes required to run CXFS.

To start up CXFS processes manually on your Solaris node, enter the following:

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

To stop CXFS processes manually, enter the following:

```
# /etc/init.d/cxfs_cluster stop
```

Software Maintenance for Solaris

This section contains the following:

- "Upgrading the CXFS for Solaris Software"
- "Modifying the CXFS for Solaris Software", page 52

Upgrading the CXFS for Solaris Software

Before upgrading CXFS software, ensure that no applications on the node are accessing files on a CXFS filesystem. You can then run the new CXFS software package, which will automatically upgrade all CXFS software.

Modifying the CXFS for Solaris Software

You can modify the CXFS client service (/usr/cxfs_cluster/bin/cxfs_client) by placing options in the /usr/cxfs_cluster/bin/cxfs_client.options file. The available options are documented in the cxfs_client(1M) man page.

⚠ **Caution:** Some of the options are intended to be used internally by SGI only for testing purposes and do not represent supported configurations. Consult your SGI service representative before making any changes.

The first line in the cxfs_client.options file must contain the options you want cxfs_client to process; you cannot include a comment as the first line.

To see if cxfs_client is using the options in cxfs_client.options, enter the following:

```
# ps -ef | grep cxfs
```

Chapter 6

Windows NT and Windows 2000 Platforms

The information in this chapter applies to both Windows NT and Windows 2000 nodes unless otherwise noted. This chapter contains the following:

- "CXFS on Windows"
- "QLogic Fibre Channel Host Bus Adapter Installation for Windows", page 71
- "Preinstallation Steps for Windows", page 75
- "Client Software Installation Steps for Windows", page 81
- "Postinstallation Steps for Windows", page 85
- "Manual CXFS Startup/Shutdown for Windows", page 93
- "Software Maintenance for Windows", page 94

CXFS on Windows

This section contains the following:

- "Requirements Specific to Windows", page 54
- "CXFS Commands Installed on Windows", page 55
- "Windows Log Files and Cluster Status", page 55
- "Functional Limitations Specific to Windows", page 56
- "Performance Considerations on a CXFS Windows Node", page 59
- "Access Controls on a Windows Node", page 60

Requirements Specific to Windows

In addition to the items listed in "Requirements", page 6, using a Windows node to support CXFS requires the insertion of a Windows host with at least the following:

- An Intel Pentium or compatible processor.
- 128 MB of RAM (more will improve performance).
- A minimum of 10 MB of free disk space.
- A QLogic 2200 or QLogic 2310 host bus adapter.
- The following QLogic software from the `http://www.qlogic.com` website:
  - QLogic 2200/2310 driver version 8.1.5.12 or later.
  - QLogic SANSurfer 2.0.15 or later. You must install both the SANblade NT Agent and the SANblade Manager.
  - QLogic 2200 BIOS version 1.76 or later, or QLogic SANblade 2310 BIOS version 1.17 or later.

    Note: If your configuration has dual HBAs, you should keep the HBA firmware and HBA driver at the same version on both cards.

  You should install the documentation associated with the software. See the SANblade Manager `README` for the default password. Follow the QLogic instructions to install the driver, the SANblade NT Agent, and the SANblade Manager software. If you do not have the correct QLogic BIOS version installed, see the procedure in "Upgrading the QLogic BIOS", page 76.

- If two QLogic HBAs are installed, you should also install the QLdirect Filter (8.1.3 or later) in order to facilitate HBA failover and load balancing. If two different model HBAs are installed, you must install drivers for both models.

  Note: If the primary HBA path is at fault during the Windows boot up (for example, if the Fibre Channel cable is disconnected), no failover to the secondary HBA path will occur. This is a limitation of the QLogic driver.

- Windows NT 4.0 Service Pack 6, or Windows 2000 Service Pack 2 or later.

CXFS Commands Installed on Windows

A single CXFS service and a single CXFS filesystem driver are installed as part of the Windows installation. The service and the CXFS filesystem driver can be configured to run automatically when the first user logs into the node.

The command `C:\Program Files\CXFS\cxfs_license` is installed to assist with license validation; see "Configuring the FLEXlm License for Windows", page 86".

The command `C:\Program Files\CXFS\cxfs_info` is installed to display in a human-readable format the current state of the node in the cluster; see "Windows Log Files and Cluster Status""Verifying the Cluster", page 105.

Windows Log Files and Cluster Status

The Windows node will log important events in the system event log. You can view these events by selecting the following:

- For Windows NT:

  Start
  > Programs
  > Administrative Tools
  > Event Viewer

- For Windows 2000:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Event Viewer

For information about the log files created on IRIX nodes, see the *CXFS Version 2 Software Installation and Administration Guide*. The CXFS Client service will also log important information to the following file:

```
C:\Program Files\CXFS\log\cxfs_client.log
```

You may also wish to keep the **CXFS** window open to check the cluster status. To open this informational window on any Windows system, select the following:

Start
  > Programs
    > CXFS
      > CXFS Info

Functional Limitations Specific to Windows

There are a number of limitations in the CXFS software that are unique to the Windows platform. These limitations are described from both a UNIX and a Windows perspective in the following sections.

UNIX Perspective of CXFS on a Windows Node

This section describes the differences and limitations of a CXFS filesystem on a Windows node from a UNIX perspective:

- Windows nodes can support multiple CXFS filesystems mounted under a single drive letter. Only one CXFS drive letter may be configured on a Windows node.

  The top-level file structure under the CXFS drive letter consists of an in-memory directory structure that mimics the mount points on the IRIX server. The CXFS software creates these directories before mounting the CXFS filesystems. This file structure supports only creating and deleting directories; there is no support for creating and deleting regular files, renaming directories, and so on. Attempts to perform unsupported actions will generally result in an invalid parameter error. You can perform normal filesystem operations on files and directories beneath the mount points.

- A Windows node can support regular files, directories, and links. However, it does not support other XFS file types.

- Symbolic links cannot be distinguished from normal files or directories on a Windows node. Opening a symbolic link will open the target of the link, or will report `file not found` if it is a dangling link.

- You can move, rename, or delete a symbolic link; however, you cannot copy a symbolic link. Copying a valid symbolic link will result in copying the file or directory that the link refers to, rather than the normal UNIX behavior that copies the link itself.

- CXFS Windows nodes cannot use the `edquota`(1M) or `repquota`(1M) commands to view or edit user and group quotas because CXFS administration must be performed on an IRIX node. However, user and group quotas will be enforced correctly, independent of the node in the CXFS cluster where files are created and deleted.

Windows Perspective of CXFS on a Windows Node

This section describes the differences and limitations of a CXFS filesystem on a Windows node in comparison to other Windows filesystems from a Windows perspective:

- CXFS filesystems should not be shared to another Windows host from a CXFS Windows node because there is no support for the opportunistic locking that is used by Windows to guarantee data coherency. If this is required, it is recommended that an IRIX server share the filesystem via Samba.
- Avoid using duplicate filenames in the same directory that vary only in case. CXFS is case sensitive, but some Windows applications may not maintain the case of all filenames, which may result in unexpected behavior.
- CXFS software does not export 8.3 alternative filenames. Older Windows applications that only support 8.3 filenames may be unable to open files with longer filenames.
- Avoid using completely uppercase 8.3 filenames. If you use completely uppercase 8.3 filenames, some applications (including Windows Explorer) may incorrectly assume that only 8.3 filenames are supported by the filesystem and will not preserve case.
- Take care when using Disk Manager and other similar Microsoft and third-party disk tools under Windows NT. These tools assume particular disk formats (FAT or NTFS) and may even write to the super block of a Fibre Channel disk and write a Windows signature on the disk. This will corrupt the XVM volume information on the disk. A CXFS filter driver prevents non-CXFS tools from writing to the disks under Windows 2000.

⚠ **Caution:** Writing a Windows signature to the disk used for CXFS will corrupt the volume and filesystem structure on that disk, potentially resulting in the loss of data on the disk.

- Install the CXFS software components onto a NTFS partition rather than a FAT partition. The security of the following files cannot be guaranteed if these files are installed onto a FAT filesystem:

  ```
  C:\Program Files\CXFS\passwd
  C:\Program Files\CXFS\group
  C:\Program Files\CXFS\fencing.conf
  ```

- There is no recycle bin; deleted files will be permanently deleted.
- There is no automatic notification of directory changes performed by other nodes in the cluster. Applications (such as Windows Explorer) will not automatically update their display if another node adds or removes files from the directory currently displayed.
- A CXFS filesystem cannot be used as the boot partition of a Windows host.
- The volume properties window for the CXFS drive letter will display the total capacity of all mounted filesystems and the largest free space on any one of those filesystems.
- The alignment block size for direct I/O will be the largest block size of the filesystems that are mounted, which may be larger than the block size of the filesystem in use.

Forced Unmount on a Windows Node

SGI recommends that you enable the forced unmount feature on CXFS filesystems. See "Recommendations", page 8, and "Forced Unmount of CXFS Filesystems", page 106, page 97.

On UNIX machines, a forced unmount using the `-k` option to the `umount`(1M) command causes all processes that have open files on the specified filesystem to be unconditionally killed. However, on Windows nodes, a forced unmount has a slightly different effect.

When issued on a Windows node, a forced unmount does not terminate user processes by default. Instead, all open file handles that refer to the filesystem enter a special "zombie" state that blocks any further I/O to the filesystem. The filesystem is then unmounted out from underneath the running processes. Processes continue to run normally with the exception that any further I/O operations on the filesystem using a stale file handle will be failed with a `STATUS_INVALID_HANDLE` error. File handles invalidated by a forced unmount will remain invalid indefinitely. If the filesystem is later remounted, a process must open a fresh file handle before it can resume I/O operations to the filesystem.

There is one exception to this behavior. If a file is memory-mapped by any process at the time of the forced unmount, then it, and every other process that has an open file handle to that file, will be unconditionally killed. This is necessary for memory mapped files on Windows nodes as it is the only way to free the reference to the filesystem and to guarantee that no stale file data remains in memory after the unmount.

Performance Considerations on a CXFS Windows Node

The following are performance considerations on a CXFS Windows node, in addition to the limitations described in "Performance Considerations", page 5:

- If you open the Windows Explorer **Properties** window on a directory, it will attempt to traverse the filesystem in order to count the number and size of all subdirectories and files; this action is the equivalent of running the UNIX `du`(1) command. This can be an expensive operation, especially if performed on directories between the drive letter and the mount points, because it will traverse all mounted filesystems.

- Virus scanners, Microsoft Find Fast, and similar tools that traverse a filesystem are very expensive on a CXFS filesystem. Such tools should be configured so that they do not automatically traverse the CXFS drive letter.

- The mapping from Windows user and group names to UNIX identifiers occurs as the CXFS software starts up. In a Windows domain environment, this process can take a number of seconds per user for usernames that do not have accounts within the domain. If the `/etc/passwd` file contains a number of such users, you should remove users who do not have accounts on the Windows nodes from the `passwd` file that is installed on the Windows nodes.

  This issue has less impact on Windows nodes in a workgroup than on those in a domain because the usernames can be quickly resolved on the node itself, rather than across the network to the domain controller.

- With 1-GB fabric to a single RAID controller, it is possible for one 32–bit 33–MHz QLogic card to reach the bandwidth limitations of the fabric, and therefore there will be no benefit from load balancing two HBAs in the same PCI bus. This can be avoided by using 2-GB fabric and/or multiple RAID controllers.

- For load balancing of two HBAs to be truly beneficial, the host must have at least one of the following three attributes:
  - A 64-bit PCI bus
  - A 66-MHz PCI bus
  - Multiple PCI buses

Access Controls on a Windows Node

The XFS filesystem used by CXFS implements and enforces UNIX mode bits and POSIX access control lists (ACLx), which are quite different to Windows file attributes and access control lists. The CXFS software attempts to map Windows access controls to the UNIX access controls for display and manipulation, but there are a number of features that are not supported, or may result in unexpected behavior, which are described here.

User Identification on a Windows Node

As noted in "User Administration for CXFS", page 4, and "Performing User Configuration", page 89, the UNIX /etc/passwd and /etc/group files must be installed on the CXFS Windows client. Windows user and group names that match entries in the /etc/passwd and /etc/group files will be mapped to those user IDs (UID) and group IDs (GIDs).

The following additional mappings are automatically applied:

- User Administrator is mapped to root
- Group Administrators is mapped to sys

A user's default UNIX GID is the default GID in the passwd file for the user and is not based on a Windows group mapped to a UNIX group name.

