CXFS™ Version 2 for CXFS Client-Only
Nodes: Installation and Configuration
Guide

New Features in this Guide

This guide contains the following new features:

- Support for Windows NT nodes in a CXFS multiOS cluster. Platform-specific information is grouped into separate chapters:

  - Chapter 5, "Solaris Platform", page 21

  - Chapter 6, "Windows NT Platform", page 53

- Support for up to four JNI HBAs in each CXFS Solaris node. See "JNI Fibre Channel Host Bus Adapter Installation and Configuration", page 26

  Note: JNI supports a maximum of four JNI HBAs in operating environments with qualified Solaris platforms.

  Note: Implementation of the mirroring feature for Solaris and Windows NT nodes is deferred in this release

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 Solaris and Windows NT platforms.

Contents

Figures

About This Guide

This publication documents the CXFS MultiOS Clients 2.1 release for IRIX 6.5.16f. This release supports Solaris nodes and Windows NT 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 Solaris platform 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 :

  - *TPSSM7 RAID Software Concepts Guide for TP9400* (`TPL link`)

  - *SGI TP9400 RAID Owner's Guide* (`TPL link`)

  - *SGI TP9400 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`

- 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 QLA22xx Board Family*

  - – *Software Installation Guide for the QLA22xx Board Family*

  See the QLogic website at:

  `http://www.qlogic.com`

- Solaris 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`

- 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 NT software documentation:

  - *Start Here Microsoft Windows NT Workstation: Basics and Installation*

- Hardware documentation for the Intel platform

- *Flexible License Manager End User Manual* from GLOBEtrotter Software, Inc.

Obtaining Publications

To obtain SGI documentation, go to the SGI Technical Publications Library at:

`http://techpubs.sgi.com.`

Conventions

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, please tell us. 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 us 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 World Wide Web page:

  `http://techpubs.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 Pkwy., M/S 535
  Mountain View, California 94043–1351

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

We value 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

- Windows NT 4.0

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 9

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 4

- "Requirements", page 5

- "Recommendations", page 7

CXFS Processes

When CXFS is started on a client-only node, a user-space daemon 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 15.

- CXFS FLEXlm license installed on every node in the cluster; see Chapter 4, "CXFS and XVM FLEXlm License Requirements", page 19.

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 premount and postmount scripts are not supported on client-only nodes. For information about using these scripts on IRIX 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 ID should not be used by two different people anywhere 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 user identifier (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: The CXFS software will detect that these files have changed and will apply the updated contents when mapping Windows NT users to UNIX UIDs. The Administration user on the Windows NT node will be mapped to the root user on the UNIX nodes for the purpose of file access controls. Therefore, access to the Administration user account on Windows NT 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 "Additional Performance Considerations Specific to Windows NT", page 55.

Requirements

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

- 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 6.5.16f or later, installed on the IRIX nodes in a multiOS cluster. In a multiOS cluster, all IRIX nodes must be at a minimum of IRIX 6.5.16f or later to allow a CXFS Solaris or Windows NT node to be present. For more details, see the platform-specific CXFS multiOS release notes.

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

Note: Implementation of the mirroring feature for Solaris and Windows NT nodes is deferred in this release.

- A Brocade Fibre Channel 2400, 2800, or 3800 switch, which is required to protect data integrity.

  IRIX nodes use serial reset lines to protect the integrity of the data stored in the cluster. (Serial reset lines are mandatory for the IRIX nodes in a cluster with only two IRIX nodes or with only two weighted IRIX nodes. Larger clusters should have an odd number of weighted IRIX nodes.)

  Serial reset lines are not available for Solaris nodes and Windows NT nodes. Instead of serial reset lines, these 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. 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.

Note: 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 one possible implementation of fencing.

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.

However, 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 no more than 16 nodes. At least one IRIX node must have weight; other nodes can be CXFS clients. All Solaris nodes and Windows NT 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 16 nodes, as many as 8 of which can run FailSafe. However, FailSafe cannot run on Solaris nodes or Windows NT nodes.

- No nodes within the cluster running Trusted IRIX. A multiOS cluster cannot contain Trusted IRIX nodes. If you want to run CXFS and Trusted IRIX, SGI recommends that all nodes in the cluster run Trusted IRIX.

Also see "Requirements Specific to Solaris", page 21, and "Requirements Specific to Windows NT", page 53.

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 weighted IRIX nodes.

- For large clusters, SGI recommends that you define only the first three weighted IRIX 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.

- Create a new cluster using nodes that have the same IRIX release installed. When the cluster is functioning, you can later upgrade selected nodes to an adjacent release.

- 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 NT.

- 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*.

- Weight nodes appropriately when defining the node with the IRIX `cmgr`(1M) command or the IRIX CXFS GUI:

  - Client-only nodes always have a weight of 0. The CXFS GUI and `cmgr`(1M) commands automatically assign a weight of 0 to these nodes and you cannot change it.

  - IRIX nodes should have a weight of 1 (for a potential metadata server) or 0 (for a CXFS client-only node).

  - Use an odd number of weighted IRIX nodes.

  - Weight potential metadata servers only (unless weighting a stable IRIX client gives you an odd number).

- Shut down cluster services before maintenance.

- Avoid relocation and recovery in the current release.

- Use the following good practices:

  - Unmount the filesystems from the metadata server, shut down the node, and remount the filesystem when possible.

  - If relocation is necessary, limit activity on a filesystem.

  - 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. 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.

– There is no forced unmount for Windows NT nodes in this release.

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

Tasks
   > Filesystems
      > Unmount a Filesystem

You can also specify this feature using the cmgr(1M) commands to define the filesystem. 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", page 10

• "Windows NT Overview", page 11

Solaris Overview

Note: For additional details, see the CXFS multiOS release notes for Solaris. If you run into problems, see Chapter 8, "Troubleshooting", page 81.

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

1. Install the Solaris 8 operating system according to the directions in the Solaris documentation (if not already done).

2. Install and verify the SGI TP9400 RAID. See Chapter 2, "SGI TP9400 RAID Firmware Verification", page 13.

3. Install and verify the Brocade Fibre Channel switch. See Chapter 3, "Brocade Fibre Channel Switch Verification", page 15.

4. Obtain and install the CXFS and (optional) XLV licenses. See Chapter 4, "CXFS and XVM FLEXlm License Requirements", page 19.

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 26.

   Note: The EZ Fibre GUI is needed for JNI configuration, and it requires that you run the JNI v5.1 driver, but the v5.1 driver does not support CXFS. Therefore, you must install the v5.1 driver to use the GUI and then install the v5.0.1 SGI JNI driver to use CXFS.

6. Prepare the Solaris node, including adding a private network. See "Preinstallation Steps for Solaris", page 39.

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 46.

8. Create the I/O fencing file. See "Postinstallation Steps for Solaris", page 48.

9. Configure the cluster to add the new Solaris node and mount filesystems. See Chapter 7, "Cluster Configuration", page 75.

Windows NT Overview

Note: For additional details, see the CXFS multiOS release notes for Windows NT. If you run into problems, see Chapter 8, "Troubleshooting", page 81.

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

1. Install the Windows NT 4.0 operating system according to the directions in the Windows NT documentation (if not already done).

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

3. Install and verify the SGI TP9400 RAID. See Chapter 2, "SGI TP9400 RAID Firmware Verification", page 13.

4. Install and verify the Brocade Fibre Channel switch. See Chapter 3, "Brocade Fibre Channel Switch Verification", page 15.

5. Obtain and install the CXFS and (optional) XLV licenses. See Chapter 4, "CXFS and XVM FLEXlm License Requirements", page 19.

6. Install and verify the QLogic host bus adapter (HBA) and driver. See "QLogic Fibre Channel Host Bus Adapter Installation", page 57.

7. Prepare the Windows NT node, including adding a private network. See "Preinstallation Steps for Windows NT", page 58.

8. Install the CXFS software. See "Client Software Installation Steps for Windows NT", page 61.

9. Perform post-installation configuration steps:

   • "Create the Windows NT I/O Fencing File", page 67

   • "Perform User Configuration", page 69

- "Create a New Hardware Profile", page 70

10. Configure the cluster to add the new Windows NT node, restart the node, and mount filesystems. See Chapter 7, "Cluster Configuration", page 75.

11. Restart the Windows NT node to see the mounted filesystems under the configured drive letter.

SGI TP9400 RAID Firmware Verification

The SGI TP9400 RAID will be initially installed and configured by SGI personnel. You can use the information in this chapter to verify the installation.

