



CXFS™ MultiOS Client-Only Guide for  
SGI® InfiniteStorage

007-4507-011

---

## CONTRIBUTORS

Written by Lori Johnson

Illustrated by Chrystie Danzer

Engineering contributions to the book by Vlad Apostolov, Neil Bannister, Dale Brantly, David Chatterton, Mark Cruciani, Tad Dolphay, Dave Ellis, Eric Epe, Andrew Gildfind, Gary Hagensen, John Keller, Dennis Kender, Aaron Mantel, Troy McCorkell, Ken McDonell, Bill Mckevitt, Terry Merth, Daniel Moore, Max Matveev, Fujio Nakajima, Barry Naujok, Mike Raskie, Eric Sandeen, Tim Sirianni, Paddy Sreenivasan, Wesley Smith, Michael Umansky, Madan Valluri, Chris Walters, Geoffrey Wehrman, Connie Woodward, Herry Wiputra

---

## COPYRIGHT

© 2002–2005 Silicon Graphics, Inc. All rights reserved; provided portions may be copyright in third parties, as indicated elsewhere herein. No permission is granted to copy, distribute, or create derivative works from the contents of this electronic documentation in any manner, in whole or in part, without the prior written permission of Silicon Graphics, Inc., in the United States and/or other countries worldwide.

---

## LIMITED RIGHTS LEGEND

The software described in this document is "commercial computer software" provided with restricted rights (except as to included open/free source) as specified in the FAR 52.227-19 and/or the DFAR 227.7202, or successive sections. Use beyond license provisions is a violation of worldwide intellectual property laws, treaties and conventions. This document is provided with limited rights as defined in 52.227-14.

---

## TRADEMARKS AND ATTRIBUTIONS

Silicon Graphics, SGI, the SGI logo, IRIX, O2, Origin, XFS, and Altix are registered trademarks and CXFS, FailSafe, IRIS FailSafe, SGI ProPack, and Trusted IRIX are trademarks of Silicon Graphics, Inc., in the United States and/or other countries worldwide.

Active Directory, Microsoft, Windows, and Windows NT are registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. AIX and IBM are registered trademarks of IBM Corporation. Brocade and SilkWorm are trademarks of Brocade Communication Systems, Inc. AMCC, FibreStar, and JNI are registered trademarks and EZ Fibre is a trademark of Applied Micro Circuits Corporation (formerly JNI Corporation). AMD, AMD Athlon, AMD Duron, and AMD Opteron are trademarks of Advanced Micro Devices, Inc. Apple, Mac, Mac OS, Power Mac, and Xserve are registered trademarks of Apple Computer, Inc. Disk Manager is a registered trademark of ONTRACK Data International, Inc. Engenio and SANshare are trademarks or registered trademarks of LSI Logic Corporation. LSI Logic is a trademarks or registered trademark of LSI Logic Corp. FLEXlm is a registered trademark of Macrovision Corporation. HP-UX is a trademark of Hewlett-Packard Company. InstallShield is a registered trademark of InstallShield Software Corporation in the United States and/or other countries. Intel, Intel Xeon, Itanium, and Pentium are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. All other trademarks are the property of their respective holder. Legato NetWorker is a registered trademark of Legato Systems, Inc. Linux is a registered trademark of Linus Torvalds. OpenLDAP is a registered trademark of OpenLDAP Foundation. Red Hat and all Red Hat-based trademarks are trademarks or registered trademarks of Red Hat, Inc. in the United States and other countries. SANsurfer and QLogic are registered trademarks and SANblade is a trademark of QLogic Corporation. Solaris, Sun, Sun Blade, Sun Fire, SunOS, and Ultra Enterprise are trademarks or registered trademarks of Sun Microsystems, Inc. SPARC is a registered trademark of SPARC International, Inc. UNIX and the X device are registered trademarks of The Open Group in the United States and other countries. All other trademarks mentioned herein are the property of their respective owners.

Screen snaps of the EZ Fibre product are provided by express permission of JNI Corporation (now Applied Micro Circuits Corporation).

The `lsOf` command is written by Victor A. Abell and is copyright of Purdue Research Foundation.

---

## New Features in this Guide

---

**Note:** Be sure to read the release notes for your platforms to learn about any late-breaking changes to the installation and configuration procedures.

---

This guide contains the following changes:

- Support for Microsoft Windows Server 2003. See Chapter 6, "Windows Platforms" on page 145.
- Support for AMD AMD64, Intel EM64T, and Intel Itanium 2 third-party Linux systems as client-only nodes. See Chapter 3, "Linux Third-Party Platforms" on page 47.
- Information about guaranteed-rate I/O (GRIO) version 2:
  - "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14
  - "GRIO on AIX" on page 38
  - "GRIO on Linux" on page 71
  - "GRIO on Mac OS X" on page 97
  - "GRIO on Solaris" on page 137
  - "GRIO on Windows" on page 201
- Information about XVM version 2 failover:
  - "XVM Failover Version 2 and CXFS" on page 17
  - "XVM Failover Version 2 on AIX" on page 38
  - "XVM Failover Version 2 on Linux" on page 72
  - "XVM Failover Version 2 on Mac OS X" on page 98
  - "XVM Failover Version 2 on Solaris" on page 137
- Platform-specific information about FLEXlm licenses and troubleshooting has been separated out into the various platform-specific chapters. General troubleshooting information remains in Chapter 8, "General Troubleshooting" on page 231.

- "Configuring the Windows XP SP2 Firewall for Windows" on page 184.
- Appendix D, "Error Messages" on page 251.
- Information about the steps required after making changes to the storage systems:
  - "Recognizing Storage Changes for AIX" on page 38
  - "Recognizing Storage Changes for Mac OS X" on page 97
  - "Recognizing Storage Changes for Solaris" on page 137
  - "Recognizing Storage Changes for Windows" on page 200
- System tunables information:
  - "System Tunables for Solaris" on page 108
  - "System Tunables for Windows" on page 168
- "SANshare License and XVM Version 2 Failover for AIX" on page 27
- "Configuring HBA Failover for Windows 2000 and Windows 2003" on page 180
- New sections to
  - "Verifying the Cluster Configuration" on page 222
  - "Verifying Connectivity in a Multicast Environment" on page 222
  - "Verifying the Cluster Status" on page 223
  - "Verifying the I/O Fencing Configuration" on page 226
  - "Verifying Access to XVM Volumes" on page 227
- Removed references to `xvmprobe`. The functionality of `xvmprobe` has been replaced by the `xvm` command.

---

## Record of Revision

<b>Version</b>	<b>Description</b>
001	March 2002 Original publication with the CXFS MultiOS Clients 2.0 release for IRIX 6.5.16f.
002	May 2002 Revised to support the CXFS MultiOS Clients 2.1 release for IRIX 6.5.16f. This release supports the Sun Microsystems Solaris and Microsoft Windows NT platforms.
003	June 2002 Revised to support the CXFS MultiOS Clients 2.1.1 release for IRIX 6.5.16f. This release supports the Sun Microsystems Solaris and Microsoft Windows NT platforms.
004	August 2002 Revised to support the CXFS MultiOS 2.2 Clients release for IRIX 6.5.17f. This release supports the Sun Microsystems Solaris, Microsoft Windows NT, and Microsoft Windows 2000 platforms.
005	November 2002 Revised to support the CXFS MultiOS Clients 2.3 release for IRIX 6.5.18f. This release supports the Sun Microsystems Solaris, Microsoft Windows NT, and Microsoft Windows 2000 platforms.
006	February 2003 Revised to support the CXFS MultiOS Clients 2.4 release for IRIX 6.5.19f. This release supports the Sun Microsystems Solaris, Microsoft Windows NT, and Microsoft Windows 2000 platforms.
007	May 2003 Revised to support the CXFS MultiOS Clients 2.5 release for IRIX 6.5.20f. This release supports the IBM AIX platform, Linux on supported 32-bit platforms, SGI ProPack for Linux on SGI Altix 3000 family of servers and superclusters, Sun Microsystems Solaris platform, Microsoft Windows NT platform, and Microsoft Windows 2000 platform.

- 008            September 2003  
Revised to support CXFS MultiOS Clients 3.0. This release supports the IBM AIX platform, Linux on supported 32-bit platforms, Sun Microsystems Solaris platform, Microsoft Windows NT platform, Microsoft Windows 2000 platform, and Microsoft Windows XP platform. The documentation for Linux 64-bit nodes supported by the CXFS 3.0 for SGI ProPack release will appear in the next version of the *CXFS Administration Guide for SGI InfiniteStorage*.
- 009            February 2004  
Revised to support CXFS MultiOS Clients 3.1. This release supports the Apple Mac OS X platform, IBM AIX platform, Linux on supported 32-bit platforms, Sun Microsystems Solaris platform, Microsoft Windows 2000 platform, and Microsoft Windows XP platform.
- 010            November 2004  
Revised to support CXFS MultiOS Clients 3.2. This release supports the Apple Mac OS X platform, IBM AIX platform, Linux on supported 32-bit platforms, Sun Microsystems Solaris platform, Microsoft Windows 2000 platform, and Microsoft Windows XP platform.
- 011            April 2005  
Revised to support CXFS MultiOS Clients 3.3. This release supports the Apple Mac OS X platform, IBM AIX platform, Linux on supported third-party platforms (x86, AMD64/EM64T, Intel Itanium 2), Sun Microsystems Solaris platform, Microsoft Windows 2000 platform, Microsoft Windows Server 2003, and Microsoft Windows XP platform.

---

# Contents

<b>About This Guide</b>	<b>xxiii</b>
Prerequisites	xxiii
Related Publications	xxiii
Obtaining Publications	xxv
Conventions	xxvi
Reader Comments	xxvii
<b>1. Introduction</b>	<b>1</b>
When to Use CXFS	2
CXFS on Client-Only Nodes	3
CXFS Processes	3
Cluster Administration	4
User Administration for CXFS	4
Performance Considerations	4
User and Group Quotas	6
CXFS Mount Scripts	6
Requirements	7
Recommendations	9
Licenses	12
Hostname Resolution and Network Configuration Rules	13
Guaranteed-Rate I/O (GRIO) and CXFS	14
GRIO Commands	15
GRIO Examples	15
List GRIO-Managed Volumes	16
Monitor a Particular Stream	16
<b>007-4507-011</b>	<b>vii</b>

XVM Failover Version 2 and CXFS . . . . .	17
Monitoring CXFS . . . . .	18
<b>2. AIX Platform . . . . .</b>	<b>19</b>
CXFS on AIX . . . . .	19
Requirements for AIX . . . . .	20
CXFS Commands on AIX . . . . .	21
Log Files on AIX . . . . .	21
CXFS Mount Scripts on AIX . . . . .	22
Limitations and Considerations for AIX . . . . .	22
Access Control Lists and AIX . . . . .	24
Installation and Configuration Overview for AIX . . . . .	25
FLEXlm License for AIX . . . . .	25
Gathering the AIX Host Information . . . . .	26
Obtaining the Licenses for AIX . . . . .	26
Installing the Licenses for AIX . . . . .	27
SANshare License and XVM Version 2 Failover for AIX . . . . .	27
HBA Installation for AIX . . . . .	28
Preinstallation Steps for AIX . . . . .	28
Adding a Private Network for AIX . . . . .	28
Verifying the Private and Public Network for AIX . . . . .	31
Client Software Installation for AIX . . . . .	31
Installing CXFS Software on AIX . . . . .	32
Verifying the AIX Installation . . . . .	34
I/O Fencing for AIX . . . . .	35
Start/Stop <code>cxfs_client</code> Daemon for AIX . . . . .	36
Maintenance for AIX . . . . .	37
Upgrading the CXFS Software for AIX . . . . .	37

Modifying the CXFS Software for AIX . . . . .	37
Recognizing Storage Changes for AIX . . . . .	38
GRIO on AIX . . . . .	38
XVM Failover Version 2 on AIX . . . . .	38
Troubleshooting for AIX . . . . .	39
Verifying the Licenses for AIX . . . . .	39
Unable to Mount Filesystems on AIX . . . . .	40
The <code>cxfs_client</code> Daemon is Not Started on AIX . . . . .	42
Filesystems Do Not Mount on AIX . . . . .	43
Panic Occurs when Executing <code>cxfs_cluster</code> on AIX . . . . .	43
A Memory Error Occurs with <code>cp -p</code> on AIX . . . . .	43
An ACL Problem Occurs with <code>cp -p</code> on AIX . . . . .	43
Large Log Files on AIX . . . . .	44
Reporting AIX Problems . . . . .	44
<b>3. Linux Third-Party Platforms . . . . .</b>	<b>47</b>
CXFS on Linux . . . . .	48
Requirements for Linux . . . . .	48
CXFS Commands on Linux . . . . .	49
Log Files on Linux . . . . .	49
CXFS Mount Scripts on Linux . . . . .	50
Limitations and Considerations for Linux . . . . .	50
Access Control Lists and Linux . . . . .	51
Installation and Configuration Overview for Linux . . . . .	51
FLEXlm License for Linux . . . . .	52
Gathering the Linux Host Information . . . . .	52
Obtaining the Licenses for Linux . . . . .	55
Installing the Licenses for Linux . . . . .	55

HBA Installation for Linux . . . . .	55
Preinstallation Steps for Linux . . . . .	57
Adding a Private Network for Linux . . . . .	57
Modifications Required for CXFS Connectivity Diagnostics for Linux . . . . .	59
Verifying the Private and Public Networks for Linux . . . . .	60
Client Software Installation for Linux . . . . .	61
Linux Installation Overview . . . . .	61
Verifying the Linux Installation . . . . .	65
I/O Fencing for Linux . . . . .	65
Start/Stop <code>cxfs_client</code> for Linux . . . . .	67
Maintenance for Linux . . . . .	69
Modifying the CXFS Software for Linux . . . . .	69
Using <code>cxfs-reprobe</code> with Linux . . . . .	69
GRIO on Linux . . . . .	71
XVM Failover Version 2 on Linux . . . . .	72
Troubleshooting for Linux . . . . .	72
Verifying the Licenses for Linux . . . . .	73
Device Filesystem Enabled for Linux . . . . .	74
The <code>cxfs_client</code> Daemon is Not Started on Linux . . . . .	74
Filesystems Do Not Mount on Linux . . . . .	74
Large Log Files on Linux . . . . .	76
Reporting Linux Problems . . . . .	76
<b>4. Mac OS X Platform . . . . .</b>	<b>79</b>
CXFS on Mac OS X . . . . .	79
Requirements for Mac OS X . . . . .	80
CXFS Commands on Mac OS X . . . . .	80
Log Files on Mac OS X . . . . .	81

Limitations and Considerations on Mac OS X . . . . .	81
Configuring Hostnames on Mac OS X . . . . .	82
Mapping User and Group Identifiers for Mac OS X . . . . .	83
Access Control Lists and Mac OS X . . . . .	84
Installation and Configuration Overview for Mac OS X . . . . .	84
FLEXlm License for Mac OS X . . . . .	85
Gathering the Mac OS X Host Information . . . . .	85
Obtaining the Licenses for Mac OS X . . . . .	86
Installing the Licenses for Mac OS X . . . . .	86
HBA Installation for Mac OS X . . . . .	87
Installing the Apple HBA . . . . .	87
Installing the Fibre Channel Utility for Mac OS X . . . . .	88
Configuring Two or More Apple HBA Ports . . . . .	88
Using <code>point-to-point</code> Fabric Setting for Apple HBAs . . . . .	88
Preinstallation Steps for Mac OS X . . . . .	89
Adding a Private Network for Mac OS X Nodes . . . . .	89
Verifying the Private and Public Networks for Mac OS X . . . . .	91
Disabling Power Save Mode for Mac OS X . . . . .	91
Client Software Installation for Mac OS X . . . . .	92
I/O Fencing for Mac OS X . . . . .	93
Start/Stop <code>cxfs_client</code> for Mac OS X . . . . .	96
Maintenance for Mac OS X . . . . .	96
Upgrading the CXFS Software for Mac OS X . . . . .	96
Modifying the CXFS Software for Mac OS X . . . . .	97
Removing the CXFS Software for Mac OS X . . . . .	97
Recognizing Storage Changes for Mac OS X . . . . .	97
GRIO on Mac OS X . . . . .	97

XVM Failover Version 2 on Mac OS X . . . . .	98
Mapping XVM Volumes to Storage Targets on Mac OS X . . . . .	98
Troubleshooting for Mac OS X . . . . .	99
Verifying the Licenses for Mac OS X . . . . .	99
The <code>cxfs_client</code> Daemon is Not Started on Mac OS X . . . . .	100
XVM Volume Name is Too Long on Mac OS X . . . . .	101
Large Log Files on Mac OS X . . . . .	101
Reporting Mac OS X Problems . . . . .	101
<b>5. Solaris Platform . . . . .</b>	<b>103</b>
CXFS on Solaris . . . . .	103
Requirements to Solaris . . . . .	104
CXFS Commands on Solaris . . . . .	105
Log Files on Solaris . . . . .	106
CXFS Mount Scripts on Solaris . . . . .	106
Limitations and Considerations on Solaris . . . . .	106
Access Control Lists and Solaris . . . . .	107
System Tunables for Solaris . . . . .	108
Installation and Configuration Overview for Solaris . . . . .	109
FLEXlm License for Solaris . . . . .	110
Gathering the Solaris Host Information . . . . .	110
Obtaining the Licenses for Solaris . . . . .	110
Installing the Licenses for Solaris . . . . .	111
HBA Installation for Solaris . . . . .	111
Installing the AMCC JNI HBA . . . . .	112
Setting the LUN Discovery Method for SGI TP9100 . . . . .	113
Protecting Data Integrity . . . . .	114
Installing and Running the EZ Fibre Configuration GUI . . . . .	114

Verifying the JNI HBA Installation . . . . .	123
Preinstallation Steps for Solaris . . . . .	125
Adding a Private Network for Solaris Nodes . . . . .	125
Verifying the Private and Public Networks for Solaris . . . . .	130
Client Software Installation for Solaris . . . . .	131
Solaris Installation Overview . . . . .	131
Verifying the Solaris Installation . . . . .	132
I/O Fencing for Solaris . . . . .	133
Start/Stop <code>cxfs_client</code> for Solaris . . . . .	135
Maintenance for Solaris . . . . .	135
Upgrading the CXFS Software for Solaris . . . . .	136
Modifying the CXFS Software for Solaris . . . . .	136
Recognizing Storage Changes for Solaris . . . . .	137
GRIO on Solaris . . . . .	137
XVM Failover Version 2 on Solaris . . . . .	137
Mapping XVM Volumes to Storage Targets on Solaris . . . . .	138
Troubleshooting for Solaris . . . . .	139
Verifying the Licenses for Solaris . . . . .	139
The <code>cxfs_client</code> Daemon is Not Started on Solaris . . . . .	140
Filesystems Do Not Mount on Solaris . . . . .	141
New Storage is Not Recognized on Solaris . . . . .	142
Large Log Files on Solaris . . . . .	142
Reporting Solaris Problems . . . . .	142
<b>6. Windows Platforms . . . . .</b>	<b>145</b>
CXFS on Windows . . . . .	146
Requirements for Windows . . . . .	147
CXFS Commands on Windows . . . . .	148

Log Files and Cluster Status for Windows . . . . .	149
Functional Limitations for Windows . . . . .	153
XVM Failover Version 2 Is Not Supported . . . . .	153
UNIX Perspective of CXFS for Windows . . . . .	153
Windows Perspective of CXFS for Windows . . . . .	154
Forced Unmount on Windows . . . . .	155
Define LUN 0 on All Storage Devices for Windows . . . . .	155
Memory-Mapping Large Files for Windows . . . . .	155
CXFS Mount Scripts for Windows . . . . .	155
Performance Considerations for Windows . . . . .	155
Access Controls for Windows . . . . .	157
User Identification for Windows . . . . .	158
User Identification Mapping Methods for Windows . . . . .	159
Enforcing Access to Files and Directories for Windows . . . . .	160
Viewing and Changing File Attributes with Windows Explorer . . . . .	161
Viewing and Changing File Permissions with Windows Explorer . . . . .	162
Viewing and Changing File Access Control Lists (ACLs) for Windows . . . . .	164
Effective Access for Windows . . . . .	165
Restrictions with file ACLs for Windows . . . . .	165
Inheritance and Default ACLs for Windows . . . . .	166
System Tunables for Windows . . . . .	168
Registry Modification . . . . .	168
Default Umask for Windows . . . . .	169
Maximum DMA Size for Windows . . . . .	169
Memory-Mapping Coherency for Windows . . . . .	170
DNLC Size for Windows . . . . .	170
Mandatory Locks for Windows . . . . .	171
User Identification Map Updates for Windows . . . . .	172

I/O Size Issues Within the QLogic HBA . . . . .	173
Command Tag Queueing (CTQ) Used by the QLogic HBA . . . . .	173
Windows Installation and Configuration Overview . . . . .	174
FLEXlm License for Windows . . . . .	175
Gathering the Windows Host Information for Windows . . . . .	175
Obtaining the Licenses for Windows . . . . .	176
Installing the Licenses for Windows . . . . .	177
HBA Installation for Windows . . . . .	177
Confirming the QLogic HBA Installation for Windows . . . . .	178
Configuring Multiple HBAs for Load Balancing on Windows . . . . .	178
Configuring HBA Failover for Windows 2000 and Windows 2003 . . . . .	180
Preinstallation Steps for Windows . . . . .	181
Adding a Private Network for Windows . . . . .	181
Procedure to Add a Private Network for Windows . . . . .	181
Ensuring Proper Hostname Configuration for Windows . . . . .	183
Verifying the Private and Public Networks for Windows . . . . .	183
Configuring the Windows XP SP2 Firewall for Windows . . . . .	184
Client Software Installation for Windows . . . . .	185
Postinstallation Steps for Windows . . . . .	191
Performing User Configuration for Windows . . . . .	192
Checking Permissions on the Password and Group Files for Windows . . . . .	193
I/O Fencing on Windows . . . . .	193
Start/Stop the CXFS Client Service for Windows . . . . .	196
Maintenance for Windows . . . . .	196
Modifying the CXFS Software for Windows . . . . .	197
Upgrading the CXFS Software for Windows . . . . .	198
Removing the CXFS Software for Windows . . . . .	199

Downgrading the CXFS Software for Windows . . . . .	200
Recognizing Storage Changes for Windows . . . . .	200
GRIO on Windows . . . . .	201
Troubleshooting for Windows . . . . .	202
Verifying the Licenses for Windows . . . . .	203
Verifying that the CXFS Software is Running Correctly for Windows . . . . .	203
Unable to Mount Filesystems on Windows . . . . .	204
Access-Denied Error when Accessing Filesystem on Windows . . . . .	206
Application Works with NTFS but not CXFS for Windows . . . . .	206
Delayed-Write Error Dialog is Generated by the Windows Kernel . . . . .	207
CXFS Client Service Does Not Start on Windows . . . . .	207
HBA Problems . . . . .	207
CXFS Client Service Cannot Map Users other than Administrator for Windows . . . . .	208
Filesystems Are Not Displayed on Windows . . . . .	209
Large Log Files on Windows . . . . .	210
Windows Failure on Restart . . . . .	210
Memory Configuration for Windows . . . . .	210
Application Cannot Create File Under CXFS Drive Letter . . . . .	211
Reporting Windows Problems . . . . .	211
Retain Windows Information . . . . .	211
Save Crash Dumps for Windows . . . . .	212
Generating a Crash Dump on a Hung Windows Node . . . . .	212
<b>7. Cluster Configuration . . . . .</b>	<b>215</b>
Defining the Client-Only Nodes . . . . .	216
Adding the Client-Only Nodes to the Cluster . . . . .	217
Defining the Switch for I/O Fencing . . . . .	218

Starting CXFS Services on the Client-Only Nodes . . . . .	219
Verifying LUN Masking . . . . .	220
Mounting Filesystems on the Client-Only Nodes . . . . .	220
Forced Unmount of CXFS Filesystems . . . . .	221
Restarting the Windows Node . . . . .	222
Verifying the Cluster Configuration . . . . .	222
Verifying Connectivity in a Multicast Environment . . . . .	222
Verifying the Cluster Status . . . . .	223
Verifying the I/O Fencing Configuration . . . . .	226
Verifying Access to XVM Volumes . . . . .	227
<b>8. General Troubleshooting . . . . .</b>	<b>231</b>
Identifying Problems . . . . .	231
Is the Client-Only Node Configured Correctly? . . . . .	231
Is the Client-Only Node in Membership? . . . . .	232
Is the Client-Only Node Mounting All Filesystems? . . . . .	232
Can the Client-Only Node Access All Filesystems? . . . . .	232
Are There Error Messages? . . . . .	233
Typical Problems and Solutions . . . . .	233
Unable to Achieve Membership . . . . .	234
Filesystem Appears to Be Hung . . . . .	235
Determining If a Client-Only Node Is Fenced . . . . .	236
No HBA WWPNs are Detected . . . . .	238
Reporting Problems to SGI . . . . .	238
<b>Appendix A. Operating System Path Differences . . . . .</b>	<b>239</b>
<b>Appendix B. Filesystem and Logical Unit Specifications . . . . .</b>	<b>245</b>

<b>Appendix C. Mount Options Support</b>	<b>247</b>
<b>Appendix D. Error Messages</b>	<b>251</b>
Windows Messages	251
CMS Error Messages	252
Mount Messages	253
Network Connectivity Messages	253
<b>Appendix E. Summary of New Features from Previous Releases</b>	<b>255</b>
CXFS MultiOS 2.0	255
CXFS MultiOS 2.1	255
CXFS MultiOS 2.1.1	255
CXFS MultiOS 2.2	256
CXFS MultiOS 2.3	256
CXFS MultiOS 2.4	256
CXFS MultiOS 2.5	257
CXFS MultiOS 3.0	258
CXFS MultiOS 3.1	258
CXFS MultiOS 3.2	258
<b>Glossary</b>	<b>261</b>
<b>Index</b>	<b>273</b>

---

## Figures

<b>Figure 5-1</b>	Example: Second Window: EZ Fibre Configuration Utility - Standalone . . .	116
<b>Figure 5-2</b>	Location of icon (+) to Display the HBA . . . . .	117
<b>Figure 5-3</b>	Example: After Clicking + to Display the HBA . . . . .	118
<b>Figure 5-4</b>	Location of the Icon to Display the Adapter Parameters . . . . .	119
<b>Figure 5-5</b>	Example: After Clicking the HBA Icon to Show the Adapter Parameters . . .	120
<b>Figure 5-6</b>	After Clicking the Adapter Information Tab . . . . .	121
<b>Figure 5-7</b>	After Clicking on LUN-Level Zoning . . . . .	122
<b>Figure 5-8</b>	Example: After Mapping the LUNs and Committing the Changes . . . . .	123
<b>Figure 6-1</b>	CXFS Info Window — Nodes Tab Display . . . . .	150
<b>Figure 6-2</b>	CXFS Info Window — Filesystems Tab . . . . .	151
<b>Figure 6-3</b>	CXFS Info Window — CXFS Client Log Tab . . . . .	152
<b>Figure 6-4</b>	Private Properties: Selecting only TCP/IP . . . . .	182
<b>Figure 6-5</b>	Choose Destination Location . . . . .	186
<b>Figure 6-6</b>	Enter CXFS Details . . . . .	187
<b>Figure 6-7</b>	Active Directory Details . . . . .	188
<b>Figure 6-8</b>	Generic LDAP Details . . . . .	189
<b>Figure 6-9</b>	Review the Settings . . . . .	190
<b>Figure 6-10</b>	Start CXFS Driver . . . . .	191
<b>Figure 6-11</b>	Modify CXFS for Windows . . . . .	197
<b>Figure 6-12</b>	Upgrading the Windows Software . . . . .	199
<b>Figure 6-13</b>	CXFS Info Display for GRIO for Windows . . . . .	201



---

## Tables

<b>Table 3-1</b>	Processor and Package Extension Examples . . . . .	61
<b>Table 6-1</b>	Permission Flags that May Be Edited . . . . .	163
<b>Table A-1</b>	AIX Paths . . . . .	239
<b>Table A-2</b>	Linux Paths . . . . .	240
<b>Table A-3</b>	Mac OS X Paths . . . . .	241
<b>Table A-4</b>	Solaris Paths . . . . .	242
<b>Table A-5</b>	Windows Paths . . . . .	243
<b>Table B-1</b>	Filesystem and Logical Unit Specifications . . . . .	246
<b>Table C-1</b>	Mount Options Support for Client-Only Platforms . . . . .	248



---

## About This Guide

This publication documents the CXFS MultiOS Clients 3.3 release. This release supports Apple Computer Mac OS X, IBM AIX, Linux on supported third-party platforms, Sun Microsystems Solaris, Microsoft Windows 2000, Microsoft Windows 2003, and Microsoft Windows XP nodes.

### Prerequisites

This guide assumes the following:

- The IRIX or SGI ProPack for Linux CXFS cluster is installed and operational.
- The CXFS client-only nodes have the appropriate platform-specific operating system software installed.
- The reader is familiar with the information presented in the *CXFS Administration Guide for SGI InfiniteStorage* and the platform's operating system and installation documentation.

### 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 Administration Guide for SGI InfiniteStorage* ([TPL link](#))
- AMCC (formerly 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 AMCC website at:

<http://www.amcc.com>

- QLogic HBA card and driver documentation:

- *SANblade 2x00 Series User's Guide*
- *SANsurfer Applications User's Guide*

See the QLogic website at:

<http://www.qlogic.com>

- AIX documentation on the IBM website at:

<http://www.ibm.com>

- Linux third-party platform documentation:

- *The Official Red Hat Linux x86 Installation Guide*
- *The Official Red Hat Linux Reference Guide*

- Mac OS X software documentation:

- *Welcome to Mac OS X*
- *Mac OS X Server Administrator's Guide*
- *Understanding and Using NetInfo*

See the Apple website at:

<http://www.apple.com>

- Solaris documentation:

- *Solaris 8 Installation Guide*
- *Solaris 8 System Administration Collection*
- *Solaris 8 Advanced Installation Guide*
- *Solaris 9 Installation Guide*
- *Solaris 9 System Administration Collection*

See the Sun Microsystems website at:

<http://www.sun.com>

- Sun Microsystems owner's guide and product notes for the Sun hardware platform
- Windows software documentation: see the Microsoft website at:

<http://www.microsoft.com>

- Hardware documentation for the Intel platform
- *Flexible License Manager End User Manual* from Macrovision Corporation.

## Obtaining Publications

You can obtain SGI documentation as follows:

- See the SGI Technical Publications Library at <http://docs.sgi.com>. Various formats are available. This library contains the most recent and most comprehensive set of online books, release notes, man pages, and other information.
- If it is installed on your IRIX SGI system, you can use InfoSearch, an online tool that provides a more limited set of online books, release notes, and man pages. On an IRIX system, enter `infosearch` at a command line or select **Help > InfoSearch** from the Toolchest.

- You can view the release notes as follows:

- On IRIX systems, use either `grelnotes` or `relnotes`
- On Linux third-party systems, see:

– `/cdrom/linux-IA32/RedHatAS3/README_CXFS_LINUX32_3.3.0.0.txt`

– `/cdrom/linux-IA32/RedHatLinux9/README_CXFS_LINUX32_3.3.0.0.txt`

- On Mac OS X systems, see `/usr/cluster/docs/ReadMe.html`
- On Solaris and AIX systems, see `/usr/cxfs_cluster/doc/relnotes`
- On Windows systems, see `%ProgramFiles%\CXFS\relnotes.rtf`. You can access this by selecting:

- Start**
- > **Programs**
- > **CXFS**
- > **Release Notes**

- On all but Windows systems, you can view man pages by typing `man title` at a command line.

## Conventions

This guide uses the following terminology abbreviations:

- *Solaris* to Solaris 8 and Solaris 9
- *Windows* to refer to Microsoft Windows 2000, Microsoft Windows 2003, and Microsoft Windows XP
- *Linux* used alone refers to the Linux operating system running on third-party hardware

The following conventions are used throughout this document:

<b>Convention</b>	<b>Meaning</b>
<code>command</code>	This fixed-space font denotes literal items such as commands, files, routines, path names, signals, messages, and programming language structures.
<i>variable</i>	Italic typeface denotes variable entries and words or concepts being defined.
<b>user input</b>	This bold, fixed-space font denotes literal items that the user enters in interactive sessions. (Output is shown in nonbold, fixed-space font.)
<b>GUI</b>	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 publication, contact SGI. Be sure to include the title and document number of the publication with your comments. (Online, the document number is located in the front matter of the publication. In printed publications, the document number is located at the bottom of each page.)

You can contact SGI in any of the following ways:

- Send e-mail to the following address:  
techpubs@sgi.com
- Use the Feedback option on the Technical Publications Library Web page:  
<http://docs.sgi.com>
- Contact your customer service representative and ask that an incident be filed in the SGI incident tracking system.
- Send mail to the following address:  
Technical Publications  
SGI  
1500 Crittenden Lane, M/S 535  
Mountain View, California 94043-1351

SGI values your comments and will respond to them promptly.



## Introduction

This guide provides an overview of the installation and configuration procedures for CXFS client-only nodes running SGI CXFS clustered filesystems. A *CXFS client-only node* has a minimal implementation of CXFS services that run a single daemon, the CXFS client daemon (`cxfs_client`).

A cluster running multiple operating systems is known as a *multiOS cluster*. It will contain potential CXFS metadata servers and client-only nodes. The following nodes are always client-only nodes:

- AIX
- Mac OS X
- Red Hat Enterprise Linux
- Sun Microsystems Solaris
- Microsoft Windows 2000, Microsoft Windows Server 2003, and Microsoft Windows XP

IRIX nodes and SGI ProPack for Linux nodes can be either server-capable administration nodes or client-only nodes. For information about IRIX and SGI ProPack for Linux nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*.

---

**Note:** This guide uses the following terminology abbreviations:

- *Solaris* refers to Solaris 8 and Solaris 9
  - *Windows* refers to Windows 2000, Windows Server 2003, and Windows XP
  - *Linux* refers to the Linux operating system running on third-party (as opposed to SGI) hardware
- 

For details, see the release notes and the platform-specific chapters in this guide.

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 Administration Guide for SGI InfiniteStorage*.

This chapter discusses the following:

- "When to Use CXFS" on page 2
- "CXFS on Client-Only Nodes" on page 3
- "Licenses" on page 12
- "Hostname Resolution and Network Configuration Rules" on page 13
- "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14
- "XVM Failover Version 2 and CXFS" on page 17
- "Monitoring CXFS" on page 18

## 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 read/write operations 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 read/write operations for multiple processes on multiple hosts.
- Large files and file accesses are being used.

Applications that perform well on a client typically do the following:

- Issue large I/O requests, rather than several smaller requests
- Use asynchronous or multithreaded I/O to have several I/O requests in flight at the same time
- Minimize the number of metadata operations they perform

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" on page 3
- "Cluster Administration" on page 4
- "User Administration for CXFS" on page 4
- "Performance Considerations" on page 4
- "User and Group Quotas " on page 6
- "CXFS Mount Scripts" on page 6
- "Requirements" on page 7
- "Recommendations" on page 9

## CXFS Processes

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

## Cluster Administration

There must be at least one server-capable administration node in the cluster that is responsible for updating that filesystem's metadata. This node is referred to as the *CXFS metadata server*. (Client-only nodes cannot be metadata servers.) Metadata servers store information in the CXFS cluster database. The CXFS cluster database is not stored on client-only nodes; only administration nodes contain the cluster database.

An administration node is required to perform administrative tasks, using either the `cmgr` command or the CXFS graphical user interface (GUI). For more information about using these tools, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## User Administration for CXFS

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

The following must be observed to achieve this consistency:

- Users must have the same usernames on all nodes in the cluster. An individual user identifier (UID) should not be used by two different people anywhere in the cluster. Ideally, group names and group identifiers (GIDs) should also be consistent on all nodes in the cluster.
- Each CXFS client and server node must have access to the same UID and GID information. The simplest way to achieve this is to maintain the same `/etc/passwd` and `/etc/group` files on all CXFS nodes, but other mechanisms may be supported.