You can display the users and groups that have been successfully mapped by running the following in a Windows command shell:

```
C:\Program Files\CXFS\cxfs_info -Dug
```

For example, with the following `passwd` and `group` files:

```
C:\> type C:\Program Files\CXFS\passwd
root::0:0:Super-User:/root:/bin/tcsh
guest::998:998:Guest Account:/usr/people/guest:/bin/csh
fred::1040:402:Fred Costello:/users/fred:/bin/tcsh
diane::1052:402:Diane Green:/users/diane:/bin/tcsh

C:\> type C:\Program Files\CXFS\group
sys::0:root,bin,sys,adm
root::0:root
guest:*:998:
video::402:fred,diane
audio::403:fred
```

The following output indicates that there are five nodes configured for the cluster, with four currently in the cluster, three filesystems mounted, a number of recognized users (`Administrator`, `fred`, `diane`, and `guest`) and three groups (`Administrators`, `video`, and `audio`). The default group for user `guest` is also named `guest` note that in this example, this group did not map to any known Windows group.

```
cxfs_client status [timestamp Nov 20 14:29:11 / generation 48]

CXFS client:
    state: stable (2), cms: up, xvm: up, fs: up
Cluster:
    cluster250 (1) - enabled
Local:
    cxfs1 (5) - enabled
Nodes:
    cxfs1 enabled up 5
    cxfs2 enabled DOWN 1
    cxfs250 enabled up 0
    cxfs3 enabled up 2
    cxfs4 enabled up 4
```

Enforcing Access to Files and Directories

Access controls are enforced on the metadata server using the mapped UID and GID of the user attempting to access the file. Therefore, a user can expect the same access on a Windows node as any other node in the cluster when mounting the same filesystem. Access is determined using the file's ACL, if one is defined, otherwise by using the file's mode bits.

ACLs that are set on any files or directories are also enforced as they would be on any IRIX node. The presentation of ACLs is customized to the interfaces of Windows Explorer, so the enforcement of the ACL may vary from an NTFS ACL that is presented in the same way. A new file will inherit the parent directory default ACL, if one is defined.

The user `Administrator` has read and write access to all files on a CXFS filesystem, in the same way that `root` has super user privileges on a UNIX node.

The following example is a directory listing on the metadata server:

```
# ls -l .
drwxr-x---    2 fred   video           6 Nov 20 13:33 dir1
-rw-r-----    1 fred   audio           0 Nov 20 12:59 file1
-rw-rw-r--    1 fred   video           0 Nov 20 12:59 file2
```

Users will have the following access to the contents of this directory:

- `file1` will be readable and writable to user `fred` and `Administrator` on a CXFS Windows node. It can also be read by other users in group `audio`. No other users, including `diane` and `guest`, will be able to access this file.

- `file2` will be readable by all users, and writable by user `fred`, `diane`, and `Administrator`.

- `dir1` will be readable, writable, and searchable by user `fred` and `Administrator`. It will be readable and searchable by other users in group `video`, and not accessible by all other users.

Viewing and Changing File Attributes with Windows Explorer

File permissions may be viewed and manipulated in two different ways when using Windows Explorer:

- By displaying the list of attributes in a detailed directory listing; this is the most limited approach

- By selecting properties on a file

The only file attribute that is supported by CXFS is the read-only attribute, other attributes will not be set by CXFS and changes to those attributes will be ignored.

If the user is not permitted to write to the file, the read-only attribute will be set. The owner of the file may change this attribute and modify the mode bits. Other users, including the user Administrator, will receive an error message if they attempt to change this attribute.

Marking a file read-only will remove the write bit from the user, group, and other mode bits on the file. Unsetting the read-only attribute will make the file writable by the owner only.

For example, selecting file properties on file1 using Windows Explorer on a CXFS Windows node will display the read-only attribute unset if logged in as Administrator or fred, and it will be set for diane and guest.

Only user fred will be able to change the attribute on these files, which will change the files under UNIX to the following:

```
-r--r-----    1 fred   audio           0 Nov 20 12:59 file1
-r--r--r--    1 fred   video           0 Nov 20 12:59 file2
```

If fred then unset these flags, only he could write to both files:

```
-rw-r-----    1 fred   audio           0 Nov 20 12:59 file1
-rw-r--r--    1 fred   video           0 Nov 20 12:59 file2
```

Viewing and Changing File Permissions with Windows Explorer

By selecting the **Security** tab in the **File Properties** window of a file, a user may view and change a file's permissions with a high level of granularity. However, the Windows Explorer interface is significantly different between Windows NT and Windows 2000.

Windows Explorer will list the permissions of the file's owner, file's group and the Everyone group, which represents the mode bits for other users. Not all Windows permission flags are supported.

Windows NT

Under Windows NT, file1 will be displayed with the following permissions:

```
audio                            Special Access(R)
Everyone                         No Access(None)
fred (Fred Costello)             Special Access(RWPO)
```

Likewise, file2 is displayed as follows:

```
fred (Fred Costello)             Special Access(RWPO)
Everyone                         Special Access(R)
video                            Special Access(R
```

Windows NT uses the term Special Access because the combination permission flags are not a standard Windows NT set, such as Read, Write, or Full Control. Selecting the group video for file2 and then **Special Access...** will raise a dialog with the Windows NT permission flags

CXFS maps the UNIX read mode bit to the Windows NT read (R) flag, write mode bit to the write (W) flag, and execute mode bit to the execute (X) flag. The Windows NT delete (D) flag is not set and is ignored because under UNIX this depends on the permissions of the parent directory. The change permissions (P) and the take ownership (O) flags are only set for the owner of the file or the user Administrator, and will be ignored if set on other users or groups.

Opening properties on dir1 will display the following permissions:

```
Everyone                         Special Access(None)*(Not Stated)
fred (Fred Costello)             Special Access(RWXPO)*(Not Stated)
video                            Special Access(RX)*(Not Stated)
```

The asterisk and the second set of brackets containing Not Stated indicates that there is no default ACL that will be inherited when creating files or subdirectories in that directory.

Windows 2000

Under Windows 2000, the permissions on file1 are displayed as follows:

```
audio (cxfs1\audio)                Allow: Read          Deny:
Everyone                           Allow:               Deny: All
Fred Costello (cxfs1\fred)         Allow: Read, Write   Deny:
```

Using the **Advanced** button, file1 is displayed as follows:

```
Deny   Everyone                    Full Control
Allow  Fred Costello (cxfs1\fred)  Special
Allow  audio (cxfs1\audio)         Read
```

User Everyone represents all users that are not explicitly listed. In this case, all other users have Deny and Full Control, which indicates that they are denied all access.

User fred is listed as having Special access because the permission flags in the next example do not exactly match the standard Windows permissions for read and write access to a file. Select Fred Costello and then click**View/Edit** to display the permission flags listed in Table 6-1. (The table displays the permissions in the order in which they appear in the **View/Edit** window). You can choose to allow or deny each flag, but some flags will be ignored as described in Table 6-1.

Table 6-1 Permission Flags that May Be Edited

Permission	Description
Traverse Folder / Execute File	Used to display and change the execute mode bit on the file or directory
List Folder / Read Data	Used to display and change the read mode bit on the file or directory
Read Attributes	Set if the read mode bit is set; changing this flag has no effect
Read Extended Attributes	Set if the read mode bit is set; changing this flag has no effect
Create Files / Write Data	Used to display and change the write mode bit on the file or directory
Create Folders / Append Data	Set if the write mode bit is set; changing this flag has no effect
Write Attributes	Set if the write mode bit is set; changing this flag has no effect
Write Extended Attributes	Set if the write mode bit is set; changing this flag has no effect
Delete Subfolders and Files	Set for directories if you have write and execute permission on the directory; changing this flag has no effect
Delete	Never set (because delete depends on the parent directory permissions); changing the flag has no effect
Read Permissions	Always set; changing the flag has no effect
Change Permissions	Always set for the owner of the file and the user Administrator; changing this flag has no effect
Take Ownership	Always set for the owner of the file and the user Administrator; changing this flag has no effect

The permissions for `file2` are displayed as follows:

```
Everyone                       Allow: Read           Deny:
video (cxfs1\video)            Allow: Read, Write    Deny:
Fred Costello (cxfs1\fred)     Allow: Read, Write    Deny:
```

The permissions for `dir1` under Windows 2000 are displayed as follows:

```
Fred Costello (cxfs1\fred)     Allow:                Deny:
Video (cxfs1\video)            Allow:                Deny:
```

Note: In this example, the permission flags for directories do not match any of the standard permission sets, therefore no allow or deny flags are set.

In general, you will need to click the **Advanced** button under Windows 2000 to see the actual permissions of directories. For example:

```
Deny   Everyone                Full Control        This folder only
Allow  Fred Costello           Special             This folder only
Allow  audio                   Read & Execute      This folder only
```

The `dir1` directory does not have a default ACL, so none of these permissions are inherited, as indicated by the `This folder only` tag, when a new subdirectory or file is created.

Viewing and Changing File Access Control Lists (ACLs)

If the file or directory has an ACL, the list may include other users and groups, and the `CXFS ACL Mask` group that represents the IRIX ACL mask. See the `chacl`(1) man page for an explanation of IRIX ACLs and the mask bits. The effective permissions of all entries except the owner and the `Everyone` group will be the intersection of the listed permissions for that user or group and the mask permissions. Therefore, changing the `CXFS ACL Mask` permissions will set the maximum permissions that other listed users and groups may have. Their access may be further constrained in the specific entries for those users and groups.

By default, files and directories do not have an ACL, only mode bits, but one will be created if changes to the permissions require an ACL to be defined. For example, granting or denying permissions to another user or group will force an ACL to be created. Once an ACL has been created for a file, the file will continue to have an ACL even if the permissions are reduced back to only the owner or group of the file. The `chacl`(1) command under IRIX can be used to remove an ACL from a file.

Windows NT

For example, `fred` grants `diane` read access to `file1` by adding user `diane` using the file properties dialogs. The access list under Windows NT will then appear as follows:

```
fred (Fred Costello)  Special Access(RWPO)
audio                 Special Access(R)
Everyone              Special Access(None)
diane (Diane Green)   Read(RX)
```

The Windows NT read permission set includes execute, so to change this to read-only you must select **Special Access ...** and deselect **Execute**. Pressing **OK** will cause a dialog to be raised claiming that you have denied everyone access to the file. Under NTFS, this ACL will deny access to everyone including the other listed users and groups, since the deny access on user `Everyone` will be applied before all other ACL entries. You can safely ignore this warning under CXFS (because ordering is not important to a UNIX ACL) and click **Yes** to continue. This will cause an ACL to be added to the file:

```
> ls -lD
-rw-r-----+   1 fred     audio        0 Nov 20 12:59 file1
[u:diane:r--,g::r--,u::rw-,o::---,m::r--]
```

Opening the file permission dialog again on `file1` will display the following:

```
audio                        Special Access(R)
CXFS ACL Mask                Special Access(R)
diane (Diane Green)          Special Access(R)
Everyone                     Special Access(None)
fred (Fred Costello)         Special Access(RWPO)
```

Windows 2000

For example, `fred` grants `diane` read access to `file1` by adding user `diane` using the file properties dialogs, and then deselecting `Read & Execute` so that only `Read` is selected. The access list under Windows 2000 now appears as follows:

```
audio (cxfs1\audio)            Allow: Read            Deny:
Diane Green (cxfs1\diane)      Allow: Read            Deny:
Everyone                       Allow:                 Deny: All
Fred Costello (cxfs1\fred)     Allow: Read, Write     Deny:
```

After clicking **OK**, the properties for `file1` will also include the `CXFS ACL Mask` displayed as follows:

```
audio (cxfs1\audio)              Allow: Read             Deny:
CXFS ACL Mask (cxfs1\CXFS...)    Allow: Read             Deny:
Diane Green (cxfs1\diane)        Allow: Read             Deny:
Everyone                         Allow:                  Deny: All
Fred Costello (cxfs1\fred)       Allow: Read, Write      Deny:
```

Effective Access

The effective access of user `diane` and group `audio` is read-only. Granting write access to user `diane` as in the following example does not give `diane` write access because the mask remains read-only. However, because user `fred` is the owner of the file, the mask does not apply to his access to `file1`.

Windows NT

For example:

```
audio                          Special Access(R)
CXFS ACL Mask                  Special Access(R)
diane (Diane Green)            Special Access(RW)
fred (Fred Costello)           Special Access(RWPO)
```

Windows 2000

For example:

```
audio (cxfs1\audio)              Allow: Read             Deny:
CXFS ACL Mask (cxfs1\CXFS...)    Allow: Read             Deny:
Diane Green (cxfs1\diane)        Allow: Read, Write      Deny:
Fred Costello (cxfs1\fred)       Allow: Read, Write      Deny:
```

Restrictions with file ACLs on Window nodes

If the users and groups listed in a file's permissions (whether mode bits and/or ACL entries) cannot be mapped to users and groups on the Windows node, attempts to display the file permissions in a file properties window will fail with an unknown user or group error. This prevents the display of an incomplete view, which could be misleading.

Both the owner of the file and the user Administrator may change the permissions of a file or directory using Windows Explorer. All other users will get a permission denied error message.

Note: A user must use a node that is **not** running Windows to change the ownership of a file because a Windows user takes ownership of a file with Windows Explorer, rather than the owner giving ownership to another user (which is supported by the UNIX access controls).