Required SGI TP9400 RAID Firmware

The TP9400 4.0 CD contains the required controller firmware and NVSRAM files for 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.

Verify the TP9400 RAID Firmware

To verify that the TP9400 RAID is properly installed and ready for use with CXFS, you should dump the TP9400 profile and verify the controller software revisions, according to the instructions provided with FCO 1056.

For More Information

The following TP9400 documentation is used to install and verify the TP9400 RAID:

- *TPSSM7 RAID Software Concepts Guide for TP9400*

- *SGI TP9400 RAID Owner's Guide*

- *SGI TP9400 RAID Administration Guide*

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

- `tp9400_sw_concepts_guide.pdf`

- `tp9400_owners_guide.pdf`

- `tp9400_admin_guide.pdf`

Brocade Fibre Channel Switch Verification

In order to protect data integrity, Solaris nodes and Windows NT 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; 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), and 3800 (2 Gbit) switches. Please refer to the release notes for further information.

All Brocade switches contained within the SAN fabric must have the appropriate Brocade license key installed and must be upgraded to the firmware level v2.2.1c or later. (This release was tested with v2.2.1c.) If the current firmware level of the switches must be upgraded, please contact your local SGI service representative or customer support center.

For the Brocade 3800 switch, Firmware level 3.0 must be installed. The Brocade switch must be configured so that its Ethernet interface is accessible from all IRIX cluster nodes using `telnet`. Either the public or private heartbeat network can be used to do this.

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

Verify 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
```

Verify 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
```

Note: Changing the Brocade FC Cable Connections:

If it is necessary to change the Brocade switch ports that one or more nodes in the CXFS cluster are using, it is required to first cleanly shutdown CXFS services on all nodes.

Once the Fiber Channel cables have been rearranged, CXFS services can be started again. If one or more Brocade switches were added or removed and I/O fencing is enabled on the cluster, it will be necessary to reconfigure I/O fencing using the CXFS GUI. This will involve adding or removing switches from the CXFS configuration as required.

After changing the I/O fencing configureation, if any CXFS client nodes are connected to a new or different Brocade switch, then those nodes will need to have CXFS services restarted to ensure the IRIX servers can correctly identify the Brocade ports used by all clients. Consult the *CXFS Version 2 Software Installation and Administration Guide* for instructions on configuring I/O fencing.

CXFS and XVM FLEXlm License Requirements

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.

Note: Implementation of the mirroring feature for Solaris and Windows NT nodes is deferred in this release.

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 NT Host Information

For FLEXlm you will need the Ethernet (MAC) address. This can be taken from the target NT machine either by opening a command prompt window and running the command `C:\>ipconfig /all` or by viewing the network adapter properties in the Windows Control Panel.

Note: The "Physical Address" printed by this command is the NT host identifier required to generate FLEXlm license. If the machine has more than one network interface, it is recommended to use a private network interface "Physical Address."

Obtain 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 25, and "FLEXlm License Verification for Windows NT", page 57.

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.

Solaris Platform

This chapter contains the following:

- "CXFS on Solaris"
- "FLEXlm License Verification for Solaris", page 25
- "JNI Fibre Channel Host Bus Adapter Installation and Configuration", page 26
- "Preinstallation Steps for Solaris", page 39
- "Client Software Installation Steps for Solaris", page 46
- "Postinstallation Steps for Solaris", page 48
- "Manual CXFS Startup/Shutdown for Solaris", page 50
- "Software Maintenance for Solaris", page 51

CXFS on Solaris

This section contains the following:

- "Requirements Specific to Solaris"
- "CXFS Commands Installed on Solaris", page 22
- "Solaris Log Files", page 23
- "Solaris Limitations", page 23

Requirements Specific to Solaris

In addition to the items listed in"Requirements", page 5, using a Solaris node to support CXFS requires the following:

- Solaris 8 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:

    – 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 46.

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

CXFS for Solaris has the following limitations:

- 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 on to an IRIX node in the cluster and make any necessary changes. If you would like to provide a viewing command such as that done by `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.

- Under heavy activity, CXFS Solaris can require more kernel threads than the default CXFS modules have allocated. To overcome this limitation, you should increase the number of threads allocated.

  The formula for the number of threads (*nthreads*) allocated is the calculated value of the following:

  *nthreads = ncpu * 10 * taskq_thread_factor*

  Where:

  – *nthreads* is the number of threads allocated

  – *ncpu* is the number of CPUs in the Sun Microsystems node

  – *taskq_thread_factor* is 2 (default)

  The default of 2 for *taskq_thread_factor* should be adequate for most usage. However, in cases of heavy activity, you should increase the value of this parameter to 3 in the /etc/system file. To do this, 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. Add the following line to /etc/system:

     ```
     set cell:taskq_thread_factor = 3
     ```

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

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

     ```
     # echo "taskq_thread_factor/X" | adb -k
     physmem 799d
     taskq_thread_factor:
     taskq_thread_factor:              3
     ```

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

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

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

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

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

    d. 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
```

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

Note: You are required to have a second network interface that must be used for the private metadata network. If you do not already have a second interface installed, you must install a network interface card. You may wish to install that card at this time.

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

- "Install the JNI HBA"

- "Install and Run the EZ Fibre Configuration GUI", page 28

- "Verify the JNI HBA Installation", page 38

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.

Install the JNI HBA

Note: The EZ Fibre GUI requires that you run the JNI v5.1 driver. The following procedure assumes that you do not have any driver installed at this point.

However, if you do already have another driver installed, you must first remove the driver package with the Solaris `pkgrm`(1M) command and then reboot the system before using the following procedure.

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"

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 JNI HBA v5.1 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 v5.1 driver and SNIA package from the following JNI website:

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

- Under "Locate Driver by PRoduct" click on "FCE-6460"

- Under the "Solaris" section, right click "JNIC146x.pkg" and "Save Link As."

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
```

Note: The v5.1 driver does not support CXFS. If you attempt to start CXFS while the v5.1 driver is installed, the system will log messages such as the following:

```
/usr/kernel/drv/sparcv9/xvm: undefined symbol 'jnic146x_pkt_comp'
/usr/kernel/drv/sparcv9/xvm: undefined symbol 'jnic146x_tran_start'
WARNING: mod_load: cannot load module 'xvm'
```

The appropriate driver for CXFS will be installed later in this procedure.

Install and Run 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 EZ Fibre 2.2d GUI from the JNI website:

      • Go to http://www.jni.com/Drivers

      • Under "Locate Driver by Product" click on "FCE-6460"

      • Under the "Solaris" section, right click "EZF_22d.tar" and "Save Link As" /var/tmp/EZF_22d.tar

b. Un-tar the saved file using the command: # tar xvf
    /var/tmp/EZF_22d.tar

c. Run the volcheck(1M) command:

    # **volcheck**

d. Change to the /cdrom directory:

    # **cd /cdrom/ez_fibre/Solaris**

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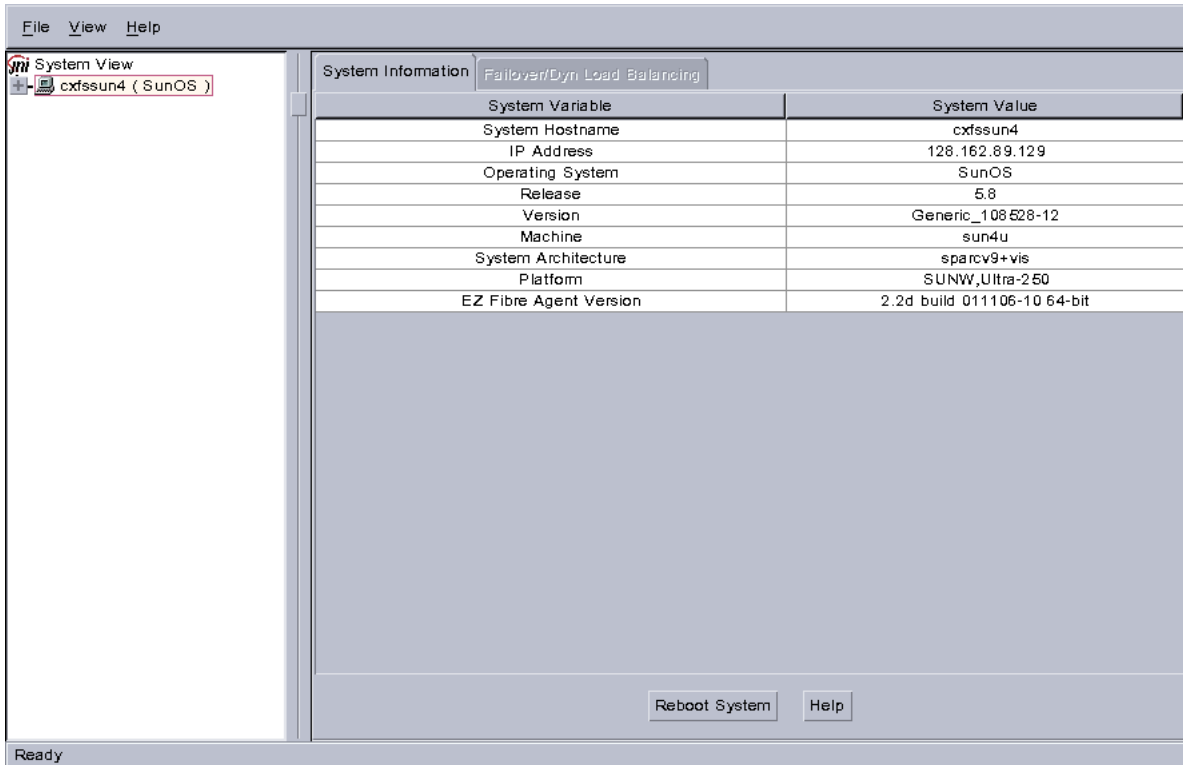
**Figure 5-1** Example: Second Window: EZ Fibre Configuration Utility - Standalone