## Performance Considerations

CXFS may not give optimal performance under the following circumstances:

- When distributed applications write to shared files that are memory mapped.
- When you are using NFS, Samba, or CIFS to export a CXFS filesystem from a CXFS client. Performance will be much better when the export is performed from the active CXFS metadata server than when it is performed from a CXFS client node. If a client is used, SGI will only support an NFS or Samba export from an IRIX or SGI ProPack for Linux client.

- When extending large highly fragmented files. The metadata traffic when growing files with a large number of extents will increase as more extents are added to the file. The following I/O patterns will cause highly fragmented files:
  - Random writes to sparse files
  - Files generated with memory mapped I/O
  - Writing files in an order other than linearly from beginning to end

Do the following to prevent highly fragmented files:

- Create files with linear I/O from beginning to end
  - Use file preallocation to allocate space for a file before writing
  - Create filesystems with sparse files disabled (`unwritten=0`)
- 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, renaming, 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.

Also see "Functional Limitations for Windows" on page 153.

## User and Group Quotas

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

To view or edit your quota information, you must log in to a CXFS administration node and make any necessary changes. If you want to provide a viewing command on the client-only node, such as `repquota`, you can construct a 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
```

## CXFS Mount Scripts

---

**Note:** NFS and Samba exports of CXFS filesystems are only supported from IRIX and SGI ProPack for Linux nodes.

---

Scripts are provided for execution prior to and after a CXFS filesystem is mounted or unmounted on the following platforms:

- AIX
- Linux
- Solaris

The scripts are not supported on Mac OS X or Windows.

These scripts are installed in the following locations:

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

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

These scripts are used by CXFS to ensure that LUN path failover works after fencing.

These scripts can be customized to suit a particular environment. For example, an application could be started when a CXFS filesystem is mounted by extending the `cxfs-post-mount` script. The application could be terminated by changing the `cxfs-pre-umount` script.

The following script is run by `cxfs_client` when it reprobes the Fibre Channel controllers upon joining or rejoining membership:

```
/var/cluster/cxfs_client-scripts/cxfs-reprobe
```

For Linux nodes, you must define a group of environment variables in the `/etc/cluster/config/cxfs_client.options` file in order for `cxfs-reprobe` to appropriately probe all of the targets on the SCSI bus. For more information, see "Using `cxfs-reprobe` with Linux " on page 69.

For information about using these scripts, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Requirements

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

- A supported storage area network (SAN) hardware configuration.

---

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

---

- At least one server-capable administration node to act as the metadata server and from which to perform cluster administration tasks. CXFS should be installed on the administration node before CXFS is installed on the client-only nodes.
- A private 100baseT (or greater) TCP/IP network connected to each node, to be dedicated to the CXFS private heartbeat and control network. This network must not be a virtual local area network (VLAN) and the Ethernet switch must not connect to other networks. All nodes must be configured to use the same subnet.

SGI recommends that 10.x.x.x 176.16.x.x or 192.168.x.x network addresses are used (RFC 1918).

- The appropriate licenses. See "Licenses" on page 12.
- A switch, which is required to protect data integrity on nodes without system controllers. See the release notes for supported switches.

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

I/O fencing differs from zoning:

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

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

---

AIX, Linux, Solaris, Mac OS X, and Windows client-only nodes must use I/O fencing to protect the data integrity of the filesystems in the cluster. Potential metadata servers should use serial reset lines.

- At least one administration node must be a server-capable administration node in order to be a CXFS metadata server; other nodes should be CXFS client-only nodes. A cluster may contain as many as 64 nodes, of which as many as 16 can be administration nodes; the rest must be client-only nodes.

A cluster in which both CXFS and FailSafe 2.1 or later are run (known as *coexecution*) is supported with a maximum of 64 nodes, as many as 8 of which can run FailSafe. FailSafe runs on IRIX nodes only.

- No nodes within the cluster running Trusted IRIX. A multiOS cluster cannot contain Trusted IRIX nodes.
- Ensure that the appropriate software is installed on IRIX potential metadata server nodes. For example, if you want to use quotas and access control lists (ACLs) on any cluster node with IRIX metadata servers, the `eo.e.sw.quotas`, `nfs.sw.acl_nfs`, and `eo.e.sw.acl` subsystems must be installed on the administration nodes listed as potential metadata servers. Likewise, if using guaranteed-rate I/O (GRIO) version 2 in the cluster, ensure that `eo.e.sw.grio2` and `cxfs.sw.grio2_cell` are installed on all IRIX nodes in the cluster.

For more information, see the following:

- *IRIX Admin: Disks and Filesystems*
- *IRIX Admin: Backup, Security and Accounting*
- *Guaranteed-Rate I/O Version 2 Guide*
- Your site's IRIX system administrator
- Set the `mtcp_nodelay` system tunable parameter to 1 on potential metadata servers in order to provide adequate performance on file deletes.

Also see "Requirements to Solaris" on page 104 and "Requirements for Windows" on page 147.

## Recommendations

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

- Fix SAN connectivity and network issues before trying to use CXFS.
- 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.
- Configure a cluster with an odd number of server-capable administration nodes.

- For large clusters, SGI recommends that you define only the first three server-capable administration nodes and then continue on with the steps to define the cluster. After you have a successful small cluster, go back and add the remaining nodes.
- Launch any task initiated using `cron` on a CXFS filesystem from a single node in the cluster, preferably from the active metadata server.

The `cron` 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` 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 other supported operating systems.
- Be very careful when running `xfs_repair` 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 Administration Guide for SGI InfiniteStorage*.
- Only those nodes that you want to be potential metadata servers should be **CXFS administration** nodes (installed with the `cluster_admin` software product). **CXFS client administration** nodes should only be used when necessary for coexecution with IRIS FailSafe. All other nodes should be **client-only** nodes (installed with `cxfs_client`). Use an odd number of server-capable administration nodes.
- Disable CXFS before maintenance (perform a forced CXFS shutdown, stop the `cxfs_client` daemon, and disable `cxfs_client` from automatically restarting).
- In this release, relocation is disabled by default and recovery is supported only when using standby nodes.

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

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

- When shutting down, resetting, or restarting a CXFS client-only node, do not stop CXFS services on the node. (Stopping CXFS services is more intrusive on other nodes in the cluster because it updates the cluster database. Stopping CXFS services is appropriate only for a CXFS administration node.) Rather, let the CXFS shutdown scripts on the node stop CXFS when the client-only node is shut down or restarted.
- Do not run power management software, which may interfere with the CXFS cluster.
- Enable the *forced unmount* feature for CXFS filesystems, which is turned off by default. Normally, an unmount operation will fail if any process has an open file on the filesystem. However, a forced unmount allows the unmount to proceed regardless of whether the filesystem is still in use.

Many sites have found that enabling this feature improves the stability of their CXFS cluster, particularly in situations where the filesystem must be unmounted. For more information, see "Forced Unmount of CXFS Filesystems" on page 221.

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 AIX nodes, a similar function is performed with the `fuser -k` command and `umount` command.
- On Linux nodes, a similar function is performed with the `fuser -m -k` command and the `umount` command.
- On Mac OS X nodes, applications will not be terminated but access to open files on the unmounted filesystem will result in I/O errors.
- On Solaris nodes, a similar function is performed with the `fuser -c -k` command and the `umount -f` command.

- On Windows nodes, all processes with open files on the CXFS filesystem are killed. For details, see "Forced Unmount on Windows" on page 155.

This feature is available on an administration node with the following CXFS GUI menu:

**Tasks**

> **Filesystems**

> **Unmount a Filesystem**

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

You can also specify this feature using the `cmgr(1M)` commands to define the filesystem. For more information, see "Forced Unmount of CXFS Filesystems" on page 221.

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

For more information, see the *CXFS Administration Guide for SGI InfiniteStorage*, and the `fuser` and `umount` man pages.

- Enable system dumps; see the operating system documentation for instructions.

## Licenses

CXFS requires the following licenses:

- Fibre Channel switch license. See the release notes.
- CXFS FLEXlm license installed on every node in the cluster.
- XVM FLEXlm license (only on Linux nodes).
- XVM FLEXlm mirroring license (only for nodes that will access a mirrored volume). For information about purchasing this license, see your SGI sales representative.

To request software keys or information about software licensing, see the following web page:

<http://www.sgi.com/support/licensing>

If you do not have access to the web, please contact your local Customer Support Center.

When you order CXFS, you will receive an entitlement identifier (ID). You must submit the system host ID, host name, and entitlement ID when requesting your permanent CXFS license. The procedures vary by platform.

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.

For platform-specific instructions, see the following:

- "FLEXlm License for AIX" on page 25
- "FLEXlm License for Linux" on page 52
- "FLEXlm License for Mac OS X" on page 85
- "FLEXlm License for Solaris" on page 110
- "FLEXlm License for Windows" on page 175

To verify that the licenses have been installed properly, use the `cxfslicense` command after installing the CXFS software. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more information on FLEXlm, you may order the *Flexible License Manager End User Manual* from Macrovision Corporation.

## Hostname Resolution and Network Configuration Rules

---



**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 all nodes:

- The first node you define (using the CXFS GUI or `cmgr`) must be an administration node.
- 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 CXFS 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 administration node (which creates the cluster database), you must recreate the cluster database.

Use the `cxfs-config -check -ping` command line on an administration node to confirm network connectivity. For more information, see *CXFS Administration Guide for SGI InfiniteStorage*.

## Guaranteed-Rate I/O (GRIO) and CXFS

CXFS supports guaranteed-rate I/O (GRIO) version 2 on all platforms. GRIO does the following:

- Enables a cluster node or a user application to reserve part of a filesystem's I/O resources for its exclusive use
- Guarantees delivery of data from a storage device at a predefined rate, regardless of any other I/O activity on the system or on other nodes in the cluster
- Ensures that the rate at which a process issues I/O does not exceed its guarantee and will throttle the I/O if necessary

GRIO includes the following features:

- Support for CXFS filesystems shared among nodes in a cluster as well as locally attached XFS filesystems
- A simple filesystem-level performance qualification model
- A range of tools for monitoring and measuring delivered bandwidth and I/O service time

GRIO is disabled by default on Linux. See "GRIO on Linux" on page 71.

See the platform-specific chapters in this guide for GRIO limitations and considerations:

- "GRIO on AIX" on page 38
- "GRIO on Linux" on page 71
- "GRIO on Mac OS X" on page 97
- "GRIO on Solaris" on page 137
- "GRIO on Windows" on page 201

For details about GRIO installation, configuration, and use, see the *Guaranteed-Rate I/O Version 2 Guide*.

## GRIO Commands

Once installed in a cluster, the following commands can be run from any node in the cluster as a superuser:

- `grioadmin` provides stream and bandwidth management
- `griogps` is the comprehensive stream quality-of-service monitoring tool

Run any of the above tools with the `-h` (help) option for a full description of all available options. See Appendix A, "Operating System Path Differences" on page 239 for the platform-specific locations of these tools.

## GRIO Examples

The following sections show example GRIO commands. The examples are taken from a cluster with the following GRIO-managed filesystems:

```
highbw  
lowbw  
largefs
```

### List GRIO-Managed Volumes

To list the GRIO-managed volumes:

```
grioadmin -l
```

For example:

```
irix# grioadmin -l
/mnt/highbw DYNAMIC 001e88dd-77a4-1027-834c-0800690d75f7 bytes: 65536 msecs: 1000
/mnt/lowbw DYNAMIC 9eb10270-7871-1027-8482-0800690d75f7 bytes: 65536 msecs: 1000
/mnt/largefs DYNAMIC de7f81d5-76af-1027-8cfa-0800690d8e03 bytes: 65536 msecs: 1000
```

### Monitor a Particular Stream

To monitor a particular GRIO stream, use the following command:

```
griogqs -I intervals stream [delay|count]
```

To get a usage statement, use the `-h` option.

For example, to monitor the NONGRIO stream associated with `highbw` filesystem:

```
irix# griogqs -I 1000ms 001e88dd-77a4-1027-834c-0800690d75f7 1
```

```
IRIX64 irix 6.5 04070317 IP27          04/06/05
```

```
Filesystem: /mnt/highbw
```

```
Dynamic    0.06 MB/s 001e88dd-77a4-1027-834c-0800690d75f7
Static     None
```

-	interval	minbw	maxbw	lastbw	minio	maxio	lastio
-	-	MB/s	MB/s	MB/s	ms	ms	ms
10:28:38	1000ms	0.00	0.20	0.19	0.3	3.3	1.6
10:28:39	1000ms	0.00	0.20	0.19	0.3	3.3	1.6
10:28:40	1000ms	0.00	0.20	0.19	0.3	3.3	1.6

To terminate `griogqs`, press `Ctrl-C` on Windows or send a `SIGINT` on other platforms.

## XVM Failover Version 2 and CXFS

CXFS supports XVM failover version 2 on the following platforms:

- IRIX
- SGI ProPack for Linux
- AIX
- Linux third-party
- Mac OS X
- Solaris

---

**Note:** CXFS supports XVM failover version 1 and version 2 on the IRIX and SGI ProPack for Linux platforms. Version 1 is not supported on the other CXFS platforms.

---

Failover version 2 requires that the RAID be configured in AVT mode. AIX also requires a SANshare license; see "SANshare License and XVM Version 2 Failover for AIX" on page 27.

To configure failover version 2, you must create and edit the `/etc/failover2.conf` file. The easiest method to generate a default `failover2.conf` file is to run the following command:

```
xvm show -v phys | fgrep affinity
```

For more information, see the comments in the `/etc/failover2.conf` file on a CXFS administration node and the *XVM Volume Manager Administrator's Guide*.

For platform-specific examples of `/etc/failover2.conf`, see the *CXFS Administration Guide for SGI InfiniteStorage* and the following section in this guide:

- "XVM Failover Version 2 on AIX" on page 38
- "XVM Failover Version 2 on Linux" on page 72
- "XVM Failover Version 2 on Mac OS X" on page 98
- "XVM Failover Version 2 on Solaris" on page 137

## Monitoring CXFS

To monitor CXFS, you can use the `cxfs_info` command on the client or view area of the CXFS GUI or the `clconf_info` command on a CXFS administration node. For more information, see "Verifying the Cluster Status" on page 223.

## AIX Platform

CXFS supports a client-only node running the AIX operating system. This chapter contains the following sections:

- "CXFS on AIX" on page 19
- "Installation and Configuration Overview for AIX" on page 25
- "FLEXlm License for AIX" on page 25
- "SANshare License and XVM Version 2 Failover for AIX" on page 27
- "HBA Installation for AIX" on page 28
- "Preinstallation Steps for AIX" on page 28
- "Client Software Installation for AIX" on page 31
- "I/O Fencing for AIX" on page 35
- "Start/Stop `cxfs_client` Daemon for AIX" on page 36
- "Maintenance for AIX" on page 37
- "GRIO on AIX" on page 38
- "XVM Failover Version 2 on AIX" on page 38
- "Troubleshooting for AIX" on page 39
- "Reporting AIX Problems" on page 44

### CXFS on AIX

This section contains the following information about CXFS on AIX:

- "Requirements for AIX" on page 20
- "CXFS Commands on AIX" on page 21
- "Log Files on AIX" on page 21
- "CXFS Mount Scripts on AIX" on page 22

- "Limitations and Considerations for AIX" on page 22
- "Access Control Lists and AIX " on page 24

## Requirements for AIX

In addition to the items listed in "Requirements" on page 7, using an AIX node to support CXFS requires the following:

- IBM AIX 5L:
  - Version 5.1 Maintenance Level 5 (64-bit mode) APAR number IY48486
  - Version 5.2 Maintenance Level 3 (64-bit mode) APAR number IY49013

To verify the operating system level, use the following command:

```
oslevel -r
```

- IBM FC 6228 or FC 6239 2-Gbit Fibre Channel host bus adapters (HBAs)
- One or more of the following IBM hardware platforms:
  - pSeries 610
  - pSeries 620
  - pSeries 630
  - pSeries 640
  - pSeries 650
  - pSeries 660
  - pSeries 670
  - pSeries 680
  - pSeries 690

---

**Note:** This release was tested with the following:

- pSeries 610 Model 6E1
  - pSeries 630 Model 6E4
-

For the latest information, see the CXFS AIX release notes.

## CXFS Commands on AIX

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

```
/usr/cxfs_cluster/bin/cxfs_client  
/usr/cxfs_cluster/bin/cxfs_info  
/usr/cxfs_cluster/bin/cxfslicense  
/usr/cxfs_cluster/bin/grioadmin  
/usr/cxfs_cluster/bin/grioqos  
/usr/cxfs_cluster/bin/xvm
```

The `cxfs_client` and `xvm` commands are needed to include an AIX node in a CXFS cluster. The `cxfs_info` command reports the current status of this node in the CXFS cluster.

You can use the `cxfslicense` command to verify a CXFS or XVM FLEXlm license.

The `lslpp` output lists all of the software added; see "Installing CXFS Software on AIX" on page 32.

For information about the GRIO commands, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and "GRIO on AIX" on page 38.

For more information on these commands, see the man pages.

## Log Files on AIX

The `cxfs_client` command creates a `/var/tmp/cxfs_client` log file. To rotate this log file, use the `-z` option in the `/usr/cxfs_cluster/bin/cxfs_client.options` file; see the `cxfs_client.options` man page for details.

Some daemons working in the kernel related to CXFS output a message in the console log, which is rotated. To see the contents of this log file, use the following command:

```
alog -o -t console
```

For information about the log files created on administration nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## CXFS Mount Scripts on AIX

AIX supports the CXFS mount scripts. See "CXFS Mount Scripts" on page 6 and the *CXFS Administration Guide for SGI InfiniteStorage*.

## Limitations and Considerations for AIX

Note the following:

- 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 JFS or NFS filesystem on top of an AIX CXFS filesystem, this is not recommended.
- There is no default access control list (ACL) in AIX. Therefore, the setup and display of the default ACL cannot be completed using the following commands:

```
aclget  
aclput  
acledit
```

If an IRIX ACL exists, the ACL becomes effective when the default ACL is set up by IRIX and a file and a directory are made under that directory in AIX.

- There is no MASK entry in AIX, but the access permissions in AIX follow those established when an ACL set up by IRIX contains a MASK entry. If the default ACL is set up for a given directory and the MASK entry exists, then that MASK entry is used when a file or a subdirectory is made by AIX. When the MASK entry does not exist, `rxw` is used.
- The ACL control of the following, which the AIX JFS filesystem has, cannot be applied to CXFS:
  - The access to a certain user or the group is rejected (`deny`)
  - When a user belongs to the specific group, access is permitted or rejected (`specify`)

If `deny` or `specify` is used, an error occurs (`EINVAL`) because these features are not in IRIX.

- Socket files cannot be copied. The following error is returned:

```
AIX:The socket does not allow the requested operation.
```

- You can use the `fuser` command to extract process information about the mounted filesystem, but you cannot extract process information about the file or the directory.
- When a CXFS mount is performed on a mirror volume created by CXFS XVM, the AIX system goes into panic status. The mirror volume cannot be mounted on the AIX CXFS system.
- The AIX node does not automatically detect the worldwide port number (WWPN). In order to use I/O fencing, you must list the WWPN in the `/etc/fencing.conf` file. See "Licenses" on page 12.
- If your users want to use a file size/offset maximum greater than 1 GB, you must change their user properties to allow files of unlimited size. To do this, use the `smit` command. For more information, see the `smit` man page.
- By default, the maximum request size for direct I/O is 512 MB (524288 KB). A direct I/O request larger than 512 MB will revert to buffered I/O. However, you can change the maximum XVM direct memory access (DMA) size to improve direct I/O performance. To do this, use the `chdev` command to modify the `xvm_maxdmasz` attribute. The actual maximum limit will always be 4 KB less than any of the supplied or displayed values (for example, the default is actually 512 MB minus 4 KB).

---

**Note:** The XVM module must be loaded if any attribute changes are to be noticed and applied.

---

To display the current setting, use the following command:

```
lsattr -E -l xvm -a xvm_maxdmasz
```

To change the current setting, use the following command:

```
chdev [-P|-T] -l xvm -a xvm_maxdmasz=NewValue
```

Legal values for *NewValue* are specified in KB units in the range 256 to 209715 (that is, 256 KB to 2 GB).

By default, using `chdev` on a running system makes a permanent change for subsequently mounted filesystems. (Running filesystems will not be changed until they are remounted, either manually or after a reboot.)

If you use `-P`, the change is deferred until the next boot and after that it is permanent. If you use `-T` (temporary), the change is immediate for subsequently mounted filesystems, but lasts only until the next boot.

For example, to change the DMA size to 2 GB for subsequently mounted filesystems on the currently running device and in the database, enter the following:

```
aix# chdev -l xvm -a xvm_maxdmasz=209715
```

For more information, see the `lsattr` and `chdev` man pages.

See also Appendix B, "Filesystem and Logical Unit Specifications" on page 245.

## Access Control Lists and AIX

All CXFS files have UNIX mode bits (read, write, and execute) and optionally an ACL. For more information, see the AIX `chmod`, `acledit`, `aclget`, and `aclput` man pages.

If you want to use an AIX node to restore a CXFS file with an ACL, you should use the `backup` and `restore` commands. If you use the `tar`, `cpio`, or `pax` command, the ACL will not be used because these tools behave "intelligently" by not calling `acl` subroutines to set an ACL. These tools will only set the file mode.

When using the `ls` command to display access permissions for a file with an ACL, the mode reported for a CXFS file follows IRIX semantics instead of AIX JFS semantics.

The IRIX model calls for reporting the ACL `MASK` for the group permission in the mode. Therefore, if the `GROUP` entry is `r-x` and the `MASK` entry is `rw-`, the group permission will be reported as `rw-`. Although it appears that the group has write permission, it does not and an attempt to write to the file will be rejected. You can obtain the real (that is, effective) group permission by using the AIX `aclget` command.

---

**Note:** Normally, AIX filesystem ACLs can have up to one memory page (4096 bytes) for a file and a directory. However, CXFS filesystems on AIX nodes in a multiOS cluster must maintain compatibility with the metadata server. The CXFS filesystems on an AIX node are limited to a maximum of 25 ACL entries converted to IRIX ACL type for a file and a directory.

---

## Installation and Configuration Overview for AIX

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

1. Read the release notes for the AIX platform to learn about any late-breaking changes in the installation procedure.
2. Install the AIX operating system according to the directions in the AIX documentation (if not already done).
3. Install and verify the RAID. See the release notes.
4. Install and verify the Fibre Channel switch. See the release notes.
5. Obtain and install the CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must purchase the mirroring software option and obtain and install a FLEXIm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See "FLEXIm License for AIX" on page 25.
6. Install and verify the host bus adapter (HBA). See "HBA Installation for AIX" on page 28.
7. Prepare the AIX node, including adding a private network. See "Preinstallation Steps for AIX" on page 28.
8. Install the CXFS software. See "Client Software Installation for AIX" on page 31.
9. Create the I/O fencing file. See "I/O Fencing for AIX" on page 35.
10. Configure the cluster to define the new AIX node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration" on page 215.

If you run into problems, see "Troubleshooting for AIX" on page 39 and Chapter 8, "General Troubleshooting" on page 231.

## FLEXIm License for AIX

This section discusses the following:

- "Gathering the AIX Host Information" on page 26

- "Obtaining the Licenses for AIX" on page 26
- "Installing the Licenses for AIX" on page 27

## Gathering the AIX Host Information

When you order CXFS, you will receive an entitlement identifier (ID). You must submit the system host ID, host name, and entitlement ID when requesting your permanent CXFS license.

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

```
/usr/bin/uname -m  
/usr/bin/hostname
```

The `uname -m` command lists the host ID as a 12-digit number. For CXFS license purposes, you will drop the first two digits and the last two digits and supply the remaining eight digits.

For example:

```
aix# /usr/bin/uname -m  
000276513100  
aix# /usr/bin/hostname  
cxfsaix3
```

When you are asked for the license manager host ID, you would provide the following information:

- Host ID: 02765131

That is, the output from the `uname -m` command minus the first two digits (00) and the last two digits (00)

- Hostname: `cxfsaix3`

You must have a separate license for each host on which CXFS is installed.

## Obtaining the Licenses for AIX

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.

## Installing the Licenses for AIX

You will install the licenses in the following location:

```
/etc/flexlm/license.dat
```

Do the following:

1. Create the license directory if necessary.

For example:

```
aix# mkdir -p /etc/flexlm
```

2. Copy the key to the `license.dat` file.

## SANshare License and XVM Version 2 Failover for AIX

AIX nodes using TP9300, TP9500, or TP9700 RAID boxes with more than two connections per command module require an Engenio SANshare license to use XVM failover version 2. AIX nodes require a primary path and one alternate path, and cannot accommodate more than two paths to a single device. To work around this issue, a SAN is used to make a single box appear to AIX as two boxes and use AVT mode.

Do the following:

1. Obtain a SANshare license from Engenio.
2. Follow the instruction in the Engenio documentation to configure the SANshare product.
3. In the LUN mapping on the RAID box, define the host type map for the worldwide node name for the group of AIX nodes as `linux` and use `SGIAVT` for a separate group for all nodes other than AIX. (Do not use the host type of AIX because that will result in RDAC mode.)

## HBA Installation for AIX

For information about installing and configuring the host bus adapter (HBA), see the IBM HBA documentation.

## Preinstallation Steps for AIX

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

- "Adding a Private Network for AIX" on page 28
- "Verifying the Private and Public Network for AIX" on page 31

## Adding a Private Network for AIX

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

1. If your system is already operational and on the network, skip to step 2. If the AIX operating system has not been installed, install it in accordance with the AIX documentation.
2. Edit the `/etc/hosts` file so that it contains entries for every node in the cluster and their private interfaces.

The `/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/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 the `/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/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/hosts` on the other nodes in the cluster.

---

**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" on page 13 and the `hosts`, `named`, and `nis` man pages.

3. (Optional) Edit the `/etc/netsvc.conf` file so that local files are accessed before either NIS or DNS. That is, the `hosts` line in `/etc/netsvc.conf` must list `local` first. For example:

```
hosts = local,nis,bind
```

(The order of `nis` and `bind` is not significant to CXFS, but `local` must be first.)

4. Determine the name of the private interface by using the `ifconfig` command as follows, to list the available networks. For example:

```
# ifconfig -l
en0 en1 lo0
```

However, if the second network interface (en1) does not appear, then the network interface must be set up in accordance with the AIX documentation.

You can set up an IP address by using `ifconfig` after restarting the system. If it is set up properly, the following information is output (line breaks added here for readability):

```
# ifconfig -a
en0: flags=4e080863<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,PSEG>
    inet 10.208.148.61 netmask 0xffffffff broadcast 10.208.148.255
en1: flags=7e080863,10<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,
    CHECKSUM_OFFLOAD,CHECKSUM_SUPPORT,RSEG>
    inet 192.168.10.61 netmask 0xffffffff broadcast 192.168.10.255
lo0: flags=e08084b<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT>
    inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
```

5. (Optional) Edit the `/.rhosts` file if you want to use remote access or if you want to use the connectivity diagnostics with CXFS. Make sure that the mode of the `/.rhosts` file is set to 600 (read and write access for the owner only).

Make sure that the `/.rhosts` file on each AIX node allows all of the nodes in the cluster to have access to each other. The connectivity tests execute a `ping` 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` as user `root`.

For example, suppose you have a cluster with three nodes: `irix0`, `aix1`, and `aix2`. The `/.rhosts` files could be as follows (where the prompt denotes the node name):

```
irix0# cat /.rhosts
aix1 root
aix2 root

aix1# cat /.rhosts
irix0 root
aix2 root

aix2# cat /.rhosts
irix0 root
aix1 root
```

## Verifying the Private and Public Network for AIX

For each private network on each AIX node in the pool, verify access with the AIX ping command. Enter the following, where *nodeIPAddress* is the IP address of the node:

```
/usr/sbin/ping -c 3 nodeIPAddress
```

For example:

```
aix# /usr/sbin/ping -c 3 192.168.10.61
PING 192.168.10.61: (192.168.10.61): 56 data data bytes
64 bytes from 192.168.10.61 icmp_seq=0 ttl=255 time=0 ms
64 bytes from 192.168.10.61 icmp_seq=1 ttl=255 time=0 ms
64 bytes from 192.168.10.61 icmp_seq=2 ttl=255 time=0 ms
----192.168.10.61 PING Statistics----
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0/0/00 ms
```

You should also execute a ping on the public networks. If that ping fails, follow these steps:

1. Verify that the network interface was configured up. For example:

```
aix# /usr/sbin/ifconfig en0
en0: flgs=4e08086<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,PSEG>
inet 10.208.148.61 netmask 0xffffffff broadcast 10.208.148.255
```

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 for AIX

The CXFS software initially will be installed and configured by SGI personnel. This section provides the following:

- "Installing CXFS Software on AIX" on page 32
- "Verifying the AIX Installation " on page 34

## Installing CXFS Software on AIX

Installing the CXFS client CD for AIX requires approximately 20 MB of space. To install the required software on an AIX node, SGI personnel will do the following:

1. Read the README file for the AIX platform to learn about any late-breaking changes in the installation procedure.
2. Verify that the node has been upgraded to AIX 5.1L according to the AIX documentation. Use the following command to display the currently installed system:

```
aix# uname -rv
1 5
```

This output indicates 1 is the revision and 5 is the version.

3. Insert the *CXFS MultiOS Client 3.3* CD-ROM.
4. Mount the CD-ROM:

```
aix# mount -v cdrfs -o ro /dev/cd0 /mnt/cdrom
```

5. Install the CXFS software (the example output below is truncated):

```
aix# installp -a -d /mnt/cdrom/aix/SGIcxfs-aix5L all
+-----+
                Pre-installation Verification...
+-----+
Verifying selections...done
Verifying requisites...done
Results...

SUCSESSES
-----
Filesets listed in this section passed pre-installation verification
and will be installed.

Selected Filesets
-----
SGIcxfs-aix5L 3.3.0.0                # CXFS CLIENT for AIX

<< End of Success Section >>
```



CXFS\_AIX license granted.

If there are errors, verify that you have obtained and installed the CXFS licenses. For more information, see "Verifying the Licenses for AIX" on page 39.

7. To start AIX CXFS services without rebooting, see "Start/Stop `cxfs_client` Daemon for AIX" on page 36. To start CXFS services automatically, reboot.

## Verifying the AIX Installation

To verify that the CXFS software has been installed properly, use the `lsllpp` command as follows:

```
aix# lsllpp -L SGIcxfs-aix5L
```

For example, the following output (showing a state of C, for "committed") indicates that the CXFS package installed properly:

```
aix# lsllpp -L SGIcxfs-aix5L
Fileset                               Level  State  Type  Description (Uninstaller)
-----
SGIcxfs-aix5L                         3.3.0.0  C     F     CXFS CLIENT for AIX
```

State codes:

- A -- Applied.
- B -- Broken.
- C -- Committed.
- O -- Obsolete. (partially migrated to newer version)
- ? -- Inconsistent State...Run `lppchk -v`.

Type codes:

- F -- Installp Fileset
- P -- Product
- C -- Component
- T -- Feature
- R -- RPM Package

## I/O Fencing for AIX

I/O fencing is required on AIX nodes in order to protect data integrity of the filesystems in the cluster. The `/etc/fencing.conf` file enumerates the worldwide port name (WWPN) for all of the host bus adapters (HBA) that will be used to mount a CXFS filesystem. The `/etc/fencing.conf` file must contain a simple list of WWPNs, as a 64-bit hexadecimal number, one per line. These HBAs will then be available for fencing.

If you want to use the `/etc/fencing.conf` file, you must update it whenever the HBA configuration changes, including the replacement of an HBA.

Do the following:

1. Follow the Fibre Channel cable on the back of the AIX host to determine the port to which it is connected in the 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` command to connect to the switch and log in as user `admin` (the password is `password` by default).
3. Execute the `switchshow` command to display the switches and their WWPNs. 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 on the AIX node and add the WWPN for the port determined in step 1. (Comment lines begin with a `#` character.) For example, if you determined that port 0 is the port connected to the switch, your `/etc/fencing.conf` file should appear as follows:

```
2000000173002c0b
```

5. After the AIX node is added to the cluster (see Chapter 7, "Cluster Configuration" on page 215), enable the fencing feature by using the CXFS GUI or `cmgr` command on a CXFS administration node. For more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Start/Stop `cxfs_client` Daemon for AIX

The `/usr/cxfs_cluster/bin/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_client` manually on your AIX node, enter the following:

```
aix# /usr/cxfs_cluster/bin/cxfs_cluster start
```

To stop `cxfs_client` manually, enter the following:

```
aix# /usr/cxfs_cluster/bin/cxfs_cluster stop
```

To stop and then start `cxfs_client` manually, enter the following:

```
aix# /usr/cxfs_cluster/bin/cxfs_cluster restart
```

## Maintenance for AIX

This section contains the following:

- "Upgrading the CXFS Software for AIX" on page 37
- "Modifying the CXFS Software for AIX" on page 37
- "Recognizing Storage Changes for AIX" on page 38

## Upgrading the CXFS Software for AIX

To upgrade the CXFS software on an AIX system, do the following:

1. Make sure that no applications on the node are accessing files on a CXFS filesystem.
2. Determine the name of the CXFS package that is installed. For example:

```
aix# lsllpp -L | grep cxfs
SGIcxfs-aix5L          3.3.0.0    C    F    CXFS CLIENT for AIX
```

3. Uninstall the old version by using the following command:

```
installp -u packagename
```

For example, given a package name of SGIcxfs-aix5L:

```
aix# installp -u SGIcxfs-aix5L
```

4. Install the new version. See "Client Software Installation for AIX" on page 31.

## Modifying the CXFS Software for AIX

You can modify the behavior of the CXFS client daemon (`/usr/cxfs_cluster/bin/cxfs_client`) by placing options in the `/usr/cxfs_cluster/bin/cxfs_client.options` file. The available options are documented in the `cxfs_client` 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.

---

## Recognizing Storage Changes for AIX

If you make changes to your storage configuration, you must rerun the HBA utilities to reprobe the storage. For more information, see the IBM HBA documentation.

## GRIO on AIX

CXFS supports guaranteed-rate I/O (GRIO) version 2 on the AIX platform.

An AIX client can mount a GRIO-managed filesystem and supports application- and node-level reservations. An AIX client will interoperate with the dynamic bandwidth allocator for all I/O outside of any reservation.

For more information, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and the *Guaranteed-Rate I/O Version 2 Guide*.

## XVM Failover Version 2 on AIX

Following is an example of the `/etc/failover2.conf` file on AIX for a configuration in which each node has dual ports connecting to separate Fibre Channel (FC) switches. Each FC switch is connected to a single RAID controller with no interconnect between the FC switches:

```
# phys/d9400_0
/dev/rhdisk136 affinity=0 preferred
/dev/rhdisk104 affinity=0
/dev/rhdisk112 affinity=1
/dev/rhdisk132 affinity=1

# phys/d9400_1
/dev/rhdisk137 affinity=1
/dev/rhdisk105 affinity=1
/dev/rhdisk113 affinity=0 preferred
/dev/rhdisk133 affinity=0

# phys/recovery
/dev/rhdisk138 affinity=0 preferred
/dev/rhdisk106 affinity=0
/dev/rhdisk114 affinity=1
/dev/rhdisk134 affinity=1
```