Inheritance and Default ACLs on a Windows node

When a new file or directory is created, normally the mode bits are set using a mask of 022. Therefore, a new file has a mode of 644 and a new directory of 755. This umask is defined in the registry of the CXFS driver and may be configured to other values (typically 000 or 002):

```
HKEY_LOCAL_MACHINE->SYSTEM->CurrentControlSet->Services->CXFS->Parameters->DefaultUmask
```

Therefore, creating a file on the IRIX metadata server results in a mode of 644 for a mask of 022:

```
% ls -lda .
drwxr-xr-x    3 fred     video           41 Nov 21 18:01 ./

> umask
22

% touch file3
% ls -l file3
-rw-r--r--    1 fred     video            0 Nov 21 18:23 file3
```

For more information, see the umask(1) man page.

Creating a file in Windows explorer on a Windows node will have the same result.

An IRIX directory ACL may include a default ACL that is inherited by new files and directories, instead of applying the umask. Default ACLs are displayed in the Windows Explorer file permission window if they have been set on a directory. Unlike a Windows inheritable ACL on an NTFS filesystem, an IRIX default ACL applies to both new files and subdirectories, there is no support for an inheritable ACL for new files and another ACL for new subdirectories.

The following example applies an ACL and a default ACL to `dir1` and then creates a file and a directory in `dir1`:

```
% chacl -b "u::rwx,g::r-x,u:diane:r-x,o::---,m::r-x" \
           "u::rwx,g::r-x,u:diane:rwx,o::---,m::rwx" dir1
% touch dir1/newfile
% mkdir dir1/newdir
% ls -D dir1
newdir [u::rwx,g::r-x,u:diane:rwx,o::---,m::r-x/
        u::rwx,g::r-x,u:diane:rwx,o::---,m::rwx]
newfile [u::rw-,g::r-x,u:diane:rwx,o::---,m::r--]
```

The permissions for `dir1` will be as follows:

- Windows NT:

```
CXFS ACL Mask            Special Access(RX)*(RWX)
diane (Diane Green)      Special Access(RX)*(RWX)
Everyone                 Special Access(None)*(None)
fred (Fred Costello)     Special Access(RWXPO)(RWXPO)
video                    Special Access(RX)*(RX)
```

- Windows 2000:

```
CXFS ACL Mask (cxfs1\CXFS...)  Allow:                        Deny:
Diane Green (cxfs1\diane)      Allow:                        Deny:
Everyone                       Allow:                        Deny: All
Fred Costello (cxfs1\fred)     Allow: Read & Exec, List,     Deny:
                                      Read, Write
Video (cxfs1\video)            Allow: Read & Exec, List,     Deny:
                                      Read
```

After clicking the **Advanced** button, the permissions displayed are as follows:,

```
Deny   Everyone        Full Control           This folder, subfolders and files
Allow  Fred Costello   Special                This folder, subfolders and files
Allow  video           Read & Execute         This folder, subfolders and files
Allow  Diane Green     Read, Write & Exec     Subfolders and files
Allow  CXFS ACL Mask   Read, Write & Exec     Subfolders and files
Allow  Diane Green     Read & Exec            This folder only
Allow  CXFS ACL Mask   Read & Exec            This folder only
```

If an ACL entry is the same in the default ACL, a single entry is generated for the `This folder, subfolders and files` entry. Any entries that are different will have both `Subfolders and files` and `This folder only` entries.

Adding the first inheritable entry to a directory will cause CXFS to generate any missing ACL entries like the owner, group, and other users. The mode bits for these entries will be generated from the umask.

The process for modifying default ACLs in Windows NT differs from Windows 2000, because the Windows Explorer file permissions window in Windows NT is limited in its support of ACLs. As with some NTFS ACLs, some CXFS ACLs will cause a dialog to be displayed by Windows Explorer under Windows NT stating that this ACL can only be manipulated by a machine running Windows NT 5.0 (that is, Windows 2000). In these cases, changes to the ACL should be performed on non-Windows NT nodes.

For example, if the directory default ACL includes a user or group entry that is not included in the main ACL for that directory, in this case user `diane`, the Windows NT will be unable to display the ACL:

```
% ls -D dir1
dir1[u::rwx,g::r-x,o::---,m::r-x/u::rwx,g::r-x,u:diane:r-x,o::---,m::r-w]
```

Adding an entry in the non-default ACL for user `diane` will remove this limitation.

Adding different `Subfolders Only` and `Files Only` entries under Windows 2000 will result in only the first entry being used because an IRIX ACL cannot differentiate between the two.

To remove a default ACL entry under Windows NT requires that all you unset all permission flags rather than selecting `None Specified`. (The latter may may result in a new entry being generated.)

QLogic Fibre Channel Host Bus Adapter Installation for Windows

The QLogic Fibre Channel host bus adapter (HBA) should be installed according to the QLogic hardware and driver installation instructions.

Information regarding large logical unit (LUN) support under Windows can be found in the QLogic documentation and also in Microsoft's support database:

`http://support.microsoft.com/default.aspx?scid=kb;en-us;Q310072`

`http://support.microsoft.com/default.aspx?scid=kb;en-us;Q245637`

This section discusses the following:

- "Confirming the QLogic HBA Installation"
- "Support for More than 8 LUNs under Windows NT ", page 72
- "Configuring Two HBAs for Failover Operation", page 73

Confirming the QLogic HBA Installation

To confirm that the QLogic HBA and driver are correctly installed, select the following to display all of the logical units (LUNs) attached to the Fibre Channel switch:

Start
> Settings
> Control Panel
> SCSI Adapters
> QLogic QLA2200
> Bus 0

Support for More than 8 LUNs under Windows NT

The Qlogic HBA will require further configuration to support more than 8 LUNs under Windows NT. This is performed by using the SANsurfer software that is distributed with the QLogic hardware or that can be downloaded from the following QLogic website:

`http://www.qlogic.com/support/drivers_software.asp`

The installation instructions for SANsurfer are also located on the QLogic website.

Upon successful installation of the HBA hardware and software, start the SANsurfer manager application.

Connect to the local host and change the LUNs per target to 0 by selecting the following:

Select NVRAM section
> Advanced NVRAM Settings
> LUNS per Target

Save this configuration and then reboot when prompted.

To confirm that the QLogic HBA and driver are correctly installed after the reboot, check the SCSI registry by selecting the following:

Start
  > Run
    > regedit

Select the following key:

```
HKEY_LOCAL_MACHINE\HARDWARE\DEVICEMAP\Scsi\Scsi Port x\Scsi Bus 0\Target Id 0\Logical Unit Id x-xxx
```

You should see all of the LUNs available in the RAID device, where `Scsi Port x` represents the `QL2xxx` driver.

If you have fewer than 72 LUNs, you should also see the HBA and LUNs in the SCSI adapters list by selecting the following:

Start
  > Settings
    > Control Panel
      > SCSI Adapters
        > QLogic QLA2xxx
          > Bus 0

The limitation of the SCSI adapters list with a large number of LUNs is further described in the following:

`http://support.microsoft.com/default.aspx?scid=kb;en-us;Q190834`

If you are unable to see the SCSI devices in the registry, check the following, as described in "Windows QLogic Problems", page 116:

- The HBA is properly seated in the host
- Cables are connected correctly
- The node is not fenced

Configuring Two HBAs for Failover Operation

Note: This procedure assumes that the CXFS driver is already installed and working properly with one host bus adapter (HBA).

To configure two HBAs for failover operation under Windows, do the following:

1. Install the QLdirect driver v8.01.03 (NT4) by following all the default settings for the installation and verify that the CXFS client still operates normally.
2. Disable fencing for this node. You can do this using the CXFS GUI or the `cmgr`(1M) command.
3. Determine the world wide port name (WWPN) of the current adapter:
   a. Install SANsurfer Qlogic SANblade NT Agent v2.0.15.
   b. Install SANsurfer Qlogic SANblade Manager v2.0.15.
   c. Run SANsurfer to determine the WWPN.
   d. Record the WWPN on paper.
4. Shut down Windows.
5. Install the second HBA and start Windows.
6. If the second HBA is a different model from the original one, install its mini port driver (for example, `ql2300.sys`).
7. Start the Qlogic SANblade Manager and verify that two HBAs are detected. Verify that both of them mirror the same devices and logical units (LUNs). Notice that both HBAs have the same world wide node name (WWNN) but different WWPNs. The original HBA can be recognized by its WWPN recorded in step 3.
8. Make the NVRAM settings (or at least LUNs per target) of the new HBA adapter the same as the original one.
9. Configure the HBA port (click on the **Configure** button).

   Note: Ignore the following message, which appears when HBA/LAN configuration is done for the first time (line breaks added here for readability):

   ```
   An invalid device and LUN configuration has been detected. Auto
   configure run automatically.
   ```

   Click on **OK** to continue.

The HBA0 devices are automatically set to be visible for Windows application (notice the open eye) and HBA1 devices are set to be invisible (notice the closed eye).

10. Select the first device in the table, right click, and then select **Configure LUN(s)**. In the new window, select the following:

    Tools
    > Load Balance
    > All LUNs

    This will statically distribute the LANs traffic load associated with this device between the two HBAs.

    Repeat this step for each of the other HBA devices.

11. Click on **Apply** to save the new configuration.

12. If fencing is used, add the new HBA WWPN to the `fencing.conf` file and update the switch port information using the CXFS GUI or the `cmgr`(1M) command. Enable fencing if required.

13. Reboot Windows.

For more information about using the CXFS GUI or the `cmgr`(1M) command to perform these tasks, see *CXFS Version 2 Software Installation and Administration Guide*.

Preinstallation Steps for Windows

When you install the CXFS software on the client-only node, you must modify certain system files. **The network configuration is critical.** Each node in the cluster must be able to communicate with every other node in the cluster by both logical name and IP address without going through any other network routing; proper name resolution is key. SGI recommends static routing.

This section provides an overview of the steps that you or a qualified Windows service representative will perform on your Windows nodes prior to installing the CXFS software. It contains the following:

- "Upgrading the QLogic BIOS"
- "Hostname Resolution and Network Configuration Rules for Windows", page 77

- "Adding a Private Network for Windows Nodes", page 77
- "Adding a Private Network for Windows 2000 Nodes", page 79
- "Verifying the Private and Public Networks for Windows", page 80

Upgrading the QLogic BIOS

If you need to upgrade the QLogic BIOS, do the following:

Note: If CXFS is already installed and running, stop the CXFS Client service and set it to manual, as described in "Manual CXFS Startup/Shutdown for Windows", page 93, and then restart the machine. You can then perform the following procedure.

---

1. Run the QLogic SANsurfer software and connect to the machine:

   Start
   > Programs
   > QLogic Management Suite
   > SANsurfer
   > Connect

   Note: If you are unable to connect to the machine, you may not have installed the QLogic NT Agent software, which is another option in the SANsurfer software installation.

   ---

2. Enable the BIOS on the HBA by selecting the following:

   Adapter 2200 ***(or Adapter 2310)***
   > NVRAM Settings
   > Enable Host Adaptor BIOS

3. Update the BIOS by selecting the following:

   Adapter 2200 ***(or Adapter 2310)***
   > Utilities
   > Update Flash

   Select the BIOS update file.

4. Mark the CXFS Client service to automatically start.
5. Reboot the machine.

Hostname Resolution and Network Configuration Rules for Windows

Caution: It is critical that you understand these rules before attempting to configure a CXFS cluster.

The following hostname resolution rules and recommendations apply to Windows nodes:

- Hostnames cannot begin with an underscore (_) or include any whitespace characters.
- The private network IP addresses on a running node in the cluster cannot be changed while cluster services are active.
- You must be able to communicate directly between every node in the cluster (including client-only nodes) using IP addresses and logical names, without routing.
- A private network must be dedicated to be the heartbeat and control network. No other load is supported on this network.
- The heartbeat and control network must be connected to all nodes, and all nodes must be configured to use the same subnet for that network.

Adding a Private Network for Windows Nodes

The steps to add a private network are platform-specific.

Adding a Private Network for Windows NT Nodes

The following procedure provides an overview of the steps required to add a private network to the Windows NT node.

Note: A private network is **required** for use with CXFS. Only the private network is used by CXFS for heartbeat/control messages.

You may skip some steps, depending upon the starting conditions at your site.

1. Install the second network adapter in the Windows node as per the network adapter vendor instructions. In some cases you must remove all network setups, restart, and then add network services to each network adapter from scratch.

2. Ensure that the Windows NT node recognizes two network adapters in the system. Select the following:

   Start
   > Settings
   > Control Panel
   > Network
   > Adapters

3. Specify the private network settings (IP address, subnet mask, default gateway) on one of the network adapters. Select the following:

   Start
   > Settings
   > Control Panel
   > Network
   > Protocols
   > TCP/IP Protocol
   > Properties
   > IP Address
   > Adapter

   Note: The private network IP address must be a fixed address and cannot be configured by DHCP.

4. Specify DNS settings as required in the **DNS** tab.

5. Select the software components to be installed. All components should be installed.

6. Skip the **WINS Address tab** (a WINS server is not required).

7. Ensure that **IP Forwarding** is *not* selected in the **Routing** tab.

8. In the **IP Address** tab, ensure that the other network adapter is configured with a different IP address and a different subnet; this second network is the public network for all other network traffic.

Note: The public network interface can be configured with DHCP.

Adding a Private Network for Windows 2000 Nodes

The following procedure provides an overview of the steps required to add a private network to the Windows 2000 node.