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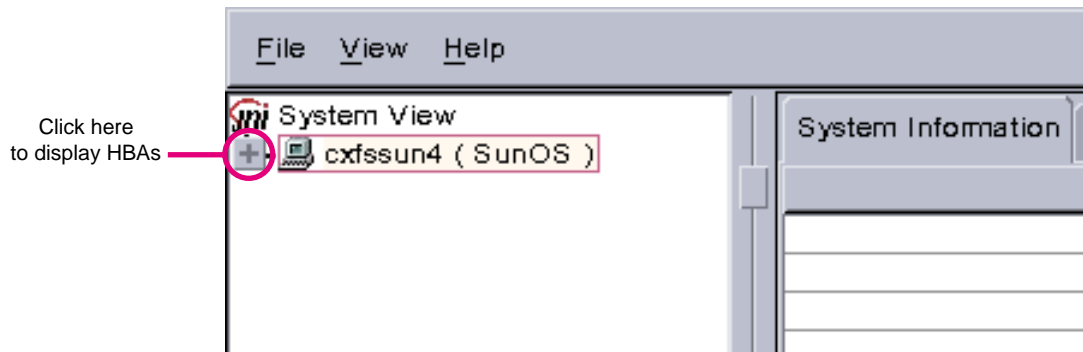
Figure 5-2 Location of icon (+) to Display the HBA

Figure 5-3 shows an example of the display after clicking on the + sign for
cxfssun4, which shows the JNI HBA.
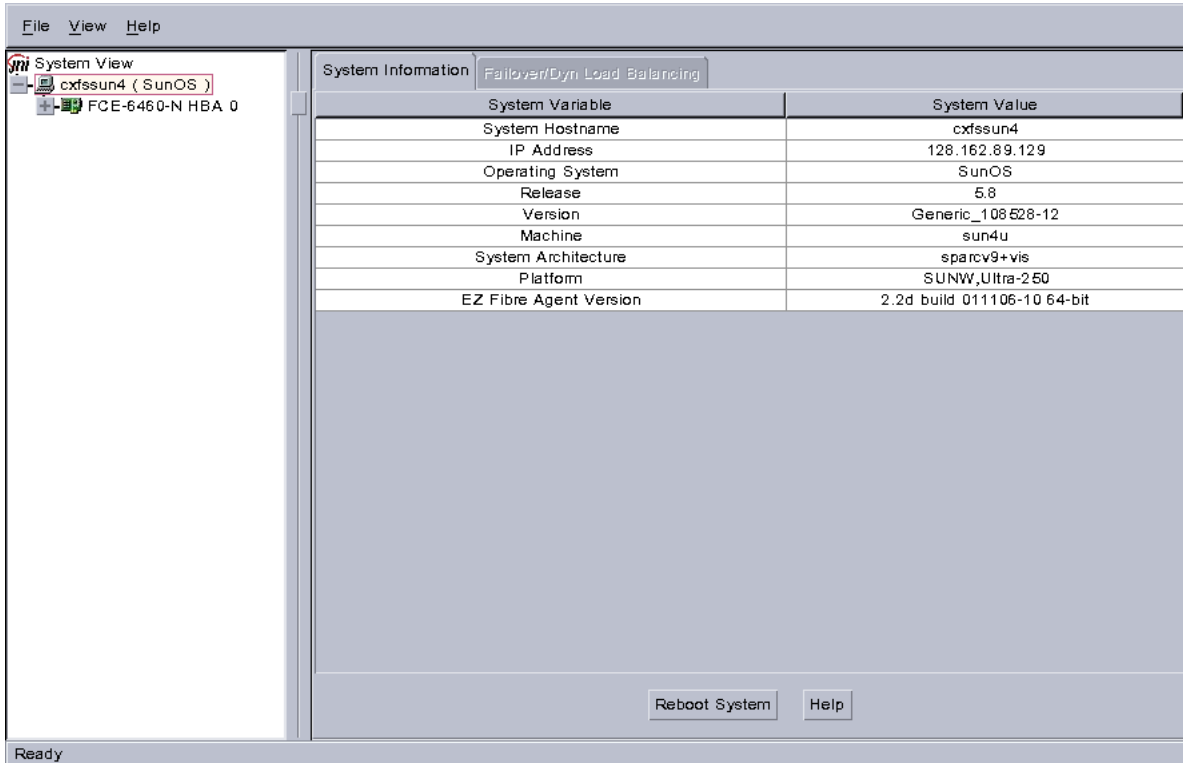
Figure 5-3 Example: After Clicking + to Display the HBA

4. Click on the icon to the right (not the + sign to the left). Figure 5-4 shows the icon.
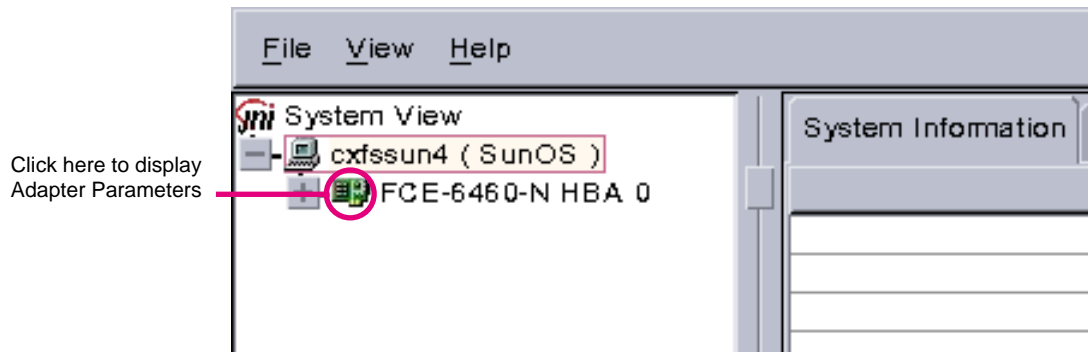
Figure 5-4 Location of the Icon to Display the Adapter Parameters