```
# phys/wally
/dev/rhdisk139 affinity=1
/dev/rhdisk107 affinity=1
/dev/rhdisk115 affinity=0 preferred
/dev/rhdisk135 affinity=0
```

For more information, see "XVM Failover Version 2 and CXFS" on page 17, the comments in the `/etc/failover2.conf` file on a CXFS administration node, and the *XVM Volume Manager Administrator's Guide*.

## Troubleshooting for AIX

This section discusses the following:

- "Verifying the Licenses for AIX" on page 39
- "Unable to Mount Filesystems on AIX" on page 40
- "The `cxfs_client` Daemon is Not Started on AIX" on page 42
- "Filesystems Do Not Mount on AIX" on page 43
- "Panic Occurs when Executing `cxfs_cluster` on AIX " on page 43
- "A Memory Error Occurs with `cp -p` on AIX" on page 43
- "An ACL Problem Occurs with `cp -p` on AIX" on page 43
- "Large Log Files on AIX" on page 44

For general troubleshooting information, see Chapter 8, "General Troubleshooting" on page 231 and Appendix D, "Error Messages" on page 251

## Verifying the Licenses for AIX

To verify that the licenses have been installed properly, use the `cxfslicense` command after installing the CXFS software. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example

- License installed properly:

```
aix# /usr/cxfs_cluster/bin/cxfslicense
aix# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_AIX license granted.
CXFS_AIX license granted.
```

- No license (error condition):

```
aix# /usr/cxfs_cluster/bin/cxfslicense
cxfs license check failed - use 'cxfslicense -d' for details
aix# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_AIX license granted.
Error on CXFS_AIX license check out.
checkout failed: No such feature exists
Feature:          CXFS_AIX
License path:    /etc/flexlm/license.dat
FLEXlm error:   -5,357
For further information, refer to the FLEXlm End User Manual,
available at "www.globetrotter.com".
```

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

```
Starting CXFS 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 '/usr/cxfs_cluster/bin/cxfs_cluster restart'.
```

An error such as the following example will appear in the SYSLOG file (line breaks added here for readability):

```
Jan 25 10:24:03 ncc1701:Jan 25 10:24:03 cxfs_client:
cis_main FATAL: cxfs_client failed the CXFS license check.
Use the cxfslicense command to diagnose the license problem
```

## Unable to Mount Filesystems on AIX

If `cxfs_info` reports that `cms` is up but XVM or the filesystem is in another state, then one or more mounts is still in the process of mounting or has failed to mount.

The CXFS node might not mount filesystems for the following reasons:

- The client may not be able to see all the LUNs. This is usually caused by misconfiguration of the HBA or the SAN fabric:
  - Check that the ports on the Fibre Channel switch connected to the HBA are active. Physically look at the switch to confirm the light next to the port is green, or remotely check by using the `switchShow` command.
  - Check that the HBA configuration is correct.
  - Check that the HBA can see all the LUNs for the filesystems it is mounting.
  - Check that the operating system kernel can see all the LUN devices.
  - If the RAID device has more than one LUN mapped to different controllers, ensure the node has a Fibre Channel path to all relevant controllers.
- The `cxfs_client` daemon may not be running. See "The `cxfs_client` Daemon is Not Started on AIX" on page 42.
- The filesystem may have an unsupported mount option. Check the `cxfs_client.log` for mount option errors or any other errors that are reported when attempting to mount the filesystem.
- The cluster membership (`cms`), `XVM`, or the filesystems are not up on the node. Execute the `/usr/cluster/bin/cxfs_info` command to determine the current state of `cms`, `XVM`, and the filesystems. If the node is not up for each of these, then check the `/var/log/cxfs_client` log to see what actions have failed.

Do the following:

- If `cms` is not up, check the following:
  - Is the node is configured on the administration node with the correct hostname? See "Configuring Hostnames on Mac OS X" on page 82.
  - Has the node been added to the cluster and enabled? See "Verifying the Cluster Status" on page 223.
- If `XVM` is not up, check that the HBA is active and can see the LUNs.
- If the filesystem is not up, check that one or more filesystems are configured to be mounted on this node and check the `/var/log/cxfs_client` file for mount errors.

## The `cxfs_client` Daemon is Not Started on AIX

Confirm that the `cxfs_client` is not running. The following command would list the `cxfs_client` process if it were running:

```
aix# ps -ef | grep cxfs_client
```

The `cxfs_client` daemon might not start for the following reasons:

- If the following message is output to the console, it means that the workstation is in 32-bit kernel mode:

```
CXFS works only in the 64 bit kernel mode
```

In this case, you must change to 64-bit mode as follows:

1. Link the following libraries:

```
aix# ln -fs /usr/lib/boot/unix_64 /unix
aix# ln -fs /usr/lib/boot/unix_64 /usr/lib/boot/unix
```

2. Create the boot image:

```
aix# bosboot -ad /dev/ipldevice
```

3. Reboot the system.

- The `cxfs_client` daemon might not start if the license has expired, which is indicated when the following message is output to the `/var/log/cxfs_client` file:

```
CXFS not properly licensed for this host
```

In this case, execute the following command to determine why the license check fails:

```
aix# /usr/cluster/bin/cxfslicense -d
```

If the license is valid, check the `cxfs_client` log file for other errors.

Restart `cxfs_client` as described in "Start/Stop `cxfs_client` Daemon for AIX" on page 36 and watch the `cxfs_client` log file for errors.

## Filesystems Do Not Mount on AIX

If the `/var/tmp` filesystem is full, CXFS cannot write logs to it and the CXFS filesystem will not be able to mount on the AIX node. In this case, you should clean out the `/var/tmp` filesystem.

If a disk is read from an AIX node and the following message is output, it means that the Fibre Channel switch has broken down:

```
no such device or address
```

In this case, you should restart the Fibre Channel switch.

## Panic Occurs when Executing `cxfs_cluster` on AIX

If the following message is output, then the `genkex` command does not exist:

```
genkex isn't found
```

In this case, you must install the `bos.perf.tools` fileset.

## A Memory Error Occurs with `cp -p` on AIX

If an error occurs when a file is copied with the `cp -p` command and the following message is output, there is a problem with NFS:

```
There is not enough memory available now
```

In this case, you must use maintenance level 5100-04+IY42428.

For more information, see:

<https://techsupport.services.ibm.com/server/aix.fdc>

## An ACL Problem Occurs with `cp -p` on AIX

If an ACL is not reflected when a file with an ACL is copied from JFS to CXFS using the `cp -p` command, there is a problem with the AIX software. (The ACL information for the file is indicated by the `aclget` command.)

In this case, you must use maintenance level 5100-04.

For more information, see:

<https://techsupport.services.ibm.com/server/aix.fdc>

## Large Log Files on AIX

The `/var/tmp/cxfs_client` log file may become quite large over a period of time if the verbosity level is increased. To manually rotate this log file, use the `-z` option in the `/usr/cxfs_cluster/bin/cxfs_client.options` file.

See the `cxfs_client.options` man page and "Log Files on AIX" on page 21.

## Reporting AIX Problems

When reporting a problem about a CXFS AIX node to SGI, you should retain the following information:

- Information about the AIX node system dump and system configuration:

```
aix# snap -a -o /dev/rmt0
```

- Console log:

```
aix# alog -o -t console
```

- Current syslog file
- The `/var/tmp/cxfs_client` CXFS log file
- Moduler debugger output from the `kdb` command:

- For panics or generated dumps, use the following commands and save the output:

```
aix# kdb /var/adm/ras/vmcore.xx[/unix]
(0)> stat
```

- For dumps from hangs:

```
aix# kdb /var/adm/ras/vmcore.xx[/unix]
(0)> th* (to find the slot value of the working process or thread)
(0)> sw slot_value
(0)> stat
```

- A list of the installed CXFS packages. Use the `ls1pp` command as follows:

```
aix# ls1pp -l SGICxfs-aix5L
```

- The version information of the operating system. Use the following `oslevel` commands:

```
aix# oslevel -r  
aix# oslevel -g | grep bos.64bit
```

- A list of the loaded AIX kernel extensions. Use the `genkex` command.

If any of these AIX tools are not currently installed on your AIX node, you should install them.



## Linux Third-Party Platforms

CXFS supports a client-only node running the Linux operating system on supported third-party platforms.

---

**Note:** The term *Linux* in this guide always refers to Linux client-only nodes on third-party platforms. For information about SGI ProPack for Linux and CXFS, see *CXFS Administration Guide for SGI InfiniteStorage*.

---

On Linux systems, the use of XVM is supported only with CXFS; XVM does not support local Linux disk volumes.

This chapter contains the following sections:

- "CXFS on Linux" on page 48
- "Installation and Configuration Overview for Linux" on page 51
- "FLEXlm License for Linux" on page 52
- "HBA Installation for Linux" on page 55
- "Preinstallation Steps for Linux" on page 57
- "Client Software Installation for Linux" on page 61
- "I/O Fencing for Linux" on page 65
- "Start/Stop `cxfs_client` for Linux" on page 67
- "Maintenance for Linux" on page 69
- "Using `cxfs-reprobe` with Linux " on page 69
- "GRIO on Linux" on page 71
- "XVM Failover Version 2 on Linux" on page 72
- "Troubleshooting for Linux" on page 72
- "Reporting Linux Problems" on page 76

## CXFS on Linux

This section contains the following information about CXFS on Linux systems:

- "Requirements for Linux"
- "CXFS Commands on Linux" on page 49
- "Log Files on Linux" on page 49
- "CXFS Mount Scripts on Linux" on page 50
- "Limitations and Considerations for Linux" on page 50
- "Access Control Lists and Linux" on page 51

## Requirements for Linux

In addition to the items listed in "Requirements" on page 7, using a Linux node to support CXFS requires the following:

- Red Hat Enterprise Linux 3



**Caution:** You **must** update the operating system with all security fixes, bug fixes, and enhancements available from Red Hat.

---

- A choice of at least one Fibre Channel host bus adapter (HBA):
  - Qlogic QLA2200, QLA2310, QLA2342, or QLA2344
  - LSI Logic LS17202XP-LC dual channel PCI-X HBA
- A CPU of the following class:
  - x86, such as:
    - Advanced Micro Devices AMD Athlon
    - AMD Duron
    - Intel Pentium 4
  - AMD64/EM64T, such as:
    - AMD Opteron

- Intel Xeon EM64T
- Intel Itanium 2

The machine must have at least the following minimum requirements:

- 256 MB of RAM memory
- Two Ethernet 100baseT interfaces
- One empty PCI slot (to receive the HBA)

For the latest information, see the CXFS Linux release notes.

## CXFS Commands on Linux

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

```
/usr/cluster/bin/cxfs-config  
/usr/cluster/bin/cxfs_client  
/usr/cluster/bin/cxfs_info  
/usr/cluster/bin/cxfsdump  
/usr/cluster/bin/cxfslicense  
/usr/sbin/grioadmin  
/usr/sbin/griooqs  
/sbin/xvm
```

The rpm output lists all software added; see "Linux Installation Overview" on page 61.

For more information, see the man pages.

## Log Files on Linux

The `cxfs_client` command creates a `/var/log/cxfs_client` log file. This file is rotated by default.

The Linux platform uses the `logrotate` system utility to rotate the CXFS logs (as opposed to other multiOS platforms, which use the `-z` option to `cxfs_client`):

- The `/etc/logrotate.conf` file specifies how often system logs are rotated
- The `/etc/logrotate.d/cxfs_client` file specifies the manner in which `cxfs_client` logs are rotated

For information about the log files created on CXFS administration nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## CXFS Mount Scripts on Linux

Linux supports the CXFS mount scripts. See "CXFS Mount Scripts" on page 6 and the *CXFS Administration Guide for SGI InfiniteStorage*.

For Linux nodes, you must define a group of environment variables in the `/etc/cluster/config/cxfs_client.options` file in order for `cxfs-reprobe` to appropriately probe all of the targets on the SCSI bus. For more information, see "Using `cxfs-reprobe` with Linux " on page 69.

## Limitations and Considerations for Linux

Note the following:

- 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 other filesystems on top of a Linux CXFS filesystem, this is not recommended.
- On Linux systems, the `mkfs.xfs` command does not discover log or realtime subvolumes. You must specify the log or realtime subvolumes on the command line. For more information, see the `mkfs.xfs(8)` man page.
- Due to Linux kernel limitations, CXFS filesystems cannot be mounted with the `inode64` mount option. For more information, see Appendix C, "Mount Options Support" on page 247.
- CXFS filesystems with XFS version 1 directory format cannot be mounted on Linux nodes.
- By default, the Linux kernel will only scan LUN 0 of a SCSI device. This can be altered by adding `max_scsi_luns=N` to the kernel boot arguments, where *N* is the number of LUNs that should be scanned. If not all devices on the fabric are found, this may resolve the issue.
- The implementation of file creates using `O_EXCL` is not complete. Multiple applications running on the same node using `O_EXCL` creates as a synchronization mechanism will see the expected behavior (only one of the creates will succeed). However, applications running between nodes will not get the `O_EXCL` behavior

they requested (creates of the same file from two or more separate nodes will all succeed).

- The Fibre Channel HBA driver must be loaded before CXFS services are started. The HBA driver could be loaded early in the initialization scripts, or added to the initial ramdisk for the kernel.

See the `mkinitrd` man page for more information.

See also Appendix B, "Filesystem and Logical Unit Specifications" on page 245.

## Access Control Lists and Linux

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

Red Hat Linux file utilities do not provide ACL support, so commands such as `ls` and `cp` will not show or preserve ACLs. However, the commands in the `acl` package will allow manipulation of the ACLs of files on CXFS filesystems.

## Installation and Configuration Overview for Linux

Following is the order of installation and configuration steps for a CXFS Linux node on supported platforms:

1. Read the release notes for the Linux platform to learn about any late-breaking changes in the installation procedure.
2. Install the Red Hat operating system according to the directions in the Red Hat documentation (if not already done).
3. Install and verify the RAID. See the release notes.
4. Install and verify the switch. See the release notes.
5. Obtain and install the XVM and CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must also purchase a mirroring software option and obtain and install a FLEXlm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See "FLEXlm License for Linux" on page 52.

6. Install and verify the host bus adapter (HBA). See "HBA Installation for Linux" on page 55.
7. Prepare the Linux node, including adding a private network. See "Adding a Private Network for Linux" on page 57.
8. Install the CXFS kernel and user space software packages. See "Client Software Installation for Linux" on page 61.
9. Configure the cluster to define the new Linux node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration" on page 215.

If you run into problems, see "Troubleshooting for Linux" on page 72 and Chapter 8, "General Troubleshooting" on page 231.

## FLEXlm License for Linux

This section discusses the following:

- "Gathering the Linux Host Information" on page 52
- "Obtaining the Licenses for Linux" on page 55
- "Installing the Licenses for Linux" on page 55

### Gathering the Linux Host Information

When you order CXFS, you will receive an entitlement identifier (ID). You must submit the system host ID, host name, and entitlement ID when requesting your permanent CXFS license.

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

```
/sbin/ifconfig eth0  
/bin/hostname
```

The host ID is the value for the `HWaddr` field, minus the colons.

For example:

```
[root@linux root]# /bin/hostname
cxfslinux
[root@linux root]# /sbin/ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:E0:81:24:77:D1
          inet addr:128.162.240.135  Bcast:128.162.240.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:17941247  errors:2  dropped:0  overruns:0  frame:2
          TX packets:16373834  errors:10  dropped:0  overruns:0  carrier:10
          collisions:0  txqueuelen:100
          RX bytes:1539518033 (1468.1 Mb)  TX bytes:1573522873 (1500.6 Mb)
          Interrupt:19  Base address:0xb880  Memory:fe3fe000-fe3fe038
```

In this case, the host ID is 00E0812477D1

To obtain the physical CPU count, find the information in the `/proc/cpuinfo` file as follows:

```
[root@linux root]# cat /proc/cpuinfo
```

To determine the number of physical CPUs, count the number of processor lines and divide by the value for `siblings`. A value of 1 for `siblings` indicates that all of the processors are physical; a value of 2 indicates that the processors are hyperthreaded, and therefore the number of physical CPUs is half of the number of processor lines displayed.

For example (with truncated output indicated by `...`) :

```
[root@linux root]# cat /proc/cpuinfo
processor      : 0
vendor_id    : GenuineIntel
cpu family   : 15
model        : 2
model name   : Intel(R) Xeon(TM) CPU 2.00GHz
stepping     : 7
cpu MHz      : 1999.795
cache size   : 512 KB
physical id  : 0
siblings    : 2
...
processor     : 1
```

```
vendor_id      : GenuineIntel
cpu family    : 15
model         : 2
model name    : Intel(R) Xeon(TM) CPU 2.00GHz
stepping     : 7
cpu MHz      : 1999.795
cache size   : 512 KB
physical id  : 0
siblings     : 2
...

processor     : 2
vendor_id    : GenuineIntel
cpu family   : 15
model       : 2
model name  : Intel(R) Xeon(TM) CPU 2.00GHz
stepping   : 7
cpu MHz    : 1999.795
cache size : 512 KB
physical id : 3
siblings   : 2
...

processor     : 3
vendor_id    : GenuineIntel
cpu family   : 15
model       : 2
model name  : Intel(R) Xeon(TM) CPU 2.00GHz
stepping   : 7
cpu MHz    : 1999.795
cache size : 512 KB
physical id : 3
siblings   : 2
...
```

The above output shows that there are four processor lines, but they are hyperthreaded because the value of `siblings` is 2. Therefore, the physical CPU count is 2 ( $4/2=2$ ).

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

## Obtaining the Licenses for Linux

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.

You must obtain the CXFS\_LINUX32 and the XVM\_STD\_LINUX32 FLEXlm licenses on all x86 nodes.

You must obtain the CXFS\_IPF and the XVM\_STD\_IPF FLEXlm license on all AMD64/EM64T and Itanium 2 nodes.

## Installing the Licenses for Linux

You will install the licenses for CXFS and XVM base (and potentially XVM mirroring) at the following location:

```
/etc/flexlm/license.dat
```

Do the following:

1. Create the license directory if necessary.

For example:

```
[root@linux cdrom]# mkdir -p /etc/flexlm
```

2. Copy the key to the license.dat file.

## HBA Installation for Linux

This section provides an overview of the Fibre Channel host bus adapter (HBA) installation information for Linux nodes.

The installation may be performed by you or by a qualified service representative for your hardware. See the Red Hat documentation and the documentation for your hardware platform.

The driver requirements are as follows:

- LSI Logic card: the drivers are supplied with the Red Hat Linux kernel. The module name is `mptscsih.o`.

- QLogic card: the drivers are supplied with the Red Hat Linux kernel.

You must ensure that the HBA driver is loaded prior to CXFS initialization by building the module into the initial RAM disk automatically or manually. For example, using the QLogic card and the `qla2200` driver:

- **Automatic** method: Add a new line such as the following:

```
alias scsi_hostadapter1 qla2200
```

When the new kernel is installed, the driver will be automatically included in the corresponding `initrd` image.

- **Manual** method: enter a command similar to the following example:

```
# mkinitrd -f -v --with=qla2200 /boot/initrd-version.img version.sg1
```

Where:

`-f` overwrites any existing image of the same name

`-v` displays verbose output

`qla2200` is the name of the module to include

`/boot/initrd-version.img` is the created image

`version` is the output of the `uname -r` command for the corresponding kernel

You should then verify the appropriate `initrd` information:

- If using the LILO loader, do the following:
  1. Verify that the following line appears in the appropriate stanza of `/etc/lilo.conf`:
- If using the GRUB loader, verify that the following line appears in the `/boot/grub/grub.conf` file:

```
initrd /initrd-version.img
```

The system must be rebooted (and when using LILO, LILO must be rerun) for the new `initrd` image to take effect.

Instead of this procedure, you could also modify the `/etc/rc.sysinit` script to load the `qla2200driver` early in the `initscript` sequence.

## Preinstallation Steps for Linux

This section provides an overview of the steps that you will perform on your Linux nodes prior to installing the CXFS software. It contains the following sections:

- "Adding a Private Network for Linux" on page 57
- "Verifying the Private and Public Networks for Linux" on page 60

## Adding a Private Network for Linux

The following procedure provides an overview of the steps required to add a private network to the Linux 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 Red Hat Linux documentation.

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

The `/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/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/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/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/hosts` on the other nodes in the cluster.

For more information, see the `hosts` and `resolver` 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" on page 13.

2. 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.)

3. Configure your private interface according to the instructions in the Network Configuration section of your Linux distribution manual. To verify that the private interface is operational, issue the following command:

```
[root@linux root]# ifconfig -a

eth0      Link encap:Ethernet  HWaddr 00:50:81:A4:75:6A
          inet addr:192.168.1.1  Bcast:192.168.1.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:13782788  errors:0  dropped:0  overruns:0  frame:0
          TX packets:60846  errors:0  dropped:0  overruns:0  carrier:0
```

```
collisions:0 txqueuelen:100
RX bytes:826016878 (787.7 Mb) TX bytes:5745933 (5.4 Mb)
Interrupt:19 Base address:0xb880 Memory:fe0fe000-fe0fe038

eth1 Link encap:Ethernet HWaddr 00:81:8A:10:5C:34
inet addr:10.0.0.10 Bcast:10.0.0.255 Mask:255.255.255.0
UP BROADCAST MULTICAST MTU:1500 Metric:1
RX packets:0 errors:0 dropped:0 overruns:0 frame:0
TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:100
RX bytes:0 (0.0 b) TX bytes:0 (0.0 b)
Interrupt:19 Base address:0xef00 Memory:febfd000-febfd038

lo Link encap:Local Loopback
inet addr:127.0.0.1 Mask:255.0.0.0
UP LOOPBACK RUNNING MTU:16436 Metric:1
RX packets:162 errors:0 dropped:0 overruns:0 frame:0
TX packets:162 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:0
RX bytes:11692 (11.4 Kb) TX bytes:11692 (11.4 Kb)
```

This example shows that two ethernet interfaces, `eth0` and `eth1`, are present and running (as indicated by `UP` in the third line of each interface description).

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.

## Modifications Required for CXFS Connectivity Diagnostics for Linux

In order to test node connectivity by using the GUI or the `cmgr` command, the `root` user on the node running the CXFS diagnostics must be able to access a remote shell using the `rsh` command (as `root`) on all other nodes in the cluster. There are several ways of accomplishing this, depending on the existing settings in the pluggable authentication modules (PAMs) and other security configuration files.

Following is one possible method that works with default settings. Do the following on all nodes in the cluster:

1. Install the `rsh-server` RPM.
2. Enable `rsh`.

3. Restart `xinted`.
4. Add `rsh` to the `/etc/securetty` file.
5. Add the hostname of the node from which you will be running the diagnostics into the `/root/.rhosts` file. Make sure that the mode of the `.rhosts` file is set to 600 (read and write access for the owner only).

After you have completed running the connectivity tests, you may wish to disable `rsh` on all cluster nodes.

For more information, see the Red Hat documentation about PAM and the `hosts.equiv` man page.

## Verifying the Private and Public Networks for Linux

For each private network on each Linux node in the pool, verify access with the `ping` command. Enter the following, where `nodeIPAddress` is the IP address of the node:

```
ping nodeIPAddress
```

For example:

```
[root@linux root]# ping 10.0.0.1
PING 10.0.0.1 (10.0.0.1) from 128.162.240.141 : 56(84) bytes of data.
64 bytes from 10.0.0.1: icmp_seq=1 ttl=64 time=0.310 ms
64 bytes from 10.0.0.1: icmp_seq=2 ttl=64 time=0.122 ms
64 bytes from 10.0.0.1: icmp_seq=3 ttl=64 time=0.127 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:

```
[root@linux root]# ifconfig eth1
eth1      Link encap:Ethernet  HWaddr 00:81:8A:10:5C:34
          inet addr:10.0.0.10  Bcast:10.0.0.255  Mask:255.255.255.0
          UP BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:100
          RX bytes:0 (0.0 b)  TX bytes:0 (0.0 b)
          Interrupt:19 Base address:0xef00 Memory:febfd000-febfd038
```

In the third 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 for Linux

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.

Table 3-1 provides examples of the differences in package extensions among the various processor classes supported by CXFS.

---

**Note:** The kernel package extensions vary by architecture. Ensure that you install the appropriate package for your processor architecture.

---

**Table 3-1** Processor and Package Extension Examples

---

Class	Example Processors	User Package Extension	Kernel Package Extension
x86	AMD Athlon	.i386.rpm	.athlon.rpm
	AMD Duron	.i386.rpm	.athlon.rpm
	Intel Pentium 4	.i386.rpm	.i686.rpm
AMD64/EM64T	AMD Opteron	.x86_64.rpm	.x86_64.rpm
	Intel Xeon EM64T	.x86_64.rpm	.ia32e.rpm
Itanium 2	Intel Itanium 2	.ia64.rpm	.ia64.rpm

---

## Linux Installation Overview

---

**Note:** Specific packages listed here are examples and may not match the released product.

---

Installing the CXFS client CD for Linux requires approximately 50–200 MB of space, depending upon the packages installed at your site.

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

1. Read the README file for the Linux platform to learn about any late-breaking changes in the installation procedure.
2. Verify that the node is running a supported Linux distribution, according to the CXFS for Linux release notes. For example:

```
[root@linux root]# cat /etc/redhat-release
Red Hat Enterprise Linux AS release 3 (Taroon Update 4)
```



---

**Caution:** You **must** update the operating system with all security fixes, bug fixes, and enhancements available from Red Hat.

---

3. Insert and mount the *CXFS MultiOS Client 3.3 Kernel distribution for Linux Client* CD-ROM.
4. Install the appropriate kernel RPM for the system, according to the information about upgrading the kernel in the operating system documentation. Use the following command, using the appropriate kernel (*smp*, *hugemem*, or *nothing* for non-*smp*), version, and architecture for your system. See Table 3-1 on page 61 for a summary of the architecture types.

```
[root@linux cdrom]# rpm -ivh kernel-kernel-version-architecture.rpm
```



---

**Caution:** You should not use the *nodeps* or *force* flag during RPM installations.

---

For example, to install the kernel on an Intel Pentium 4 dual-processor machine, enter the following:

```
[root@linux cdrom]# rpm -ivh kernel-smp*.i686.rpm
Preparing...                               ##### [100%]
 1:kernel-smp                               ##### [100%]
```

5. Install the `xfs-modules` RPM. The `xfs-modules` package architecture and type must match the kernel package. See Table 3-1 on page 61 for a summary of the architecture types. For example, for an Intel Pentium 4 dual-processor machine:

```
[root@linux cdrom]# rpm -ivh xfs-modules-kernel*.i686.rpm
Preparing... ##### [100%]
 1:xfs-modules ##### [100%]
```

6. Install the `dmapi-modules` RPM. The `dmapi-modules` package architecture and type must match the kernel package. See Table 3-1 on page 61 for a summary of the architecture types. For example, for an Intel Pentium 4 dual-processor machine:

```
[root@linux cdrom]# rpm -ivh dmapi-modules-kernel*.i686.rpm
Preparing... ##### [100%]
 1:dmapi-modules ##### [100%]
```

7. Configure the new kernel into the bootloader according to the instructions in the operating system documentation and make it the default kernel:

- If you are using the LILO loader, verify that `/etc/lilo.conf` contains a stanza with the correct entries for the newly installed kernel and then run the following command to install the boot loader and verify that all information is placed appropriately:

```
[root@linux cdrom]# lilo -v
```

- If you are using the GRUB loader, verify that `/boot/grub/grub.conf` contains a stanza with the correct entries for the newly installed kernel.

8. Install the XFS filesystem packages. For example:

```
[root@linux cdrom]# rpm -Uvh dmapi* xfsprogs*
Preparing... ##### [100%]
 1:dmapi ##### [ 50%]
 2:xfsprogs ##### [100%]
```

9. Insert and mount the *CXFS MultiOS Client 3.3* CD-ROM.
10. Install the CXFS modules that correspond to the kernel installed in step 4. The `cxfs-modules` package architecture and type must match the kernel package.

See Table 3-1 on page 61 for a summary of the architecture types. For example, for an Intel Pentium 4 dual-processor machine:

```
[root@linux cdrom]# rpm -ivh cxfsm-modules-kernel*.i686.rpm
Preparing... ##### [100%]
 1:cxfsm-modules ##### [100%]
```

The following is an example of installing user-space packages:

```
[root@linux cdrom]# rpm -Uvh cxfsm_client* cxfsm_util* xvm_cmds* cxfsm-doc*
Preparing... ##### [100%]
 1:cxfsm_util ##### [25%]
 2:xvm_cmds ##### [50%]
 3:cxfsm_client ##### [75%]
 4:cxfsm-doc ##### [100%]
```

11. If you want to use quotas on a CXFS filesystem, you must install an updated quota package. For example, for an Intel Pentium 4 dual-processor machine:

```
[root@linux cdrom]# rpm -Uvh quota-*
Preparing...##### [100%]
 1:quota ##### [100%]
```

12. Edit the `/etc/cluster/config/cxfsm_client.options` file as necessary. See the "Maintenance for Linux" on page 69 and the `cxfsm_client(1M)` man page.
13. Use the `cxfsmlicense` command to verify that the license is installed in `/etc/flexlm/license.dat`. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example, the following shows a properly installed license:

```
[root@linux root]# /usr/cluster/bin/cxfsmlicense -d
XVM_STD_LINUX32 license granted.
XVM_PLEX_LINUX32 license granted.
CXFS_LINUX32 license granted.
```

If there are errors, verify that you have obtained and installed the CXFS licenses. For more information, see "Verifying the Linux Installation" on page 65.

14. Reboot the system with the newly installed kernel:

```
[root@linux root]# reboot
```

## Verifying the Linux Installation

Use the `uname -r` command to ensure the kernel installed above is running. For example:

```
[root@linux root]# uname -r  
2.4.18-27.7.x.sgismp
```

To verify that the CXFS software has been installed properly, use the `rpm` command to query the packages.

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

```
[root@linux root]# rpm -q cxfs-modules cxfs_util cxfs_client xvm-cmds  
cxfs-modules-3.3-sgi  
cxfs_util-3.3-sgi  
cxfs_client-3.3-sgi  
xvm-cmds-3.3-sgi
```

## I/O Fencing for Linux

I/O fencing is required on Linux nodes in order to protect data integrity of the filesystems in the cluster. The `cxfs_client` software automatically detects the world wide port names (WWPNs) of any supported host bus adapters (HBAs) for Linux nodes that are connected to a switch that is configured in the cluster database. These HBAs are available for fencing.

However, if no WWPNs are detected, there will be messages about loading the HBA/SNIA library logged to the `/var/log/cxfs_client` file.

If no WWPNs are detected, you can manually specify the WWPNs in the fencing file.

---

**Note:** This method does not work if the WWPNs are partially discovered.

---

The `/etc/fencing.conf` file enumerates the WWPN for all of the HBAs that will be used to mount a CXFS filesystem. There must be a line for the HBA WWPN as a 64-bit hexadecimal number.

---

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

---

If used, `/etc/fencing.conf` must contain a simple list of WWPNs, one per line. You must update it whenever the HBA configuration changes, including the replacement of an HBA.

Do the following:

1. Set up the switch and HBA. See the release notes for supported hardware.
2. Determine the HBA WWPN: Follow the Fibre Channel cable on the back of the node to determine the port to which it is connected in the switch. Ports are numbered beginning with 0. (For example, if there are 8 ports, they will be numbered 0 through 7.)
3. Use the `telnet` command to connect to the switch and log in as user `admin`. (On Brocade switches, the password is `password` by default).
4. 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 fencing file).

5. Edit or create `/etc/fencing.conf` and add the WWPN for the port determined in step 2. (Comment lines begin with #.)

For dual-ported HBAs, you must include the WWPNs of any ports that are used to access cluster disks. This may result in multiple WWPNs per HBA in the file; the numbers will probably differ by a single digit.

For example, if you determined that port 0 is the port connected to the switch, your fencing file should contain the following:

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

6. After the node is added to the cluster (see Chapter 7, "Cluster Configuration" on page 215), enable the fencing feature by using the CXFS GUI or `cmgr` command on a server-capable administration node; for more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Start/Stop `cxfs_client` for Linux

The `/etc/init.d/cxfs_client` script will be invoked automatically during normal system startup and shutdown procedures. This script starts and stops the `cxfs_client` daemon.

To start up `cxfs_client` manually on your Linux node, enter the following:

```
[root@linux root]# /etc/init.d/cxfs_client start
Loading cxfs modules: [ OK ]
```

```
Mounting devfs filesystems: [ OK ]
Starting cxfs client: [ OK ]
```

To stop `cxfs_client` manually, enter the following:

```
[root@linux root]# /etc/init.d/cxfs_client stop
Stopping cxfs client: [ OK ]
```

To stop and then start `cxfs_client` manually, enter the following:

```
[root@linux root]# /etc/init.d/cxfs_client restart
Stopping cxfs client: [ OK ]
```

To see the current status, use the `status` argument. For example:

```
[root@ceara root]# /etc/init.d/cxfs_client status
cxfs_client status [timestamp Apr 20 14:54:30 / generation 4364]
```

```
CXFS client:
  state: stable (5), cms: up, xvm: up, fs: up
Cluster:
  connies_cluster (707) - enabled
Local:
  ceara (7) - enabled
Nodes:
  aiden      enabled up    12
  brenna     enabled DOWN 10
  brigid     enabled up    11
  ceara      enabled up    7
  chili      enabled up    4
  cxfsibm2   enabled up    9
  cxfssun4   enabled up    5
  daghada    enabled up    8
  flynn      enabled up    2
  gaeth      enabled up    0
  minnesota  enabled up    6
  rowan      enabled up    3
  rylie      enabled up    1
Filesystems:
  concatfs   enabled mounted          concatfs          /concatfs
  stripefs   enabled mounted          stripefs         /stripefs
  tp9300_stripefs enabled forced mounted  tp9300_stripefs /tp9300_stripefs
cxfs_client is running.
```

For example, if `cxfs_client` is stopped:

```
[root@linux root]# /etc/init.d/cxfs_client status
cxfs_client is stopped
```

## Maintenance for Linux

This section contains information about maintenance procedures for CXFS on Linux.

### Modifying the CXFS Software for Linux

You can modify the behavior of the CXFS client daemon (`/usr/cluster/bin/cxfs_client`) by placing options in the `/etc/cluster/config/cxfs_client.options` file. The available options are documented in the `cxfs_client` man page.



---

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

---

The first line in the `cxfs_client.options` file must contain the options you want `cxfs_client` to process.

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

```
[root@linux root]# ps -ax | grep cxfs_client
3612 ?          S          0:00 /usr/cluster/bin/cxfs_client -i cxfs3-5
3841 pts/0      S          0:00 grep cxfs_client
```

### Using `cxfs-reprobe` with Linux

When `cxfs_client` needs to rescan disk buses, it executes the `/var/cluster/cxfs_client-scripts/cxfs-reprobe` script. This requires the use of parameters in Linux due to limitations in the Linux SCSI layer. You can export these parameters from the `/etc/cluster/config/cxfs_client.options` file.

The script detects the presence of the SCSI and/or XSCSI layers on the system and defaults to probing whichever layers are detected. This decision can be overridden by setting `CXFS_PROBE_SCSI` (for Linux SCSI) or `CXFS_PROBE_XSCSI` (for Linux XSCSI) to either 0 (to disable the probe) or 1 (to force the probe).

When an XSCSI scan is performed, all buses are scanned by default. This can be overridden by specifying a space-separated list of buses in `CXFS_PROBE_XSCSI_BUSES`. (If you include space, you must enclose the list within single quotation marks.) For example:

```
export CXFS_PROBE_XSCSI_BUSES='/dev/xscsi/pci01.03.0-1/bus /dev/xscsi/pci02.01.0-2/bus'
```

When a SCSI scan is performed, a fixed range of buses/channels/IDs and LUNs are scanned; these ranges may need to be changed to ensure that all devices are found. The ranges can also be reduced to increase scanning speed if a smaller space is sufficient.