Note: A private network is **required** for use with CXFS. Only the private network is used by CXFS for heartbeat/control messages.

You may skip some steps, depending upon the starting conditions at your site.

1. Install the second network adapter in the Windows node as per the network adapter vendor instructions. In some cases you must remove all network setups, restart, and then add network services to each network adapter from scratch.

2. Ensure that the Windows 2000 node recognizes two network adapters in the system. Select the following:

   Start
       > Settings
           > Network and Dial-up Connections

3. Specify the private network settings (IP address, subnet mask, default gateway) on one of the network adapters. Select the following:

   Start
       > Settings
           > Network and Dial-up Connections

Then right-mouse click on **Properties** and select the following:

Internet Protocol (TCP/IP)
> Properties

Note: The private network IP address must be a fixed address and cannot be configured by DHCP.

---

4. Specify the static IP address and DNS server.

Verifying the Private and Public Networks for Windows

You can confirm that the previous procedures to add private networks were performed correctly by using the `ipconfig` command in a DOS command shell. In the following example, the 10 network is the private network and the 192.168.63 network is the public network on a Windows system:

```
> ipconfig /all
Windows NT IP Configuration
        Host Name . . . . . . . . . : cxfs1
        DNS Servers . . . . . . . . :
        Node Type . . . . . . . . . : Hybrid
        NetBIOS Scope ID. . . . . . :
        IP Routing Enabled. . . . . : No
        WINS Proxy Enabled. . . . . : No
        NetBIOS Resolution Uses DNS : No

Ethernet adapter El90x1:
        Description . . . . . . . . : 3Com EtherLink PCI
        Physical Address. . . . . . : 00-01-03-46-2E-09
        DHCP Enabled. . . . . . . . : No
        IP Address. . . . . . . . . : 10.0.0.201
        Subnet Mask . . . . . . . . : 255.0.0.0
        Default Gateway . . . . . . : 10.0.0.255

Ethernet adapter El90x2:
        Description . . . . . . . . : 3Com EtherLink PCI
        Physical Address. . . . . . : 00-B0-D0-31-22-7C
        DHCP Enabled. . . . . . . . : No
        IP Address. . . . . . . . . : 192.168.63.201
```

```
Subnet Mask . . . . . . . . . : 255.255.255.0
Default Gateway . . . . . . . : 192.168.63.254
Primary WINS Server . . . . . : 192.168.63.254
```

Client Software Installation Steps for Windows

The CXFS software will be initially installed and configured by SGI personnel. This section provides an overview of those procedures. You can use the information in this section to verify the installation.

Note: This procedure assumes that the CXFS software is installed under the default path `C:\Program Files\CXFS`. If a different path is selected, then that path should be used in its place in the following instructions.

To install the CXFS client software on a Windows node, do the following:

1. Log onto the Windows node as `Administrator`.
2. Verify that the node has been updated to the correct service pack:
   - Windows NT 4.0 Service Pack 6:

     Start
     > Programs
     > Administrative Tools
     > Windows NT Diagnostics
     > Version

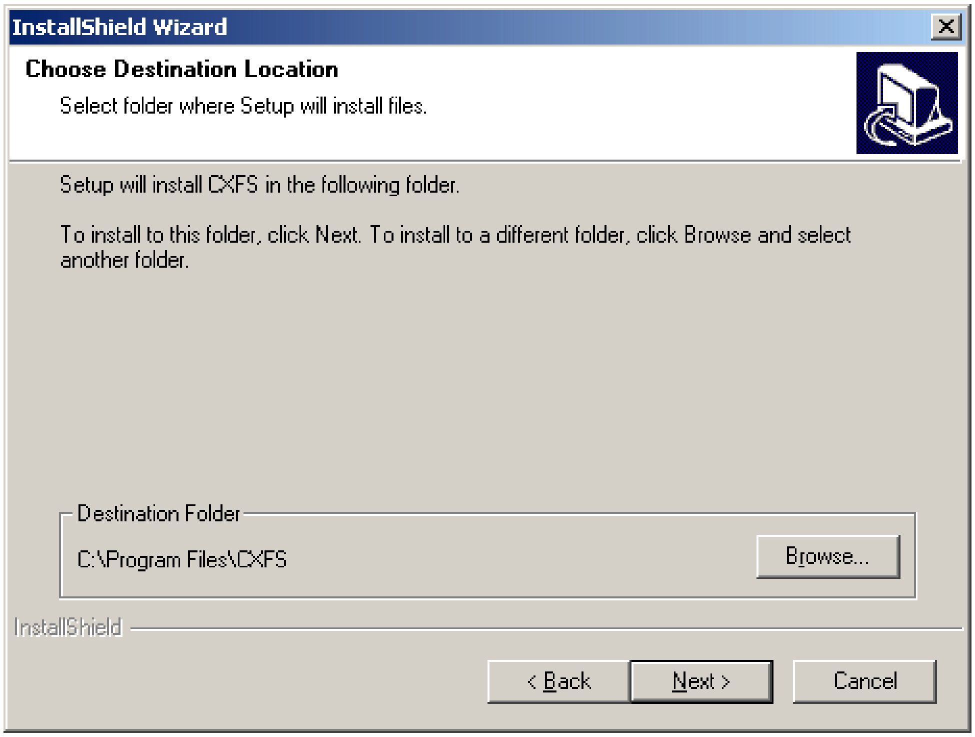
   - Windows 2000 Service Pack 2:

     Start
     > Programs
     > Accessories
     > System Tools
     > System Information

3. Insert the CXFS for Windows CD-ROM into the Windows host. Normally the setup program will automatically run, otherwise run `winnt/setup.exe` from the CD-ROM.

4. Acknowledge the software license agreement when prompted. The release notes will be displayed automatically.

5. Install the CXFS software, as shown in Figure 6-1. If the software is to be installed in a nondefault directory, click on the **Browse** button to select another directory. Click on **Next** when finished.



InstallShield Wizard

Choose Destination Location

Select folder where Setup will install files.

Setup will install CXFS in the following folder.

To install to this folder, click Next. To install to a different folder, click Browse and select another folder.

Destination Folder

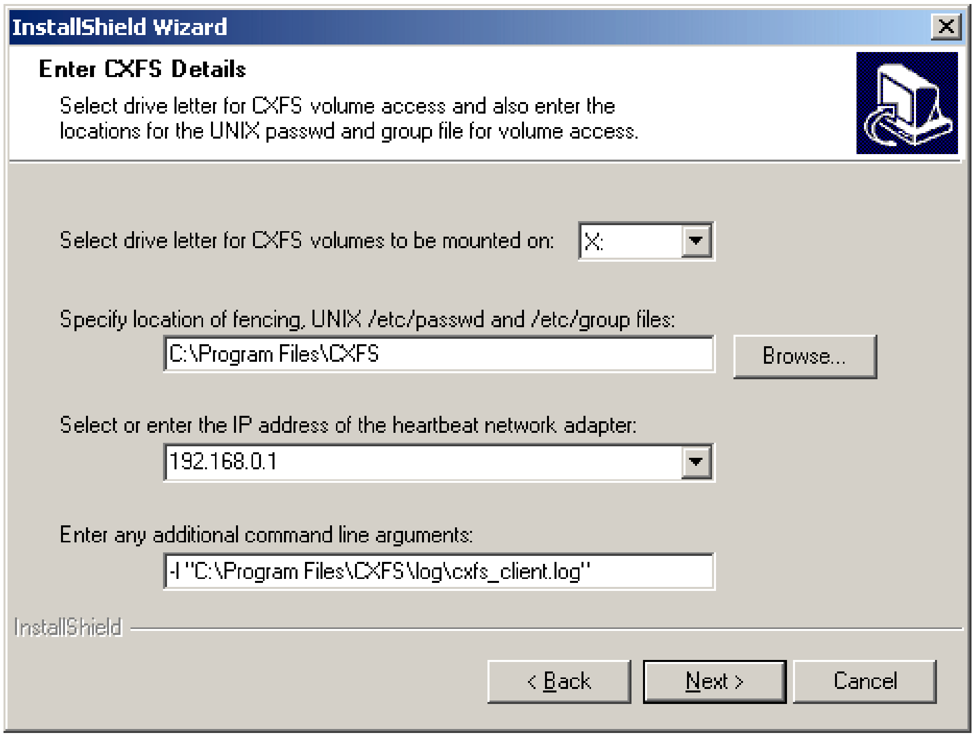
C:\Program Files\CXFS

Figure 6-1 Select All Software Components

6. Enter details for the following fields as shown in Figure 6-2 and click the **Next** button when finished:

   - **Select drive letter for CXFS volumes to be mounted on**: specify the **drive letter** under which all CXFS filesystems will be mounted. You cannot select a drive letter that is currently in use.

- **Specify location of fencing, UNIX /etc/passwd and /etc/group files:** specify the path where the configuration files will be installed. The default is the same location as the software under `C:\Program Files\CXFS`.
- **Select or enter the IP address of the heartbeat network adapter**: specify the IP address of the private network adapter on the Windows node.
- **Enter any additional command line arguments**: enter arguments that may be passed to the CXFS Client service. For most configurations, this can be left empty. See "Modifying the CXFS for Solaris Software", page 52.



Select drive letter for CXFS volume access and also enter the locations for the UNIX passwd and group file for volume access.

Select drive letter for CXFS volumes to be mounted on: X:

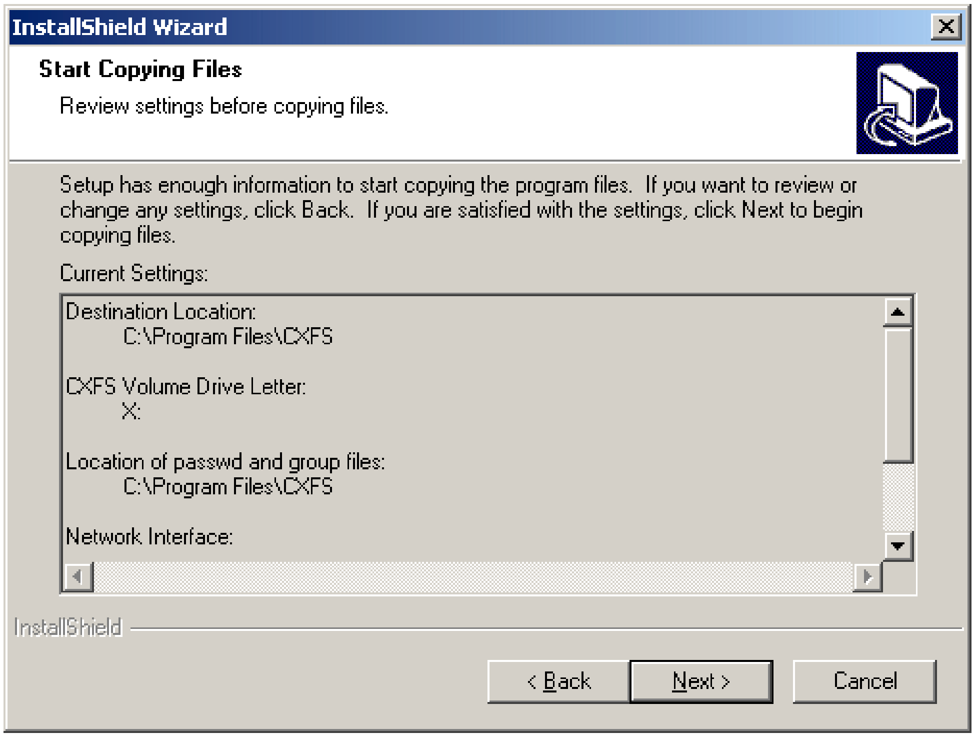
Specify location of fencing, UNIX /etc/passwd and /etc/group files: C:\Program Files\CXFS

Select or enter the IP address of the heartbeat network adapter: 192.168.0.1

Enter any additional command line arguments: -l "C:\Program Files\CXFS\log\cxfs_client.log"

Figure 6-2 Enter CXFS Details

7. Review the settings, as shown in Figure 6-3. If they appear as you intended, click the **Next** button. If you need to make corrections, click the **Back** button.



Start Copying Files

Review settings before copying files.

Setup has enough information to start copying the program files. If you want to review or change any settings, click Back. If you are satisfied with the settings, click Next to begin copying files.