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

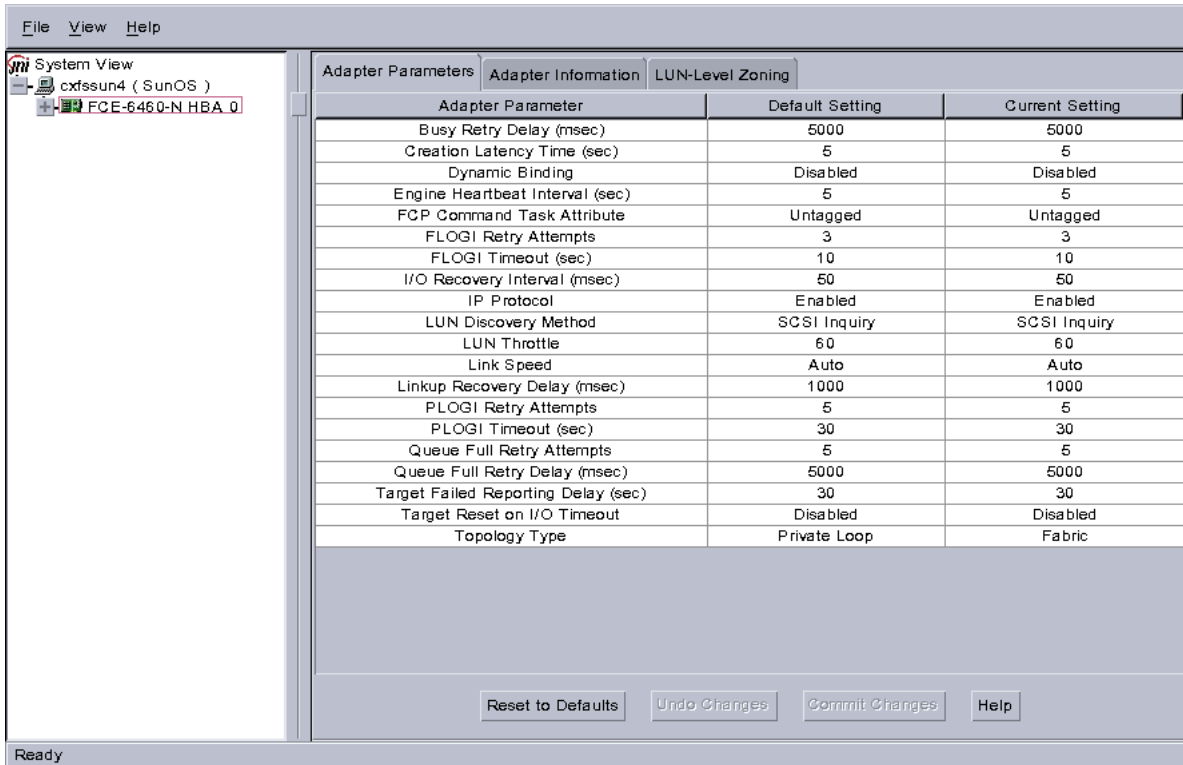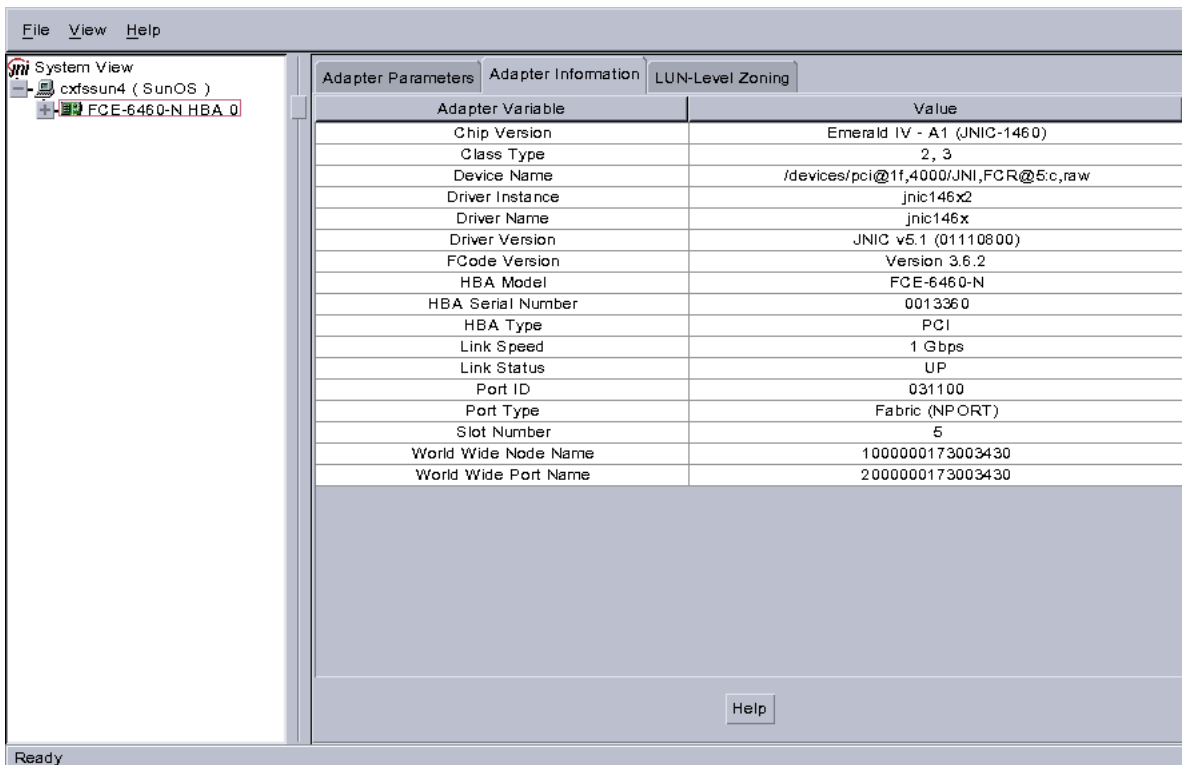| Adapter Parameter	Default Setting	Current Setting
Busy Retry Delay (msec)	5000	5000
Creation Latency Time (sec)	5	5
Dynamic Binding	Disabled	Disabled
Engine Heartbeat Interval (sec)	5	5
FCP Command Task Attribute	Untagged	Untagged
FLOGI Retry Attempts	3	3
FLOGI Timeout (sec)	10	10
I/O Recovery Interval (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 "Create the Solaris I/O Fencing File", page 48.

File View Help	

Adapter Parameters	Adapter Information	LUN-Level Zoning

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

System View
cxfssun4 (SunOS)
FCE-6460-N HBA 0

Help

Ready

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.
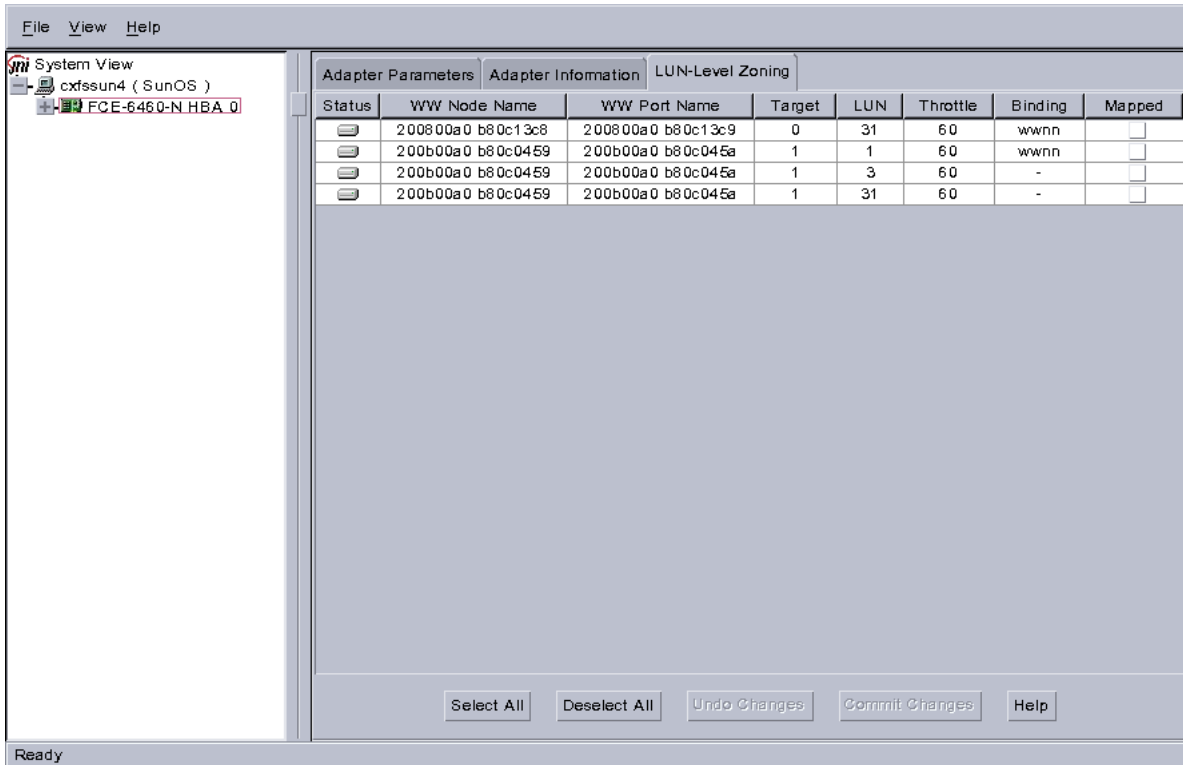
Figure 5-7 After Clicking on LUN-Level Zoning