The following summarizes the environment variables (separate multiple values by white space and enclose withing single quotation marks):

`CXFS_PROBE_SCSI=0|1`

Stops (0) or forces (1) a SCSI probe. Default: 1 if SCSI

`CXFS_PROBE_SCSI_BUSES=BusList`

Scans the buses listed. Default: 0 1 2

`CXFS_PROBE_SCSI_CHANNELS=ChannelList`

Scans the channels listed. Default: 0

`CXFS_PROBE_SCSI_IDS=IDList`

Scans the IDS listed. Default: 0 1 2 3

`CXFS_PROBE_SCSI_LUNS=LunList`

Scans the LUNs listed. Default: 0 1 2 3 4 5 6 7 8 9 10 11 12  
13 14 15

`CXFS_PROBE_XSCSI=0|1`

Stops (0) or forces (1) an XSCSI probe. Default: 1 if XSCSI

```
CXFS_PROBE_XSCSI_BUSES=BusList
```

Scans the buses listed. Default: all XSCSI buses

For example, the following would only scan the first two SCSI buses:

```
export CXFS_PROBE_SCSI_BUSES='0 1'
```

The following would scan 16 LUNs on each bus, channel, and ID combination (all on one line):

```
export CXFS_PROBE_SCSI_LUNS='0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15'
```

Other options within the `/etc/cluster/config/cxfs_client.options` file begin with a `-` character. Following is an example `cxfs_client.options` file:

```
# Example cxfs_client.options file
#
-Dnormal -serror
export CXFS_PROBE_SCSI_BUSES=1
export CXFS_PROBE_SCSI_LUNS='0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20'
```

---

**Note:** The `-` character or the term `export` must start in the first position of each line in the `cxfs_client.options` file; otherwise, they are ignored by the `/etc/init.d/cxfs_client` script.

---

## GRIO on Linux

CXFS supports guaranteed-rate I/O (GRIO) version 2 on the Linux platform. However, GRIO is disabled by default on Linux.

To enable GRIO, change the following line in `/etc/cluster/config/cxfs_client.options` from:

```
export GRIO2=off
```

to:

```
export GRIO2=on
```

For more information, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and the *Guaranteed-Rate I/O Version 2 Guide*.

## XVM Failover Version 2 on Linux

Following is an example of the `/etc/failover2.conf` file on Linux for a configuration in which each node has dual ports connecting to separate Fibre Channel (FC) switches. Each FC switch is connected to a single RAID controller with no interconnect between the FC switches:

```
# phys/d9400_0
/dev/xscsi/pci01.02.1/node200500a0b80cedb3/port1/lun0/disc affinity=0 preferred
/dev/xscsi/pci01.03.1/node200400a0b80cedb3/port1/lun0/disc affinity=1

# phys/d9400_1
/dev/xscsi/pci01.03.1/node200400a0b80cedb3/port1/lun1/disc affinity=0 preferred
/dev/xscsi/pci01.02.1/node200500a0b80cedb3/port1/lun1/disc affinity=1

# phys/wally
/dev/xscsi/pci01.03.1/node200400a0b80cedb3/port1/lun2/disc affinity=0 preferred
/dev/xscsi/pci01.02.1/node200500a0b80cedb3/port1/lun2/disc affinity=1

# phys/recovery
/dev/xscsi/pci01.02.1/node200500a0b80cedb3/port1/lun3/disc affinity=0 preferred
/dev/xscsi/pci01.03.1/node200400a0b80cedb3/port1/lun3/disc affinity=1
```

For more information, see "XVM Failover Version 2 and CXFS" on page 17, the comments in the `/etc/failover2.conf` file on a CXFS administration node, and the *XVM Volume Manager Administrator's Guide*.

## Troubleshooting for Linux

This section discusses the following:

- "Verifying the Licenses for Linux" on page 73
- "Device Filesystem Enabled for Linux" on page 74
- "The `cxfs_client` Daemon is Not Started on Linux" on page 74
- "Filesystems Do Not Mount on Linux" on page 74
- "Large Log Files on Linux" on page 76

For general troubleshooting information, see Chapter 8, "General Troubleshooting" on page 231 and Appendix D, "Error Messages" on page 251.

## Verifying the Licenses for Linux

To verify that the licenses have been installed properly, use the `cxfslicense` command after installing the CXFS software. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example

- License installed properly:

```
[root@linux cdrom]# /usr/cluster/bin/cxfslicense
[root@linux cdrom]# /usr/cluster/bin/cxfslicense -d
no xvm user license required.
XVM_STD_LINUX32 license granted.
CXFS_LINUX32 license granted.
```

- No license (error condition):

```
[root@linux cdrom]# /usr/cluster/bin/cxfslicense
cxfs license check failed - use 'cxfslicense -d' for details
[root@linux cdrom]# /usr/cluster/bin/cxfslicense -d
no xvm user license required.
XVM_STD_LINUX32 license granted.
Error on CXFS_LINUX32 license check out.
checkout failed: No such feature exists
Feature:          CXFS_LINUX32
License path:    /etc/flexlm/license.dat
FLEXlm error:   -5,357
For further information, refer to the FLEXlm End User Manual,
available at "www.globetrotter.com".
```

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

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

An error such as the following example will appear in the SYSLOG file (line breaks added here for readability):

```
Jan 25 10:24:03 ncc1701:Jan 25 10:24:03 cxfs_client:  
cis_main FATAL: cxfs_client failed the CXFS license check.  
Use the cxfslicense command to diagnose the license problem
```

## Device Filesystem Enabled for Linux

The kernels provided for the Linux client have the Device File System (`devfs`) enabled. This can cause problems with locating system devices in some circumstances.

See the `devfs` FAQ at the following location:

<http://www.atnf.csiro.au/people/rgooch/linux/docs/devfs.html>

## The `cxfs_client` Daemon is Not Started on Linux

Confirm that the `cxfs_client` is not running. The following command would list the `cxfs_client` process if it were running:

```
linux# ps -ax | grep cxfs_client
```

The `cxfs_client` daemon might not start if the license has expired, which is indicated when the following message is output to the `/var/log/cxfs_client` file:

```
CXFS not properly licensed for this host
```

In this case, execute the following command to determine why the license check fails:

```
linux# /usr/cluster/bin/cxfslicense -d
```

If the license is valid, check the `cxfs_client` log file for other errors.

Restart `cxfs_client` as described in "Start/Stop `cxfs_client` for Linux" on page 67 and watch the `cxfs_client` log file for errors.

## Filesystems Do Not Mount on Linux

If `cxfs_info` reports that `cms` is up but XVM or the filesystem is in another state, then one or more mounts is still in the process of mounting or has failed to mount.

The CXFS node might not mount filesystems for the following reasons:

- The client may not be able to see all the LUNs. This is usually caused by misconfiguration of the HBA or the SAN fabric:
  - Check that the ports on the Fibre Channel switch connected to the HBA are active. Physically look at the switch to confirm the light next to the port is green, or remotely check by using the `switchShow` command.
  - Check that the HBA configuration is correct.
  - Check that the HBA can see all the LUNs for the filesystems it is mounting.
  - Check that the operating system kernel can see all the LUN devices.
  - If the RAID device has more than one LUN mapped to different controllers, ensure the node has a Fibre Channel path to all relevant controllers.
- The `cxfs_client` daemon may not be running. See "The `cxfs_client` Daemon is Not Started on Linux" on page 74.
- The filesystem may have an unsupported mount option. Check the `cxfs_client.log` for mount option errors or any other errors that are reported when attempting to mount the filesystem.
- The cluster membership (`cms`), XVM, or the filesystems are not up on the node. Execute the `/usr/cluster/bin/cxfs_info` command to determine the current state of `cms`, XVM, and the filesystems. If the node is not up for each of these, then check the `/var/log/cxfs_client` log to see what actions have failed.

Do the following:

- If `cms` is not up, check the following:
  - Is the node is configured on the administration node with the correct hostname?
  - Has the node been added to the cluster and enabled? See "Verifying the Cluster Status" on page 223.
- If XVM is not up, check that the HBA is active and can see the LUNs.
- If the filesystem is not up, check that one or more filesystems are configured to be mounted on this node and check the `/var/log/cxfs_client` file for mount errors.

## Large Log Files on Linux

The `/var/log/cxfs_client` log file may become quite large over a period of time if the verbosity level is increased.

See the `cxfs_client.options` man page and "Log Files on Linux" on page 49.

## Reporting Linux Problems

When reporting a problem about a Linux node to SGI, you should retain the following information:

- The kernel you are running:

```
[root@linux root]# uname -a
```

- The CXFS packages you are running:

```
[root@linux root]# rpm -q cxfs_client cxfs-modules cxfs_utils xvm-cmds
```

- The numbers and types of the processors on your machine:

```
[root@linux root]# cat /proc/cpuinfo
```

- The hardware installed on your machine:

```
[root@linux root]# /sbin/lspci
```

- Modules that are loaded on your machine:

```
[root@linux root]# /sbin/lsmmod
```

- The `/var/log/cxfs_client` log file
- Any messages that appeared in the system logs immediately before the system exhibited the problem.
- Output about the cluster obtained from the `cxfsdump` utility run on a CXFS administration node. The `cxfsdump` command transfers all of the information back to the node where the command was issued. When run in local mode on a Linux node, it stores information in `/var/cluster/cxfsdump-data/nodename.tar.gz`.
- After a system kernel panic, the debugger information from the `kdb` built-in kernel debugger.



---

**Caution:** When the system enters the debugger after a panic, it will render the system unresponsive until the user exits from the debugger. Also, if `kdb` is entered while the system is in graphical (X) mode, the debugger prompt cannot be seen. For these reasons, `kdb` is turned off by default.

---

You can temporarily enable `kdb` by entering the following:

```
[root@linux root]# echo 1 > /proc/sys/kernel/kdb
```

To enable `kdb` at every boot, place the following entry in the `/etc/sysctl.conf` file:

```
# Turn on KDB
kernel.kdb = 1
```

For more information, see the `sysctl` man page.

When `kdb` is enabled, a system panic will cause the debugger to be invoked and the keyboard LEDs will blink. The `kdb` prompt will display basic information. To obtain a stack trace, enter the `bt` command at the `kdb` prompt:

```
kdb> bt
```

To get a list of current processes, enter the following:

```
kdb> ps
```

To backtrace a particular process, enter the following, where `PID` is the process ID:

```
kdb> btp PID
```

To exit the debugger, enter the following:

```
kdb> go
```

If the system will be run in graphical mode with `kdb` enabled, SGI highly recommends that you use `kdb` on a serial console so that the `kdb` prompt can be seen.



## Mac OS X Platform

CXFS supports a client-only node running the Mac OS X operating system. This chapter contains the following sections:

- "CXFS on Mac OS X" on page 79
- "Installation and Configuration Overview for Mac OS X " on page 84
- "FLEXlm License for Mac OS X" on page 85
- "HBA Installation for Mac OS X" on page 87
- "Preinstallation Steps for Mac OS X" on page 89
- "Client Software Installation for Mac OS X" on page 92
- "I/O Fencing for Mac OS X" on page 93
- "Start/Stop `cxfs_client` for Mac OS X" on page 96
- "Maintenance for Mac OS X" on page 96
- "GRIO on Mac OS X" on page 97
- "XVM Failover Version 2 on Mac OS X" on page 98
- "Troubleshooting for Mac OS X" on page 99
- "Reporting Mac OS X Problems" on page 101

### CXFS on Mac OS X

This section contains the following information about CXFS on Mac OS X:

- "Requirements for Mac OS X" on page 80
- "CXFS Commands on Mac OS X" on page 80
- "Log Files on Mac OS X" on page 81
- "Limitations and Considerations on Mac OS X" on page 81
- "Configuring Hostnames on Mac OS X" on page 82

- "Mapping User and Group Identifiers for Mac OS X" on page 83
- "Access Control Lists and Mac OS X" on page 84

## Requirements for Mac OS X

In addition to the items listed in "Requirements" on page 7, using a Mac OS X node to support CXFS requires the following:

- Mac OS X operating system 10.3.7
- One of the following single- or dual-processor Apple Computer hardware platforms:
  - Power Mac G4
  - Xserve G4
  - Power Mac G5
  - Xserve G5
- Apple Fibre Channel PCI and PCI-X host bus adapter (HBA)

For the latest information, see the CXFS Mac OS X release notes.

## CXFS Commands on Mac OS X

The following commands are shipped as part of the CXFS Mac OS X package:

```
/usr/cluster/bin/cxfs_client  
/usr/cluster/bin/cxfs_info  
/usr/cluster/bin/cxfslicense  
/usr/cluster/bin/cxfsdump  
/usr/cluster/bin/install-cxfs  
/usr/cluster/bin/uninstall-cxfs  
/Library/StartupItems/cxfs/cxfs  
/usr/sbin/grioadmin  
/usr/sbin/grioqos  
/usr/cluster/bin/xvm
```

The `cxfs_client` and `xvm` commands are needed to include a Mac OS X node in a CXFS cluster. The `cxfs_info` command reports the current status of this node in the CXFS cluster.

You can use the `cxfslicense` command to verify a CXFS or XVM FLEXlm license.

The installation package uses `install-cxfs` to install or update all of the CXFS files. You can use the `uninstall-cxfs` command to uninstall all CXFS files; `uninstall` is not an installation package option.

The `/Library/StartupItems/cxfs/cxfs` command is run by the operating system to start and stop CXFS on the Mac OS X node.

For more information on these commands, see the man pages.

For information about the GRIO commands, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and "GRIO on Mac OS X" on page 97.

## Log Files on Mac OS X

The `cxfs_client` command creates a `/var/log/cxfs_client` log file. To rotate this log file, use the `-z` option in the `/usr/cluster/bin/cxfs_client.options` file; see the `cxfs_client` man page for details.

The CXFS installation process (`install-cxfs` and `uninstall-cxfs`) appends to `/var/log/cxfs_inst.log`.

For information about the log files created on CXFS administration nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Limitations and Considerations on Mac OS X

CXFS for Mac OS X has the following limitations and considerations:

- Mac OS X does not support the `inode64` mount option. For more information, see Appendix C, "Mount Options Support" on page 247.
- Mac OS X is unable to memory map a file larger than 2 GB.
- XVM volume names are limited to 31 characters and subvolumes are limited to 26 characters. For more information about XVM, see *XVM Volume Manager Administrator's Guide*.

- Mac OS X does not support the CXFS mount scripts.

See also Appendix B, "Filesystem and Logical Unit Specifications" on page 245.

## Configuring Hostnames on Mac OS X

A Mac OS X node may use a combination of methods for obtaining the node's hostname, depending on if it is in a NetInfo domain or is standalone.

The hostname is normally specified using the following menu selection:

```
System Preferences
  > Sharing
    > Computer Name
```

The hostname entry in `/etc/hostconfig` is `HOSTNAME=-AUTOMATIC-`, which will normally result in the hostname specified for the machine with the following domain:

```
.local.
```

For example, if the hostname was specified as `cxfsmacl`, then you would see the following when requesting the hostname:

```
macosx# /bin/hostname
cxfsmacl.local.
```

The full hostname including `.local.` is the hostname that the CXFS software will use to determine its identity in the cluster, not `cxfsmacl`.

Therefore, you must configure the node as `cxfsmacl.local.` or specify the fully qualified hostname in `/etc/hostconfig`. For example:

```
HOSTNAME=cxfsmacl.sgi.com
```

Specifying the hostname in this way may impact some applications, most notably Rendezvous, and should be researched and tested carefully. There are also known issues with the hostname being reported as `localhost` on some reboots after making such a change.

SGI recommends that other hosts in the cluster are specified in the Mac OS X node's `/etc/hosts` file.

## Mapping User and Group Identifiers for Mac OS X

To ensure that the correct access controls are applied to users on Mac OS X nodes when accessing CXFS filesystems, you must ensure that the user IDs (UIDs) and group IDs (GIDs) are the same on the Mac OS X node as on all other nodes in the cluster, in particular any CXFS administration nodes.

---

**Note:** A user does not have to have user accounts on all nodes in the cluster. However, all access control checks are performed by CXFS administration nodes, so any administration nodes must be configured with the superset of all users in the cluster.

---

Users can quickly check that their UID and GID settings are correct by using the `id` command on both the Mac OS X node and the CXFS administration node, and the `groups` command on the administration node. For example:

```
macosx% id
uid=1113(fred) gid=999(users) groups=999(users), 20(staff)
```

```
irix% id
uid=1113(fred) gid=999(users)
irix% groups
users staff
```

If the UID and/or GID do not match, or if the user is not a member of the same groups, then the user may unexpectedly fail to access some files.

To change the user's UID, GID, or other groups requires changes to the NetInfo domain, whether local or distributed. Do the following:

- Run the NetInfo Manager tool:

```
Applications
  > Utilities
    > NetInfo Manager
```

- Select the domain (if not the local domain):

```
Domain
  > Open....
```

- Select the user in question:

```
users  
  > username
```

- Modify the `uid`, `gid`, or `group` fields as required.

---

**Note:** Changing a user's primary UID and/or GID will also require modifying all files owned by the user to the new UID and GID. Ideally, users should be created with the correct values.

---

Alternatively, you can change the UID and GID on the CXFS administration nodes and CXFS filesystems.

## Access Control Lists and Mac OS X

All CXFS files have UNIX mode bits (read, write, and execute) and optionally an access control list (ACL). For more information, see the `chmod` and `chac1` man pages on a metadata server.

Mac OS X supports UNIX mode bits but has no native support for ACLs, so they cannot be viewed or modified from a Mac OS X node. However, they are enforced by the Mac OS X node if applied to files by other nodes in the cluster.

When using `tar`, `cpio`, or other third-party backup tools from a Mac OS X node, the ACL will be lost because those tools are unaware of the ACL on the file.

## Installation and Configuration Overview for Mac OS X

Following is the order of installation and configuration steps for a CXFS Mac OS X node:

1. Read the release notes for the Mac OS X platform to learn about any late-breaking changes in the installation procedure.
2. Install the Mac OS X operating system according to the directions in the Mac OS X documentation (if not already done).
3. Install and verify the RAID. See the release notes.
4. Install and verify the switch. See the release notes.

5. Obtain and install the CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must purchase a mirroring software option and obtain and install a FLEXIm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See "FLEXIm License for Mac OS X" on page 85.
6. Install and verify the host bus adapter (HBA). See "HBA Installation for Mac OS X" on page 87.
7. Prepare the Mac OS X node, including adding a private network. See "Preinstallation Steps for Mac OS X" on page 89..
8. Install the CXFS software. See "Client Software Installation for Mac OS X" on page 92.
9. Configure the cluster to define the new Mac OS X node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration" on page 215.

If you run into problems, see "Troubleshooting for Mac OS X" on page 99 and Chapter 8, "General Troubleshooting" on page 231.

## **FLEXIm License for Mac OS X**

This section discusses the following:

- "Gathering the Mac OS X Host Information" on page 85
- "Obtaining the Licenses for Mac OS X" on page 86
- "Installing the Licenses for Mac OS X" on page 86

### **Gathering the Mac OS X Host Information**

When you order CXFS, you will receive an entitlement identifier (ID). You must submit the system host ID, host name, and entitlement ID when requesting your permanent CXFS license.

To obtain the hostname of the system on which you run CXFS, execute the following Mac OS X command in a terminal to obtain the hostname and host ID from the onboard Ethernet interface:

```
/bin/hostname -s  
/sbin/ifconfig en0 | grep ether
```

For example:

```
macosx# /bin/hostname -s  
cxfsmacl  
macosx# /sbin/ifconfig en0 | grep ether  
ether 00:03:93:cc:3a:e4
```

The Ethernet address is 00:03:93:cc:3a:e4, so the host ID is 000393cc3ae4.

Alternatively, the Ethernet address is also displayed with the interface by using the following menu selection:

**System Preferences**  
**> Network**

---

**Note:** Ensure that the address is for the onboard Ethernet interface and not any additional interface cards.

---

## Obtaining the Licenses for Mac OS X

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.

## Installing the Licenses for Mac OS X

You will install the licenses in the following location:

```
/etc/flexlm/license.dat
```

Do the following:

1. Create the license directory if necessary.

For example:

```
macosx# mkdir -p /etc/flexlm
```

2. Copy the key to the `license.dat` file.

## HBA Installation for Mac OS X

CXFS for Mac OS X supports Apple Computer, Inc. host bus adapters (HBAs).

---

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

---

This section discusses the following:

- "Installing the Apple HBA" on page 87
- "Installing the Fibre Channel Utility for Mac OS X" on page 88
- "Configuring Two or More Apple HBA Ports" on page 88
- "Using `point-to-point` Fabric Setting for Apple HBAs" on page 88

### Installing the Apple HBA

Do the following:

1. Install the Apple HBA into a spare PCI or PCI-X slot in the Mac OS X node, according to the manufacturer's instructions. Do not connect the HBA to the Fibre Channel switch at this time.

---

**Note:** Apple HBAs are normally shipped with copper SFPs and copper cables, so additional optic SFPs and optic cables may be required.

---

2. Reboot the node.

## Installing the Fibre Channel Utility for Mac OS X

Do the following:

1. Install the configuration utility from the CD distributed with the Apple HBA. To do this, copy **Mac OS X Utilities/Fibre Channel Utility** from the CD to your **Application** directory.
2. Run the Fibre Channel Utility after it is copied to the node. The tool will list the HBA on the left-hand side of the window. Select the **Apple FC card** item to display the status of the ports via a pull-down menu. Initially, each port will report that it is up (even though it is not connected to the switch), and the speed and port topology will configure automatically.
3. Connect one of the HBA ports to the switch via a Fibre Channel cable. After a few seconds, close and relaunch the Fibre Channel Utility. Select the **Apple FC card** item and then the connected port from the drop down list to display the speed of the link.

Repeat these steps for the second HBA port if required.

4. (*Optional*) Configure each port to a specific speed and topology, if desired. However this may cause problems if the fabric configuration is changed in the future.

## Configuring Two or More Apple HBA Ports

The Fibre Channel Utility does not support masking logical units (LUNs) on specific ports. However, if the first port can see all of the LUNs, the default is that all I/O will go through a single port. To avoid this, configure the switch so that each port can see a different set of LUNs. This may be achieved by zoning the switch or by using multiple switches, with different controllers and HBA ports to each switch.

## Using point-to-point Fabric Setting for Apple HBAs

SGI recommends that you use the manual point-to-point fabric setting rather than rely on automatic detection, which can prove unreliable after a reboot.

## Preinstallation Steps for Mac OS X

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

- "Adding a Private Network for Mac OS X Nodes"
- "Verifying the Private and Public Networks for Mac OS X" on page 91
- "Disabling Power Save Mode for Mac OS X" on page 91

### Adding a Private Network for Mac OS X Nodes

The following procedure provides an overview of the steps required to add a private network to the Mac OS X 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 Mac OS X system documentation.

1. Install Mac OS X and configure the machine's hostname (see "Configuring Hostnames on Mac OS X" on page 82) and IP address on its public network interface.
2. Decide if the Mac OS X node will be part of a NetInfo domain or a standalone machine. If part of a NetInfo domain, configure the node into the domain before proceeding further.
3. Add the IP addresses and hostnames of other machines in the cluster to the NetInfo database and/or the `/etc/hosts` file. You should be consistent about specifying the hostname or the fully qualified domain name for each host. A common convention is to name the CXFS private network address for each host as `hostname-priv`.
4. Install a second network interface card as per the manufacturer's instructions.

5. Configure the second network interface card by using the following menu selection:

**System Preferences**  
    > **Network**  
        > **Show**

Select the second network interface (most likely PCI Ethernet Slot 1), and specify the IP address, subnet mask, and router. The private network interface should not require a DNS server because the private network address of other cluster nodes should be explicitly listed in the NetInfo database and/or in the `/etc/hosts` file. Relying on a DNS server for private network addresses introduces another point of failure into the cluster and must be avoided.

6. Confirm the configuration using `ifconfig` to list the network interfaces that are up:

```
macosx# ifconfig -u
```

In general this should include `en0` (the onboard Ethernet) and `en1` (the additional PCI interface), but the names of these interface may vary.

For more information, see the `ifconfig` man page.

7. (Optional) Enable the `rshd` daemon and edit the `~root/.rhosts` file if you want to use the connectivity diagnostics provided with CXFS.



**Caution:** Mac OS X by default enables `ssh` connections and not `rsh` connections, because `rsh` connections are less secure. You should consider the security implications of enabling `rsh` access to the Mac OS X node before performing this action.

---

To enable `rshd` on Mac OS X, edit the `/etc/inetd.conf` file and uncomment the following lines:

```
ftp      stream  tcp      nowait  root    /usr/libexec/tcpd      ftpd -l
login    stream  tcp      nowait  root    /usr/libexec/tcpd      rlogind
shell    stream  tcp      nowait  root    /usr/libexec/tcpd      rshd
telnet   stream  tcp      nowait  root    /usr/libexec/tcpd      telnetd
```

Then restart the `inetd` process by entering the following:

```
macosx% sudo kill -HUP `cat /var/run/inetd.pid`
```

Edit the root user's `.rhosts` file to allow root users on other CXFS nodes, notably CXFS administration nodes, to access this machine using `rsh` without a password. For example, if `cxfsmds1` is the metadata server, add the following to the `.rhosts` file:

```
macosx% sudo vi ~root/.rhosts cxfsmds1 root
macosx% sudo chmod 600 ~root/.rhosts
```

The connectivity tests execute a `ping` 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` as user `root`. Likewise, the CXFS diagnostic tool `cxfsdump` uses `rsh` from an administration node to collect information from the Mac OS X node.

## Verifying the Private and Public Networks for Mac OS X

Verify each interface using the `ping` command to connect to the public and private network addresses of the other nodes that are in the CXFS pool.

For example:

```
macosx# grep cxfsmac2 /etc/hosts
134.14.55.115 cxfsmac2
macosx# ping -c 3 134.14.55.115
PING 134.14.55.115 (134.14.55.115): 56 data bytes
64 bytes from 134.14.55.115: icmp_seq=0 ttl=64 time=0.247 ms
64 bytes from 134.14.55.115: icmp_seq=1 ttl=64 time=0.205 ms
64 bytes from 134.14.55.115: icmp_seq=2 ttl=64 time=0.197 ms

--- 134.14.55.115 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.197/0.216/0.247 ms
```

## Disabling Power Save Mode for Mac OS X

CXFS does not support the energy saving mode on Mac OS X. If this mode is enabled, the Mac OS X node will lose CXFS membership and unmount the CXFS filesystem whenever it is activated.

Select the following to disable the power save mode:

**System Preferences**  
    > **Energy Saver**  
        > **Put the computer to sleep when it is inactive for**  
            > **Never**

## Client Software Installation for Mac OS X

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.

Installing the CXFS client CD for Mac OS X requires approximately 30 MB of space.

To install the required software on a Mac OS X node, SGI personnel will do the following:

1. Read the `ReadMe` file for the Mac OS X platform to learn about any late-breaking changes in the installation procedure.
2. Verify that the node is running Mac OS X operating system 10.3.5 according to the Mac OS X installation guide.

Use the following command to display the currently installed system:

```
macosx# uname -r
```

This command should return a value of 7.3 because the Mac OS X kernel, Darwin, is version 7.X in 10.3.X.

3. Insert the *CXFS MultiOS Client 3.3* CD-ROM.
4. Using the **Finder**, open `macosx/cxfs.pkg` from the CD-ROM. This will launch the installation application, which will do the following:
  - Display the `ReadMe` file
  - Display the license agreement and request acceptance
  - Force you to select the boot disk if multiple local disk partitions are installed

Before starting the actual file installation, you may use the following menu selection to view the installation process in more detail:

```
File
  > Show Log
```

This information is also appended to the `/var/log/cxfs_inst.log` file.

5. Use the `cxfslicense` command to verify that the license is installed in `/etc/flexlm/license.dat`. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example, the following shows a properly installed license:

```
macosx# /usr/cluster/bin/cxfslicense -d
no xvm user license required.
XVM_MACOSX license granted.
CXFS_MACOSX license granted.
```

If there are errors, verify that you have obtained and installed the CXFS licenses. For more information, see "Verifying the Licenses for Mac OS X" on page 99.

6. Restart the machine.

## I/O Fencing for Mac OS X

I/O fencing is required on Mac OS X nodes in order to protect data integrity of the filesystems in the cluster. The `cxfs_client` software automatically detects the world wide port names (WWPNs) of any supported host bus adapters (HBAs) for Mac OS X nodes that are connected to a switch that is configured in the cluster database. These HBAs are available for fencing.

However, if no WWPNs are detected, the following messages will be logged to the `/var/log/cxfs_client` file:

```
hba_wwpn_list warning: No WWPN found from IO Registry
cis_get_hbas warning: Not able to find WWN (err=Device not
configured). Falling back to "/etc/fencing.conf".
cis_config_swports_set error fetching hbas
```

If no WWPNs are detected, you can manually specify the WWPNs in the fencing file.

---

**Note:** This method does not work if the WWPNs are partially discovered.

---

The `/etc/fencing.conf` file enumerates the WWPN for all of the HBAs that will be used to mount a CXFS filesystem. There must be a line for the HBA WWPN as a 64-bit hexadecimal number.

---

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

---

If used, `/etc/fencing.conf` must contain a simple list of WWPNs, one per line. You must update it whenever the HBA configuration changes, including the replacement of an HBA.

Do the following:

1. Set up the switch and HBA. See the release notes for supported hardware.
2. Follow the Fibre Channel cable on the back of the node to determine the port to which it is connected in the switch. Ports are numbered beginning with 0. (For example, if there are 8 ports, they will be numbered 0 through 7.)
3. Use the `telnet` command to connect to the switch and log in as user `admin`. (On Brocade switches, the password is `password` by default).
4. 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 fencing file).

5. Edit or create `/etc/fencing.conf` and add the WWPN for the port determined in step 2. (Comment lines begin with #.)

For dual-ported HBAs, you must include the WWPNs of any ports that are used to access cluster disks. This may result in multiple WWPNs per HBA in the file; the numbers will probably differ by a single digit.

For example, if you determined that port 0 is the port connected to the switch, your fencing file should contain the following:

```

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

```

6. After the node is added to the cluster (see Chapter 7, "Cluster Configuration" on page 215), enable the fencing feature by using the CXFS GUI or `cmgr` command on a server-capable administration node; for more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Start/Stop `cxfs_client` for Mac OS X

The `/Library/StartupItems/cxfs/cxfs` script will be invoked automatically during normal system startup and shutdown procedures. This script starts and stops the `cxfs_client` daemon.

To start `cxfs_client` manually on your Mac OS X node, enter the following:

```
macosx# sudo /Library/StartupItems/cxfs/cxfs start
```

To stop `cxfs_client` manually, enter the following:

```
macosx# sudo /Library/StartupItems/cxfs/cxfs stop
```

To stop and start `cxfs_client` manually, enter the following:

```
macosx# sudo /Library/StartupItems/cxfs/cxfs restart
```

To prevent the automatic startup of `cxfs_client` on boot, move the `/Library/StartupItems/cxfs` directory out of `/Library/StartupItems`.

## Maintenance for Mac OS X

This section contains the following:

- "Upgrading the CXFS Software for Mac OS X"
- "Modifying the CXFS Software for Mac OS X" on page 97
- "Removing the CXFS Software for Mac OS X" on page 97
- "Recognizing Storage Changes for Mac OS X" on page 97

## Upgrading the CXFS Software for Mac OS X

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 Software for Mac OS X

You can modify the behavior of the CXFS client daemon (`/usr/cluster/bin/cxfs_client`) by placing options in the `/usr/cluster/bin/cxfs_client.options` file. The available options are documented in the `cxfs_client` man page.



---

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

---

The first line in the `cxfs_client.options` file must contain the options you want `cxfs_client` to process.

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

```
macosx# ps -auxwww | grep cxfs
```

## Removing the CXFS Software for Mac OS X

After terminating any applications that access CXFS filesystems on the Mac OS X node, execute the following:

```
macosx# sudo /usr/cluster/bin/uninstall-cxfs
```

Restart the system to unload the CXFS module from the Mac OS X kernel.

## Recognizing Storage Changes for Mac OS X

If you make changes to your storage configuration, you may have to reboot your machine because there is currently no mechanism in Mac OS X to reprobe the storage.

## GRIO on Mac OS X

CXFS supports guaranteed-rate I/O (GRIO) version 2 on the Mac OS X platform.

For more information, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and the *Guaranteed-Rate I/O Version 2 Guide*.

## XVM Failover Version 2 on Mac OS X

Following is an example of the `/etc/failover2.conf` file on Mac OS X for a configuration in which each node has dual ports connecting to separate Fibre Channel (FC) switches. Each FC switch is connected to a single RAID controller with no interconnect between the FC switches:

```
# phys/d9400_0
/dev/rxvm-200500a0b80cedb3-000 affinity=0 preferred
/dev/rxvm-200400a0b80cedb3-000 affinity=1

# phys/d9400_1
/dev/rxvm-200400a0b80cedb3-001 affinity=0 preferred
/dev/rxvm-200500a0b80cedb3-001 affinity=1

# phys/wally
/dev/rxvm-200400a0b80cedb3-002 affinity=0 preferred
/dev/rxvm-200500a0b80cedb3-002 affinity=1

# phys/recovery
/dev/rxvm-200500a0b80cedb3-003 affinity=0 preferred
/dev/rxvm-200400a0b80cedb3-003 affinity=1
```

For more information, see "XVM Failover Version 2 and CXFS" on page 17, the comments in the `/etc/failover2.conf` file on a CXFS administration node, and the *XVM Volume Manager Administrator's Guide*.

## Mapping XVM Volumes to Storage Targets on Mac OS X

XVM volumes can be mapped to devices and to targets on RAID controllers using the output from the `xvm` command and the device entries in the filesystem. The `xvm` command can be used to display the device names:

```
macosx# /usr/cxfs_cluster/bin/xvm show -e -t vol
vol/stripel                0 online,open
  subvol/stripel/data       2292668416 online,open
    stripe/stripel         2292668416 online,open (unit size: 1024)
      slice/d9400_0s0       1146334816 online,open
(d9400_0:/dev/rxvm-200400a0b80cd5fe-000)
      slice/d9400_1s0       1146334816 online,open
(d9400_1:/dev/rxvm-200500a0b80cd5fe-001)
```

These devices include the controller WWN and LUN number in their name.

For more information, see "Verifying Access to XVM Volumes" on page 227 and the *XVM Volume Manager Administrator's Guide*.

## Troubleshooting for Mac OS X

This section discusses the following:

- "Verifying the Licenses for Mac OS X" on page 99
- "The `cxfs_client` Daemon is Not Started on Mac OS X" on page 100
- "XVM Volume Name is Too Long on Mac OS X" on page 101
- "Large Log Files on Mac OS X" on page 101

For general troubleshooting information, see Chapter 8, "General Troubleshooting" on page 231 and Appendix D, "Error Messages" on page 251

## Verifying the Licenses for Mac OS X

To verify that the licenses have been installed properly, use the `cxfslicense` command after installing the CXFS software. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example

- License installed properly:

```
macosx# /usr/cxfs_cluster/bin/cxfslicense
macosx# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_MACOSX license granted.
CXFS_MACOSX license granted.
```

- No license (error condition):

```
macosx# /usr/cxfs_cluster/bin/cxfslicense
cxfs license check failed - use 'cxfslicense -d' for details
macosx# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_MACOSX license granted.
```

```
Error on CXFS_MACOSX license check out.
checkout failed: No such feature exists
Feature:          CXFS_MACOSX
License path:    /etc/flexlm/license.dat
FLEXlm error:   -5,357
For further information, refer to the FLEXlm End User Manual,
available at "www.globetrotter.com".
```

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

```
Starting CXFS 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 '/usr/cxfs_cluster/bin/cxfs_cluster restart'.
```

An error such as the following example will appear in the SYSLOG file (line breaks added here for readability):

```
Jan 25 10:24:03 ncc1701:Jan 25 10:24:03 cxfs_client:
cis_main FATAL: cxfs_client failed the CXFS license check.
Use the cxfslicense command to diagnose the license problem
```

## The `cxfs_client` Daemon is Not Started on Mac OS X

Confirm that the `cxfs_client` is not running. The following command would list the `cxfs_client` process if it were running:

```
macosx# ps -auxww | grep cxfs_client
```

The `cxfs_client` daemon might not start if the license has expired, which is indicated when the following message is output to the `/var/log/cxfs_client` file:

```
CXFS not properly licensed for this host
```

In this case, execute the following command to determine why the license check fails:

```
macosx# /usr/cluster/bin/cxfslicense -d
```

If the license is valid, check the `cxfs_client` log file for other errors.