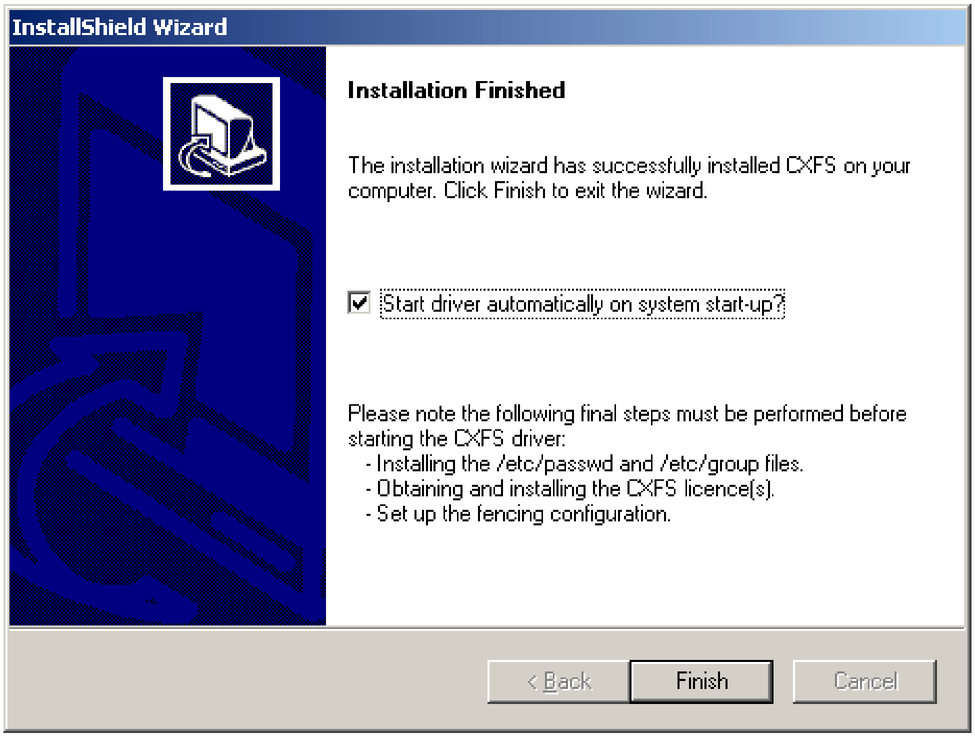
Current Settings:

```
Destination Location:
        C:\Program Files\CXFS

CXFS Volume Drive Letter:
        X:

Location of passwd and group files:
        C:\Program Files\CXFS

Network Interface:
```

Figure 6-3 Review the Settings

After you click the **Next** button, the CXFS software will be installed.

8. You will be given the option to start the driver at system start-up or now, as show in Figure 6-4.



Installation Finished

The installation wizard has successfully installed CXFS on your computer. Click Finish to exit the wizard.

☑ Start driver automatically on system start-up?

Please note the following final steps must be performed before starting the CXFS driver:
- Installing the /etc/passwd and /etc/group files.
- Obtaining and installing the CXFS licence(s).
- Set up the fencing configuration.

Figure 6-4 Start CXFS Driver

Because there are some important postinstallation steps, do not start the CXFS driver now. Choose **Start driver automatically on system start-up** and click the **Finish** button.

Postinstallation Steps for Windows

This section discusses the configuration steps that you should perform after installing CXFS software but before restarting a Windows node.

The following postinstallation steps are required to ensure the correct operation of the CXFS software:

- "Configuring the FLEXlm License for Windows"

- "Creating the Windows I/O Fencing File", page 87
- "Performing User Configuration", page 89
- "Checking Permissions on the Password, Group, and Fencing Files", page 90
- "Creating a New Hardware Profile", page 90

Configuring the FLEXlm License for Windows

Note: Windows NT licenses cannot be used under Windows 2000 and vice versa. If you are upgrading a Windows NT node to Windows 2000, you must obtain a new license.

You must configure a FLEXlm license before you restart the Windows node by following these steps:

1. Add the mandatory CXFS license and the optional XVM license to the following file:

   ```
   C:\Program Files\CXFS\lib\license.dat
   ```

   For more information, see Chapter 4, "Obtaining CXFS and XVM FLEXlm Licenses", page 21.

2. Validate these licenses by running the `cxfslicense` command in a DOS command shell.

   Create a DOS command shell with the following sequence:

   - Windows NT:

     Start
     > Programs
     > Command Prompt

- Windows 2000:

  Start
  > Programs
  > Accessories
  > Command Prompt

To run `cxfslicense`, enter the following command:

```
C:\Program Files\CXFS\cxfslicense.exe
```

If a valid license has been correctly specified, the following will be displayed:

```
Found valid license for feature CXFS_NT version 2.000
The CPU count specified in the license is OK.
```

If the Windows node has the optional XVM mirroring license, you will also see the following:

```
Found valid license for feature XVM_NT version 3.000
The CPU count specified in the license is OK.
```

Note: Licenses for Windows 2000 have the feature names `CXFS_W2K` and `XVM_W2K`.

Creating the Windows I/O Fencing File

Note: Windows nodes use the *I/O fencing* feature, which isolates a problem node so that it cannot access I/O devices and therefore cannot corrupt data in the shared CXFS filesystem. Windows nodes do not have reset lines and therefore require I/O fencing to protect data integrity. I/O fencing can only be used with an SGI-sold and supported Brocade Fibre Channel switch; therefore, the Brocade switch is a required piece of hardware in a multiOS cluster.

To use I/O fencing, you must create the Windows fencing file, which enumerates the worldwide port name for all of the QLogic host bus adapters (HBA) that will be used to mount a CXFS filesystem. The path to the I/O fencing file is as follows:

```
C:\Program Files\CXFS\fencing.conf
```

The I/O fencing file must contain a line for the QLogic HBA worldwide port name (WWPN) as a 64-bit hexadecimal number. You must update the fencing file whenever the QLogic HBA configuration changes, including the replacement of a QLogic HBA.

To determine the HBA WWPN, do the following:

1. Set up the Brocade Fibre Channel switch and QLogic HBA according to the directions in Chapter 3, "Brocade Fibre Channel Switch Verification", page 17, and "QLogic Fibre Channel Host Bus Adapter Installation for Windows", page 71.
2. Follow the Fibre Channel cable on the back of the Windows host to determine the port to which it is connected in the Brocade Fibre Channel switch. Ports are numbered beginning with 0. (For example, if there are 8 ports, they will be numbered 0 through 7.)
3. Use the `telnet`(1) command to connect to the Brocade Fibre Channel switch and log in as user `admin` (the password is `password` by default).
4. Execute the `switchshow` command to display the switches and their WWPN numbers.

   ```
   brocade04:admin> switchshow
   switchName:     brocade04
   switchType:     2.4
   switchState:    Online
   switchRole:     Principal
   switchDomain:   6
   switchId:       fffc06
   switchWwn:      10:00:00:60:69:12:11:9e
   switchBeacon:   OFF
   port  0: sw  Online        F-Port  20:00:00:01:73:00:2c:0b
   port  1: cu  Online        F-Port  21:00:00:e0:8b:02:36:49
   port  2: cu  Online        F-Port  21:00:00:e0:8b:02:12:49
   port  3: sw  Online        F-Port  20:00:00:01:73:00:2d:3e
   port  4: cu  Online        F-Port  21:00:00:e0:8b:02:18:96
   port  5: cu  Online        F-Port  21:00:00:e0:8b:00:90:8e
   port  6: sw  Online        F-Port  20:00:00:01:73:00:3b:5f
   port  7: sw  Online        F-Port  20:00:00:01:73:00:33:76
   port  8: sw  Online        F-Port  21:00:00:e0:8b:01:d2:57
   port  9: sw  Online        F-Port  21:00:00:e0:8b:01:0c:57
   port 10: sw  Online        F-Port  20:08:00:a0:b8:0c:13:c9
   port 11: sw  Online        F-Port  20:0a:00:a0:b8:0c:04:5a
   port 12: sw  Online        F-Port  20:0c:00:a0:b8:0c:24:76
   ```

```
port 13: sw  Online        L-Port  1 public
port 14: sw  No_Light
port 15: cu  Online        F-Port  21:00:00:e0:8b:00:42:d8
```

The WWPN is the hexadecimal string to the right of the port number. For example, the WWPN for port 0 is `2000000173002c0b` (you must remove the colons from the WWPN reported in the `switchshow` output to produce the string to be used in the `C:\Program Files\CXFS\fencing.conf` file).

5. Edit the following file to add the WWPN for the port determined in step 2 (comment lines begin with #):

   ```
   C:\Program Files\CXFS\fencing.conf
   ```

   For example, if you determined that port 0 is the port connected to the Brocade Fibre Channel switch, the fencing file should appear as follows:

   ```
   C\> type C:\Program Files\CXFS\fencing.conf
   #
   # WWPN of the JNI HBA installed on this system
   #
   2000000173002c0b
   ```

6. After the Windows node is added to the cluster (see Chapter 7, "Cluster Configuration", page 99), enable the fencing feature by using the CXFS GUI or `cmgr` command on an IRIX node; for more information, see the *CXFS Version 2 Software Installation and Administration Guide*.

Performing User Configuration

Note: If you do not install the `passwd` and `group` files properly, the CXFS software will treat all filesystem operations as user `nobody`.

After installing the CXFS software onto the Windows node, but before the CXFS node is restarted, you must install the `/etc/passwd` and `/etc/group` files to the location specified during installation, which defaults to the following:

- `/etc/passwd` as `C:\Program Files\CXFS\passwd`
- `/etc/group` as `C:\Program Files\CXFS\group`

Note: The default passwd and group files that are installed are invalid files containing comments; these invalid files will cause CXFS Client to generate warnings in its log file. You must remove the comments in these files when you install the /etc/passwd and /etc/group files.

Checking Permissions on the Password, Group, and Fencing Files

The permissions on the fencing.conf, passwd, and group files must restrict access so that only the system administrator can modify the files. This can be done by right-clicking on the file names in Windows Explorer and selecting the following:

- Windows NT:

  Properties
  > Security
  > Permissions

- Windows 2000:

  Properties
  > Security

Verify that the permissions are Read for Everyone and Full Control for Administrators.

Caution: Failure to set permissions on the passwd and group files would allow users to change their UID/GUI at will and even gain superuser access to the files on the CXFS filesystem. Failure to set permissions on the fencing.conf̧ file would allow users to misconfigure and even disable I/O fencing, which could result in data corruption on the CXFS filesystems.

Creating a New Hardware Profile

It is strongly recommended that you create a new hardware profile and that you disable the CXFS software in the current hardware profile, in order to have a backup profile available. If the CXFS software causes the host to crash on startup, you can easily switch back to the original hardware profile and successfully return to the configuration before the CXFS software was installed.

To create a new hardware profile, right-click the **My Computer** icon and select the following:

- Windows NT:

  Properties
  > Hardware Profiles
  > Copy

- Windows 2000:

  Properties
  > Hardware
  > Hardware Profiles
  > Copy

This action copies the current hardware profile, most likely called the **Original Configuration (Windows NT)** or **Profile 1 (Windows 2000)**. You should call this new profile **CXFS Configuration** to distinguish it from other profiles. You can make the **CXFS Configuration** the default profile chosen on startup by selecting the up arrow button and moving the **CXFS Configuration** profile to the top of the list.

To remove the CXFS driver from the current hardware profile, which should be the original profile, select the following:

- Windows NT:

  Start
  > Settings
  > Control Panel
  > Devices
  > CXFS
  > HW Profile
  > Disable

- Windows 2000:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Computer Management
  > System Tools
  > Device Manager

  To show non-plug-and-play devices, select the following:

  View
  > Show hidden devices

To show the CXFS driver, select the following:

Non-Plug and Play Devices
> CXFS
> Properties
> Device Usage
> Do not use this device in the current hardware profile

You should also disable the CXFS Client service for the current profile by selecting the following:

- Windows NT:

  Start
  > Settings
  > Control Panel
  > Services
  > CXFS Client
  > HW Profile
  > Disable

- Windows 2000:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Services
  > CXFS Client
  > Properties
  > Log On
  > Disable

When the Windows host boots, you may choose **CXFS Configuration** to automatically start CXFS or choose the previous profile (most likely **Original Configuration**) to start without CXFS.

Manual CXFS Startup/Shutdown for Windows

The CXFS processes are automatically started when a Windows node is restarted. This behavior may be altered by changing the configuration of the CXFS filesystem driver and the CXFS Client service.

By default, the driver is configured to start manually and the Client service is configured to start automatically. Because the CXFS Client service depends on the CXFS filesystem driver, the driver will be started by the service.

It is recommended that the CXFS driver configuration remains manual.

You can change the CXFS Client service configuration to start manually, so that CXFS does not automatically start, by selecting the following:

- Windows NT:

  Start
  > Settings
  > Control Panel
  > Services

- Windows 2000:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Services

Change **CXFS Client** to manual rather than automatic. CXFS can then be started and stopped manually by the `Administrator` using the same selection sequence.