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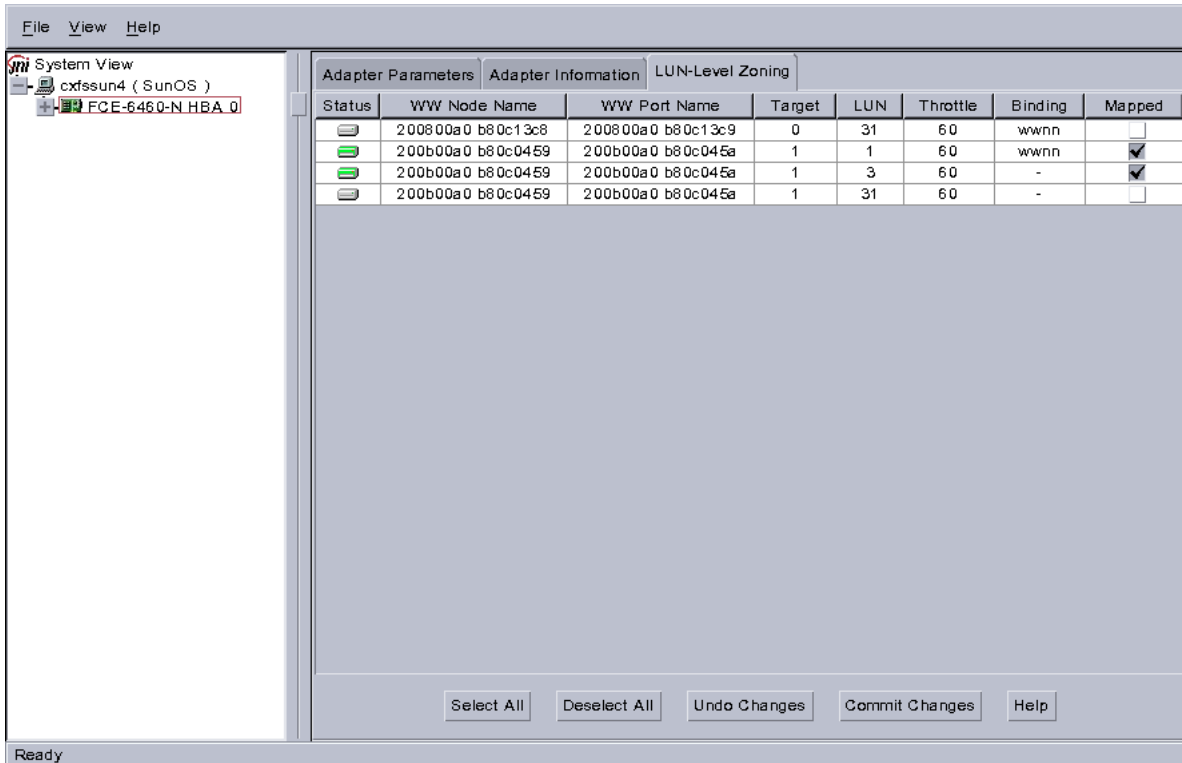
Figure 5-8 Example: After Mapping the LUNs and Committing the Changes

6. Save the `/kernel/drv/jnic146x.conf` to a backup file. For example:

   # **`cp /kernel/drv/jnic146x.conf /jnic146x.conf.save`**

7. Remove the v5.1 JNI driver package:

   # **`pkgrm JNIC146x`**

8. Insert the CXFS MultiOS CD-ROM.

9. Install the v5.0.1.SGI JNI driver (which is required for use with CXFS):

   # **`pkgadd -d /cdrom/solaris JNIC146x.pkg`**

10. Restore the `/kernel/drv/jnic146x.conf` file from the backup copy. For example:

    # **cp /jnic146x.conf.save /kernel/drv/jnic146x.conf**

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

    # **init 6**

Verify 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 "Solaris JNI Problems", page 84.

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"

*   "Add a Private Network for Solaris Nodes", page 40

*   "Verify the Private and Public Networks for Solaris", page 45

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.

Add 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 39.

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
```

Verify 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 3 128.162.89.128
PING 128.162.89.128: 56 data bytes
64 bytes from cxfssun3.americas.sgi.com (128.162.89.128): icmp_seq=0. time=0. ms
64 bytes from cxfssun3.americas.sgi.com (128.162.89.128): icmp_seq=1. time=0. ms
64 bytes from cxfssun3.americas.sgi.com (128.162.89.128): icmp_seq=2. time=0. ms
64 bytes from cxfssun3.americas.sgi.com (128.162.89.128): 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 128.162.89.127 netmask ffffff00 broadcast 128.162.89.255
        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) according to the *Solaris 8 Advanced Installation Guide*. Use the following command to display the currently installed system:

   # **uname -r**

   This command should return a value of 5.8.

2. Do the following:

   a. Insert the CXFS multiOS CD-ROM.

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

      # **pkgadd -d /cdrom SGIcxfs.pkg**

      For example (ellipses (. . .) indicate truncated output):

      pkgadd -d /cdrom SGIcxfs.pkg The following packages are available:

      • 1 SGIcxfs SGI CXFS client software (sparc) release 2.1.0.

```
# Select packages you wish to process (or "all" to process all packages). (default: all)  [?,??,q]


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

```
...

This package contains scripts which will be executed with super-user
permission during the process of installing this package.

Do you want to continue with the installation of <SGIcxfs> [y,n,?] y

Installing SGI CXFS client software as <SGIcxfs>

## Installing part 1 of 1.
/etc/init.d/cxfs_cluster
/etc/rc0.d/K77cxfs_cluster <symbolic link>
/etc/rc2.d/S77cxfs_cluster <symbolic link>
/usr/cxfs_cluster/bin/cxfs_client
/usr/cxfs_cluster/bin/cxfslicense
/usr/cxfs_cluster/bin/xvmprobe
/usr/cxfs_cluster/doc/relnotes_2.0.0.23
/usr/kernel/drv/cell.conf
/usr/kernel/drv/sparcv9/cell
/usr/kernel/drv/sparcv9/xvm
/usr/kernel/drv/xvm.conf
/usr/kernel/fs/sparcv9/cxfs
/usr/sbin/clmount
/usr/share/man/man1/xvmprobe.1
/usr/share/man/man1m/cxfs_client.1m
[ verifying class <none> ]
## Executing postinstall script.
Starting CXFS services . . .

Installation of <SGIcxfs> successful.
```

    c.   Verify that the CXFS license key has been installed. See Chapter 4, "CXFS and XVM FLEXlm License Requirements", page 19.

For example:

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

Verify 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 2.1.0
   BASEDIR:  /
    VENDOR:  Silicon Graphics Inc.
```

Postinstallation Steps for Solaris

This section discusses the configuration steps that you should perform after installing the CXFS software on the Solaris node.

Create the Solaris 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; 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 15, and "JNI Fibre Channel Host Bus Adapter Installation and Configuration", page 26.

At this point, you might be able to determine the HBA WWPN by running the EZ Fibre Configuration GUI (see "Install and Run the EZ Fibre Configuration GUI", page 28); see Figure 5-6, page 35. 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:

```
# cat /etc/fencing.conf
#
# 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 75), 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", page 51
- "Modifying the CXFS for Solaris Software", page 51

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 /etc/cxfs_cluster/bin/cxfs_cluster.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.

Windows NT Platform

This chapter contains the following:

- "CXFS on Windows NT"
- "FLEXlm License Verification for Windows NT", page 57
- "QLogic Fibre Channel Host Bus Adapter Installation", page 57
- "Preinstallation Steps for Windows NT", page 58
- "Client Software Installation Steps for Windows NT", page 61
- "Postinstallation Steps for Windows NT", page 66
- "Manual CXFS Startup/Shutdown for Windows NT", page 71
- "Software Maintenance for Windows NT", page 71

CXFS on Windows NT

This section contains the following:

- "Requirements Specific to Windows NT", page 53
- "CXFS Commands Installed on Windows NT", page 54
- "Windows NT Log Files", page 54
- "Windows NT Filesystem Administration", page 54
- "Additional Performance Considerations Specific to Windows NT", page 55
- "Additional Recommendations for Windows NT", page 56
- "Windows NT Limitations", page 56

Requirements Specific to Windows NT

In addition to the items listed in "Requirements", page 5, using a Windows NT node to support CXFS requires the insertion of a Windows NT host with at least the following:

- A 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.

- QLogic 2200/2310 Driver version 8.1.3 or later, Bios version 1.76 or later, with QLconfig and QLview installed. The latest drivers can be downloaded from the QLogic web site.