Restart `cxfs_client` as described in "Start/Stop `cxfs_client` for Mac OS X" on page 96 and watch the `cxfs_client` log file for errors.

## XVM Volume Name is Too Long on Mac OS X

On Mac OS X nodes, the following error message in the `system.log` file indicates that the volume name is too long and must be shortened so that the Mac OS X node can recognize it:

```
devfs: volumename name slot allocation failed (Errno=63)
```

See "Limitations and Considerations on Mac OS X" on page 81.

## Large Log Files on Mac OS X

The `/var/log/cxfs_client` log file may become quite large over a period of time if the verbosity level is increased.

To manually rotate this log file, use the `-z` option in the `/usr/cxfs_cluster/bin/cxfs_client.options` file.

See the `cxfs_client.options` man page and "Log Files on Mac OS X" on page 81.

## Reporting Mac OS X Problems

When reporting a problem about a CXFS Mac OS X node to SGI, you should gather the following information:

- Panic log: `/Library/Logs/panic.log` (if there is a system panic)
- CXFS client log: `/var/log/cxfs_client`
- System log: `/var/log/system.log`
- Console log: `/var/tmp/console.log`



## Solaris Platform

CXFS supports a client-only node running the Solaris operating system. This chapter contains the following sections:

- "CXFS on Solaris" on page 103
- "Installation and Configuration Overview for Solaris" on page 109
- "FLEXlm License for Solaris" on page 110
- "HBA Installation for Solaris" on page 111
- "Preinstallation Steps for Solaris" on page 125
- "Client Software Installation for Solaris" on page 131
- "I/O Fencing for Solaris" on page 133
- "Start/Stop `cxfs_client` for Solaris" on page 135
- "Maintenance for Solaris" on page 135
- "GRIO on Solaris" on page 137
- "XVM Failover Version 2 on Solaris" on page 137
- "Mapping XVM Volumes to Storage Targets on Solaris" on page 138
- "Troubleshooting for Solaris" on page 139
- "Reporting Solaris Problems" on page 142

### CXFS on Solaris

This section contains the following information about CXFS on Solaris:

- "Requirements to Solaris" on page 104
- "CXFS Commands on Solaris" on page 105
- "Log Files on Solaris" on page 106
- "CXFS Mount Scripts on Solaris" on page 106

- "Limitations and Considerations on Solaris" on page 106
- "Access Control Lists and Solaris" on page 107
- "System Tunables for Solaris" on page 108

## Requirements to Solaris

In addition to the items listed in "Requirements" on page 7, using a Solaris node to support CXFS requires the following:

- Solaris operating system:
  - Solaris 8 and patch 108528–22 (July 7, 2003) or later patch
  - Solaris 9 and patch 112233–06 (March 03) or later patch
- One to four AMCC JNI FibreStar FCE-6460-N (PCI) 2-Gbit or JNI FCX-6562-N 2 Gb 133 MHz PCI-X-to-Fibre Channel host bus adapters (HBAs). (1-Gbit HBAs and Sbus HBAs are not supported.)



**Caution:** The AMCC JNI HBA that is provided by Sun Microsystems **does not function with CXFS** and cannot be configured to do so. You must purchase the JNI HBA directly from AMCC.

SGI only supports the AMCC JNI HBA that is purchased directly from AMCC.

---

- One or more of the following Sun Microsystems hardware platform series:
  - Sun Blade 2000
  - Sun Fire 280R
  - Sun Fire V210
  - Sun Fire V250
  - Sun Fire V480
  - Sun Fire V880
  - Sun Fire 4800/4810 (PCI slots only, cPCI is not supported)
  - Sun Fire 6800 (PCI slots only, cPCI is not supported)

- Sun Fire 12K
- Sun Fire 15K
- Ultra Enterprise 220R
- Ultra Enterprise 250
- Ultra Enterprise 450
- Ultra Enterprise 3500
- Ultra Enterprise 4000
- Ultra Enterprise 4500
- Ultra Enterprise 5000
- Ultra Enterprise 5500
- Ultra Enterprise 6000
- Ultra Enterprise 6500
- Ultra Enterprise 10000

---

**Note:** CXFS supports a Solaris client only on the SPARC platforms listed here. It is not supported on other hardware platforms.

---

For the latest information, see the CXFS Solaris release notes.

## CXFS Commands on Solaris

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

```
/usr/cxfs_cluster/bin/cxfs_client  
/usr/cxfs_cluster/bin/cxfs_info  
/usr/cxfs_cluster/bin/cxfsdump  
/usr/cxfs_cluster/bin/cxfslicense  
/usr/sbin/grioadmin  
/usr/sbin/grioqos  
/usr/cxfs_cluster/bin/xvm
```

The `pkgadd` output lists all software added; see "Solaris Installation Overview" on page 131.

For more information, see the man pages.

For information about the GRIIO commands, see "Guaranteed-Rate I/O (GRIIO) and CXFS" on page 14 and "GRIIO on Solaris" on page 137.

## Log Files on Solaris

The `cxfs_client` command creates a `/var/log/cxfs_client` log file. To rotate this log file, use the `-z` option in the `/usr/cxfs_cluster/bin/cxfs_client.options` file; see the `cxfs_client` man page for details.

For information about the log files created on CXFS administration nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## CXFS Mount Scripts on Solaris

Solaris supports the CXFS mount scripts. See "CXFS Mount Scripts" on page 6 and the *CXFS Administration Guide for SGI InfiniteStorage*.

## Limitations and Considerations on Solaris

Note the following:

- 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.
- After a crash, attempts to reclaim locks and commit asynchronous writes to a CXFS filesystem from an NFS client may result in a stale file handle.
- For optimal performance, you should set the value of the Solaris system tunable parameter `maxphys` in the `/etc/system` file. Do the following:
  1. Make a backup copy of the `/etc/system` file.

---

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

---

2. Change the value of `maxphys` to `0x800000` (hexadecimal) in the `/etc/system` file.
3. Reboot the Solaris node. This causes the change to take effect.
4. Verify that the new value for `maxphys` is in effect by running the following command:

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

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

See also Appendix B, "Filesystem and Logical Unit Specifications" on page 245.

## Access Control Lists and Solaris

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

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

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

- **When using other backup/restore utilities:** A mask will be added to the ACL if the application calls `acl` for every file.

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

---

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

---

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

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

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

## System Tunables for Solaris

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

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

---

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

---

2. Change the value of `maxphys` to `0x800000` (hexadecimal) in the `/etc/system` file.
3. Reboot the Solaris node. This causes the change to take effect.
4. Verify that the new value for `maxphys` is in effect by running the following command:

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

## Installation and Configuration Overview for Solaris

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

1. Read the release notes for the Solaris platform to learn about any late-breaking changes in the installation procedure.
2. Install the Solaris 8 or Solaris 9 operating system according to the directions in the Solaris documentation (if not already done).
3. Install and verify the RAID. See the release notes.
4. Install and verify the switch. See the release notes.
5. Obtain and install the CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must purchase a mirroring software option and obtain and install a FLEXlm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See "FLEXlm License for Solaris" on page 110.
6. Install and verify the host bus adapter (HBA). See "HBA Installation for Solaris" on page 111.
7. Prepare the Solaris node, including adding a private network. See "Preinstallation Steps for Solaris" on page 125.
8. Install the CXFS software. See "Client Software Installation for Solaris" on page 131.

9. Configure the cluster to define the new Solaris node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration" on page 215.

If you run into problems, see "Troubleshooting for Solaris" on page 139 and Chapter 8, "General Troubleshooting" on page 231.

## FLEXIm License for Solaris

This section discusses the following:

- "Gathering the Solaris Host Information" on page 110
- "Obtaining the Licenses for Solaris" on page 110
- "Installing the Licenses for Solaris" on page 111

### Gathering the Solaris Host Information

When you order CXFS, you will receive an entitlement identifier (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.

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

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

For example:

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

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

## Obtaining the Licenses for Solaris

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.

## Installing the Licenses for Solaris

You will install the licenses in the following location:

```
/etc/flexlm/license.dat
```

Do the following:

1. Create the license directory if necessary.

For example:

```
solaris# mkdir -p /etc/flexlm
```

2. Copy the key to the license.dat file.

## HBA Installation for Solaris

This section provides an overview of the AMCC JNI Fibre Channel host bus adapter (HBA) installation and verification for Solaris nodes:

- "Installing the AMCC JNI HBA"
- "Setting the LUN Discovery Method for SGI TP9100" on page 113
- "Protecting Data Integrity" on page 114
- "Installing and Running the EZ Fibre Configuration GUI" on page 114
- "Verifying the JNI HBA Installation" on page 123

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.

---

**Note:** CXFS is not compatible with the firmware installed on a Sun OEM JNI HBA. JNI HBAs should be sourced from AMCC or an AMCC reseller.

---

## Installing the AMCC JNI HBA

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

To install the AMCC 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 AMCC JNI host bus adapter (HBA) into the Solaris system. Perform the steps in the following chapter:

- “Hardware Installation”

---

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

---

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

- “Unix Server DriverSuite”
  - “Solaris Driver”
    - “Verifying Hardware in OpenBoot PROM”
    - “Verifying Hardware in Solaris”

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



---

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

---

3. Install the AMCC JNI HBA driver software (JNIC146x) v5.3.x.x according to the instructions in the following “Installing” section:

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

Do the following:

- a. Retrieve the driver package from the following AMCC website:

<http://www.amcc.com/drivers>

- b. Install the driver package with the following command:

```
solaris# pkgadd -d package
```

Choose to install both the HBA driver and the SNIA HBA library.

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:

```
solaris# init 6
```

## Setting the LUN Discovery Method for SGI TP9100

If you are using an SGI TP9100 controller and the v5.3 AMCC JNI driver, you must set the LUN discovery method to 0 (meaning "SCSI Inquiry") in the `/kernel/drv/jnic146x.conf` file. Change the following line from:

```
# LunDiscoveryMethod = 1;
```

to:

```
LunDiscoveryMethod = 0;
```

Failure to do this will limit access to only LUN 0 on each target.

## Protecting Data Integrity

It is possible to put data integrity at risk if Fibre Channel cables are disconnected or fail. To protect against this issue, you must uncomment the `FailoverDelay` parameter in the `/kernel/drv/jnic146x.conf` file and change it to 0. It then should appear as follows:

```
FailoverDelay = 0;
```

This change ensures data integrity in the CXFS filesystem in the event when the Solaris client machine is unable to see the filesystems.

## Installing and Running the EZ Fibre Configuration GUI

After you have verified the installation of the HBA and the driver's attachment to it, you are ready to install and run the 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. Download the latest EZ Fibre GUI from the following website:

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

2. Install the EZ Fibre GUI:

- a. Extract the saved file using the `tar` command. For example:

```
solaris# tar xvf EZF_xxx.tar
```

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

```
solaris# cd EZF_xx
```

- c. Run the `install.sh` script:

```
solaris# ./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**.

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

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

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

```
solaris# ./ezf
```

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

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

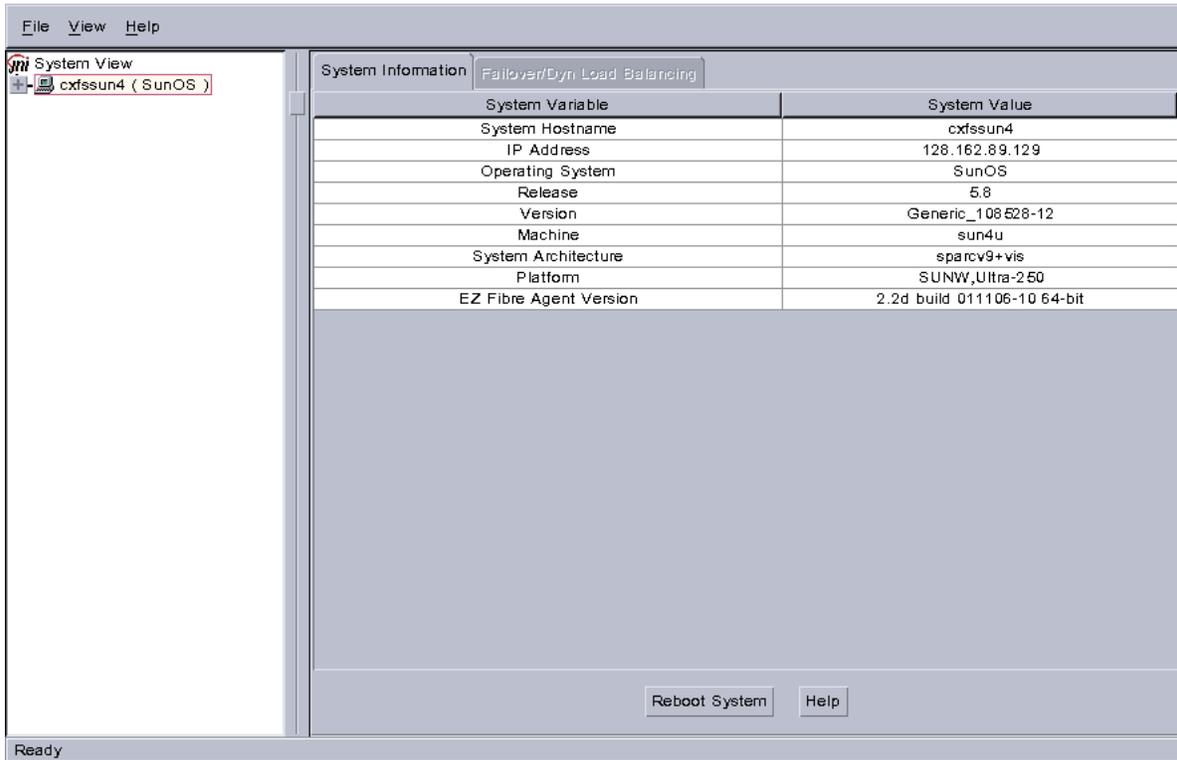
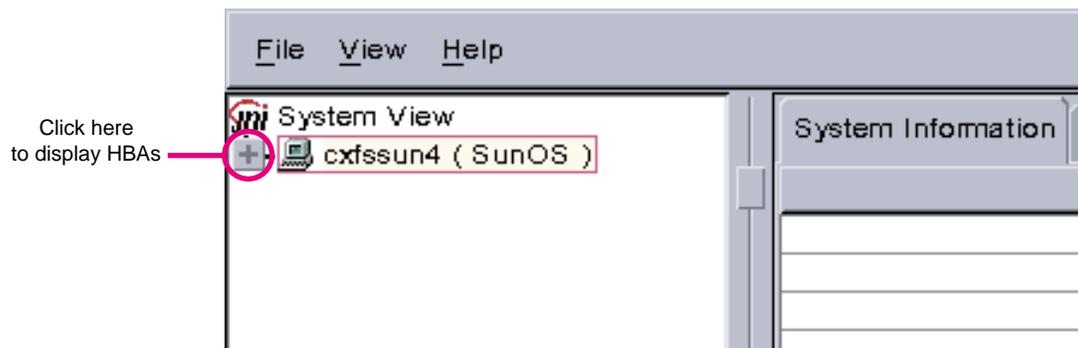


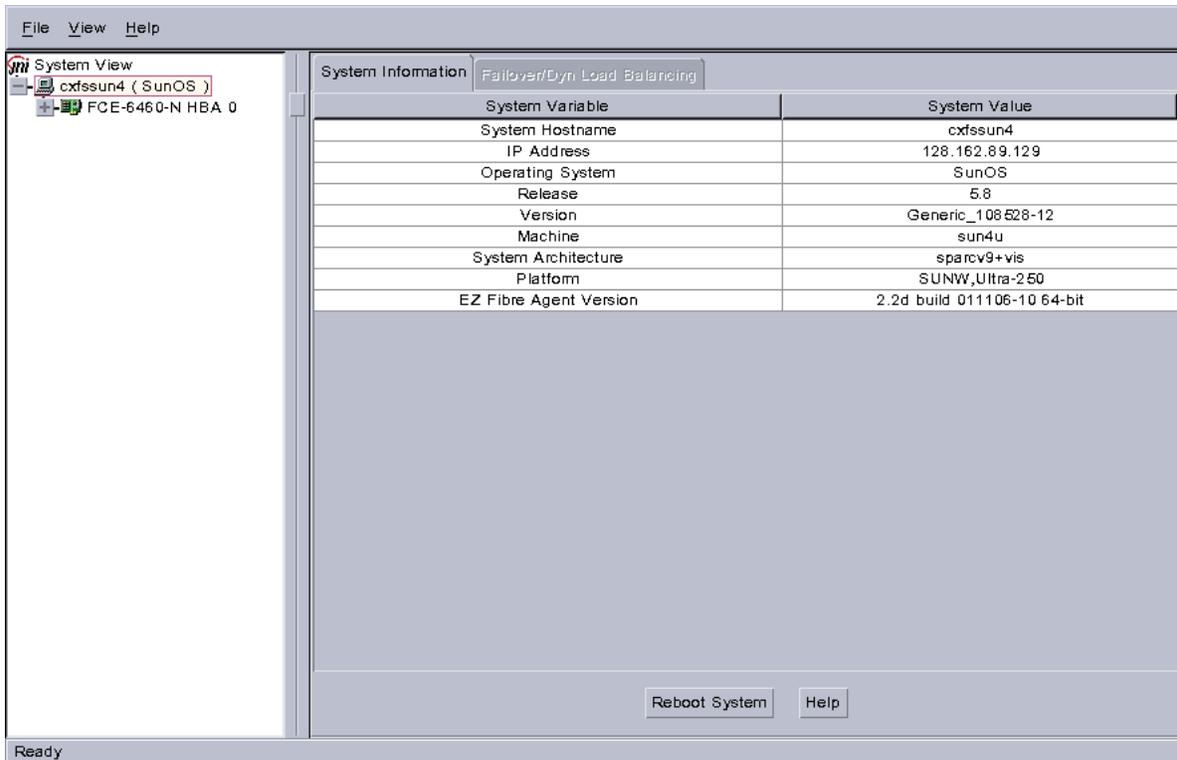
Figure 5-1 Example: Second Window: EZ Fibre Configuration Utility - Standalone

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.



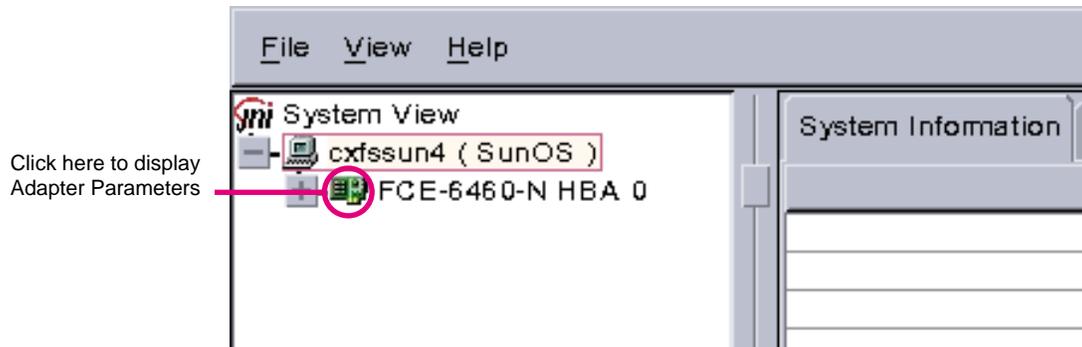
**Figure 5-2** Location of icon (+) to Display the HBA

Figure 5-3 shows an example of the display after clicking on the + sign for `cxfsun4`, which shows the JNI HBA.



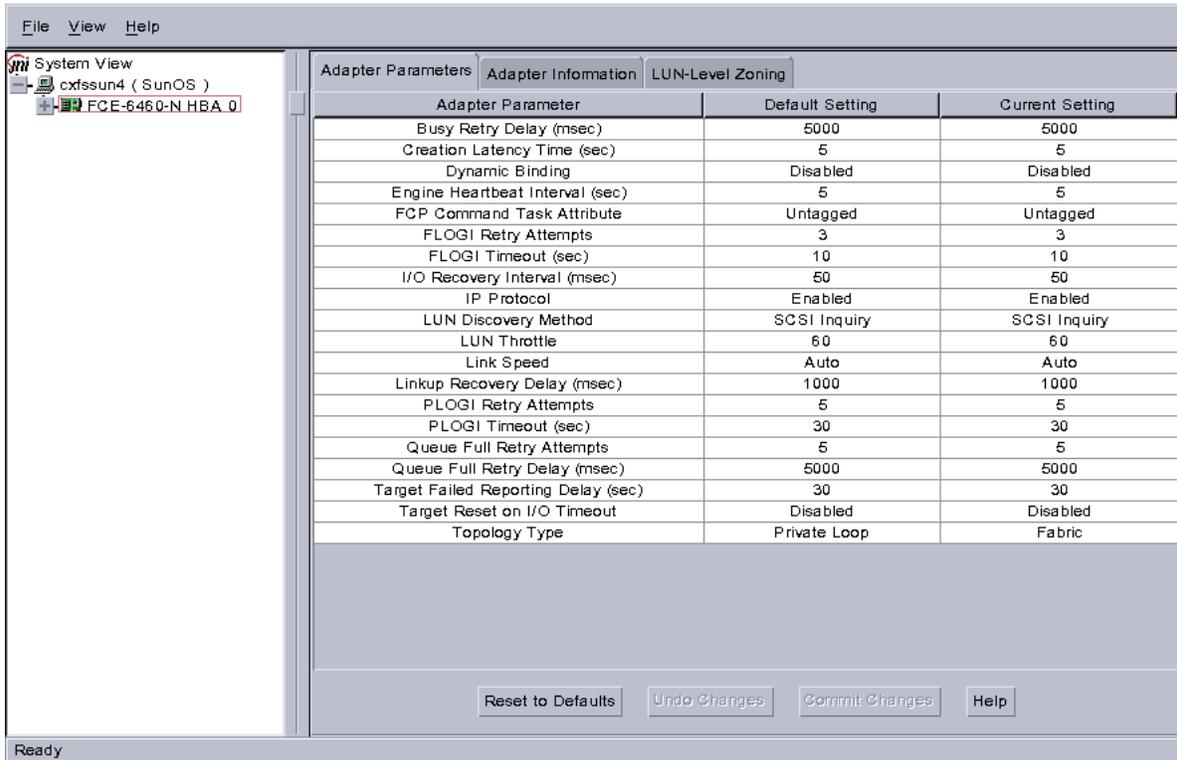
**Figure 5-3** Example: After Clicking + to Display the HBA

5. 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

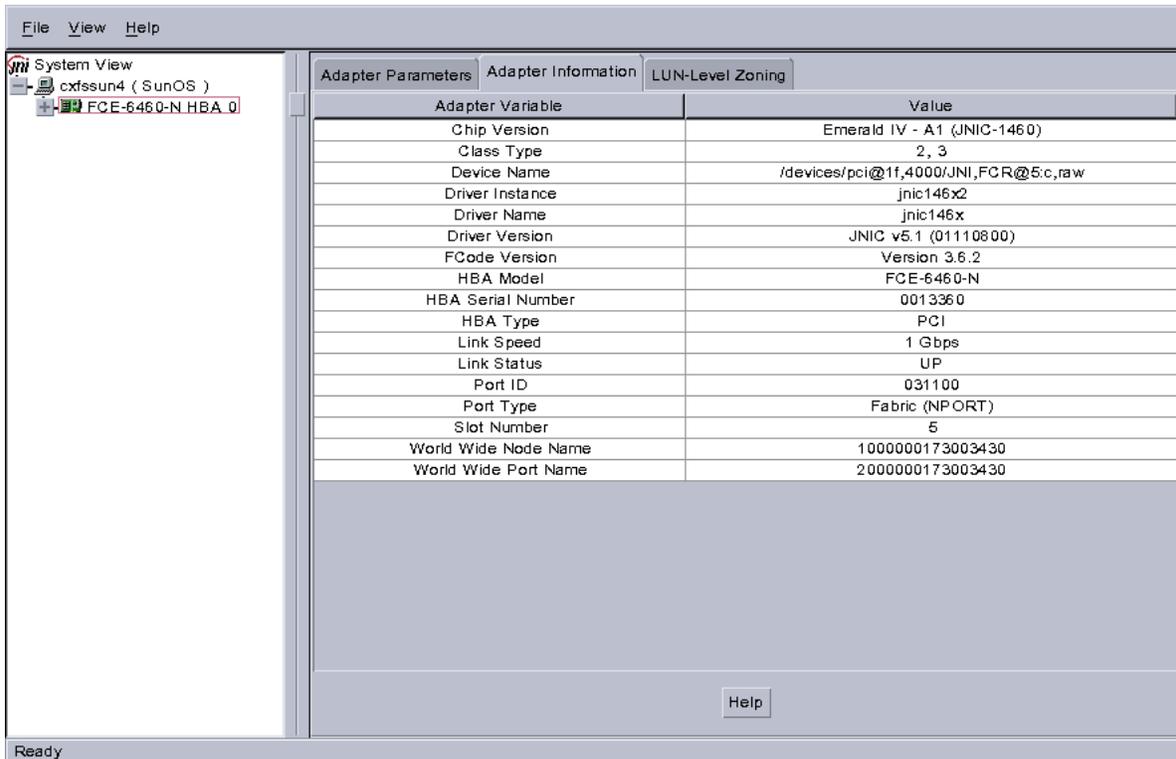
The right-hand pane will change to show **Adapter Parameters** for the selected HBA, as shown in Figure 5-5.



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

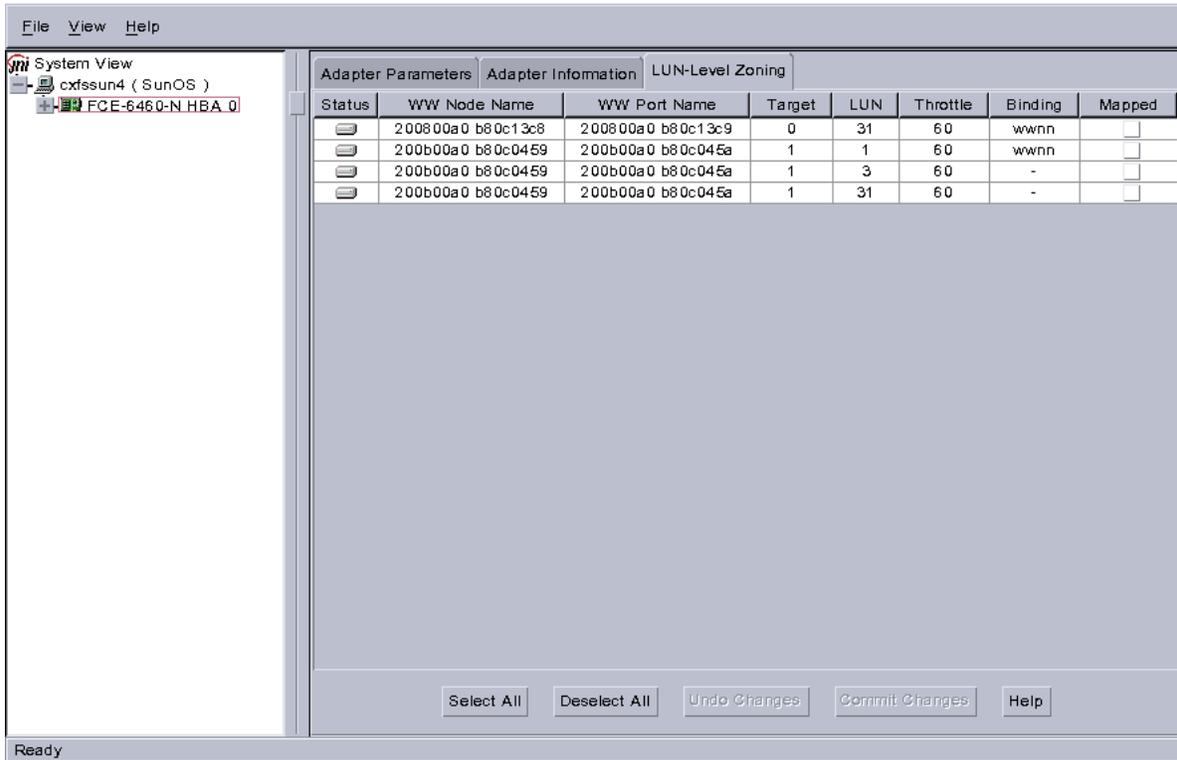
- a. Click on the **Adapter Information** tab to see the information in Figure 5-6.

The last two lines show the WWNN and WWPN of the JNI HBA. These may be used in the the `/etc/fencing.conf` file, if automatic detection does not work; see "No HBA WWPNs are Detected" on page 238.



**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.



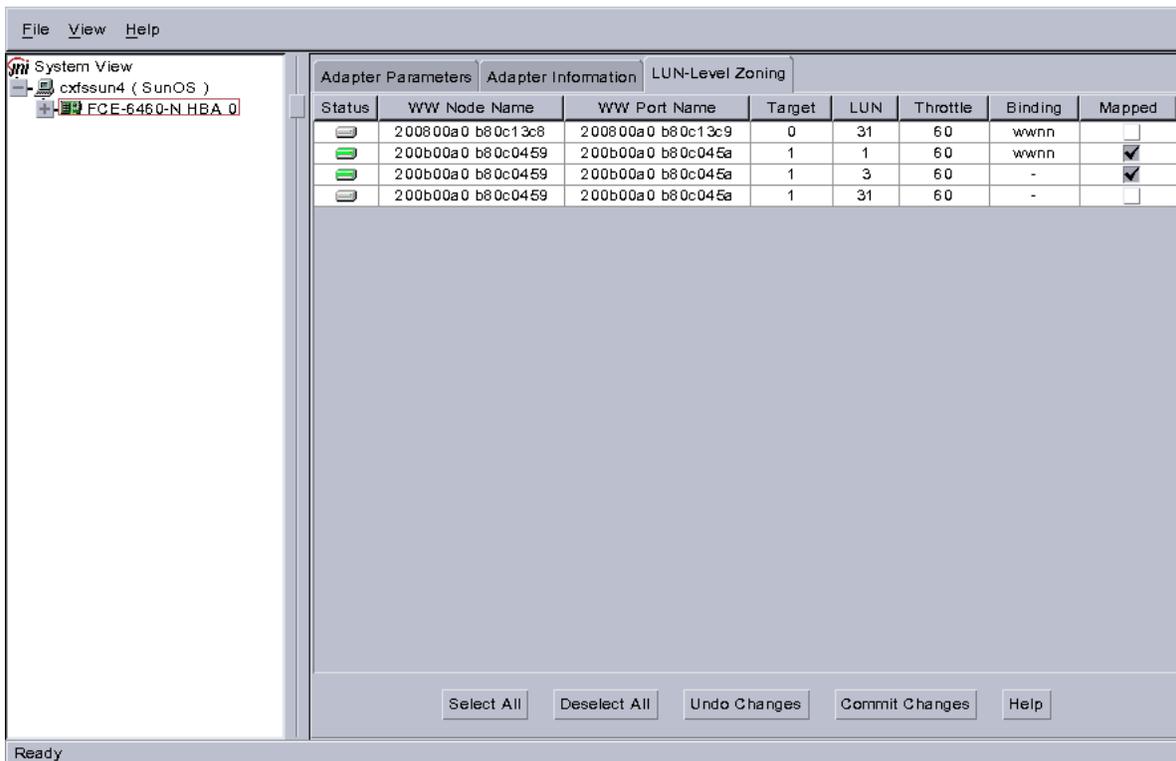
**Figure 5-7** After Clicking on LUN-Level Zoning

6. 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 TP9x00. 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

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

```
solaris# init 6
```

## Verifying the JNI HBA Installation

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

For example:

```
solaris# 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 "New Storage is Not Recognized on Solaris" on page 142.

## Preinstallation Steps for Solaris

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:

- "Adding a Private Network for Solaris Nodes" on page 125
- "Verifying the Private and Public Networks for Solaris" on page 130

### Adding a Private Network for Solaris Nodes

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

---

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

---

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

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

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

```
> boot -s
```

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

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

The `/etc/inet/ipnodes` 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/ipnodes` 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/ipnodes` 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/ipnodes` 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/ipnodes` on the other nodes in the cluster.

For more information, see the `hosts`, `named`, and `nis` 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" on page 13.

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

For example:

```
ipnodes:      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` command as follows:

```
solaris# 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:

```
solaris# 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 ffffffff 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` 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:

```
solaris# 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` command as follows, where *interface* is the value determined in step b:

```
ifconfig interface plumb
```

For example:

```
solaris# 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:

```
solaris# 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 ffffffff 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:

```
solaris# 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:

```
solaris# cat /etc/hostname.hme0
cxfssun3
solaris# 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 mode of the `.rhosts` file is set to 600 (read and write access for the owner only).

Make sure 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` 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` as user `root`.

For example, suppose you have a cluster with three nodes: `irix0`, `solaris1`, and `solaris2`. The `/.rhosts` files could be as follows (the prompt denotes the node name):

```
irix0# cat /.rhosts
solaris1 root
solaris1-priv root
solaris2 root
solaris2-priv root
```

```
solaris1# cat /.rhosts
irix0 root
irix0-priv root
solaris2 root
solaris2-priv root
```

```
solaris2# cat /.rhosts
irix0 root
irix0-priv root
solaris1 root
solaris1-priv root
```

8. Reboot the Solaris system:

```
solaris# 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 ffffffff 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 ffffffff broadcast 10.1.1.255
    ether 8:0:20:d2:29:c5
```

## Verifying the Private and Public Networks for Solaris

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

```
solaris# /usr/sbin/ping -s -c 3 nodeIPAddress
```

For example:

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

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

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

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

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

2. Verify that the cables are correctly seated.

Repeat this procedure on each node.

## Client Software Installation 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.

### Solaris 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. Read the `README` file for the Solaris platform to learn about any late-breaking changes in the installation procedure.
2. Verify that the node has been upgraded to Solaris 8 (also known as SunOS 5.8) or Solaris 9 (also known as SunOS 5.9) according to the Solaris installation guide. Use the following command to display the currently installed system:

```
solaris# uname -r
```

This command should return a value of 5.8 or 5.9.

3. Insert the *CXFS MultiOS Client 3.3* CD-ROM.
4. Read the already inserted CD-ROM as follows:

- Solaris 8:

```
solaris# pkgadd -d /cdrom/cdrom01/solaris/SGIcxfssol8.pkg
```

- Solaris 9

```
solaris# pkgadd -d /cdrom/cdrom01/solaris/SGIcxfssol9.pkg
```

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

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

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

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

5. Use the `cxfslicense` command to verify that the license is installed in `/etc/flexlm/license.dat`. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example, the following shows a properly installed license:

```
solaris# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_SOLARIS license granted.
CXFS_SOLARIS license granted.
```

If there are errors, verify that you have obtained and installed the CXFS licenses. For more information, see "Verifying the Linux Installation" on page 65.

6. If CXFS previously failed to load because of a license error, restart CXFS now.

## Verifying the Solaris Installation

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

```
pkginfo -l SGIcxfs
```

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

```
% pkginfo -l SGIcxfs
PKGINST:  SGIcxfs
NAME:     SGI CXFS MultiOS client software
```

```
CATEGORY: system
ARCH: sparc
VERSION: release 2.4
BASEDIR: /
VENDOR: Silicon Graphics Inc.
```

## I/O Fencing for Solaris

I/O fencing is required on Solaris nodes in order to protect data integrity of the filesystems in the cluster. The `cxfs_client` software automatically detects the world wide port names (WWPNs) of any supported host bus adapters (HBAs) for Solaris nodes that are connected to a switch that is configured in the cluster database. These HBAs are available for fencing.