Software Maintenance for Windows

This section contains the following:

- "Modifying the CXFS for Windows Software", page 94
- "Upgrading the CXFS for Windows Software", page 96
- "Removing the CXFS for Windows Software", page 96
- "Downgrading the CXFS for Windows Software", page 97

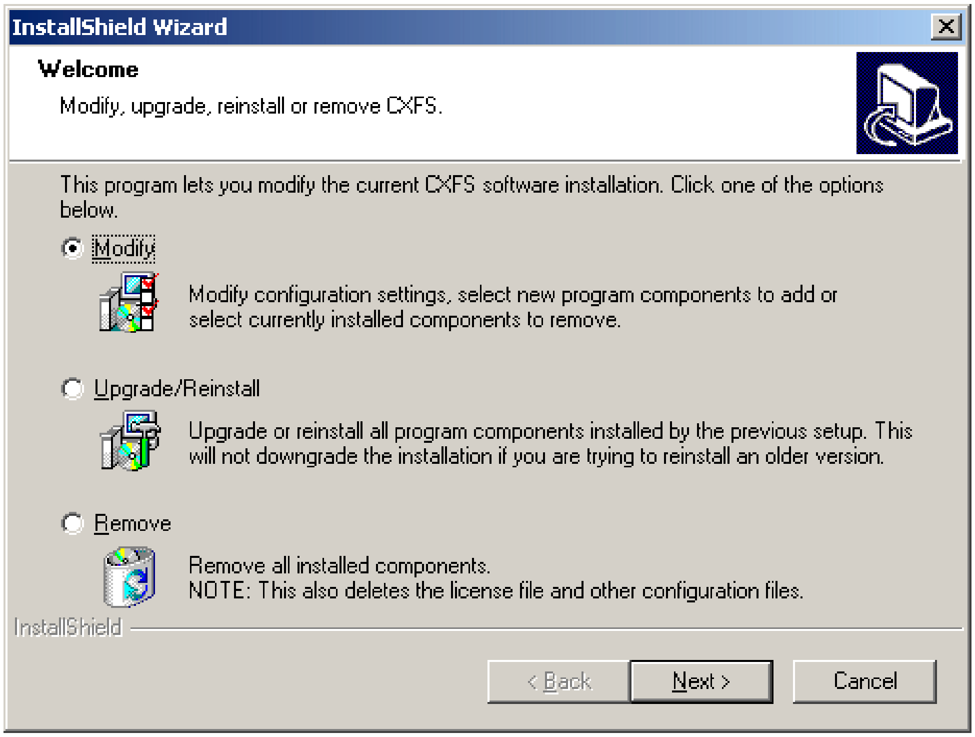
Modifying the CXFS for Windows Software

To change the location of the software and other configuration settings that were requested in "Client Software Installation Steps for Windows", page 81, perform the following steps:

1. Select the following:

   Start
   > Settings
   > Control Panel
   > Add/Remove Programs
   > CXFS
   > Add/Remove
   > Modify

   Figure 6-5, page 95, shows the screen that lets you modify the software.



InstallShield Wizard

Welcome

Modify, upgrade, reinstall or remove CXFS.

This program lets you modify the current CXFS software installation. Click one of the options below.

● Modify

Modify configuration settings, select new program components to add or select currently installed components to remove.

○ Upgrade/Reinstall

Upgrade or reinstall all program components installed by the previous setup. This will not downgrade the installation if you are trying to reinstall an older version.

○ Remove

Remove all installed components.
NOTE: This also deletes the license file and other configuration files.

Figure 6-5 Modify the CXFS for Windows

2. Make the necessary configuration changes.

   You can display the list of possible command line arguments supported by the CXFS Client service by running the service from a DOS command shell as follows:

   ```
   > C:\Winnt\system32\cxfs_client.exe -h
   ```

3. Restart the Windows node, which causes the changes to take effect.

Upgrading the CXFS for Windows Software

To upgrade the CXFS for Windows software, perform the following steps:

1. Insert the CD-ROM containing the upgraded software to run the setup program. If the setup program does not automatically start, run `winnt/setup.exe` from the CD-ROM.
2. Select **Upgrade/Reinstall** and follow the prompts as discussed in "Client Software Installation Steps for Windows", page 81.

   Figure 6-5, page 95, shows the screen that lets you modify the software.
3. Restart the Windows node. The upgraded software will not activate until the Windows node is restarted.

Removing the CXFS for Windows Software

To remove the CXFS for Windows software, first ensure that no applications on this host are accessing files on a CXFS filesystem. Then, select the following sequence to remove all installed files and registry entries:

Start
    > Settings
        > Control Panel
            > Add/Remove Programs
                > CXFS
                    > Add/Remove
                        > Remove

Figure 6-5, page 95, shows the screen that lets you remove the software.

Note: The `passwd`, `group`, and `fencing.conf` files will be removed.

You should then restart the Windows node. This will cause the changes to take effect.

Downgrading the CXFS for Windows Software

To downgrade the CXFS software, follow the instructions to remove the software in "Removing the CXFS for Windows Software", page 96, and then install the older version of the software as directed in "Client Software Installation Steps for Windows", page 81.

Note: The removal process may remove the configuration and license files. You should back up these files before removing the CXFS software so that you can easily restore them after installing the downgrade.

Chapter 7

Cluster Configuration

This chapter provides an overview of the procedures to add the client-only nodes to an established IRIX cluster. It assumes that you already have an IRIX cluster installed and running with mounted filesystems. These procedures will be performed by you or by SGI service personnel.

All CXFS administrative tasks other than restarting the Windows node must be performed on an IRIX node, using either the CXFS GUI (invoked by the `cxtask` command) or the `cmgr`(1M) command. The GUI also provides a guided configuration for defining a cluster.

This section discusses the following tasks in cluster configuration:

- "Defining the Client-Only Nodes"
- "Adding the Client-Only Nodes to the Cluster", page 101
- "Defining the Switch for I/O Fencing", page 101
- "Starting CXFS Services on the Client-Only Nodes", page 103
- "Mounting Filesystems on the Client-Only Nodes", page 104
- "Restarting the Windows Node", page 105
- "Verifying the Cluster", page 105
- "Forced Unmount of CXFS Filesystems", page 106

For detailed configuration instructions, see the *CXFS Version 2 Software Installation and Administration Guide*.

Defining the Client-Only Nodes

To add a client-only node to a CXFS cluster, you must define it as a node in the pool. You can do this on an IRIX node using the CXFS GUI or `cmgr`(1M) command.

Do the following to determine the value for the hostname field in the GUI:

- Solaris: use the value displayed by `/etc/nodename`, which must match the node's primary hostname in the `/etc/inet/hosts` (or `/etc/hosts`) file; that is, the first field after the node's IP address in /etc/inet/hosts (or /etc/hosts). This field can be either the hostname or the fully qualified domain name.

- Windows NT:

  Start
  > Settings
  > Control Panel
  > Network
  > Identification
  > Computer Name

- Windows 2000:

  Start
  > Settings
  > Network and Dial-up Connections
  > Advanced
  > Network Identification

When you specify that a node is running an operating system other than IRIX, the node will automatically be given a weight of 0, and you cannot change it. (These nodes cannot be potential metadata servers, and always have a weight of 0 so that they are not counted when calculating the CXFS membership quorum.) For client-only nodes, you must specify a unique node ID.

For example, the following shows the entries used to define a Solaris node named solaris1 using the cmgr command in prompting mode:

```
# /usr/cluster/bin/cmgr -p
Welcome to SGI Cluster Manager Command-Line Interface

cmgr> define node solaris1
Enter commands, you may enter "done" or "cancel" at any time to exit

Hostname[optional] ?
Is this a FailSafe node <true|false> ? false
Is this a CXFS node <true|false> ? true
Operating System <IRIX|Solaris|Windows> ? solaris
Node ID ? 7
Do you wish to define failure hierarchy[y/n]:y
Hierarchy option 0 <System|Fence|Shutdown>[optional] ? fence
Hierarchy option 1 <System|Fence|Shutdown>[optional] ? shutdown
Hierarchy option 2 <System|Fence|Shutdown>[optional] ?
Number of Network Interfaces ? (1)
NIC 1 - IP Address ? 163.154.18.172
NIC 1 - Heartbeat HB (use network for heartbeats) <true|false> ? true
NIC 1 - (use network for control messages)
NIC 1 - (use network for control messages) true|false> ? true
NIC 1 - Priority <1,2,...> ? 1
```

For details about these commands, see the “Define a Node” sections of the GUI or cmgr reference chapters in the *CXFS Version 2 Software Installation and Administration Guide*.

Adding the Client-Only Nodes to the Cluster

After you define all of the client-only nodes, you must add them to the cluster using either the CXFS GUI or the cmgr(1M) command on an IRIX node.

For example, if you have already defined a cluster named cxfscluster and want to add the Solaris nodes solaris1 and solaris2, you could use the following cmgr command:

```
cmgr> modify cluster cxfscluster

cxfscluster ? add node solaris1
cxfscluster ? add node solaris2
cxfscluster ? done
```

For details, see the “Modify a Cluster” sections of the GUI or cmgr reference chapters in the *CXFS Version 2 Software Installation and Administration Guide*.

Depending upon your filesystem configuration, you may also need to add the node to the list of clients that have access to the volume. See "Mounting Filesystems on the Client-Only Nodes", page 104.

Defining the Switch for I/O Fencing

You are required to use I/O fencing on client-only nodes in order to protect data integrity. I/O fencing requires a Brocade Fibre Channel switch. To define the switch for the cluster database, use either the CXFS GUI or the cmgr(1M) command on an IRIX node.

For example:

```
cmgr> define switch ptg-brocade username admin password password
```

Note: The masking feature applies only to IRIX nodes. (IRIX nodes automatically discover the available HBAs, and the masking feature is used to restrict the HBAs that are eligible for fencing.) For Solaris nodes and Windows nodes, the fencing file determines the ports that may be fenced. For more information, see "Postinstallation Steps for Solaris: Creating the I/O Fencing File", page 49, and "Creating the Windows I/O Fencing File", page 87.

After you have defined the switch, you must ensure that all of the Brocade ports that are connected to the cluster nodes are enabled. To determine port status, enter the following on a CXFS administration node:

```
# hafence -v
```

If there are disabled ports that are connected to cluster nodes, you must enable them. Log into the switch as user `admin` and use the following command:

```
# portEnable portnumber
```

You must then update the switch port information using the GUI or `cmgr`(1M).

For example, suppose that you have a cluster with port 0 connected to the node `blue`, port 1 connected to the node `green`, and port 5 connected to the node `yellow`, all of which are defined in cluster `colors`. The following output shows that the status of port 0 and port 1 is `disabled` and that the host is `UNKNOWN` (as opposed to port 5, which has a status of `enabled` and a host of `yellow`). Ports 2, 3, 4, 6, and 7 are not connected to nodes in the cluster and therefore their status does not matter.

```
# hafence -v
  Switch[0] "ptg-brocade" has 8 ports
    Port 0 type=FABRIC status=disabled hba=0000000000000000 on host UNKNOWN
    Port 1 type=FABRIC status=disabled hba=0000000000000000 on host UNKNOWN
    Port 2 type=FABRIC status=enabled  hba=210000e08b05fecf on host UNKNOWN
    Port 3 type=FABRIC status=enabled  hba=210000e08b01fec5 on host UNKNOWN
    Port 4 type=FABRIC status=enabled  hba=210000e08b01fec3 on host UNKNOWN
    Port 5 type=FABRIC status=enabled  hba=210000e08b019ef0 on host yellow
    Port 6 type=FABRIC status=enabled  hba=210000e08b0113ce on host UNKNOWN
    Port 7 type=FABRIC status=enabled  hba=210000e08b027795 on host UNKNOWN
```

In this case, you could need to enable ports 0 and 1.

Logged in to the switch:

```
# portEnable 0
# portEnable 1
```

Logged in to a CXFS administration node:

```
# hafence -v
  Switch[0] "ptg-brocade" has 8 ports
    Port 0 type=FABRIC status=disabled hba=210000e08b0103b8 on host UNKNOWN
    Port 1 type=FABRIC status=disabled hba=210000e08b0102c6 on host UNKNOWN
    Port 2 type=FABRIC status=enabled  hba=210000e08b05fecf on host UNKNOWN
    Port 3 type=FABRIC status=enabled  hba=210000e08b01fec5 on host UNKNOWN
    Port 4 type=FABRIC status=enabled  hba=210000e08b01fec3 on host UNKNOWN
    Port 5 type=FABRIC status=enabled  hba=210000e08b019ef0 on host yellow
    Port 6 type=FABRIC status=enabled  hba=210000e08b0113ce on host UNKNOWN
    Port 7 type=FABRIC status=enabled  hba=210000e08b027795 on host UNKNOWN

# cmgr -c admin fence update

# hafence -v
  Switch[0] "ptg-brocade" has 8 ports
    Port 0 type=FABRIC status=disabled hba=210000e08b0103b8 on host blue
    Port 1 type=FABRIC status=disabled hba=210000e08b0102c6 on host green
    Port 2 type=FABRIC status=enabled  hba=210000e08b05fecf on host UNKNOWN
    Port 3 type=FABRIC status=enabled  hba=210000e08b01fec5 on host UNKNOWN
    Port 4 type=FABRIC status=enabled  hba=210000e08b01fec3 on host UNKNOWN
    Port 5 type=FABRIC status=enabled  hba=210000e08b019ef0 on host yellow
    Port 6 type=FABRIC status=enabled  hba=210000e08b0113ce on host UNKNOWN
    Port 7 type=FABRIC status=enabled  hba=210000e08b027795 on host UNKNOWN
```

For details, see the “Define a Switch” and “Update Switch Port Information” sections of the GUI or `cmgr` reference chapters in the *CXFS Version 2 Software Installation and Administration Guide*.

Starting CXFS Services on the Client-Only Nodes

After adding the client-only nodes to the cluster, you must start CXFS services on them. You can do this using either the CXFS GUI or the `cmgr`(1M) command on an IRIX node.