- Windows NT 4.0 Service Pack 6.

CXFS Commands Installed on Windows NT

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

Windows NT Log Files

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

Start
  > Programs
    > Administrative Tools
      > Event Viewer

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

Windows NT Filesystem Administration

Windows NT nodes can support multiple CXFS filesystems mounted under a single drive letter. Only one CXFS drive letter may be configured on a Windows NT 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, with certain restrictions; for more information, see "Windows NT Limitations", page 56.

Operations that look up or modify the directory structure (including creating files) will be slower than a local FAT or NTFS filesystem, because the operation is often performed on the CXFS metadata server. However, reading and/or writing to large files, which may be larger than 2 GB, should be significantly faster because the Windows NT host has a direct Fibre Channel connection to the disks.

Additional Performance Considerations Specific to Windows NT

In addition to the items listed in "Performance Considerations", page 4, the following issues also apply under Windows NT:

- CXFS will not properly support sharing a CXFS filesystem with another Windows NT host. It is recommended, if this is required, that an IRIX server shares the filesystem via Samba.

- 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 NT usernames to UNIX identifiers occurs as the CXFS software starts up. In a Windows NT 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 NT nodes from the passwd file that is installed on the Windows NT nodes.

This issue has less impact on Windows NT 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.

Additional Recommendations for Windows NT

In addition to the items listed in "Performance Considerations", page 4, the following are additional recommendations for Windows NT nodes:

- 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.

- 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. These tools assume particular disk formats (FAT or NTFS) and may even write to the super block of a Fibre Channel disk, just by selecting the disk. This will corrupt the XFS filesystem on the disk.

- Install the CXFS software components onto a NTFS partition rather than a FAT partition. The security of the `C:\Program Files\CXFS\passwd` and `C:\Program Files\CXFS\group` files cannot be guaranteed if these files are installed onto a FAT filesystem.

Windows NT Limitations

The following limitations are unlikely to be addressed in any future release of CXFS for Windows NT:

- User names on the Windows NT node must match the equivalent user names on other nodes in the cluster that run operating systems based on UNIX.

- 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.

- CXFS software does not export 8.3 alternative filenames. Older Windows NT applications that only support 8.3 filenames may be unable to open files with longer filenames.

- A CXFS filesystem cannot be used as the boot partition of a Windows NT host.

FLEXlm License Verification for Windows NT

The Windows NT license can be found in:

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

To verify the FLEXlm license on the target Windows NT machine, go to:

`C:\Program Files\cxfslicense`

Double click on "cxfslicense" and the following should come up:

`Found valid license for feature CXFS_NT version 2.000`

`The CPU count specified in the license is OK`

`Found valid license for feature XVM_NT version 3.000`

`The CPU count specified in the license is OK.`

QLogic Fibre Channel Host Bus Adapter Installation

Note: You are required to have a second network interface that must be used for the private metadata network. If you do not already have a second interface installed, you must install a network interface card. You may wish to install that card at this time.

The installation procedure for the QLogic Fibre Channel Host Bus Adapter (HBA) may be performed by you or by a qualified Windows NT service representative. For instructions, see the documentation included with the QLogic software.

Support for more than 8 LUNs should also be enabled as described in the Qlogic drvier README file and at the following links:http://support.microsoft.com/default.aspx?scid=kb;en-us;Q310072http://support.microsoft.com/default.aspx?scid=kb;en-us;Q245637

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

Preinstallation Steps for Windows NT

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 NT service representative will perform on your Windows NT nodes prior to installing the CXFS software. It contains the following:

- "Hostname Resolution and Network Configuration Rules for Windows NT", page 58

- "Add a Private Network for Windows NT Nodes", page 59

- "Verify the Private and Public Networks for Windows NT", page 60

Hostname Resolution and Network Configuration Rules for Windows NT

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 NT 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.

Add 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 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.