However, if no WWPNs are detected, there will be messages about loading the HBA/SNIA library logged to the `/var/log/cxfs_client` file.

If no WWPNs are detected, you can manually specify the WWPNs in the fencing file.

---

**Note:** This method does not work if the WWPNs are partially discovered.

---

The `/etc/fencing.conf` file enumerates the WWPN for all of the HBAs that will be used to mount a CXFS filesystem. There must be a line for the HBA WWPN as a 64-bit hexadecimal number.

---

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

---

If used, `/etc/fencing.conf` must contain a simple list of WWPNs, one per line. You must update it whenever the HBA configuration changes, including the replacement of an HBA.

Do the following:

1. Set up the switch and HBA. See the release notes for supported hardware.
2. Follow the Fibre Channel cable on the back of the node to determine the port to which it is connected in the switch. Ports are numbered beginning with 0. (For example, if there are 8 ports, they will be numbered 0 through 7.)

3. Use the `telnet` command to connect to the switch and log in as user `admin`. (On Brocade switches, the password is `password` by default).
4. 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 fencing file).

5. Edit or create `/etc/fencing.conf` and add the WWPN for the port determined in step 2. (Comment lines begin with `#`.)

For dual-ported HBAs, you must include the WWPNs of any ports that are used to access cluster disks. This may result in multiple WWPNs per HBA in the file; the numbers will probably differ by a single digit.

For example, if you determined that port 0 is the port connected to the switch, your fencing file should contain the following:

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

6. After the node is added to the cluster (see Chapter 7, "Cluster Configuration" on page 215), enable the fencing feature by using the CXFS GUI or `cmgr` command on a server-capable administration node; for more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Start/Stop `cxfs_client` 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 `cxfs_client` daemon.

To start `cxfs_client` manually on your Solaris node, enter the following:

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

To stop `cxfs_client` manually, enter the following:

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

To stop and then start `cxfs_client` manually, enter the following:

```
solaris# /etc/init.d/cxfs_cluster restart
```

## Maintenance for Solaris

This section contains the following:

- "Upgrading the CXFS Software for Solaris" on page 136
- "Modifying the CXFS Software for Solaris" on page 136
- "Recognizing Storage Changes for Solaris" on page 137

## Upgrading the CXFS Software for Solaris

---

**Note:** Before upgrading CXFS software, ensure that no applications on the node are accessing files on a CXFS filesystem.

---

To upgrade CXFS on a Solaris system, do the following:

1. Remove the current package:

```
solaris# pkgrm SGIcxfs
```

2. Reboot the Solaris system:

```
solaris# reboot
```

3. Follow the installation instructions to install the new package. See "Client Software Installation for Solaris" on page 131.

## Modifying the CXFS Software for Solaris

You can modify the behavior of the CXFS client daemon (`/usr/cxfs_cluster/bin/cxfs_client`) by placing options in the `/usr/cxfs_cluster/bin/cxfs_client.options` file. The available options are documented in the `cxfs_client` man page.



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

---

The first line in the `cxfs_client.options` file must contain the options you want `cxfs_client` to process.

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

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

## Recognizing Storage Changes for Solaris

If you make changes to your storage configuration, you must rerun the HBA utilities to reprobe the storage.

See "HBA Installation for Solaris" on page 111.

## GRIO on Solaris

CXFS supports guaranteed-rate I/O (GRIO) version 2 on the Solaris platform. Application bandwidth reservations on Solaris must be explicitly released by the application before exit. If the application terminates unexpectedly or is killed, its bandwidth reservations are not automatically released and will cause a bandwidth leak.

For more general information about GRIO, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and the *Guaranteed-Rate I/O Version 2 Guide*.

## XVM Failover Version 2 on Solaris

Following is an example of the `/etc/failover2.conf` file on Solaris for a configuration in which each node has dual ports connecting to separate Fibre Channel (FC) switches. Each FC switch is connected to a single RAID controller with no interconnect between the FC switches:

```
# phys/d9400_0
pci@9,600000/JNI,FCR@2,1/sd@2,0 affinity=0 preferred
pci@9,600000/JNI,FCR@2/sd@2,0 affinity=1

# phys/d9400_1
pci@9,600000/JNI,FCR@2/sd@2,1 affinity=0 preferred
pci@9,600000/JNI,FCR@2,1/sd@2,1 affinity=1

# phys/wally
pci@9,600000/JNI,FCR@2/sd@2,2 affinity=0 preferred
pci@9,600000/JNI,FCR@2,1/sd@2,2 affinity=1

# phys/recovery
pci@9,600000/JNI,FCR@2,1/sd@2,3 affinity=0 preferred
pci@9,600000/JNI,FCR@2/sd@2,3 affinity=1
```

For more information, see "XVM Failover Version 2 and CXFS" on page 17, the comments in the `/etc/failover2.conf` file on a CXFS administration node, and the *XVM Volume Manager Administrator's Guide*.

## Mapping XVM Volumes to Storage Targets on Solaris

XVM volumes can be mapped to devices and to targets on RAID controllers using the output from the `xvm` command, the EZ Fibre utility, and the device entries in the filesystem.

You can use the `xvm` command to display the device names:

```
solaris# /usr/cxfs_cluster/bin/xvm show -e -t vol
vol/stripel                0 online,open
  subvol/stripel/data       2292668416 online,open
    stripe/stripel         2292668416 online,open (unit size: 1024)
      slice/d9400_0s0      1146334816 online,open
(d9400_0:pci@9,600000/JNI,FCR@2,1/sd@2,0)
      slice/d9400_1s0      1146334816 online,open
(d9400_1:pci@9,600000/JNI,FCR@2/sd@2,1)
```

These devices are created in the `/devices` directory (note the `*:c,raw`):

```
solaris# ls /devices/pci@9,600000/JNI,FCR@2,1/sd@2,0*:c,raw
/devices/pci@9,600000/JNI,FCR@2,1/sd@2,0:c,raw
```

These devices are linked from the disk devices under `/dev/rdisk`:

```
solaris# ls -l /dev/rdisk/c*s2 /dev/dsk/c*s2 | grep
'pci@9,600000/JNI,FCR@2,1/sd@2,0'
lrwxrwxrwx  1 root    root          51 May 14  2004 /dev/rdisk/c2t2d0s2
-> ../../devices/pci@9,600000/JNI,FCR@2,1/sd@2,0:c,raw
```

Using the EZ Fibre utility, it is possible to find the LUN and therefore the target WWN by exploring the LUNs to find the matching `/dev/rdisk` entry. See "Installing and Running the EZ Fibre Configuration GUI" on page 114.

For more information, see "Verifying Access to XVM Volumes" on page 227 and the *XVM Volume Manager Administrator's Guide*.

## Troubleshooting for Solaris

This section contains the following:

- "Verifying the Licenses for Solaris" on page 139
- "The `cxfs_client` Daemon is Not Started on Solaris" on page 140
- "Filesystems Do Not Mount on Solaris" on page 141
- "New Storage is Not Recognized on Solaris" on page 142
- "Large Log Files on Solaris" on page 142

For general troubleshooting information, see Chapter 8, "General Troubleshooting" on page 231 and Appendix D, "Error Messages" on page 251

### Verifying the Licenses for Solaris

To verify that the licenses have been installed properly, use the `cxfslicense` command after installing the CXFS software. If the license is installed properly, there will be no output. If it is not installed, there will be errors.

For more verbose output, use the `-d` option. For example

- License installed properly:

```
solaris# /usr/cxfs_cluster/bin/cxfslicense
solaris# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_SOLARIS license granted.
CXFS_SOLARIS license granted.
solaris#
```

- No license (error condition):

```
solaris# /usr/cxfs_cluster/bin/cxfslicense
cxfs license check failed - use 'cxfslicense -d' for details
solaris# /usr/cxfs_cluster/bin/cxfslicense -d
no xvm user license required.
XVM_SOLARIS license granted.
Error on CXFS_SOLARIS license check out.
checkout failed: No such feature exists
Feature:          CXFS_SOLARIS
```

```
License path: /etc/flexlm/license.dat
FLEXlm error: -5,357
For further information, refer to the FLEXlm End User Manual,
available at "www.globetrotter.com".
solaris#
```

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

```
Starting CXFS 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 '/usr/cxfs_cluster/bin/cxfs_cluster restart'.
```

An error such as the following example will appear in the SYSLOG file (line breaks added here for readability):

```
Jan 25 10:24:03 ncc1701:Jan 25 10:24:03 cxfs_client:
cis_main FATAL: cxfs_client failed the CXFS license check.
Use the cxfslicense command to diagnose the license problem
```

### The `cxfs_client` Daemon is Not Started on Solaris

Confirm that the `cxfs_client` is not running. The following command would list the `cxfs_client` process if it were running:

```
solaris# ps -ef | grep cxfs_client
```

The `cxfs_client` daemon might not start if the license has expired, which is indicated when the following message is output to the `/var/log/cxfs_client` file:

```
CXFS not properly licensed for this host
```

In this case, execute the following command to determine why the license check fails:

```
linux# /usr/cluster/bin/cxfslicense -d
```

If the license is valid, check the `cxfs_client` log file for other errors.

Restart `cxfs_client` as described in "Start/Stop `cxfs_client` for Linux" on page 67 and watch the `cxfs_client` log file for errors.

## Filesystems Do Not Mount on Solaris

If `cxfs_info` reports that `cms` is up but XVM or the filesystem is in another state, then one or more mounts is still in the process of mounting or has failed to mount.

The CXFS node might not mount filesystems for the following reasons:

- The client may not be able to see all the LUNs. This is usually caused by misconfiguration of the HBA or the SAN fabric:
  - Can the HBA see all of the LUNs for the filesystems it is mounting?
  - Can the operating system kernel see all of the LUN devices? Verify that the LUNs used by this filesystem have been mapped in the EZ Fibre configuration utility. See "Installing and Running the EZ Fibre Configuration GUI" on page 114.

See "New Storage is Not Recognized on Solaris" on page 142.

- The `cxfs_client` daemon may not be running. See "The `cxfs_client` Daemon is Not Started on Solaris" on page 140.
- The filesystem may have an unsupported mount option. Check the `cxfs_client.log` for mount option errors or any other errors that are reported when attempting to mount the filesystem.
- The cluster membership (`cms`), XVM, or the filesystems are not up on the node. Execute the `/usr/cluster/bin/cxfs_info` command to determine the current state of `cms`, XVM, and the filesystems. If the node is not up for each of these, then check the `/var/log/cxfs_client` log to see what actions have failed.

Do the following:

- If `cms` is not up, check the following:
  - Is the node is configured on the administration node with the correct hostname?
  - Has the node been added to the cluster and enabled? See "Verifying the Cluster Status" on page 223.
- If XVM is not up, check that the HBA is active and can see the LUNs.
- If the filesystem is not up, check that one or more filesystems are configured to be mounted on this node and check the `/var/log/cxfs_client` file for mount errors.

## New Storage is Not Recognized on Solaris

If you have a problem with an HBA, verify that you enabled fabric mode. See step 4 in "Installing the AMCC JNI HBA" on page 112.

Also see "Recognizing Storage Changes for Solaris" on page 137.

## Large Log Files on Solaris

The `/var/log/cxfs_client` log file may become quite large over a period of time if the verbosity level is increased. To manually rotate this log file, use the `-z` option in the `/usr/cxfs_cluster/bin/cxfs_client.options` file.

For information about the log files created on CXFS administration nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Reporting Solaris Problems

When reporting a problem about a CXFS Solaris node to SGI, you should retain the following information:

- If there is a system panic, retain the system core file in `/var/crash/hostname` on a Solaris node.
- Output from the `crash` utility.
- `mdb(1M)` modular debugger output:
  - For panics or generated dumps, use the following commands and save the output:
    - `$c` (or `$C`)
    - `$r`
    - `$<msgbuf`
  - For dumps from hangs:
    - `$<threadlist`
    - `$c` (or `$C`)
    - `$r`

- `$<msgbuf`
- A list of the installed CXFS packages. Use the `pkginfo` command as follows:  

```
# pkginfo -l SGIcxf
```
- A list of the Solaris patches that have been installed. Use the `showrev` 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` command.
- Output about the cluster obtained from the `cxfsdump` utility run on a CXFS administration node. When run in local mode on a Solaris node, it stores information in `/var/cluster/cxfsdump-data/nodename.tar.gz`.

If any of the above Solaris tools are not currently installed on your Solaris system, you should install them.



## Windows Platforms

CXFS supports a client-only node running the Windows 2000, Windows Server 2003, or Windows XP operating system. The information in this chapter applies to all of these versions of Windows unless otherwise noted.

This chapter contains the following sections:

- "CXFS on Windows" on page 146
- "Windows Installation and Configuration Overview" on page 174
- "FLEXlm License for Windows" on page 175
- "HBA Installation for Windows" on page 177
- "Preinstallation Steps for Windows" on page 181
- "Client Software Installation for Windows" on page 185
- "Postinstallation Steps for Windows" on page 191
- "I/O Fencing on Windows" on page 193
- "Start/Stop the CXFS Client Service for Windows" on page 196
- "Maintenance for Windows" on page 196
- "GRIO on Windows" on page 201
- "Troubleshooting for Windows" on page 202
- "Reporting Windows Problems" on page 211

---

**Note:** Your Windows XP or Windows Server 2003 **Start** menu may differ from the examples shown in this guide, depending upon your start menu preferences. For example, this guide describes selecting the control panel as follows:

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

However, on your system this menu could be as follows:

**Start**  
    > **Control Panel**

---

## CXFS on Windows

This section contains the following information about CXFS on Windows:

- "Requirements for Windows" on page 147
- "CXFS Commands on Windows" on page 148
- "Log Files and Cluster Status for Windows" on page 149
- "Functional Limitations for Windows" on page 153
- "Performance Considerations for Windows" on page 155
- "Access Controls for Windows" on page 157
- "System Tunables for Windows" on page 168

## Requirements for Windows

In addition to the items listed in "Requirements" on page 7, CXFS requires at least the following:

- An Intel Pentium or compatible processor
- Minimum RAM requirements (more will improve performance):
  - Windows 2000: 128 MB
  - Windows XP, Windows Server 2003: 256 MB
- A minimum of 10 MB of free disk space
- A Qlogic QLA2200, QLA2310, QLA2342, or QLA2344 host bus adapter (HBA)
- The following QLogic software from the <http://www.qlogic.com> website:
  - QLA2200:
    - Windows 2000, Windows Server 2003: v8.1.5.15
    - Windows XP: v8.1.5.12
  - QLA2310, QLA2342 and QLA2344:
    - Windows 2000, Windows XP, Windows Server 2003: v9.01.12
  - SANblade Agent and Manager 2.0.30
  - QLDirect (required for HBA failover):
    - Windows 2000 and Windows Server 2003: v8.01.12
    - Windows XP: (not supported)

---

**Note:** At the time of this release, Qlogic did not support Windows XP. SGI has successfully tested the Windows Server 2003 drivers under Windows XP, with the exception of the QLDirect driver that is required for HBA failover. SGI is attempting to rectify these issues with QLogic.

---

You should install the documentation associated with the software. See the SANblade Manager README for the default password. Follow the QLogic instructions to install the driver, the SANblade NT Agent, and the SANblade Manager software.

- If two QLogic HBAs are installed for Windows 2000 or Windows Server 2003, you should also install the QLogic Direct Filter (8.01.12) in order to facilitate HBA failover and load balancing. If two different model HBAs are installed, you must install drivers for both models.

---

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

---

- Windows versions:
  - Windows 2000 Service Pack 3 or Service Pack 4
  - Windows XP Service Pack 1 or Service Pack 2
  - Windows Server 2003

---

**Note:** CXFS will not work on 64-bit versions of Windows.

---

For the latest information, see the CXFS Windows release notes.

## CXFS Commands on Windows

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

The command `%ProgramFiles%\CXFS\cxfslicense.exe` assists with license validation. See "Verifying the Licenses for Windows" on page 203.

The command `%ProgramFiles%\CXFS\cxfs_info.exe` displays the current state of the node in the cluster in a graphical user interface. See "Log Files and Cluster Status for Windows" and "Verifying the Cluster Status" on page 223.

For information about the GRIO commands, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and "GRIO on Windows" on page 201.

## Log Files and Cluster Status for Windows

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

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

For information about the log files created on CXFS administration nodes, see the *CXFS Administration Guide for SGI InfiniteStorage*. The CXFS Client service will also log important information to the following file:

```
%ProgramFiles%\CXFS\log\cxfs_client.log
```

When CXFS is first installed, the log file is automatically rotated when it grows to 10 MB. This is set by the `-z` option in the CXFS Client service **Additional arguments** window during installation (see Figure 6-6 on page 187) and may be adjusted by following the steps described in "Modifying the CXFS Software for Windows" on page 197.

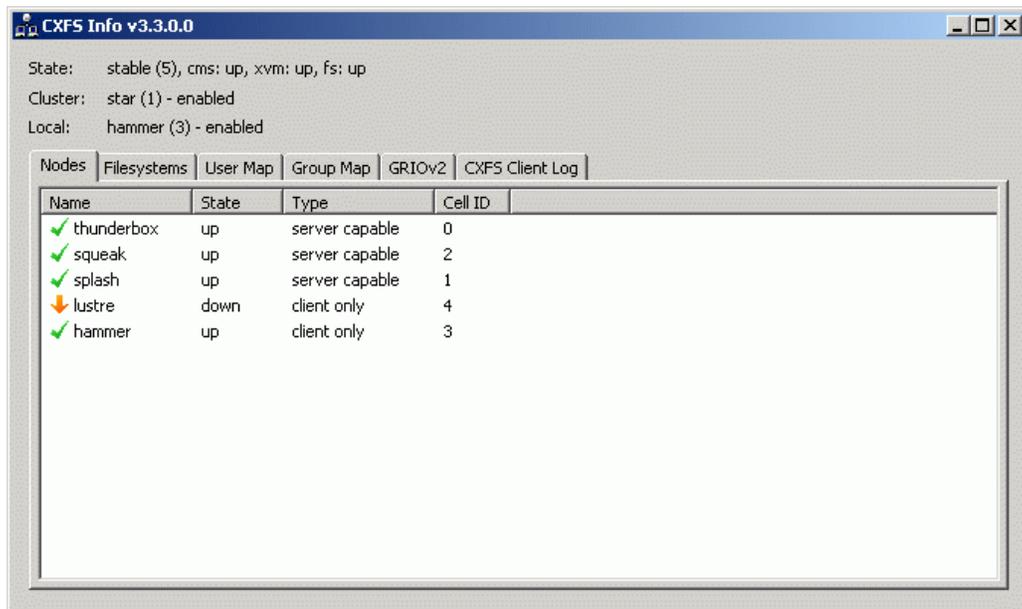
You may also wish to keep the **CXFS Info** window open to check the cluster status and view the log file. To open this informational window on any Windows system, select the following:

```
Start
  > Programs
      > CXFS
          > CXFS Info
```

The top of **CXFS Info** window displays the overall state of the cluster environment:

- Number of stable nodes
- Status of the cms cluster membership daemon
- Status of XVM
- Status of filesystems
- Status of the cluster
- Status of the local node

Figure 6-1 shows an example of the **CXFS Info** window.



**Figure 6-1** CXFS Info Window — Nodes Tab Display

The **CXFS Info** window also provides the following tabs to access further information:

- **Nodes** displays each node in the cluster, its state, and its cell ID number. For more information, see "Verifying the Cluster Status" on page 223.

- **Filesystems** displays each CXFS filesystem, its state, size, and other statistics. Figure 6-2 shows an example.

State: stable (5), cms: up, xvm: up, fs: up  
Cluster: star (1) - enabled  
Local: hammer (3) - enabled

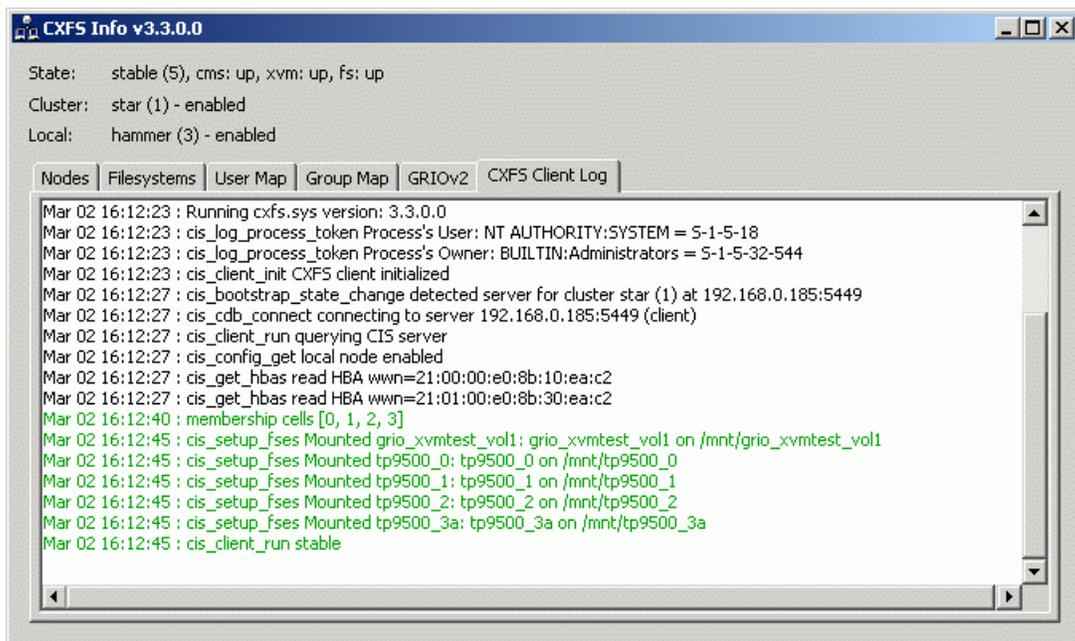
Name	State	Total size	Free space	Use%	Mount point	Mount option
✓ tp9500_3a	mounted	67.9 GB	61.6 GB	9%	/mnt/tp9500_3a	
✓ tp9500_2	forced mounted	135 GB	98.1 GB	28%	/mnt/tp9500_2	rw,inode64,r
✓ tp9500_1	forced mounted	543 GB	19.8 GB	96%	/mnt/tp9500_1	dmapi
✓ tp9500_0	mounted	543 GB	280 GB	48%	/mnt/tp9500_0	
⊗ grio_xvmtest...	disabled				/mnt/grio_xvmtest_vol2	
✓ grio_xvmtest...	forced mounted	1.38 GB	1.25 GB	9%	/mnt/grio_xvmtest_vol1	

**Figure 6-2** CXFS Info Window — Filesystems Tab

- **User Map** displays the usernames that are mapped to UNIX user identifiers.
- **Group Map** displays the groups that are mapped to UNIX group identifiers.
- **GRIOv2 Status** displays each guaranteed-rate I/O (GRIO) stream, its reservation size, and other statistics. See "GRIO on Windows" on page 201.
- **CXFS Client log** displays the log since the CXFS Client service last rebooted. It highlights the text in different colors based on the severity of the output:
  - Red indicates an error, which is a situation that will cause a problem and must be fixed
  - Orange indicates a warning, which is a situation that might cause a problem and should be examined

- Black indicates general log information that can provide a frame of reference
- Green indicates good progress in joining membership and mounting filesystems

Figure 6-3 shows an example.



**Figure 6-3** CXFS Info Window — CXFS Client Log Tab

The **CXFS Info** icon in the task bar will change from green to yellow or red depending on the state of the client in the cluster:

- Green indicates that the client is in the membership, everything is fully functional, and all enabled filesystems are mounted
- Yellow indicates an in-between state (neither inactive nor stable state)
- Red indicates that CXFS is not running (inactive state)

Also see Figure 6-13 on page 201.

## Functional Limitations for Windows

There are a number of limitations in the CXFS software that are unique to the Windows platform. See also Appendix B, "Filesystem and Logical Unit Specifications" on page 245.

### XVM Failover Version 2 Is Not Supported

Windows does not support XVM failover version 2. However, it does support a mechanism to configure the automatic failover to another HBA port if the configured port is no longer able to see the target. See "Configuring HBA Failover for Windows 2000 and Windows 2003" on page 180.

### UNIX Perspective of CXFS for Windows

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

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

The top-level file structure under the CXFS drive letter consists of an in-memory directory structure that mimics the mount points on the CXFS administration node. The CXFS software creates these directories before mounting the CXFS filesystems. For example, a CXFS filesystem with a mount point of `/mnt/cxfs` on a CXFS Windows node configured to use drive letter `X`, will create `X:\mnt\cxfs` during filesystem mount process.

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, but an application that must write to the directory directly under the CXFS drive letter will fail.

---

**Note:** A CXFS mount point or directory beneath a mount point can be mapped to another drive letter by using the `subst` command from a command shell to which the application can write. See "Application Cannot Create File Under CXFS Drive Letter" on page 211.

---

- A Windows node can support regular files, directories, and links. However, it does not support other XFS file types.
- Symbolic links cannot be distinguished from normal files or directories on a Windows node. Opening a symbolic link will open the target of the link, or will report `file not found` if it is a dangling link.
- You can move, rename, or delete a symbolic link; however, you cannot copy a symbolic link. Copying a valid symbolic link will result in copying the file or directory that the link refers to, rather than the normal UNIX behavior that copies the link itself.

### Windows Perspective of CXFS for Windows

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

- Avoid using duplicate filenames in the same directory that vary only in case. CXFS is case-sensitive, but some Windows applications may not maintain the case of all filenames, which may result in unexpected behavior.
- CXFS software does not export 8.3 alternative filenames. Older Windows applications that only support 8.3 filenames may be unable to open files with longer filenames.
- Avoid using completely uppercase 8.3 filenames. If you use completely uppercase 8.3 filenames, some applications (including Windows Explorer) may incorrectly assume that only 8.3 filenames are supported by the filesystem and will not preserve case.
- Install the CXFS software components onto a NTFS partition rather than a FAT partition. The security of the following files cannot be guaranteed if these files are installed onto a FAT filesystem:

```
%ProgramFiles%\CXFS\passwd  
%ProgramFiles%\CXFS\group
```

- There is no recycle bin; deleted files are permanently deleted.
- There is no automatic notification of directory changes performed by other nodes in the cluster. Applications (such as Windows Explorer) will not automatically update their display if another node adds or removes files from the directory currently displayed.

- A CXFS filesystem cannot be used as the boot partition of a Windows node.
- The volume properties window in Windows Explorer for the CXFS drive letter will display the total capacity of all mounted filesystems and the largest free space on any one of those filesystems.

### Forced Unmount on Windows

SGI recommends that you enable the forced unmount feature on CXFS filesystems. See "Recommendations" on page 9 and "Forced Unmount of CXFS Filesystems" on page 221.

A forced unmount causes all processes that have open files on the specified filesystem to be unconditionally killed and therefore permit the filesystem to be unmounted without delay.

### Define LUN 0 on All Storage Devices for Windows

Windows, and therefore CXFS, may not detect any LUNs on a storage device if LUN 0 is not defined on the storage device. This problem may occur when **CXFS Info** reports that XVM is up, but one or more filesystems are not mounted and CXFS therefore retries the mount continuously. This problem has occurred on Windows XP CXFS clients. For more information about this issue, see:

<http://support.microsoft.com/default.aspx?scid=kb;en-us;821666>

### Memory-Mapping Large Files for Windows

You can memory-map a file much larger than 2 GB under Windows, but only up to 2 GB of that file in one or more parts can be mapped into a process at any one time on a 32-bit platform. See the Windows Platform Software Development Kit for more details.

### CXFS Mount Scripts for Windows

Windows does not support the CXFS mount scripts.

### Performance Considerations for Windows

The following are performance considerations on a CXFS Windows node, in addition to the limitations described in "Performance Considerations" on page 4:

- Using CIFS to share a CXFS filesystem from a CXFS Windows node to another Windows host is not recommended for the following reasons:
  - Metadata operations sent to the Windows node must also be sent to the CXFS metadata server causing additional latency
  - CXFS Windows does not support opportunistic locking, which CIFS uses to improve performance

SGI recommends that you use Samba on the CXFS metadata server to export CXFS filesystems to other nodes that are not running CXFS.
- Windows supplies autonotification APIs for informing applications when files or directories have changed on the local client. With each notification, Windows Explorer will do a full directory lookup. Under CXFS, directory lookups can require multiple RPCs to the server (about 1 per 30 files in the directory), resulting in a linear increase in network traffic. This can grow to megabytes per second for directories with large numbers of files. For better performance, do one of the following:
  - Select the destination folder itself
  - Close the drive tree or mount point folder by clicking on the |+| on the drive icon or mount point folder
- 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` command. This can be an expensive operation, especially if performed on directories between the drive letter and the mount points, because it will traverse all mounted filesystems.
- Virus scanners, Microsoft Find Fast, and similar tools that traverse a filesystem are very expensive on a CXFS filesystem. Such tools should be configured so that they do not automatically traverse the CXFS drive letter.
- The mapping from Windows user and group names to UNIX identifiers occurs as the CXFS software starts up. In a Windows domain environment, this process can take a number of seconds per user for usernames that do not have accounts within the domain. If you are using a `passwd` file for user identification and the file contains a number of unknown users on the Windows node, you should remove users who do not have accounts on the Windows nodes from the `passwd` file that is installed on the Windows nodes.

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

- With 1-GB fabric to a single RAID controller, it is possible for one 32-bit 33-MHz QLogic card to reach the bandwidth limitations of the fabric, and therefore there will be no benefit from load balancing two HBAs in the same PCI bus. This can be avoided by using 2-GB fabric and/or multiple RAID controllers.
- For load balancing of two HBAs to be truly beneficial, the host must have at least one of the following three attributes:
  - A 64-bit PCI bus
  - A 66-MHz PCI bus
  - Multiple PCI buses
- Applications running on a CXFS Windows client should perform well when their I/O access patterns are similar to those described in "When to Use CXFS" on page 2.
- The maximum I/O size issued by the QLogic HBA to a storage target and the command tag queue length the HBA maintains to each target can be configured in the registry. See "System Tunables for Windows" on page 168.

## Access Controls for Windows

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

- "User Identification for Windows" on page 158
- "User Identification Mapping Methods for Windows" on page 159
- "Enforcing Access to Files and Directories for Windows" on page 160
- "Viewing and Changing File Attributes with Windows Explorer" on page 161
- "Viewing and Changing File Permissions with Windows Explorer" on page 162

- "Viewing and Changing File Access Control Lists (ACLs) for Windows" on page 164
- "Effective Access for Windows" on page 165
- "Restrictions with file ACLs for Windows" on page 165
- "Inheritance and Default ACLs for Windows" on page 166

### User Identification for Windows

The CXFS software supports several user identification mechanisms, which are described in "User Identification Mapping Methods for Windows" on page 159. Windows user and group names that match entries in the configured user list will be mapped to those user IDs (UIDs) and group IDs (GIDs).

The following additional mappings are automatically applied:

- **User Administrator** is mapped to `root` (UID = 0)
- **Group Administrators** is mapped to `sys` (GID = 0)

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

You can display the users and groups that have been successfully mapped by looking at the tables for the **User Map** and **Group Map** tabs in the **CXFS Info** window.

The following sections assume that a CXFS Windows node was configured with the following `passwd` and `group` files:

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

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

## User Identification Mapping Methods for Windows

User identification can be performed by one choosing one of the following methods for the **User ID mapping lookup sequence** item of the **Enter CXFS Details** window:

- **files:** /etc/passwd and /etc/group files from the metadata server copied onto the clients. If you select this method, you must install the /etc/passwd and /etc/group files immediately after installing the CXFS software, as described in "Performing User Configuration for Windows" on page 192.
- **ldap\_actedir:** Windows Active Directory server with Services for UNIX (SFU) installed, which uses lightweight directory access protocol (LDAP).

The **ldap\_actedir** method configures the CXFS Windows software to communicate with the Active Directory for the CXFS node's domain. With the Windows Services for UNIX (SFU) extensions, the Active Directory User Manager lets you define UNIX identifiers for each user and export these identifiers as an LDAP database.

Permissions on the Active Directory server must allow Authenticated Users to read the SFU attributes from the server. Depending on the installation and configuration of the server, LDAP clients may or may not be able to access the SFU attributes. For more information, see "CXFS Client Service Cannot Map Users other than Administrator for Windows" on page 208.

This configuration requires a domain controller that is installed with the following:

- Windows 2000 Server or Windows 2003 Server with Active Directory.
- Windows Services for UNIX (SFU) version 2 or later with the NFS server component installed. SGI recommends SFU version 3.5.

---

**Note:** The domain controller does not have to be a CXFS node.

---

- **ldap\_generic:** Generic LDAP lookup for UNIX users and groups from another LDAP server.

The **ldap\_generic** method configures the CXFS software to communicate with an LDAP database that maps user names and group names to UNIX identifiers.

For an example of the window, see Figure 6-6 on page 187.

You must select one of these as the primary mapping method during installation, but you can change the method at a later time, as described in "Modifying the CXFS Software for Windows" on page 197.

Optionally, you can select a secondary mapping method that will be applied to users that are not covered by the first method. If you choose a primary and a secondary mapping method, one of them must be **files**.

For example, suppose the user has selected **ldap\_generic** as the primary method and **files** as the secondary method. A user mapping will be created for all suitable **ldap\_generic** users and this mapping will be extended with any additional users found in the secondary method (**files**). The primary method will be used to resolve any duplicate entries.

Suppose the primary method (**ldap\_generic**) has users for UIDs 1, 2 and 3, and the secondary method (**files**) has users for UIDs 2 and 4. The username for UIDs 1, 2 and 3 will be determined by the **ldap\_generic** method and the username for UID 4 will be determined by the **files** method. If the LDAP lookup failed (such as if the LDAP server was down), a user mapping for UIDs 2 and 4 would be generated using the **files** method.

The default behavior is to use the **files** method to map Windows usernames to UNIX UIDs and GIDs, with no secondary method selected.

Regardless of the method used, the consistent mapping of usernames is a requirement to ensure consistent behavior on all CXFS nodes. Most platforms can be configured to use an LDAP database for user identification.

### Enforcing Access to Files and Directories for Windows

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

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

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

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

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

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

- `file1` will be readable and writable to user `fred` and `Administrator` on a CXFS Windows node. It can also be read by other users in group `audio`. No other users, including `diane` and `guest`, will be able to access this file.
- `file2` will be readable by all users, and writable by user `fred`, `diane` (because she is in group `video`), and `Administrator`.
- `dir1` will be readable, writable, and searchable by user `fred` and `Administrator`. It will be readable and searchable by other users in group `video`, and not accessible by all other users.

## Viewing and Changing File Attributes with Windows Explorer

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

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

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

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

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

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

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

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

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

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

### Viewing and Changing File Permissions with Windows Explorer

By selecting the **Security** tab in the **File Properties** window of a file, a user may view and change a file's permissions with a high level of granularity.

Windows Explorer will list the permissions of the file's owner and the file's group. The Everyone group, which represents the mode bits for other users, will also be displayed if other users have any access to the file. Not all Windows permission flags are supported.

The permissions on `file1` are displayed as follows:

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

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

```
Allow   Fred Costello (cxfs1\fred)   Special
Allow   audio (cxfs1\audio)          Read
```

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

**Table 6-1** Permission Flags that May Be Edited

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

The permissions for file2 are displayed as follows:

```

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

```

The permissions for dir1 are displayed as follows:

```

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

```

---

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

---

In general, you must click the **Advanced** button to see the actual permissions of directories. For example:

Allow	Fred Costello	Special	This folder only
Allow	video	Read & Execute	This folder only

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

### Viewing and Changing File Access Control Lists (ACLs) for Windows

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

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

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

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

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

<code>audio (cxfs1\audio)</code>	Allow: Read
<code>CXFS ACL Mask (cxfs1\CXFS...)</code>	Allow: Read
<code>Diane Green (cxfs1\diane)</code>	Allow: Read
<code>Fred Costello (cxfs1\fred)</code>	Allow: Read, Write

---

**Note:** You should select and deselect entries in the Allow column only, because UNIX ACLs do not have the concept of Deny. Using the Deny column will result in an ACL that allows everything that is not denied, even if it is not specifically selected in the Allow column, which is usually not what the user intended.