For example:

```
cmgr> start cx_services on node solaris1 for cluster cxfscluster
cmgr> start cx_services on node solaris2 for cluster cxfscluster
```

For details, see the “Start CXFS Services” sections of the GUI or `cmgr` reference chapters in the *CXFS Version 2 Software Installation and Administration Guide*.

Mounting Filesystems on the Client-Only Nodes

If you have specified that the filesystems are to be automatically mounted on any newly added nodes, then you do not need to specifically mount the filesystems on the new client-only nodes that you added to the cluster.

Otherwise, you can mount the filesystems on the new client-only nodes by unmounting the currently active filesystems, enabling the mount on the required nodes, and then performing the actual mount. You can do this using the GUI or the `cmgr`(1M) command on an IRIX node.

For example, to mount the `fs1` filesystem on all nodes in the cluster except `solaris2`, you could use the following commands:

```
cmgr> admin cxfs_unmount cxfs_filesystem fs1 in cluster cxfscluster
cmgr> modify cxfs_filesystem fs1 in cluster cxfscluster

cxfs_filesystem fs1 ? set dflt_local_status to enabled
cxfs_filesystem fs1 ? add disabled_node solaris2
cxfs_filesystem fs1 ? done
```

Note: SGI recommends that you enable the *forced unmount* feature for CXFS filesystems, which is turned off by default; see "Recommendations", page 8, and "Forced Unmount of CXFS Filesystems", page 106.

For details, see the “Define a Filesystem” and “Mount a Filesystem” sections of the GUI or the`cmgr` reference chapters in the *CXFS Version 2 Software Installation and Administration Guide*.

Restarting the Windows Node

After completing the steps in "Postinstallation Steps for Windows", page 85, and this chapter, you should restart the Windows node. This will automatically start the driver and the Client service.

When you log into the node after restarting it, Windows Explorer will list the CXFS drive letter, which will contain the CXFS filesystems configured for this node.

Verifying the Cluster

To verify that the client-only nodes have been properly added to the cluster and that filesystems have been mounted, use the view area of the CXFS GUI, the `clconf_info` command, and and the `cluster_status` command on an IRIX node.

For example:

```
irix# /var/cluster/cmgr-scripts/cluster_status

+ Cluster=cxfscluster  FailSafe=Not Configured CXFS=ACTIVE              15:15:33
   Nodes =   cxfs6    cxfs7    cxfs8    solaris1    solaris2
FailSafe =
    CXFS =      UP       UP       UP          UP          UP

CXFS              DevName           MountPoint          MetaServer     Status
fs1        /dev/cxvm/fs1                 /fs1                cxfs7         UP
fs2        /dev/cxvm/fs2                 /fs2                cxfs6         UP
```

On client-only nodes, the `cxfs_info` command serves a similar purpose. The command path is as follows:

- Solaris: `/usr/cxfs_cluster/bin/cxfs_info`
- Windows: `\program files\CXFS\cxfs_info.exe`

On Solaris nodes, you can use the `-e` option to wait for events, which keeps the command running until you kill the process and the `-c` option to clear the screen between updates.

On Windows nodes, these options are enabled by default, and the window will stay up until you close it. To disable these options on Windows and get the standard UNIX behavior, use the `-D` option.

For example, on a Solaris node:

```
solaris# /usr/cxfs_cluster/bin/cxfs_info

cxfs_client status [timestamp Jul 19 13:30:22 / generation 21604]

Cluster:
    zagato (1) - enabled
Local:
    thump (2) - enabled, state: stable, cms: up, xvm: up, fs: up
Nodes:
    leesa      enabled  up    0
    thump      enabled  up    2
    thunderbox enabled  up    1
Filesystems:
    bigstripe0 enabled  mounted          bigstripe0           /mnt/bigstripe0
    concat0    enabled  mounted          concat0              /mnt/concat0
    mirror0    enabled  mounted          mirror0              /mnt/mirror0
    r0lun0s0   enabled  mounted          r0lun0s0             /mnt/cxfs0
    r0lun0s1   enabled  mounted          r0lun0s1             /mnt/cxfs1
    r0lun0s2   enabled  mounted          r0lun0s2             /mnt/cxfs2
    stripe0    enabled  mounted          stripe0              /mnt/stripe0
```

Forced Unmount of CXFS Filesystems

Normally, an unmount operation will fail if any process has an open file on the filesystem. However, a *forced unmount* allows the unmount to proceed regardless of whether the filesystem is still in use. To enable forced unmount, define or modify the filesystem to unmount with force and then unmount the filesystem, using the following `cmgr`(1M) commands:

```
define cxfs_filesystem logical_filesystem_name [in cluster clustername]
   set force to true

modify cxfs_filesystem logical_filesystem_name [in cluster clustername]
   set force to true

admin cxfs_unmount cxfs_filesystem filesystemname [on node nodename] [in cluster clustername]
```

For example, the following set of commands modifies the `fs1` filesystem to allow forced unmount, then unmounts the filesystem on all nodes in the `cxfscluster` cluster:

```
cmgr> modify cxfs_filesystem fs1 in cluster cxfscluster
Enter commands, when finished enter either "done" or "cancel"cmgr>

cxfs_filesystem fs1 ? set force to true
cxfs_filesystem fs1 ? done
Successfully defined cxfs_filesystem fs1

cmgr> admin cxfs_unmount cxfs_filesystem fs1 in cluster cxfscluster
```

For details, see the “CXFS Filesystems Tasks with the GUI” sections of the GUI or the `cmgr` reference chapters in the *CXFS Version 2 Software Installation and Administration Guide*.

Chapter 8

Troubleshooting

This chapter contains the following:

- "Identifying Problems on Solaris Nodes"
- "Identifying Problems on Windows Nodes", page 110
- "Common Problems and Solutions", page 115
- "Reporting Problems", page 118

Identifying Problems on Solaris Nodes

The following sections will help you identify problems with Solaris client-only nodes:

- "Is the Solaris Node in the Cluster?"
- "Are there Error Messages for the Solaris Node?"

Is the Solaris Node in the Cluster?

To determine if the node is in the cluster, use the `cluster_status` command or the CXFS GUI on an IRIX node. See "Verifying the Cluster", page 105.

Are there Error Messages for the Solaris Node?

Look at the `/var/log/cxfs_client` log to see if there are any error or warning messages. These include any messages containing the words `ERROR` or `Warning`. Specific cases in which error messages will occur include the following:

- The fencing file was not found, therefore the fencing configuration will not be updated on the server. For example:

```
cxfs_client: cis_get_hba_wwns warning: fencing configuration file "fencing.conf" not found
```

- A filesystem mount has failed and will be retried. For example:

  ```
  cxfs_client:op_failed ERROR: Mount failed for concat0
  ```

For more information about these files, see "Solaris Log Files", page 25. Also see the log files on the IRIX node; for more information, see the *CXFS Version 2 Software Installation and Administration Guide*.

Identifying Problems on Windows Nodes

The following sections will help you identify problems with Windows client-only nodes:

- "Is the CXFS Software Running Correctly on the Windows Node?"
- "Is the Windows Node in the Cluster?", page 112
- "Are There Error Messages for the Windows Node?", page 112

Is the CXFS Software Running Correctly on the Windows Node?

The methods used to verify that the CXFS software is running correctly varies by Windows platform.

Windows NT CXFS Software Verification

To verify that the CXFS software is running correctly on a Windows NT node, do the following:

- Verify that the CXFS driver has started by selecting the following:

  Start
  > Settings
  > Control Panel
  > Devices

- Verify that the CXFS Client service has started by selecting the following:

  Start
  > Settings
  > Control Panel
  > Services

Windows 2000 CXFS Software Verification

To verify that the CXFS software is running correctly on a Windows 2000 node, do the following:

- Verify that the CXFS driver has started by selecting the following:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Computer Management
  > System Tools
  > Device Manager

  To show non-plug-and-play devices, select the following:

  View
  > Show hidden devices

  To show the CXFS driver, select the following:

  Non-Plug and Play Devices
  > CXFS
  > Properties

- Verify that the CXFS Client service has started by selecting the following:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Services

Is the Windows Node in the Cluster?

To determine if the Windows node is in the cluster, use the `cluster_status` command or the CXFS GUI on an IRIX node, and the `cxfs_info` command on the Windows node. See "Verifying the Cluster", page 105.

Are There Error Messages for the Windows Node?

Look in the following file to see if there are any error or warning messages:

```
C:\Program Files\CXFS\log\cxfs_client.log
```

You can also view the **System Event** log by selecting the following:

- Windows NT:

  Start
  > Programs
  > Administrative Tools
  > Event Viewer

- Windows 2000:

  Start
  > Settings
  > Control Panel
  > Administrative Tools
  > Event Viewer

Windows Error Message Explanations

Following are typical Windows error messages and their meanings:

```
cis_get_hba_wwns warning: fencing configuration file "fencing.conf" not found
```

The fencing file fencing.conf as not found, therefore the fencing configuration will not be updated on the server.

```
op_failed ERROR: Mount failed for concat0
```

A filesystem mount has failed and will be retried.

```
cis_generate_userid_map warning: could not open passwd file
```

The passwd file could not be found.

```
cis_generate_userid_map warning: could not open group file
```

The group file could not be found.

Even with passwd and group warnings above, filesystem mounts should proceed; however, all users will be given nobody credentials and will be unable to view or modify files on the CXFS filesystems. For more information about these files, see "Solaris Log Files", page 25, and "Windows Log Files and Cluster Status", page 55. Also see the log files on the IRIX node; for more information, see the *CXFS Version 2 Software Installation and Administration Guide*.

```
could not get location of passwd/group files
could not retreving fencing configuration file name from registry
error retrieving passwd filename
error retrieving group filename
error retrieving fencing filename
```

The registry entries for the location of the passwd, group, or fencing.conf files may be missing, or the path provided on the command line to the CXFS Client service is badly formed. Reset these values by modifying the current installation as described in "Modifying the CXFS for Windows Software", page 94.

```
could not open passwd file
could not open group file
fencing configuration file not found
```

Check that the passwd, group and fencing.conf files are in the configured location and are accessible as described in "Checking Permissions on the Password, Group, and Fencing Files", page 90.

```
Unix user is something other than a user on the NT domain/workgroup
Unix group is something other than a group on the NT domain/workgroup
```

This warning indicates that a username or groupname is not a valid user or group on the Windows node, which may be confusing when examining file permissions.

```
no valid users configured in passwd file
```

No users in the `passwd` file could be matched to users on the Windows node. All users will be treated as user `nobody` for the purpose of all access control checks.

```
no valid groups configured in group file
```

No groups in the `group` file could be matched to groups on the Windows node. Attempts to display file permissions will most likely fail with the message `Unknown Group Errors`.

```
cis_driver_init() failed: could not open handle to driver
cis_driver_init() failed: could not close handle to CXFS driver
```

The CXFS driver may not have successfully started. Check the system event log for errors.

```
unable to create mount point
Configured drive letter may already be in use
```

Check that the configured drive letter is not already in use by a physical or mapped drive.

```
unable to join multicast group on interface
unable to create multicast socket
unable to allocate interface list
unable query interfaces
failed to configure any interfaces
unable to create multicast socket
unable to bind socket
```

Check the network configuration of the Windows node, ensuring that the private network is working and the Windows node can at least reach the metadata server by using the `ping` command from a command shell.

Common Problems and Solutions

This section contains the following common problems and their solutions:

- "Incorrect Configuration"
- "Determining If A Client-Only Node Is Fenced", page 115
- "Common Solaris Problems: JNI Problems", page 115
- "Common Windows Problems", page 116

Incorrect Configuration

To avoid having trouble with the CXFS client-only node, ensure you have the correct configuration. See "Requirements", page 6.

Determining If A Client-Only Node Is Fenced

To determine if a client-only node is fenced, log in to IRIX and use the `hafence`(1M) command. For more details, see the *CXFS Version 2 Software Installation and Administration Guide*.

Common Solaris Problems: JNI Problems

If you have difficulty with the JNI verification steps, consult the following checklist to help you identify the problem:

- Is the HBA firmly seated in its PCI slot?
- Are all cables undamaged and connected?
- Is power applied to all devices?
- Do the link lights illuminate on all units?
- Is the problem confined to just one TP9400 unit? If so, check the cabling between the switch and the unit; if no units are being shown, suspect cabling from the HBA.
- Is the Brocade switch properly licensed?
- Did you enable fabric mode? See step 4, in "Installing the JNI HBA", page 29.

For more information, see the *Installation Guide, FCE-6460 and FCE2-6460 PCI-to-Fibre Channel Host Bus Adapters (Solaris, Windows NT/2000, Novell, AIX, HP-UX, Mac OS, Linux) JNI FibreStar* or contact your SGI service representative.