1. Install the second network adapter in the Windows NT 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.

Verify the Private and Public Networks for Windows NT

You can confirm that the previous procedure ("Add a Private Network for Windows NT Nodes", page 59) was performed correctly by using the `ipconfig` command in a DOS command shell. In this example, the 10.0.0 network is the private network and the 192.168.63 network is the public network:

```
> 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 NT

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 NT node, do the following:

1. Log onto the Windows NT node as Administrator.

2. Verify that the node has been updated to Windows NT 4.0 Service Pack 6. Select the following:

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

3. Insert the CXFS for Windows NT CD-ROM into the Windows NT 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. Select all software components, as shown in Figure 6-1. If the software is to be installed in a nondefault directory, press the **Browse** button to select another directory. Press **Next** when finished.
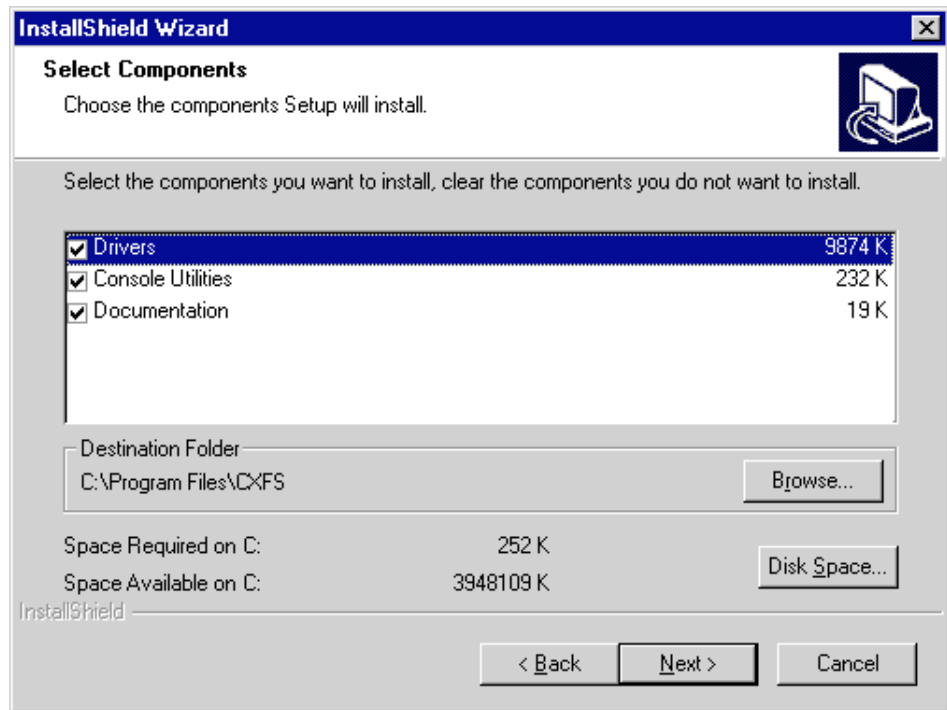
Figure 6-1 Select All Software Components

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

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

- The **Specify location of UNIX/etc/passwd and /etc/group files** field. Specify the path where the utilities will be installed, and where you must install the `passwd` and `group` file. The default is `C:\Program Files\CXFS`.

- The **Select or enter the IP address of the heartbeat network adapter** field. Specify the IP address of the private network adapter on the Windows NT node.

- The **any additional command line arguments** field. 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 51.
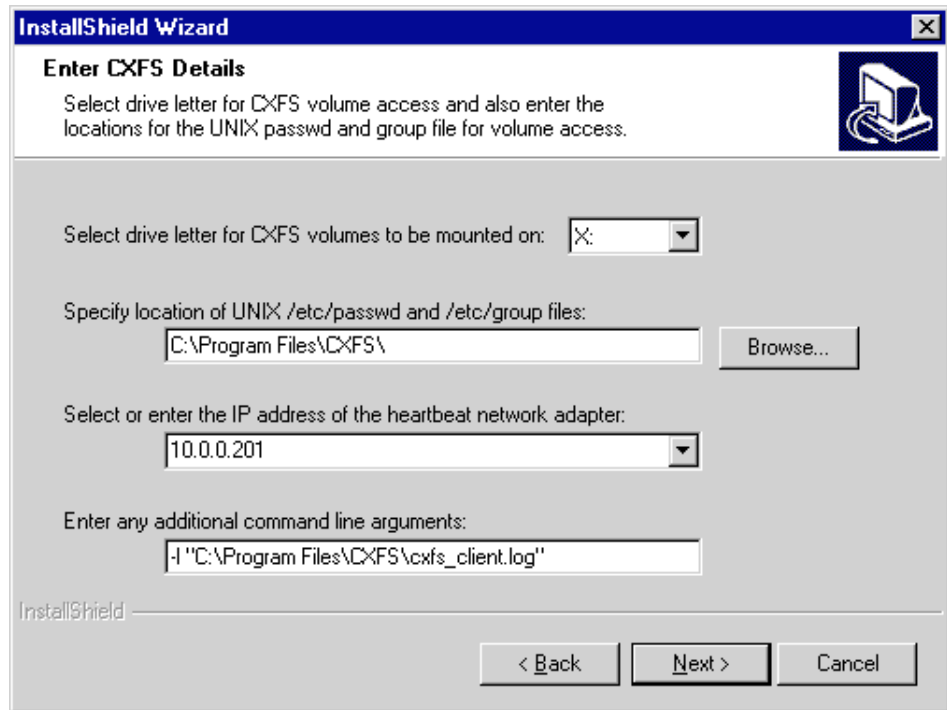
Figure 6-2 Enter CXFS Details

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.
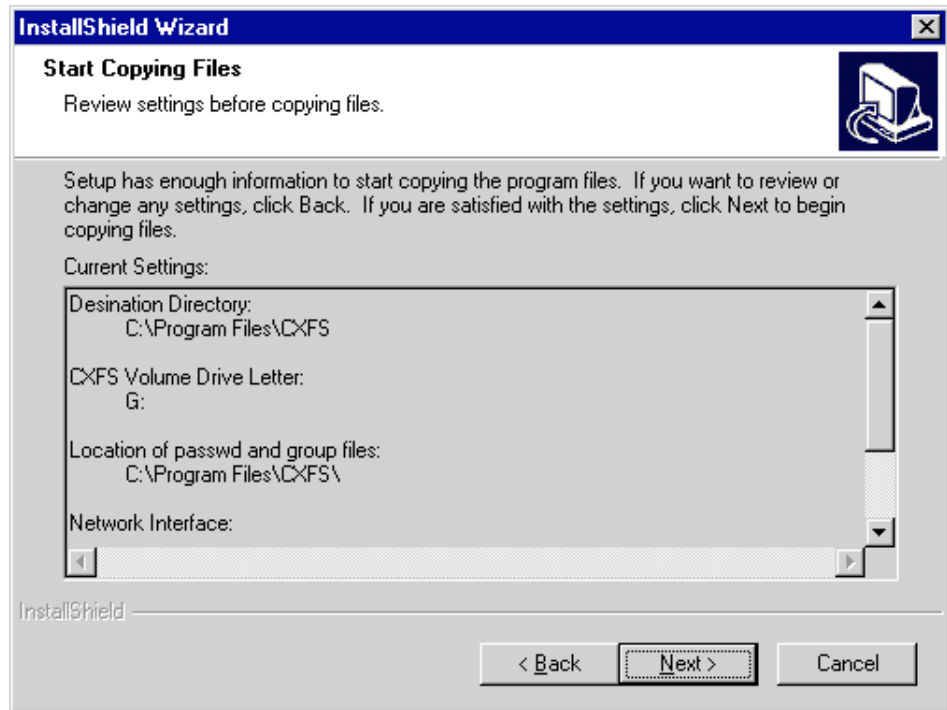
Figure 6-3 Review the Settings

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.
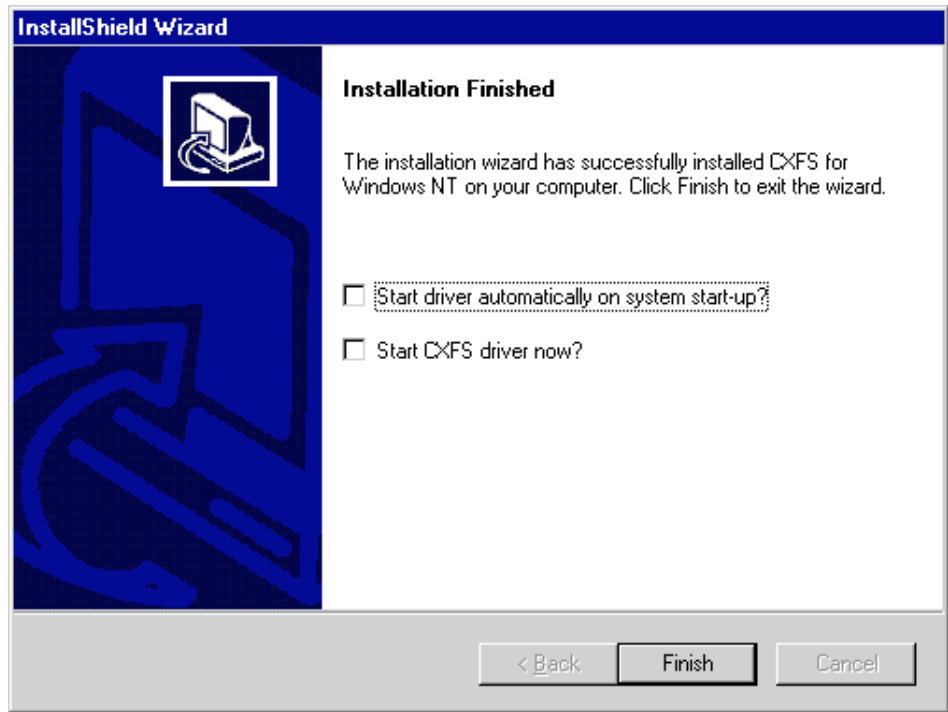
Figure 6-4 Start CXFS Driver

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 NT

This section discusses the configuration steps that you should perform after installing CXFS software on a Windows NT node.

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

- "Create the Windows NT I/O Fencing File", page 67

- "Perform User Configuration", page 69
- "Create a New Hardware Profile", page 70

Create the Windows NT I/O Fencing File

Note: Windows NT 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 NT 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; therefore, the Brocade switch is a required piece of hardware in a multiOS cluster.

To use I/O fencing, you must create the Windows NT 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 I/O fencing file can be found in:

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

There must be 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, you must first set up the Brocade Fibre Channel switch and QLogic HBA according to the directions in Chapter 3, "Brocade Fibre Channel Switch Verification", page 15, and "QLogic Fibre Channel Host Bus Adapter Installation", page 57.

Then do the following:

1. Follow the Fibre Channel cable on the back of the Windows NT 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.

```
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).

4. Edit or create the `C:\Program Files\CXFS\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 fencing file should appear as follows:

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

5. After the Windows NT node is added to the cluster (see Chapter 7, "Cluster Configuration", page 75), 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*.

Perform User Configuration

After installing the CXFS software onto the Windows NT node, but before the CXFS node is restarted, you must install the /etc/passwd and /etc/group files as follows:

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

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

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

The permissions on these files must be changed so that only the system administrator can modify the files. This can be done by right-clicking on the passwd and group files in Windows Explorer and selecting the following:

   Properties
      > Security
         > Permissions

Change the permissions to **Read** for **Everyone** and to **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.

Create 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:

> **Properties**
> > **> Hardware Profiles**
> > > **> Copy**

This action copies the current hardware profile, most likely called **Original Configuration**, to a new profile. 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, select the following and disable CXFS from the current profile:

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

You should also remove the CXFS Client service. Select the following and disable CXFS Client service from the current profile:

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

When the Windows NT host boots, you may choose the **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 NT

The CXFS processes are automatically started when a Windows NT 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:

> **Start**
>   **> Settings**
>   **> Control Panel**
>   **> Services**

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

Software Maintenance for Windows NT

This section contains the following:

- "Upgrading the CXFS for Windows NT Software"

- "Modifying the CXFS for Windows NT Software", page 73

- "Removing the CXFS for Windows NT Software", page 73

Upgrading the CXFS for Windows NT Software

To upgrade the CXFS for Windows NT 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 **Reinstall** and follow the prompts as discussed in "Client Software Installation Steps for Windows NT", page 61.