---

### Effective Access for Windows

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

For example:

<code>audio (cxfs1\audio)</code>	Allow: Read
<code>CXFS ACL Mask (cxfs1\CXFS...)</code>	Allow: Read
<code>Diane Green (cxfs1\diane)</code>	Allow: Read, Write
<code>Fred Costello (cxfs1\fred)</code>	Allow: Read, Write

### Restrictions with file ACLs for Windows

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

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

---

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

---

### Inheritance and Default ACLs for Windows

When a new file or directory is created, normally the mode bits are set using a mask of 022. Therefore, a new file has a mode of 644 and a new directory of 755, which means that only the user has write access to the file or directory.

You can change this mask during CXFS installation or later by modifying the installation. For more information, see "Client Software Installation for Windows" on page 185 and "Inheritance and Default ACLs for Windows" on page 166.

The four umask options available during installation or modification correspond to the following umask values:

- 000        Everyone can write
- 002        User and group can write
- 022        User only can write (default)
- 222        Read only (no one can write)

Therefore, creating a file on a UNIX CXFS client results in a mode of 644 for a mask of 022:

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

irix% umask
22

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

For more information, see the umask man page.

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

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

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

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

The permissions for `dir1` will be as follows:

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

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

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

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

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

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

## System Tunables for Windows

This section discusses the following topics:

- "Registry Modification" on page 168
- "Default Umask for Windows" on page 169
- "Maximum DMA Size for Windows" on page 169
- "Memory-Mapping Coherency for Windows" on page 170
- "DNLC Size for Windows" on page 170
- "Mandatory Locks for Windows" on page 171
- "User Identification Map Updates for Windows" on page 172
- "I/O Size Issues Within the QLogic HBA" on page 173
- "Command Tag Queueing (CTQ) Used by the QLogic HBA" on page 173

---

**Note:** These system tunables are removed when the software is removed. They may need to be reset when downgrading the CXFS for Windows software.

---

## Registry Modification

In order to configure system tuning settings, you must to modify the registry. Do the following:

1. Back up the registry before making any changes.
2. Click **Start**, select **Run**, and open the `Regedit.exe` program.
3. Select **HKEY\_LOCAL\_MACHINE** and follow the tree structure down to the parameter you wish to change.
4. After making the change, reboot the system so that the change takes affect.



**Caution:** Only the parameters documented here may be changed to modify the behavior of CXFS. All other registry entries for CXFS must not be modified or else the software may no longer function.

---

### Default Umask for Windows

The default umask that is set up during installation can be configured to a value not supported by the installer. For more information on the umask, see "Inheritance and Default ACLs for Windows" on page 166.

In **regedit**, navigate and edit the following value:

```
HKEY_LOCAL_MACHINE
> SYSTEM
  > CurrentControlSet
    > Services
      > CXFS
        > Parameters
          > DefaultUMask
```

This value specifies the umask in hexadecimal (and decimal), not its normal octal representation used on UNIX platforms.

### Maximum DMA Size for Windows

CXFS for Windows prior to CXFS 3.2 broke down large direct I/O requests into requests no larger than 4 MB, which would result in additional network traffic to the metadata server and potentially multiple extents on disk when it could allocate a single extent. This limit has been increased to 16 MB and can be configured by modifying a new registry key in CXFS 3.2 and later.

In **regedit**, navigate and edit the following value:

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

Create a new `DWORD` key called `MaxDMASize` and specify the maximum I/O request size in bytes. If this parameter is not defined, it defaults to `0x1000000`, which is 16 MB. The upper bound for Windows is just under 64 MB.

### Memory-Mapping Coherency for Windows

By default, a CXFS Windows client enforces memory-mapping coherency by preventing other clients and the CXFS metadata server access to the file while it is mapped. This can cause problems for some applications that do not expect this behavior.

Microsoft Office applications use memory-mapped I/O to read and write files, but use byte-range locks to prevent two people from accessing the same file at the same time. The CXFS behavior causes the second Office application to hang until the file is closed by the first application, without displaying a dialog that the file is in use.

Backup applications that search the filesystem for modified files will stall when they attempt to back up a file that has been memory-mapped on a CXFS Windows node.

In `regedit`, navigate and edit the following value:

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

You can disable this behavior in CXFS by changing the `DisableMemMapCoherency` parameter from 0 to 1 to avoid these problems. However, CXFS can no longer ensure data coherency if two applications memory-map the same file at the same time on different nodes in the cluster.



**Caution:** Use this option with extreme caution with multiple clients concurrently accessing the same files.

---

### DNLC Size for Windows

The Directory Name Lookup Cache (DNLC) in a CXFS Windows client allows repetitive lookups to be performed without going to the metadata server for each

component in a file path. This can provide a significant performance boost for applications that perform several opens in a deep directory structure.

In **regedit**, navigate and edit the following value:

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

The `DnLcSize` parameter is set to 4096 by default. You can change it to a value from 0 (which disables the DNLC) to 100000. Values outside this range will be reset to 4096.

---

**Note:** Increasing the DNLC size can have a significant memory impact on the Windows node and the metadata server because they maintain data structures for every actively opened file on the CXFS clients. You should monitor the memory usage on these nodes before and after changing this parameter because placing nodes under memory pressure is counter-productive to increasing the DNLC size.

---

## Mandatory Locks for Windows

By default, byte-range locks across the cluster are advisory locks, which do not prevent a rogue application from reading and writing to locked regions of a file.

---

**Note:** Windows filesystems (NTFS and FAT) implement a mandatory locking system that prevents applications from reading and writing to locked regions of a file. Mandatory locks are enabled within a Windows client.

---

In **regedit**, navigate and edit the following value:

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

To enable mandatory byte-range locks across the cluster, set the `ForceMandatoryLocks` parameter to 1. Setting this parameter will adversely affect performance of applications using these locks.

### User Identification Map Updates for Windows

User identification maps are updated automatically by the following triggers:

- An unmapped user logs into the system
- The `passwd` and/or `group` file is modified when the primary mapping method is **files**
- An LDAP database change is detected when the primary mapping method is **ldap\_activatedir** or **ldap\_generic**

The most common trigger in a typical environment is when an unmapped user logs into the system; the other two triggers are generally static in nature.

Updating the map can be a resource-intensive operation in a domain environment. Therefore, by default, an update is triggered only when an unmapped user logs in and not more often than every 5 minutes.

To configure the minimum update interval, select the following:

```
HKEY_LOCAL_MACHINE
> SYSTEM
  > CurrentControlSet
    > Services
      > CXFS_Client
        > Parameters
```

In the `regedit` menu:

```
Edit
  > New
    > DWORD Value
```

Enter `MinMapGenTime` for the name. Press **Enter** to edit the value, which is the minimum time between updates in minutes. The minimum time is 1 minute.

### I/O Size Issues Within the QLogic HBA

The maximum size of I/O issued by the QLogic HBA defaults to only 256 KB. Many applications are capable of generating much larger requests, so you may want to increase this I/O size to the HBA's maximum of 1 MB.

In **regedit**, navigate and edit the following value:

```
HKEY_LOCAL_MACHINE
> SYSTEM
  > CurrentControlSet
    > ql2xxx
      > Parameters
        > Device
```

### Command Tag Queueing (CTQ) Used by the QLogic HBA

Command Tag Queueing (CTQ) is used by HBAs to manage the number of outstanding requests each adapter port has to each target. Adjusting this value (up or down) can improve the performance of applications, depending on the number of clients in the cluster and the number of I/O requests they require to meet the required quality of service.

You should only modify this setting for HBA ports that are to be used by CXFS. Do not modify ports used for local storage.

While it is possible to change this value with the volume mounted, I/O will halt momentarily and there may be problems if the node is under a heavy load.

---

**Note:** The Windows QLogic HBA will not recognize the CTQ setting placed on the disk by IRIX nodes.

---

To configure the CTQ for the QLogic HBA, do the following:

1. Start the **SANblade Manager** program and connect.
2. Click the first adapter, for example **Adapter QLA2342**.
3. Click the **Settings** tab.
4. Click the **Select Settings section** drop down list and select **Advanced Adapter Settings**.

5. Enter a value in the range 1 through 256 in the **Execution Throttle** up-down edit control (the default is 16). The value describes how many commands will be queued by the HBA.
6. Repeat step 5 for each HBA port used for CXFS.
7. Click **Save**, enter the password (`qllogic` by default), and click **OK**.
8. Close the **SANblade Manager** program.
9. Reboot the system.

If you do not have SANblade Manager installed, you can also set the execution throttle in the Qlogic BIOS during boot-up. To do this, press `ctrl-q` when you see the Qlogic BIOS message. See the QLogic HBA card and driver documentation:

- *SANblade 2x00 Series User's Guide*
- *SANsurfer Applications User's Guide*

---

**Note:** Unlike CTQ, you cannot have separate depths per LUN. Execution throttle limits the number of simultaneous requests for **all** targets in the specified port.

---

## Windows Installation and Configuration Overview

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

1. Read the release notes for the Windows platform to learn about any late-breaking changes in the installation procedure.
2. Install Windows software according to the directions in the Windows documentation (if not already done).
3. Install and verify the RAID. See the release notes.
4. Install and verify the switch. See the release notes.
5. Obtain the CXFS license. If you want to access an XVM mirrored volume from a given node in the cluster, you must purchase a mirroring software option and obtain and install a FLEXlm license. Only those nodes that will access the mirrored volume must be licensed. For information about purchasing this license, see your sales representative. See "FLEXlm License for Windows" on page 175.

6. Install and verify the host bus adapter (HBA) and driver. See "HBA Installation for Windows" on page 177.
7. Prepare the Windows node, including adding a private network. See "Preinstallation Steps for Windows" on page 181.
8. Install the CXFS software. See "Client Software Installation for Windows" on page 185.
9. Perform post-installation configuration steps:
  - "Verifying the Licenses for Windows" on page 203
  - "Performing User Configuration for Windows" on page 192
10. Configure the cluster to define the new Windows node in the pool, add it to the cluster, start CXFS services, and mount filesystems. See Chapter 7, "Cluster Configuration" on page 215.
11. Start CXFS services on the Windows node to see the mounted filesystems under the configured drive letter.

If you run into problems, see "Troubleshooting for Windows" on page 202 and Chapter 8, "General Troubleshooting" on page 231.

## **FLEXIm License for Windows**

This section discusses the following:

- "Gathering the Windows Host Information for Windows" on page 175
- "Obtaining the Licenses for Windows" on page 176
- "Installing the Licenses for Windows" on page 177

### **Gathering the Windows Host Information for Windows**

When you order CXFS, you will receive an entitlement identifier (ID). You must submit the system host ID, host name, and entitlement ID when requesting your permanent CXFS license.

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

- Use the Macrovision `lmutil` tool to determine the host ID (which can be confusing if you have multiple NICs installed.) For example:

```
C:\> lmutil lmhostid  
lmutil - Copyright (c) 1989-2004 by Macrovision Corporation.  
All rights reserved.  
The FLEXlm host ID of this machine is "00065be62926"
```

You can download this tool from the Macrovision website.

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

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

If the machine has more than one network interface, you should use the first interface listed by `ipconfig /all`.

## Obtaining the Licenses for Windows

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.

All Windows licenses for CXFS are valid for all versions of Windows that support CXFS. Typically, the following feature names are supported for all versions of Windows:

- CXFS: `CXFS_NT`, `CXFS_W2K`, `CXFS_WXP`, and `CXFS_WNET`
- XVM Mirroring: `XVM_NT`, `XVM_W2K`, `XVM_WXP`, and `XVM_WNET`.

This allows a client running Windows 2000 to be upgraded to Windows XP or Windows Server 2003 without purchasing a new license from SGI. (The `hostid` cannot be changed; the `hostid` in the license must match the `hostid` on the Windows node.)

## Installing the Licenses for Windows

You will install the licenses in the following location before you restart the Windows node:

```
%ProgramFiles%\CXFS\lib\license.dat
```

1. Create the license directory `%ProgramFiles%\CXFS\lib` and the file `license.dat` if necessary.
2. Add the key to the `license.dat` file.

## HBA Installation for Windows

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

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

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

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

This section discusses the following:

- "Confirming the QLogic HBA Installation for Windows" on page 178
- "Configuring Multiple HBAs for Load Balancing on Windows" on page 178
- "Configuring HBA Failover for Windows 2000 and Windows 2003" on page 180

## Confirming the QLogic HBA Installation for Windows

To confirm that the QLogic HBA and driver are correctly installed, select the following to display all of the logical units (LUNs) visible to the HBA and listed within the Device Manager :

- Start
  - > Settings
    - > Control Panel
      - > Administrative Tools
        - > Computer Management
          - > Device Manager
            - > View
              - > Devices by connection

The Windows Device Manager hardware tree will differ from one configuration to another, so the actual location of the QLogic HBA within the Device Manager may differ. After it is located, any LUNS attached will be listed beneath it.

## Configuring Multiple HBAs for Load Balancing on Windows

The QLogic HBA can be configured to mask particular targets so that I/O to one target will always use one HBA port, while I/O to another target will use another HBA port. This procedure assumes that the CXFS driver is already installed and working properly with one HBA.

---

**Note:** QLogic only supports load balancing of two or more HBAs when all the HBAs have Fibre Channel connections to the LUNs on startup. If the connection to one of the HBAs is not present upon boot, this feature may not function correctly.

---

To configure two HBAs for static load balancing, do the following:

1. Disable fencing for this node. You can do this using the CXFS GUI or the `cmgr(1M)` command.
2. Determine the worldwide port name (WWPN) of the current adapter:
  - a. Install SANsurfer Qlogic SANblade Agent and Manager v2.0.29.
  - b. Run SANsurfer to determine the WWPN.
  - c. Record the WWPN on paper.

3. Shut down Windows.
4. Install the second HBA and start Windows.
5. If the second HBA is a different model from the original one, install its mini port driver (for example, `ql2300.sys`).
6. Start the Qlogic SANblade Manager and verify that two HBAs are detected. Verify that both of them mirror the same devices and logical units (LUNs). Notice that both HBAs have the same worldwide node name (WWNN) but different WWPNs. The original HBA can be recognized by its WWPN recorded in step 2.
7. Verify the driver parameters for the Qlogic HBAs. Run **regedit** and go to the following key:

```
HKEY_LOCAL_MACHINE
> SYSTEM
  > CurrentControlSet
    > Services
      > ql2xxx
        > Parameters
          > Device
```

There should be a value named `DriverParameters`. This must contain at least the following semicolon-separated parameters:

```
Buschange=0;FixupInquiry=1
```

It will typically include `UseSameNN=1` as well. If the `Buschange` and `FixupInquiry` values are not there or are incorrect, edit the parameter list to correct them. Do not delete any other parameters.

8. Configure the HBA port (click **Configure**).

---

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

```
An invalid device and LUN configuration has been detected. Auto
configure run automatically. click OK to continue.
```

---

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

9. Select the first device in the table, right click, and then select **Configure LUN(s)**.

In the new window, select the following:

**Tools**

- > **Load Balance**
- > **All LUNs**

This will statically distribute the LAN's traffic load that is associated with this device between the two HBAs.

Repeat step 9 for each of the other HBA devices.

10. Click **Apply** to save the new configuration.
11. Update the switch port information using the CXFS GUI or the `cmgr` command. Reenable fencing.
12. Reboot Windows.

For more information about using the CXFS GUI or the `cmgr` command to perform these tasks, see *CXFS Administration Guide for SGI InfiniteStorage*.

## Configuring HBA Failover for Windows 2000 and Windows 2003

The QLogic HBA on Windows 2000 and Windows 2003 also supports a mechanism to configure the automatic failover to another HBA port if the configured port is no longer able to see the target.

---

**Note:** QLogic only supports failover of two or more HBAs when all the HBAs have Fibre Channel connections to the LUNs on startup. If the connection to one of the HBAs is not present upon boot, this feature may not function correctly.

---

To configure two HBAs for failover, do the following:

1. Install the QLdirect driver v8.01.12 by following all the default settings for the installation and verify that the CXFS client still operates normally.
2. Perform the procedure in "Configuring Multiple HBAs for Load Balancing on Windows" on page 178. With QLdirect installed, the targets can be masked but will also failover to another port if a connection is lost.

## Preinstallation Steps for Windows

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

- "Adding a Private Network for Windows" on page 181
- "Verifying the Private and Public Networks for Windows" on page 183
- "Configuring the Windows XP SP2 Firewall for Windows" on page 184

### Adding a Private Network for Windows

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

#### Procedure to Add a Private Network for Windows

The following procedure provides an overview of the steps required to add a private network to the Windows node. You may skip some steps, depending upon the starting conditions at your site.

Do the following:

1. Install the second network adapter in the Windows node as per the network adapter vendor instructions. In some cases you must remove all network setups, restart, and then add network services to each network adapter from scratch.
2. Ensure that the node recognizes two network adapters in the system. Select the following:

**Start**

> **Settings**

> **Network and Dial-up Connections**

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

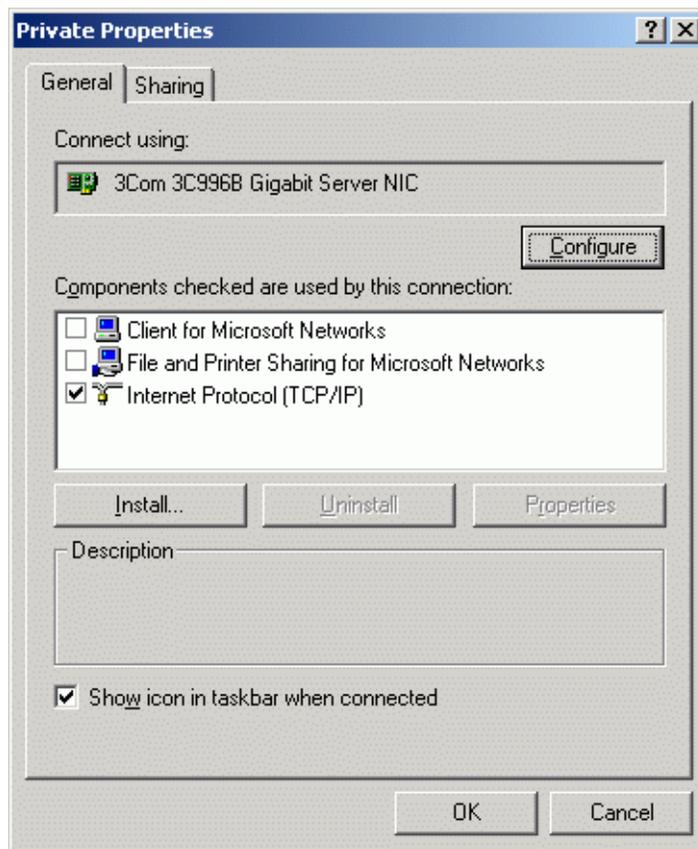
**Start**

> **Settings**

> **Network and Dial-up Connections**

Then right-mouse click the private network adapter and click **Properties**.

Uncheck all check boxes except **Internet Protocol (TCP/IP)**, thereby enabling only **Internet Protocol (TCP/IP)**, as shown in Figure 6-4.



**Figure 6-4** Private Properties: Selecting only TCP/IP

4. Select Internet Protocol (TCP/IP) and then click **Properties**. Specify the static IP address and DNS server. The private network IP address must be a fixed address and cannot be configured by DHCP.

The location of the host information is:

```
%SystemRoot%\system32\drivers\etc\hosts
```

### Ensuring Proper Hostname Configuration for Windows

It is very important, especially in the presence of private network failover, to ensure that the hostnames and IP addresses for the various network interfaces are properly configured on both the Windows node and the CXFS metadata servers. (For more information about configuring private network failover, see the *CXFS Administration Guide for SGI InfiniteStorage*.)

For example, problems may occur if the cluster is configured using hostnames and the primary network interface on the Windows node is used for the CXFS private network. In this situation, the Windows node may associate the computer name to the primary network interface rather than the private network name configured in DNS.

To avoid such problems, SGI recommends that you specify the private network addresses for the Windows node using IP addresses rather than hostnames.

### Verifying the Private and Public Networks for Windows

You can confirm that the previous procedures to add private networks were performed correctly by using the `ipconfig` command in a DOS command shell.

Create a DOS command shell with the following sequence:

```
Start
  > Programs
    > Accessories
      > Command Prompt
```

In the following example, the 10 network is the private network and the 192.168.0 network is the public network on a Windows system:

```
C:\> ipconfig /all
Windows IP Configuration

Host Name . . . . . : cxfs1
```

```
Primary Dns Suffix . . . . . : cxfs-domain.sgi.com
Node Type . . . . . : Unknown
IP Routing Enabled. . . . . : No
WINS Proxy Enabled. . . . . : No
DNS Suffix Search List. . . . . : cxfs-domain.sgi.com
                                   sgi.com
```

Ethernet adapter Public:

```
Connection-specific DNS Suffix . : cxfs-domain.sgi.com
Description . . . . . : 3Com EtherLink PCI
Physical Address. . . . . : 00-01-03-46-2E-09
Dhcp Enabled. . . . . : No
IP Address. . . . . : 192.168.0.101
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . : 192.168.0.1
DNS Servers . . . . . : 192.168.0.x
```

Ethernet adapter Private:

```
Connection-specific DNS Suffix . :
Description . . . . . : 3Com EtherLink PCI
Physical Address. . . . . : 00-B0-D0-31-22-7C
Dhcp Enabled. . . . . : No
IP Address. . . . . : 10.0.0.101
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . :
```

### Configuring the Windows XP SP2 Firewall for Windows

The Windows XP firewall will prevent a CXFS Windows node from achieving membership unless several ports are opened using the following applet:

- Start
- > Settings
- > Control Panel
- > Windows Firewall

In the **Exceptions** tab, add the following **Ports** :

- UDP on port 5449

- TCP on port 5450
- TCP on port 5451
- UDP on port 5453

## Client Software Installation for Windows

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

---

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

---

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

1. Read the release notes for the Windows platform to learn about any late-breaking changes in the installation procedure.
2. Log onto the Windows node as *Administrator* or as an account with administrative privileges.
3. Verify that the node has been updated to the correct service pack:

**Start**

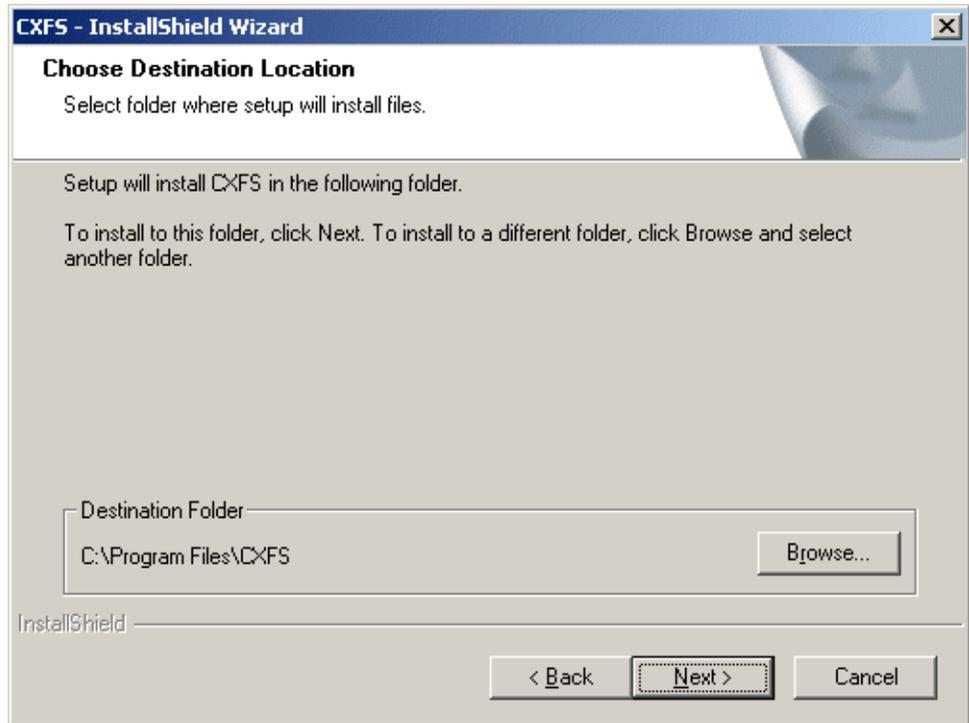
> **Programs**  
> **Accessories**  
> **System Tools**  
> **System Information**

4. Insert the *CXFS MultiOS Client 3.3* CD-ROM into the Windows host. Normally, the setup program will automatically run, otherwise run the following program from the CD-ROM:

*CD\_drive:Windows/setup.exe*

5. Acknowledge the software license agreement when prompted and read the release notes, which may contain corrections to this guide.

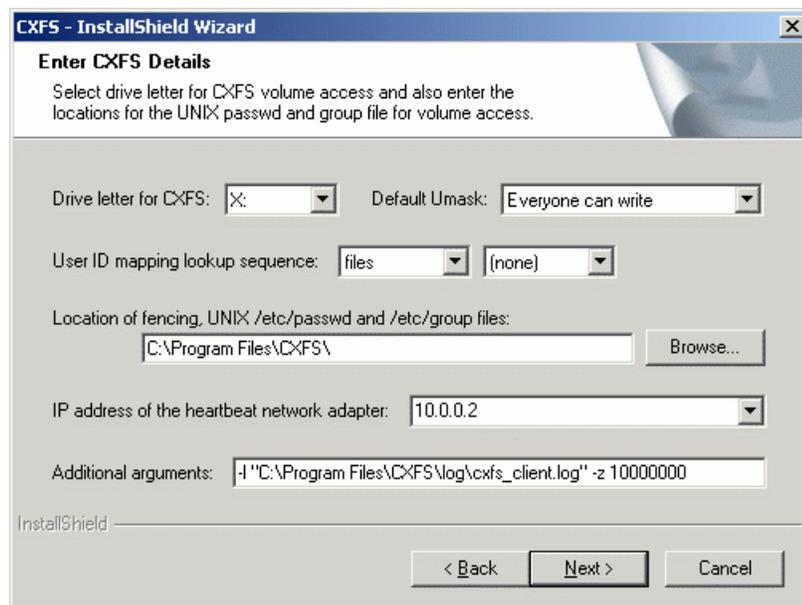
6. Install the CXFS software, as shown in Figure 6-5. If the software is to be installed in a nondefault directory, click **Browse** to select another directory. Click **Next** when finished.



**Figure 6-5** Choose Destination Location

7. Enter details for the following fields as shown in Figure 6-6 and click **Next** when finished:
  - **Drive letter for CXFS:** specify the drive letter under which all CXFS filesystems will be mounted. You cannot select a drive letter that is currently in use.
  - **Default Umask:** choose the default umask. For more information on the umask, see "Inheritance and Default ACLs for Windows" on page 166.

- **User ID mapping lookup sequence:** choose the appropriate primary and (optionally) secondary method. See "User Identification Mapping Methods for Windows" on page 159.
- **Location of fencing, UNIX /etc/passwd and /etc/group files:** specify the path where the configuration files will be installed and accessed by the CXFS software if required. The default is the same location as the software under %ProgramFiles%\CXFS.
- **IP address of the heartbeat network adapter:** specify the IP address of the private network adapter on the Windows node.
- **Additional arguments:** contains parameters that are used by the CXFS Client service when it starts up. For most configurations, this should be left alone. To get a list of options, from the command line type `cxfs_client -h`.



**Figure 6-6** Enter CXFS Details

8. If you select **ldap\_activedir** as the user ID mapping method, the dialog in Figure 6-7 is displayed after you click **Next**.

**CIFS for Windows Setup**

**Enter LDAP Details**  
Enter details for creating Windows/UNIX user ID mappings from an LDAP server.

Server Details: Host name:  Port:

Bind details:  Simple  Auth. User name:  Password:

Base DN to search from:

Search Settings: Services for UNIX defaults:

User filter:  Group filter:

Attributes: User Name:  Windows SID:  Unix UID:  Unix GID:  Grp Members:

InstallShield

< Back 

**Figure 6-7** Active Directory Details

If you have a standard Active Directory configuration with Windows Services for UNIX (SFU), you need only to select the version of SFU and **Auth** (authenticated) for **Bind details**; doing so will then define the correct Active Directory defaults. The other server details can normally remain blank.

- If you select **ldap\_generic** as the user ID mapping method, the dialog in Figure 6-8 is displayed after you click **Next**. You must provide entries for the **Host name** and the **Base DN to search from** fields. For a standard OpenLDAP server, you can select a simple anonymous bind (default settings with the **User name** and **Password** fields left blank) and select the standard search settings by clicking **Posix**.

**CXFS for Windows Setup**

**Enter LDAP Details**

Enter details for creating Windows/UNIX user ID mappings from an LDAP server.

Server Details: Host name:  Port:

Bind details:  Simple  Auth. User name:  Password:

Base DN to search from:

**Search Settings:** Generic LDAP defaults:

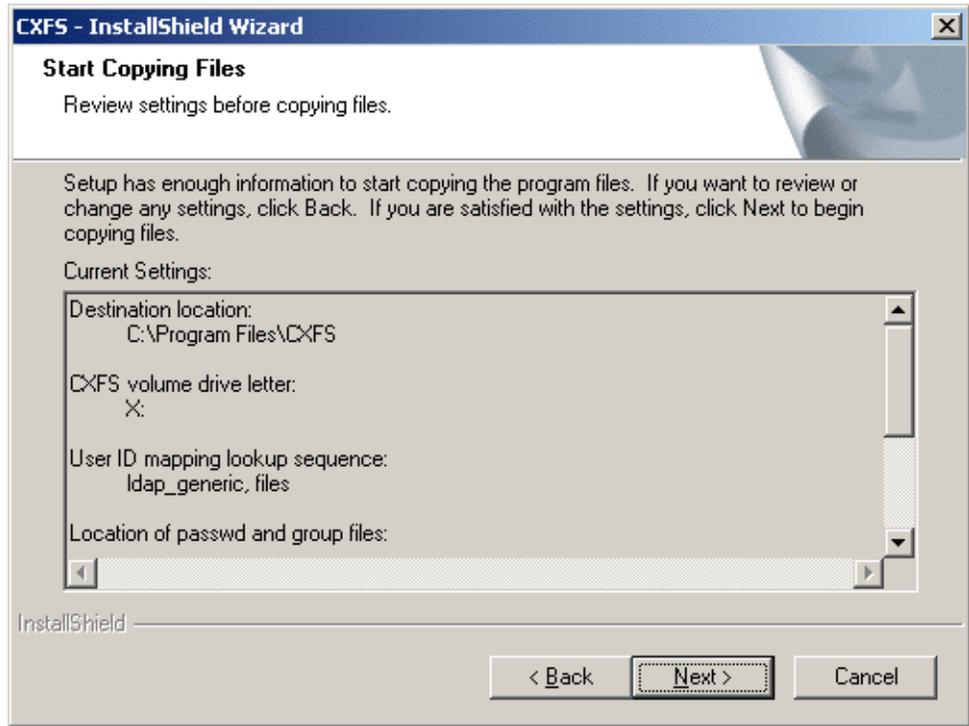
User filter:  Group filter:

Attributes: User Name:  Unix UID:  Group Name:  Unix GID:  Grp Members:

InstallShield

**Figure 6-8** Generic LDAP Details

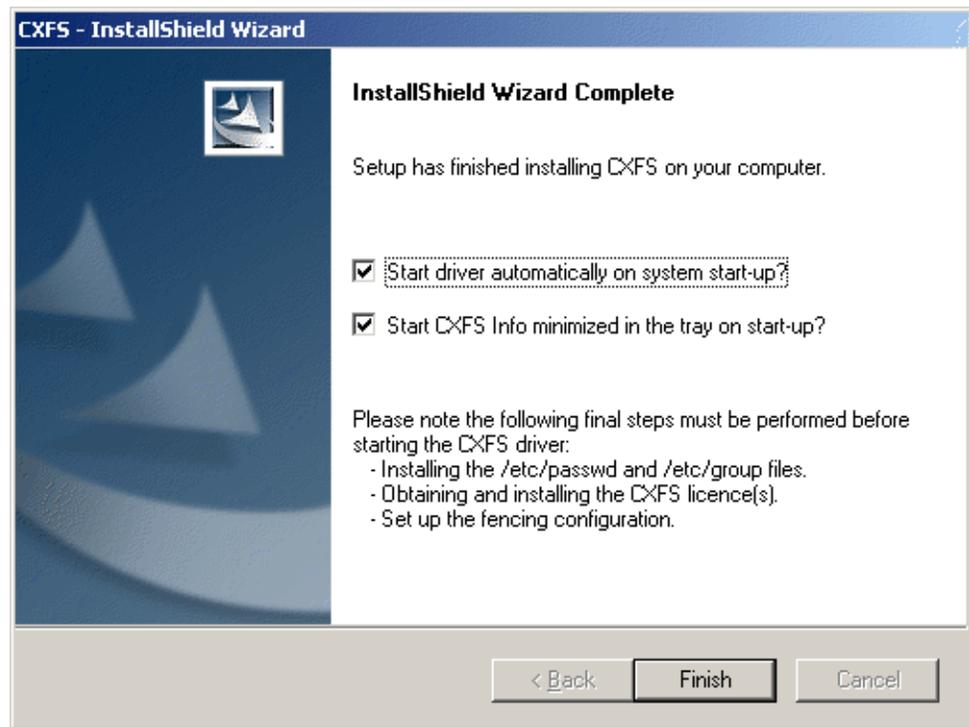
10. Review the settings, as shown in Figure 6-9. If they appear as you intended, click **Next**. If you need to make corrections, click **Back**.



**Figure 6-9** Review the Settings

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

11. You will be given the option to start the driver at system start-up, as shown in Figure 6-10. By checking the boxes, you will start the driver automatically when the system starts up and invoke the **CXFS Info** window minimized to an icon.



**Figure 6-10** Start CXFS Driver

12. Use the `cxfslicense` command to verify that the CXFS license is installed in `%ProgramFiles%\CXFS\lib\license.dat`.

If there are errors, verify that you have obtained and installed the CXFS licenses. For more information, see "FLEXlm License for Windows" on page 175.

## Postinstallation Steps for Windows

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

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

- "Performing User Configuration for Windows" on page 192
- "Checking Permissions on the Password and Group Files for Windows" on page 193

## Performing User Configuration for Windows

If the user mapping is not correctly configured, all filesystem operations will be as user nobody.

If you selected the **passwd and group files** user ID mapping method, you must install the `passwd` and `group` files. The default `passwd` and `group` files that are installed are invalid files containing comments; these invalid files will cause the CXFS Client service to generate warnings in its log file and users may not be correctly configured. You must remove the comments in these files when you install the `passwd` and `group` files.

After installing the CXFS software onto the Windows node but before restarting it, you must install the `/etc/passwd` and `/etc/group` files from a CXFS administration node to the location on the Windows node specified during installation. The defaults are as follows:

- `/etc/passwd` as `%ProgramFiles%\CXFS\passwd`
- `/etc/group` as `%ProgramFiles%\CXFS\group`

If you selected the **Active Directory** method, you must specify the UNIX identifiers for all users of the CXFS node. On the domain controller, run the following to specify the UNIX UID and GID of a given user:

```
Start
  > Program Files
    > Administrative Tools
      > Active Directory Users and Computers
        > Users
```

Select a user and then select:

**Properties**  
> **UNIX Attributes**

The CXFS software will check for changes to the LDAP database every 5 minutes.

After the CXFS software has started, you can use **CXFS Info** to confirm the user configuration, regardless of the user ID mapping method chosen. See "User Identification for Windows" on page 158.