Common Windows Problems

This section contains the following common Windows problems:

- "Windows QLogic Problems"
- "Filesystems Are Not Displayed on a Windows Node", page 117
- "Large Log Files on Windows", page 117
- "Windows Failure on Restart", page 118
- "Memory Configuration of the Windows Node", page 118

Windows QLogic Problems

If you have difficulty with the QLogic verification steps, consult the following checklist to help you identify the problem:

- Is the HBA firmly seated in its PCI slot?
- Are all cables undamaged and connected?
- Is power applied to all devices?
- Do the link lights illuminate on all units?
- Is the problem confined to just one TP9400 unit? If so, check the cabling between the switch and the unit; if no units are being shown, suspect cabling from the HBA.
- Is the Brocade switch properly licensed?
- Check the QLogic management tool event and alarm logs. Select the following:

  Start
  > Programs
  > QLogic Management Suite
  > SANsurfer

For more information, see the following QLogic documentation or contact your SGI service representative:

- *Hardware Installation Guide for the QLA2xxx Board Family*
- *Software Installation Guide for the QLA2xxx Board Family*

Also see the QLogic website at:

```
http://www.qlogic.com
```

Filesystems Are Not Displayed on a Windows Node

If the CXFS drive letter is visible in Windows Explorer but no filesystems are mounted, do the following:

- Run `C:\Program Files\CXFS\cxfs_info` to ensure that the filesystems have been configured for this node.
- Verify the filesystems that should be mounted by using the `cmgr`(1M) command on an IRIX node. For more information, see "Mounting Filesystems on the Client-Only Nodes", page 104.
- Ensure that the CXFS metadata server is up and that the Windows node is in the cluster membership; see "Verifying the Cluster", page 105.
- Check that the CXFS Client service has started. See "Is the CXFS Software Running Correctly on the Windows Node?", page 110, and "Manual CXFS Startup/Shutdown for Windows", page 93.
- Check the following file for warnings and errors regarding licenses or mounting filesystems:

  ```
  C:\Program Files\CXFS\log\cxfs_client.log
  ```

- Check the cluster configuration to ensure that this node is configured to mount one or more filesystems.

Large Log Files on Windows

The CXFS Client service creates the following log file:

```
C:\Program Files\CXFS\log\cxfs_client.log
```

This log file may become quite large over a period of time if the verbosity level is increased. The service does not perform any automatic log rotation, so the service must be stopped in order to move or truncate this file, then restarted. See "Manual CXFS Startup/Shutdown for Windows", page 93, on how to stop and start the CXFS Client Service.

Windows Failure on Restart

If the CXFS Windows node fails to start and terminates in a blue screen, restart your computer, and select the backup hardware profile (with CXFS disabled). Alternatively, pressing `L` at the **Hardware Profile** menu will select the last configuration that was successfully started and shut down. If the node has only one hardware profile, press the spacebar after selecting the boot partition to get to the **Hardware Profile** menu.

Memory Configuration of the Windows Node

A Windows problem may affect Windows CXFS nodes performing large asynchronous I/O operations. If the Windows node crashes with a `NO_MORE_SYSTEM_PTES` message, the work-around described in the following link should be considered (line break added here for readability):

```
http://www.microsoft.com/technet/treeview/default.asp?url=/technet/prodtechnol/
winxppro/reskit/prmd_stp_fztl.asp
```

Reporting Problems

When reporting a problem with a client-only node, it is important to retain the appropriate information; having access to this information will greatly assist SGI in the process of diagnosing and fixing problems. The methods used to collect required information for problem reports are platform-specific.

Reporting Solaris Problems

When reporting a problem about a CXFS Solaris node to SGI, you should retain the following information:

- System core files in `/var/crash/`*hostname* on Solaris nodes.
- Output from the `crash`(1M) utility.

- `mdb`(1M) modular debugger output:
  - For panics or generated dumps, use the following commands and save the output:
    - `$c` (or `$C`)
    - `$r`
    - `$<msgbuf`
  - For dumps from hangs:
    - `$<threadlist`
    - `$c` (or `$C`)
    - `$r`
    - `$<msgbuf`
- A list of the installed CXFS packages. Use the `pkginfo`(1) command as follows:

  ```
  # pkginfo -l SGIcxfs
  ```

- A list of the Solaris patches that have been installed. Use the `showrev`(1M) command. The `showrev` command without options prints a summary and the `-p` option lists the revision information about patches.
- A list of the loaded Solaris kernel modules and versions. Use the `modinfo`(1M) command.
- Output about the cluster obtained from the IRIX `cdump` utility run on an IRIX node. This utility is available from your SGI service representative.

If any of the above Solaris tools are not currently installed on your Solaris system, you should install them.

Reporting Windows Problems

To report problems about a Windows node, you should retain platform-specific information and save crash dumps.

Retain Windows NT Information

When reporting a problem about a CXFS Windows NT node to SGI, you should retain the following information:

- The version and memory of the current diagnostics. Select the following:

  Start
  > Programs
  > Administrative Tools
  > Windows NT Diagnostics

  Record the version displayed in the **Version** tab and the total physical memory in the **Memory** tab.

- The build date and firmware versions. Using Windows Explorer, open the following directory:

  ```
  C:\Winnt\system32\drivers
  ```

  Then do the following:

  – Right click on `cxfs.sys` and select the following:

    Properties
    > Version

    Record the values of **BuildDate** and **Product Version**.

  – Right click on `ql2200.sys` and select the following:

    Properties
    > Version

    Record the values of **Firmware** and **Product Version**.

- The contents of the following file:

  ```
  C:\Program Files\CXFS\log\cxfs_client.log
  ```

  Compress this file with `winzip` if it is large.

- The contents of the crash dump if one was generated, Compress this file with `winzip`. For more information, see "Save Crash Dumps for Windows", page 122.

Retain Windows 2000 Information

When reporting a problem about a CXFS Windows 2000 node to SGI, you should retain the following information:

- The configuration of the machine. Select the following:

  Start
  > Programs
  > Accessories
  > System Tools
  > System Information
  > Action
  > Save As System Information File

  This will create a file that describes all of the installed hardware and configured drivers on the machine.

  Alternatively, you could dump information about each item in the hardware tree to a text file by using the following selection:

  Action
  > Save As Text File

  However, you must repeat this action for each item.

- The build date and firmware versions. Using Windows Explorer, open the following directory:

  ```
  C:\Winnt\system32\drivers
  ```

  Then do the following:

  – Right click on `cxfs.sys` and select the following:

    Properties
    > Version

    Record the values of **BuildDate** and **Product Version**.

- Right click on `ql2200.sys` and select the following:

  Properties
  > Version

  Record the values of **Firmware** and **Product Version**.

- The contents of the following file:

  ```
  C:\Program Files\CXFS\log\cxfs_client.log
  ```

  Compress this file with `winzip` if it is large.

- The contents of the crash dump if one was generated, Compress this file with `winzip`. For more information, see "Save Crash Dumps for Windows", page 122.

Save Crash Dumps for Windows

If you are having problems, you should configure the Windows node to save crash dumps to a filesystem that is not a CXFS filesystem. To do this, click the right mouse button on the **My Computer** icon and select the following:

- Windows NT:

  Properties
  > Startup
  > Shutdown
  > Write debugging information to

- Windows 2000:

  Properties
  > Advanced
  > Startup and Recovery
  > Write debugging information to

Enter a path on a filesystem other than a CXFS filesystem. This change will take affect only after the node is restarted.

Glossary

active metadata server

A weighted IRIX node chosen from the list of potential metadata servers. There can be multiple active metadata servers, one for each file system.

client

See *CXFS client*.

cluster

A cluster is the set of systems (nodes) configured to work together as a single computing resource. A cluster is identified by a simple name and a cluster ID. A cluster running multiple operating systems is known as a multiOS cluster.

There is only one cluster that may be formed from a given pool of nodes.

Disks or logical units (LUNs) are assigned to clusters by recording the name of the cluster on the disk (or LUN). Thus, if any disk is accessible (via a Fibre Channel connection) from machines in multiple clusters, then those clusters must have unique names. When members of a cluster send messages to each other, they identify their cluster via the cluster ID. Thus, if two clusters will be sharing the same network for communications, then they must have unique cluster IDs. In the case of multiOS clusters, both the names and IDs must be unique if the clusters share a network.

Because of the above restrictions on cluster names and cluster IDs, and because cluster names and cluster IDs cannot be changed once the cluster is created (without deleting the cluster and recreating it), SGI advises that you choose unique names and cluster IDs for each of the clusters within your organization.

cluster database

Contains configuration information about nodes, filesystems, and the cluster. The database is managed by the `fs2d` daemon and is stored on IRIX nodes in the pool of nodes that are running the `fs2d` daemon.

cluster ID

A unique number within your network in the range 1 through 128. The cluster ID is used by the IRIX kernel to make sure that it does not accept cluster information from any other cluster that may be on the network. The kernel does not use the database for communication, so it requires the cluster ID in order to verify cluster communications. This information in the kernel cannot be changed after it has been initialized; therefore, you must not change a cluster ID after the cluster has been defined. Clusters that share a network must have unique names and IDs.

cluster node

A node that is defined as part of the cluster.

control messages

Messages that cluster software sends between the cluster nodes to request operations on or distribute information about cluster nodes. Control messages and heartbeat messages are sent through a node's network interfaces that have been attached to a control network.

A node's control networks should not be set to accept control messages if the node is not a dedicated CXFS node. Otherwise, end users who run other jobs on the machine can have their jobs killed unexpectedly when CXFS resets the node.

control network

The network that connects nodes through their network interfaces (typically Ethernet) such that CXFS can send heartbeat messages and control messages through the network to the attached nodes. CXFS uses the highest priority network interface on the control network; it uses a network interface with lower priority when all higher-priority network interfaces on the control network fail.

CXFS client

A node that is part of the cluster and is a potential metadata server, but is currently not acting as the active metadata server. See also *IRIX node* and *CXFS client-only node*.

CXFS client-only node

A node that is part of the cluster but is not a potential metadata server. Solaris nodes and Windows nodes are client-only nodes. See also *IRIX node*.

CXFS database

See *cluster database*.

CXFS membership

The group of CXFS nodes that can share filesystems in the cluster, which may be a subset of the nodes defined in a cluster. During the boot process, a node applies for CXFS membership. Once accepted, the node can share the filesystems of the cluster. (Also known as *kernel-space membership*.)

database

See *cluster database*.

GUI

Graphical user interface.

heartbeat messages

Messages that cluster software sends between the nodes that indicate a node is operational. Heartbeat messages and control messages are sent through the node's network interfaces that have been attached to a control network.

I/O fencing

The failure action that isolates a problem node so that it cannot access I/O devices, and therefore cannot corrupt data in the shared CXFS filesystem. I/O fencing can be applied to any node in the cluster (CXFS clients and metadata servers). The rest of the cluster can begin immediate recovery.

IRIX node

A CXFS node that is running the IRIX operating system. An IRIX node can be a potential metadata server if it is configured as such and has weight, or it can be a CXFS client.

membership

See *CXFS membership*.

membership weight

A number (usually 0 or 1) that is assigned to a node for purposes of calculating the CXFS membership quorum. 1 indicates that the node is eligible to be a potential metadata server. IRIX nodes may have a weight of 0 or 1. CXFS client-only nodes always have a weight of 0.

metadata

Information that describes a file, such as the file's name, size, location, and permissions.

metadata server

The IRIX node that coordinates updating of metadata on behalf of all nodes in a cluster. There can be multiple potential metadata servers, but only one is chosen to be the active metadata server for any one filesystem. See also *active metadata server* and *potential metadata server*.

multiOS cluster

A cluster that is running multiple operating systems, such as IRIX and Solaris.

node

A node is an operating system (OS) image, usually an individual computer. (This use of the term node does not have the same meaning as a node in an SGI Origin 3000 or SGI 2000 system.)

A given node can be a member of only one pool (and therefore) only one cluster.

A node can run the IRIX operating system or another operating system, such as Solaris, as defined in the CXFS Version 2 for CXFS Client-Only Nodes: Installation and Configuration Guide.

node membership

The list of nodes that are active (have CXFS membership) in a cluster.

pool

The pool is the set of nodes from which a particular cluster may be formed. Only one cluster may be configured from a given pool, and it need not contain all of the available nodes. (Other pools may exist, but each is disjoint from the other. They share no node or cluster definitions.)

A pool is formed when you connect to a given node and define that node in the cluster database using the CXFS GUI or cmgr(1M) command. You can then add other nodes to the pool by defining them while still connected to the first node, or to any other node that is already in the pool. (If you were to connect to another node and then define it, you would be creating a second pool).

potential metadata server

A weighted IRIX node that is listed in the metadata server list when defining a filesystem; there can be multiple potential metadata servers, but only one node in the list will be chosen as the active metadata server for one filesystem.

recovery

The process by which the metadata server moves from one node to another due to an interruption in services on the first node.

relocation

The process by which the metadata server moves from one node to another due to an administrative action; other services on the first node are not interrupted.

SAN

Storage area network, a high-speed, scalable network of servers and storage devices that provides storage resource consolidation, enhanced data access/availability, and centralized storage management.

standby node

A server-capable administration node that is configured as a potential metadata server for a given filesystem, but does not currently run any applications that will use that filesystem.

tree view

The portion of the CXFS GUI window that displays components graphically.

quorum

The number of nodes required to form a cluster.

weight

See *membership weight*.

Index