   Figure 6–5 shows the screen that lets you modify the software.
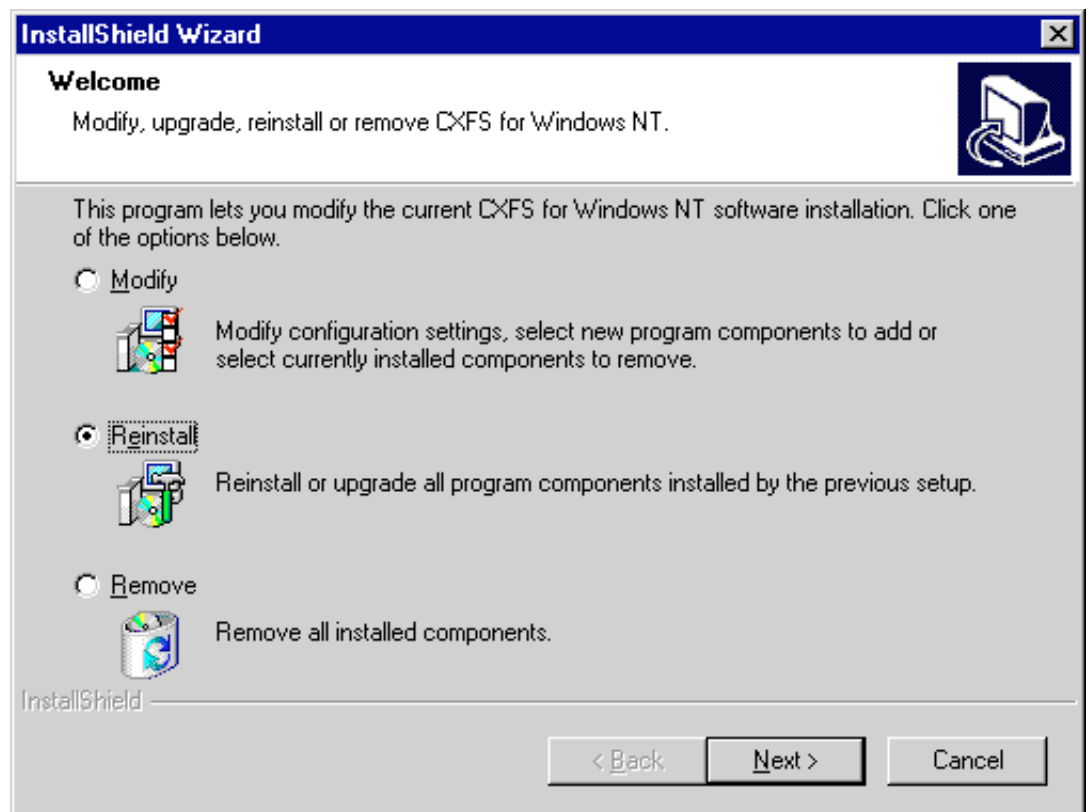
Figure 6-5 Upgrade the CXFS for Windows

3. Restart the Windows NT node. The upgraded software will not activate until the Windows NT node is restarted.

Modifying the CXFS for Windows NT Software

To change the location of the software and other configuration settings that were requested in "Client Software Installation Steps for Windows NT", page 61, perform the following steps:

1. Select the following:

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

   Figure 6-5, page 72 shows the screen that lets you modify the software.

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 NT node, which causes the changes to take effect.

Removing the CXFS for Windows NT Software

To remove the CXFS for Windows NT 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**

shows the screen that lets you remove the software.

Note: The CXFS Client service log, passwd, group and fencing.conf files will not be removed.

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

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 NT 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:

- "Define the Client-Only Nodes"
- "Add the Client-Only Nodes to the Cluster", page 77
- "Define the Switch for I/O Fencing", page 77
- "Start CXFS Services on the Client-Only Nodes", page 78
- "Mount Filesystems on the Client-Only Nodes", page 78
- "Restart the Windows NT Node", page 79
- "Verify the Cluster", page 79

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

Define 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.

For the hostname field in the GUI for a Solaris node, 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.

For the hostname field in the GUI for a Windows NT node, use the output from the following selection:

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

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|NT> ? 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*.

Add 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 "Mount Filesystems on the Client-Only Nodes", page 78.

Define 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 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 NT nodes, the fencing file determines the ports that may be fenced. For more information, see "Create the Solaris I/O Fencing File", page 48, and "Create the Windows NT I/O Fencing File", page 67.

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

Start 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*.

Mount 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
```

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*.

Restart the Windows NT Node

After completing the steps in "Postinstallation Steps for Windows NT", page 66, and this chapter, you should restart the Windows NT node. This will cause the changes to take effect.

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.

Verify the Cluster

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

```
# /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
          /dev/cxvm/fs1                  /fs1                  cxfs7        UP
```

Troubleshooting

This chapter contains the following:

- "Identifying Problems on Solaris Nodes"
- "Identifying Problems on Windows NT Nodes", page 82
- "Common Problems and Solutions", page 83
- "Reporting Problems", page 86

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?", page 81
- "Are there Error Messages for the Solaris Node?", page 81

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 "Verify the Cluster", page 79.

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 worlds "ERROR" or "Warning." Some specific cases where error messages will occur include:

- The fencing file was not found, fencing configuration will not be updated on the server. `cxfs_client: cis_get_hba_wwns warning: fencing configuration file "fencing.conf" not found`

- A filesystem mount has failed and will be retried.

  `cxfs_client:op_failed ERROR: Mount failed for concat0`

For more information about these files, see "Solaris Log Files", page 23. 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 NT Nodes

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

- "Is the CXFS Software Running Correctly on the Windows NT Node?"

- "Is the Windows NT Node in the Cluster?", page 82

- "Are there Error Messages for the Windows NT Node?", page 83

Is the CXFS Software Running Correctly on the Windows NT Node?

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**

Is the Windows NT 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 "Verify the Cluster", page 79.

Are there Error Messages for the Windows NT Node?

To see if there are any error or warning messages, look at the `C:\Program Files\CXFS\log\cxfs_client.log` file and the **System Event** log by selecting the following:

> **Start**
> > **> Programs**
> > > **> Administrative Tools**
> > > > **> Event Viewer**

Some specific warnings that may come up include:

- The fencing file was not found, fencing configuration will not be updated on the server. `cxfs_client: cis_get_hba_wwns warning: fencing configuration file "fencing.conf" not found`

- A filesystem mount has failed and will be retried.

  `cxfs_client:op_failed ERROR: Mount failed for concat0`

- The password file could not be found. `CXFS_Client:cis_generate_userid_map warning: could not open passwd file.`

- The group files could not be found.`CXFS_Client:cis_generate_userid_map warning: could not open group file.`

For these last two warnings, filesystem mounts should proceed; however, all users will be given "nobody" credentials and will be unable to modify or see files on the CXFS filesystems.

For more information about these files, see "Solaris Log Files", page 23, and "Windows NT Log Files", page 54. Also see the log files on the IRIX node; for more information, see the *CXFS Version 2 Software Installation and Administration Guide*.

Common Problems and Solutions

This section contains the following common problems and their solutions:

- "Incorrect Configuration"

- "Determining If Your Client-Only Node Is Fenced", page 84

- "Common Solaris Problems", page 84

- "Common Windows NT Problems", page 85

Incorrect Configuration

To avoid having trouble with the CXFS client-only node, ensure you have the correct configuration. See "Requirements", page 5.

Determining If Your Client-Only Node Is Fenced

To determine if your Client-only node is fenced, go into IRIX and use the command `hafence`(1M).For more details see the *CXFS Version 2 Software Installation and Administration Guide*.

Common Solaris Problems

This section contains the following common Solaris problems:

- "Solaris JNI Problems", page 84

Solaris 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 "Install the JNI HBA", page 26.

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 NT Problems

This section contains the following common Windows NT problems:

- "Filesystems Are Not Displayed on a Windows NT Node"

- "Large Log Files on Windows NT"

- "Windows NT Failure on Restart", page 86

Filesystems Are Not Displayed on a Windows NT Node

If the CXFS drive letter is visible in Windows Explorer but no filesystems are mounted, do the following:

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

  Start
  > **Settings**
  > **Control Panel**
  > **Services**

- Check the `C:\Program Files\CXFS\log\cxfs_client.log` file for warnings and errors regarding licenses or mounting filesystems.

- Check the cluster configuration to ensure that this node is configured to mount one or more filesystems.

Large Log Files on Windows NT

The CXFS Client service creates the following log file:

`C:\ProgramFiles\CXFS\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 needs to be stopped to move or truncate this file, then restarted. See "Manual CXFS Startup/Shutdown for Windows NT", page 71 on how to stop and start the CXFS Client Service.

Windows NT Failure on Restart

If the CXFS Windows NT 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.

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.

- 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

- Output from the crash(1M) utility.

- 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 NT Problems

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 `C:\Program Files\CXFS\log\cxfs_client.log` file. 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 NT", page 88.

Save Crash Dumps for Windows NT

If you are having problems, you should configure the Windows NT 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:

> **Properties**
> > **> Startup**
> > > **> Shutdown**
> > > > **> 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.

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 NT 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. In a multiOS cluster, all IRIX nodes must be running a minimum of 6.5.16f.

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.

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