If only the Administrator user is mapped, see "CXFS Client Service Cannot Map Users other than Administrator for Windows" on page 208.

## Checking Permissions on the Password and Group Files for Windows

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

**Properties**  
> **Security**

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



**Caution:** Failure to set permissions on the `passwd` and `group` files would allow users to change their UID/GID at will and even gain superuser access to the files on the CXFS filesystem.

---

## I/O Fencing on Windows

I/O fencing is required on Windows nodes in order to protect data integrity of the filesystems in the cluster. The CXFS Client software automatically detects the worldwide port names (WWPNs) of any supported host bus adapters (HBAs) for Windows nodes that are connected to a switch that is configured in the cluster database. These HBAs are available for fencing.

However, if no WWPNs are detected, there will be messages about loading the HBA/SNIA library logged to the %ProgramFiles%\CXFS\log\cxfs\_client.log file.

If no WWPNs are detected, you can manually specify the WWPNs in the fencing file.

---

**Note:** This method does not work if the WWPNs are partially discovered.

---

The %ProgramFiles%\CXFS\fencing.conf file enumerates the WWPN for all of the HBAs that will be used to mount a CXFS filesystem. There must be a line for the HBA WWPN as a 64-bit hexadecimal number.

---

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

---

If used, %ProgramFiles%\CXFS\fencing.conf must contain a simple list of WWPNs, one per line. You must update it whenever the HBA configuration changes, including the replacement of an HBA.

Do the following:

1. Set up the switch and HBA. See the release notes for supported hardware.
2. Determine the HBA WWPN:
  - a. Install SANsurfer Qlogic SANblade Agent and Manager v2.0.30.
  - b. Run SANsurfer to determine the WWPN.
  - c. Record the WWPN on paper.
3. Use the telnet command to connect to the switch and log in as user admin. (On Brocade switches, the password is password by default).
4. 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 fencing file).

5. Edit or create `%ProgramFiles%\CXFS\fencing.conf` and add the WWPN for the port determined in step 2. (Comment lines begin with #.)

For dual-ported HBAs, you must include the WWPNs of any ports that are used to access cluster disks. This may result in multiple WWPNs per HBA in the file; the numbers will probably differ by a single digit.

For example, if you determined that port 0 is the port connected to the switch, your fencing file should contain the following:

```

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

```

6. After the node is added to the cluster (see Chapter 7, "Cluster Configuration" on page 215), enable the fencing feature by using the CXFS GUI or `cmgr` command

on a server-capable administration node; for more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Start/Stop the CXFS Client Service for Windows

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

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

It is recommended that the CXFS driver configuration remains manual.

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

```
Start
  > Settings
      > Control Panel
          > Administrative Tools
              > Services
```

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

## Maintenance for Windows

This section contains the following:

- "Modifying the CXFS Software for Windows"
- "Upgrading the CXFS Software for Windows" on page 198
- "Removing the CXFS Software for Windows" on page 199
- "Downgrading the CXFS Software for Windows" on page 200
- "Recognizing Storage Changes for Windows" on page 200

## Modifying the CXFS Software for Windows

To change the location of the software and other configuration settings that were requested in "Client Software Installation for Windows" on page 185, perform the following steps:

1. Select the following:

**Start**

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

Figure 6-11 shows the screen that lets you modify the software.



**Figure 6-11** Modify CXFS for Windows

2. Make the necessary configuration changes.

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

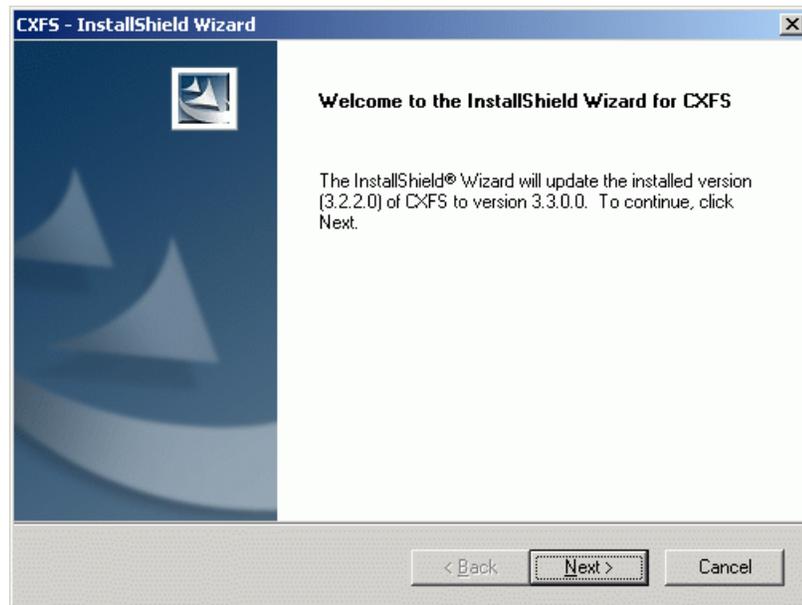
```
C:\> %SystemRoot%\system32\cxfs_client.exe -h
```

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

## Upgrading the CXFS Software for Windows

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

1. Insert the CD-ROM containing the upgraded software to run the setup program. If the setup program does not automatically start, run `CD_drive:Windows/Setup.exe` from the CD-ROM.
2. A welcome screen will appear that displays the version you are upgrading from and the version you are upgrading to. Figure 6-12 shows the screen that appears when you are upgrading the software. All the configuration options are available to update as discussed in "Client Software Installation for Windows" on page 185.



**Figure 6-12** Upgrading the Windows Software

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

## Removing the CXFS Software for Windows

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

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

Figure 6-11 on page 197 shows the screen that lets you remove the software.

---

**Note:** By default, the `passwd`, `group`, `license` and `log` files will not be removed. To remove these other files, check the following box:

**Delete all config files, license file, logs, etc**

Then click **Next**.

---

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

## Downgrading the CXFS Software for Windows

To downgrade the CXFS software, follow the instructions to remove the software in "Removing the CXFS Software for Windows" on page 199 and then install the older version of the software as directed in "Client Software Installation for Windows" on page 185.

---

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

---

## Recognizing Storage Changes for Windows

If you make changes to your storage configuration, you must rerun the HBA utilities to reprobe the storage. See "HBA Installation for Windows" on page 177.

If new storage devices are added to the cluster, you must reboot the Windows node in order to discover those devices.

To verify the driver parameters for the Qlogic HBAs, run `regedit` and go to the following key:

```
HKEY_LOCAL_MACHINE
> SYSTEM
  > CurrentControlSet
    > Services
      > ql2xxx
        > Parameters
          > Device
```

There should be a value named `DriverParameters`. This must contain at least the following semicolon-separated parameters:

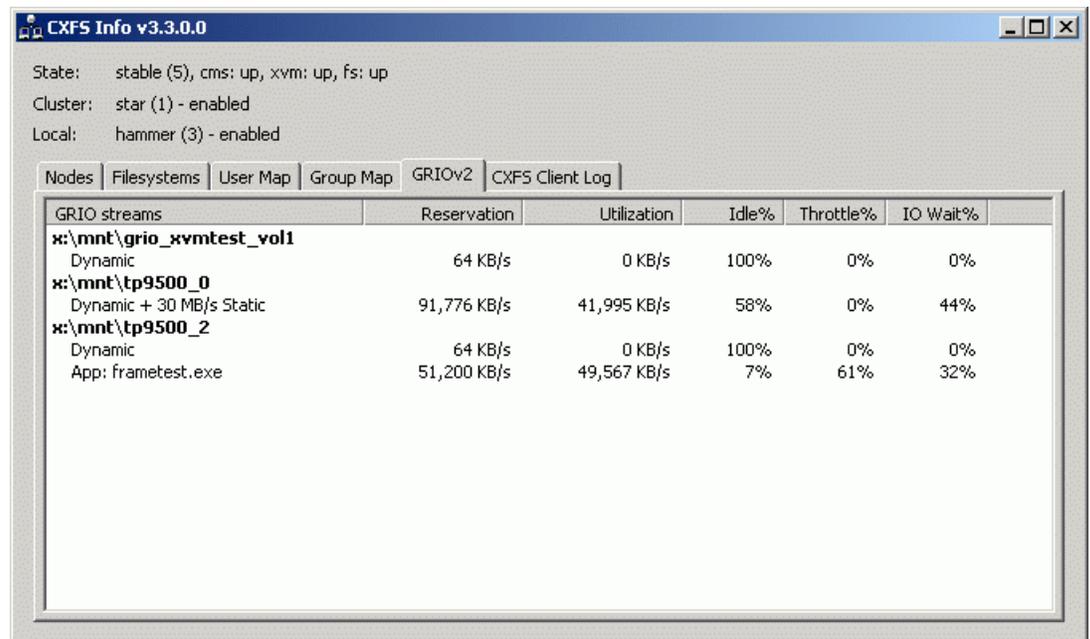
```
Buschange=0;FixupInquiry=1
```

It will typically include `UseSameNN=1` as well. If the `Buschange` and `FixupInquiry` values are not there or are incorrect, edit the parameter list to correct it. Do not delete any other parameters.

## GRIO on Windows

CXFS supports guaranteed-rate I/O (GRIO) version 2 on the Windows platform.

Figure 6-13 shows the **CXFS Info** display for GRIO.



The screenshot shows the CXFS Info v3.3.0.0 application window. The title bar reads "CXFS Info v3.3.0.0". The main window displays the following information:

```
State: stable (5), cms: up, xvm: up, fs: up
Cluster: star (1) - enabled
Local: hammer (3) - enabled
```

Below this information are several tabs: "Nodes", "Filesystems", "User Map", "Group Map", "GRIOv2", and "CXFS Client Log". The "GRIOv2" tab is selected, showing a table of GRIO streams with the following columns: "GRIO streams", "Reservation", "Utilization", "Idle%", "Throttle%", and "IO Wait%".

GRIO streams	Reservation	Utilization	Idle%	Throttle%	IO Wait%
<b>x:\mnt\grio_xvmtest_vol1</b>					
Dynamic	64 KB/s	0 KB/s	100%	0%	0%
<b>x:\mnt\tp9500_0</b>					
Dynamic + 30 MB/s Static	91,776 KB/s	41,995 KB/s	58%	0%	44%
<b>x:\mnt\tp9500_2</b>					
Dynamic	64 KB/s	0 KB/s	100%	0%	0%
App: frametest.exe	51,200 KB/s	49,567 KB/s	7%	61%	32%

**Figure 6-13** CXFS Info Display for GRIO for Windows

For more general information about GRIO, see "Guaranteed-Rate I/O (GRIO) and CXFS" on page 14 and the *Guaranteed-Rate I/O Version 2 Guide*.

## Troubleshooting for Windows

This section contains the following common Windows problems:

- "Verifying the Licenses for Windows" on page 203
- "Verifying that the CXFS Software is Running Correctly for Windows" on page 203
- "Unable to Mount Filesystems on Windows" on page 204
- "Access-Denied Error when Accessing Filesystem on Windows" on page 206
- "Application Works with NTFS but not CXFS for Windows" on page 206
- "Delayed-Write Error Dialog is Generated by the Windows Kernel" on page 207
- "CXFS Client Service Does Not Start on Windows" on page 207
- "HBA Problems" on page 207
- "CXFS Client Service Cannot Map Users other than Administrator for Windows" on page 208
- "Filesystems Are Not Displayed on Windows" on page 209
- "Large Log Files on Windows" on page 210
- "Windows Failure on Restart" on page 210
- "Memory Configuration for Windows" on page 210
- "Application Cannot Create File Under CXFS Drive Letter" on page 211

Also see:

- The Windows `cxfsdump` documentation located at `%ProgramFiles%\CXFS\cxfsdump.html`
- Chapter 8, "General Troubleshooting" on page 231

## Verifying the Licenses for Windows

To verify that the licenses have been installed properly, use the `cxfslicense` command in a DOS command shell after installing the CXFS software.

Create a DOS command shell with the following sequence:

```
Start
  > Programs
    > Accessories
      > Command Prompt
```

To run `cxfslicense`, enter the following command:

```
C:\> cxfslicense
```

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

```
Checking all supported Windows licenses
Found a valid license for CXFS for Windows
Found a valid license for XVM mirror access for Windows
```

## Verifying that the CXFS Software is Running Correctly for Windows

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

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

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

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

**View**  
    > **Show hidden devices**

To show the CXFS driver, select the following:

**Non-Plug and Play Devices**  
    > **CXFS**  
        > **Properties**

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

**Start**  
    > **Settings**  
        > **Control Panel**  
            > **Administrative Tools**  
                > **Services**

## Unable to Mount Filesystems on Windows

If **CXFS Info** reports that `cms` is up but `XVM` or the filesystem is in another state, then one or more mounts is still in the process of mounting or has failed to mount.

The CXFS node might not mount filesystems for the following reasons:

- The client may not be able to see all the LUNs. This is usually caused by misconfiguration of the HBA or the SAN fabric:
  - Check that the ports on the Fibre Channel switch connected to the HBA are active. Physically look at the switch to confirm the light next to the port is green, or remotely check by using the `switchShow` command.
  - Check that the HBA configuration is correct. For information specific to Windows, see "HBA Problems" on page 207.
  - Check that the HBA can see all the LUNs for the filesystems it is mounting.
  - Check that the operating system kernel can see all the LUN devices. For example, for Windows:

**Start**> **Settings**> **Control Panel**> **Administrative Tools**> **ComputerManagement**> **Device Manager**> **View**> **Devices by connection**

- Use debugview to monitor the CXFS driver when it probes the disk devices. You should see it successfully probe each of the LUN devices.
- If the RAID device has more than one LUN mapped to different controllers, ensure the node has a Fibre Channel path to all relevant controllers.
- The CXFS Client service may not be running. To verify that it is running, open the **Task Manager** by pressing the `Ctrl+Shift+Esc`, or right-mouse click on an empty area of the taskbar and select **Task Manager** from the popup menu. In the **Processes** tab, search for `cxfs_client.exe` in the **Image Name** column. You can sort the processes by name by clicking the heading of the column.
- The filesystem may have an unsupported mount option. Check the `cxfs_client.log` for mount option errors or any other errors that are reported when attempting to mount the filesystem.
- The cluster membership (`cms`), XVM, or the filesystems may not be up on the node. Use **CXFS Info** to determine the current state of `cms`, XVM, and the filesystems. Do the following:
  - If `cms` is not up, check the following:
    - Is the node is configured on the administration node with the correct hostname? See "Configuring Hostnames on Mac OS X" on page 82.
    - Has the node been added to the cluster and enabled? See "Verifying the Cluster Status" on page 223.
  - If XVM is not up, check that the HBA is active and can see the LUNs.
  - If the filesystem is not up, check that one or more filesystems are configured to be mounted on this node and

Also, check the **CXFS Client Log** in **CXFS Info** for mount errors. They will be highlighted in red.

## Access-Denied Error when Accessing Filesystem on Windows

If an application reports an access-denied error, do the following:

- Check the list of users and groups that **CXFS Info** has mapped to a UNIX UID and GID. If the current user is not listed as one of those users, check that the user mapping method that was selected is configured correctly, that there is an LDAP server running (if you are using LDAP), and that the user is correctly configured.
- Increase the verbosity of output from the CXFS Client service so that it shows each user as it is parsed and mapped.
- Use Sysinternals Filemon to monitor the application and verify that there is no file that has been created below a mount point under the CXFS drive letter. An error may be caused by attempting to create a file below the drive letter but above the mount point. For more information about Filemon, see:

<http://www.sysinternals.com>

## Application Works with NTFS but not CXFS for Windows

The Windows filesystem APIs are far more extensive than the UNIX POSIX APIs and there are some limitations in mapping the native APIs to POSIX APIs (see "Functional Limitations for Windows" on page 153). Sometimes these limitations may affect applications, other times the applications that have only ever been tested on NTFS make assumptions about the underlying filesystem without querying the filesystem first.

If an application does not behave as expected, and retrying the same actions on an NTFS filesystem causes it to behave as was expected, then third-party tools like SysInternals Filemon can be used to capture a log of the application when using both NTFS and CXFS. Look for differences in the output and try to determine the action and/or result that is different. Using the same filenames in both places will make this easier. For more information about Filemon, see:

<http://www.sysinternals.com>

---

**Note:** There are some problems that will not be visible in a Sysinternals Filemon log. For example, some older applications use only a 32-bit number when computing filesystem or file size. Such applications may report out of disk space errors when trying to save a file to a large (greater than 1 TB) filesystem.

---

## Delayed-Write Error Dialog is Generated by the Windows Kernel

A delayed-write error is generated by the Windows kernel when it attempts to write file data that is in the cache and has been written to disk, but the I/O failed. The write call made by the application that wrote the data may have completed successfully some time ago (the application may have even exited by now), so there is no way for the Windows kernel to notify the application that the I/O failed.

This error can occur on a CXFS filesystem if CXFS has lost access to the disk due to the following:

- Loss of membership resulting in the Windows client being fenced and the filesystem being unmounted. Check that the Windows client is still in membership and that there are no unmount messages in the `cxfs_client.log` file.
- Loss of Fibre Channel connection to the Fibre Channel switch or RAID. Check the Fibre Channel connections and use the SanManager tool to verify that the HBA can still see all of the LUNs. Make sure the filesystems are still mounted.
- The metadata server returned an I/O error. Check the system log on the metadata server for any I/O errors on the filesystem and take corrective action on the server if required.

## CXFS Client Service Does Not Start on Windows

The following error may be seen when the CXFS Client service attempts to start:

```
Error 10038: An operation was attempted on something that is not a socket.
```

Check the **CXFS Client Log** in **CXFS Info** for information on why the CXFS client failed to start. The most common cause is an invalid license file.

## HBA Problems

If you have a problem with an HBA, check the following:

- Has plug-and-play been disabled?

*Plug-and-play functionality*, which would normally discover new devices, is disabled by the QLogic HBA software so that it can perform path failover without Windows attempting to manage the change in available devices. Disabling the plug-and-play feature also enables CXFS to map CXFS volumes to the same devices if a Fibre Channel path was lost and then reestablished. If HBA path

failover or CXFS rediscovering XVM volumes and filesystems does not appear to work, verify that plug-and-play is disabled.

- Are there QLogic management tool event and alarm log messages? Select the following:

**Start**  
    > **Programs**  
        > **QLogic Management Suite**  
            > **SANsurfer**

Also see "Recognizing Storage Changes for Windows" on page 200 and "Unable to Mount Filesystems on Windows" on page 204.

### **CXFS Client Service Cannot Map Users other than Administrator for Windows**

If the CXFS Client service cannot map any users other than Administrator and there are no LDAP errors in the `cxfs_client` log file (and you are using LDAP), you must change the configuration to allow reading of the attributes.

Do the following:

1. Select the following:

**Start**  
    > **Settings**  
        > **Control Panel**  
            > **Administrative Tools**  
                > **Active Directory Users and Computers**

2. Select the following:

**View**  
    > **Advanced Features**

3. Right-mouse click the **Users** folder under the domain controller you are using and select the following:

**Properties**  
    > **Security**  
        > **Advanced**  
            > **Add**

4. Select **Authenticated Users** from the list and click **OK**.
5. Select **Child Objects Only** from the **Apply onto** drop-down list and check **Read All Properties** from the list of permissions.
6. Click **OK** to complete the operation.

If the above configuration is too broad security-wise, you can enable the individual attributes for each user to be mapped.

## Filesystems Are Not Displayed on Windows

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

- Run `%ProgramFiles%\CXFS\cxfs_info` to ensure that the filesystems have been configured for this node.
- Verify the filesystems that should be mounted by using the `cmgr` command on a CXFS administration node. For more information, see "Mounting Filesystems on the Client-Only Nodes" on page 220.
- Ensure that the CXFS metadata server is up and that the Windows node is in the cluster membership; see "Verifying the Cluster Status" on page 223.
- Check that the CXFS Client service has started. See "Start/Stop the CXFS Client Service for Windows" on page 196 and "Verifying that the CXFS Software is Running Correctly for Windows" on page 203.
- Check the **CXFS Client Log** in **CXFS Info** 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

The CXFS Client service creates the following log file:

```
%ProgramFiles%\CXFS\log\cxfs_client.log
```

On an upgraded system, this log file may become quite large over a period of time if the verbosity level is increased. (New installations perform automatic log rotation when the file grows to 10MB.)

To verify that log rotation is enabled, check the **Addition** arguments by modifying the installation (see "Modifying the CXFS Software for Windows" on page 197) and append the following if the `-z` option is not present:

```
-z 10000000
```

You must restart the CXFS Client service for the new settings to take effect. See "Start/Stop the CXFS Client Service for Windows" on page 196 for information on how to stop and start the CXFS Client service.

## Windows Failure on Restart

If the CXFS Windows node fails to start and terminates in a blue screen, restart your computer and select the backup hardware profile (with CXFS disabled). Alternatively, pressing `L` at the **Hardware Profile** menu will select the last configuration that was successfully started and shut down. If the node has only one hardware profile, press the spacebar after selecting the boot partition to get to the **Hardware Profile** menu.

## Memory Configuration for Windows

A Windows problem may affect Windows CXFS nodes performing large asynchronous I/O operations. If the Windows node crashes with a `NO_MORE_SYSTEM_PAGES` message, the work-around described in the following link should be considered (line break added here for readability):

[http://www.microsoft.com/technet/treeview/default.asp?url=/technet/prodtechnol/winxpro/reskit/prmd\\_stp\\_fztl.asp](http://www.microsoft.com/technet/treeview/default.asp?url=/technet/prodtechnol/winxpro/reskit/prmd_stp_fztl.asp)

## Application Cannot Create File Under CXFS Drive Letter

If an application requires that it be able to create files and/or directories in the root of the CXFS drive, you must create a virtual drive for the system that maps to a mounted filesystem directory.

This can be performed using the `subst` command from the command prompt. For example, to use the CXFS filesystem `X:\mnt\tp9500_0` to the free drive letter `V`, you would enter the following:

```
C:\> subst V: X:\mnt\tp9500_0
```

To remove the mapping, run:

```
C:\> subst V: /D
```

## Reporting Windows Problems

To report problems about a Windows node, you should retain platform-specific information and save crash dumps.

### Retain Windows Information

When reporting a problem about a CXFS Windows node to SGI, run the following:

```
Start
  > Program Files
    > CXFS
      > CXFS Dump
```

This will collect the following information:

- System information
- CXFS registry settings
- CXFS client logs
- License information
- CXFS version information
- Network settings

- Event log
- (optionally) Windows crash dump, as described in "Save Crash Dumps for Windows" on page 212

You can obtain information about the cluster by running the `cxfsdump` utility on a CXFS administration node.

## Save Crash Dumps for Windows

If you are experiencing crashes or if the Windows node hangs, you should configure the Windows node to save crash dumps to a filesystem that is not a CXFS filesystem. This crash dump can then be analyzed by SGI.

To do this, click the right mouse button on the **My Computer** icon and select the following:

**Properties**  
    > **Advanced**  
        > **Startup and Recovery**  
            > **Write debugging information to**

Enter a path on a filesystem other than a CXFS filesystem. You may also select a **Kernel Memory Dump**, which is a smaller dump that typically contains enough information regarding CXFS problems.

These changes will take affect only after the node is restarted.

## Generating a Crash Dump on a Hung Windows Node

If user applications on a Windows node are no longer responsive and cannot be killed, you should attempt to generate a crash dump by forcing the node to crash.

After configuring the crash dump location (see "Save Crash Dumps for Windows" on page 212), you can modify the registry so that a combination of key strokes will cause the Windows node to crash. This will only work on machines with a PS/2 keyboard.

To do this, run the registry editor as follows:

**Start**  
    > **Run**  
        > **regedit**

Then navigate to:

HKEY\_LOCAL\_MACHINE\SYSTEM\CurrentControlSet\Services\i8042prt\Parameters

Add a new entry by selecting the following:

**Edit**  
**> Add Value**

Enter the following information:

- **Value Name:** CrashOnCtrlScroll
- **Data Type:** REG\_DWORD
- **Value:** 1

These changes will take affect only after the node is restarted.

To generate a crash on the node after applying these changes, hold the right CTRL key and press SCROLL LOCK twice. See the following for more information:

<http://support.microsoft.com/?kbid=244139>



## Cluster Configuration

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

All CXFS administrative tasks other than restarting the Windows node must be performed using either the CXFS GUI (invoked by the `cxfsmgr` command and connected to a CXFS administration node) or the `cmgr` command on a CXFS administration node. The GUI also provides a guided configuration for defining a cluster.

This section discusses the following tasks in cluster configuration:

- "Defining the Client-Only Nodes" on page 216
- "Adding the Client-Only Nodes to the Cluster" on page 217
- "Defining the Switch for I/O Fencing" on page 218
- "Starting CXFS Services on the Client-Only Nodes" on page 219
- "Verifying LUN Masking" on page 220
- "Mounting Filesystems on the Client-Only Nodes" on page 220
- "Forced Unmount of CXFS Filesystems" on page 221
- "Restarting the Windows Node" on page 222
- "Verifying the Cluster Configuration" on page 222
- "Verifying Connectivity in a Multicast Environment" on page 222
- "Verifying the Cluster Status" on page 223
- "Verifying the I/O Fencing Configuration" on page 226
- "Verifying Access to XVM Volumes" on page 227

For detailed configuration instructions, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Defining the Client-Only Nodes

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

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

- AIX: use the value displayed by `/usr/bin/hostname`
- Linux: use the value displayed by `/bin/hostname`
- Mac OS X: use the value displayed by `/bin/hostname`
- Solaris: use the value displayed by `/bin/hostname`
- Windows: select the following:

**Start**

**> Settings**

**> Network and Dial-up Connections**

**> Advanced**

**> Network Identification**

When you specify that a node is running an operating system other than IRIX or Linux, the node will automatically be defined as a client-only node and you cannot change it. (These nodes cannot be potential metadata servers and are not counted when calculating the CXFS kernel 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|Linux32|Linux64|AIX|HPUX|MacOSX|Solaris|Windows> ? solaris
Node ID ? 7
```

```
Do you wish to define failure hierarchy[y/n]:y
Hierarchy option 0 <System|Fence|Shutdown>[optional] ? fence
Hierarchy option 1 <System|Fence|Shutdown>[optional] ? shutdown
Hierarchy option 2 <System|Fence|Shutdown>[optional] ?
Number of Network Interfaces ? (1)
NIC 1 - IP Address ? 163.154.18.1722
NIC 1 - Heartbeat HB (use network for heartbeats) <true|false> ? true
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 Administration Guide for SGI InfiniteStorage*.

## Adding the Client-Only Nodes to the Cluster

After you define all of the client-only nodes, you must add them to the cluster using either the CXFS GUI or the `cmgr` command.

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 Administration Guide for SGI InfiniteStorage*.

Depending upon your filesystem configuration, you may also need to add the node to the list of clients that have access to the volume. See “Mounting Filesystems on the Client-Only Nodes” on page 220.

## Defining the Switch for I/O Fencing

You are required to use I/O fencing on client-only nodes in order to protect data integrity. I/O fencing requires a switch; see the release notes for supported switches. To define the switch for the cluster database, use either the CXFS GUI or the `cmgr` command.

For example:

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

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

```
irix# hafence -v
```

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

```
switch# portEnable portnumber
```

You must then update the switch port information using the GUI or `cmgr`.

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

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

In this case, you would need to enable ports 0 and 1:

Logged in to the switch:

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

Logged in to a CXFS administration node:

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

```
irix# cmgr -c admin fence update
```

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

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

## Starting CXFS Services on the Client-Only Nodes

After adding the client-only nodes to the cluster, you must start CXFS services for them, which enables the node by setting a flag for the node in the cluster database. You can do this using either the CXFS GUI or the `cmgr` command.

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 Administration Guide for SGI InfiniteStorage*.

## Verifying LUN Masking

You should verify that the HBA has logical unit (LUN) masking configured such that the LUNs are visible to all the nodes in the cluster after you connect the HBA to the switch and before configuring the filesystems with XVM. For more information, see the RAID documentation.

## Mounting Filesystems on the Client-Only Nodes

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

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

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

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

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

---

**Note:** SGI recommends that you enable the *forced unmount* feature for CXFS filesystems, which is turned off by default; see "Recommendations" on page 9 and "Forced Unmount of CXFS Filesystems" on page 221.

---

For details, see the “Define a Filesystem” and “Mount a Filesystem” sections of the GUI or the `cmgr` reference chapters in the *CXFS Administration Guide for SGI InfiniteStorage*.

## Forced Unmount of CXFS Filesystems

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

For example:

```
define cxfs_filesystem logical_filesystem_name [in cluster clustername]  
    set force to true  
  
modify cxfs_filesystem logical_filesystem_name [in cluster clustername]  
    set force to true  
  
admin cxfs_unmount cxfs_filesystem filesystemname [on node nodename] [in cluster clustername]
```

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

```
cmgr> modify cxfs_filesystem fs1 in cluster cxfscluster  
Enter commands, when finished enter either "done" or "cancel"cmgr>  
  
cxfs_filesystem fs1 ? set force to true  
cxfs_filesystem fs1 ? done  
Successfully defined cxfs_filesystem fs1  
  
cmgr> admin cxfs_unmount cxfs_filesystem fs1 in cluster cxfscluster
```

For details, see the “CXFS Filesystems Tasks with the GUI” sections of the GUI or the `cmgr` reference chapters in the *CXFS Administration Guide for SGI InfiniteStorage*.

## Restarting the Windows Node

After completing the steps in "Postinstallation Steps for Windows" on page 191 and this chapter, you should restart the Windows node. This will automatically start the driver and the CXFS Client service.

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

## Verifying the Cluster Configuration

To verify that the client-only nodes have been properly added to the cluster, run the `cxfs-config` command on the metadata server. For example:

```
irix# /usr/cluster/bin/cxfs-config -all -check
```

This command will dump the current cluster nodes, private network configuration, filesystems, XVM volumes, failover hierarchy, and switches. It will check the configuration and report any common errors. You should rectify these error before starting CXFS services.

For more information, see the section about checking the cluster configuration with `cxfs-config` in the *CXFS Administration Guide for SGI InfiniteStorage*.

## Verifying Connectivity in a Multicast Environment

To verify general connectivity in a multicast environment, you can execute a UNIX `ping` command on the `224.0.0.1` IP address.

To verify the CXFS heartbeat, use the `224.0.0.250` IP address. The `224.0.0.250` address is the default CXFS heartbeat multicast address (because it is the default, this address does not have to appear in the `/etc/hosts` file).

---

**Note:** A node is capable of responding only when the administration daemons (`fs2d`, `cmond`, `cad`, and `crsd`) or the `cxfs_client` daemon is running.

---

For example, to see the response for two packets sent from Solaris IP address 128.162.240.27 to the multicast address for CXFS heartbeat and ignore loopback, enter the following:

```
solaris# ping -i 128.162.240.27 -s -L 224.0.0.250 2
```

To override the default address, you can use the `-c` and `-m` options or make the name `cluster_mcast` resolvable on all nodes (such as in the `/etc/hosts` file). For more information, see the `cxfs_client` man page.

## Verifying the Cluster Status

To verify that the client-only nodes have been properly added to the cluster and that filesystems have been mounted, use the view area of the CXFS GUI or the `clconf_info` command (on a CXFS administration node) and the `cxfs_info` command (on a client-only node).

For example:

```
irix# /usr/cluster/bin/clconf_info
```

```
Event at [2004-05-04 19:00:33]
```

```
Membership since Tue May 4 19:00:33 2004
```

Node	NodeID	Status	Age	CellID
cxfsirix4	1	up	27	2
cxfsirix5	2	up	26	1
cxfsirix6	3	up	27	0
cxfswin4	5	up	1	5
cxfssun3	6	up	0	6
cxfsmac3.local.	17	up	0	7

```
2 CXFS FileSystems
```

```
/dev/cxvm/vol0 on /mnt/vol0 enabled server=(cxfsirix4) 5
```

```
client(s)=(cxfsirix6,cxfsirix5,cxfswin4,cxfssun3,cxfsmac3.local.) status=UP
```

```
/dev/cxvm/vol1 on /mnt/vol1 enabled server=(cxfsirix5) 5
```

```
client(s)=(cxfsirix6,cxfsirix4,cxfswin4,cxfssun3,cxfsmac3.local.) status=UP
```

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

- AIX, IRIX, Linux, and Solaris: `/usr/cxfs_cluster/bin/cxfs_info`
- Mac OS X: `/usr/cluster/bin/cxfs_info`
- Windows: `%ProgramFiles%\CXFS\cxfs_info.exe`

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

For example, on a Solaris node:

```
solaris# /usr/cxfs_cluster/bin/cxfs_info
cxfs_client status [timestamp Jun 03 03:48:07 / generation 82342]

CXFS client:
  state: reconfigure (2), cms: up, xvm: up, fs: up
Cluster:
  performance (123) - enabled
Local:
  cxfssun3 (9) - enabled
Nodes:
  cxfsirix4 enabled up    2
  cxfsirix5 enabled up    1
  cxfsirix6 enabled up    0
  cxfswin4  enabled up    5
  cxfssun3  enabled up    6
  cxfsmac3.local. enabled up    7
Filesystems:
  vol0      enabled mounted      vol0      /mnt/vol0
  vol1      enabled mounted      vol1      /mnt/vol1
```

The `CXFS client` line shows the state of the client in the cluster, which can be one of the following states:

- |                        |   |
|------------------------|---|
| <code>bootstrap</code> | Initial state after starting <code>cxfs_client</code> , while listening for bootstrap packets from the cluster. |
| <code>connect</code>   | Connecting to the CXFS metadata server.   |

query	The client is downloading the cluster database from the metadata server.
reconfigure	The cluster database has changed, so the client is reconfiguring itself to match the cluster database.
stable	The client has been configured according to what is in the cluster database.
stuck	The client is unable to proceed, usually due to a configuration error. Because the problem may be transient, the client periodically reevaluates the situation. The number in parenthesis indicates the number of seconds the client will wait before retrying the operation. With each retry, the number of seconds to wait is increased; therefore, the higher the number the longer it has been stuck. See the log file for more information.
terminate	The client is shutting down.

The `cms` field has the following states:

unknown	Initial state before connecting to the metadata server.
down	The client is not in membership.
fetal	The client is joining membership.
up	The client is in membership.
quiesce	The client is dropping out of membership.

The `xvm` field has the following states:

unknown	Initial state before connecting to the metadata server.
down	After membership, but before any XVM information has been gathered.
fetal	Gathering XVM information.
up	XVM volumes have been retrieved.

The `fs` field has the following states:

unknown	Initial state before connecting to the metadata server.
down	One or more filesystems are not in the desired state.
up	All filesystems are in the desired state.

`retry` One or more filesystems cannot be mounted/unmounted, and will retry. See the "Filesystem" section of `cxfs_info` output to see the affected filesystems.

## Verifying the I/O Fencing Configuration

To determine if a node is correctly configured for I/O fencing, log in to a CXFS administration node and use the `cxfs-config(1M)` command. For example:

```
irix# /usr/cluster/bin/cxfs-config
```

The failure hierarchy for a client-only node should be listed as `Fence, Shutdown`, as in the following example:

Machines:

```
node cxfswin2: node 102 cell 1 enabled Windows client_only
hostname: cxfswin2.melbourne.sgi.com
fail policy: Fence, Shutdown
nic 0: address: 192.168.0.102 priority: 1
```

See "Defining the Client-Only Nodes" on page 216 to change the failure hierarchy for the node if required.

The HBA ports should also be listed in the switch configuration:

Switches:

```
switch 1: 16 port brocade admin@asg-fcsw7 <no ports masked>
port 5: 210200e08b51fd49 cxfswin2
port 15: 210100e08b32d914 cxfsirix2
switch 2: 16 port brocade admin@asg-fcsw8 <no ports masked>
port 5: 210300e08b71fd49 cxfswin2
port 14: 210000e08b12d914 cxfsirix2
```

No warnings or errors should be displayed regarding the failure hierarchy or switch configuration.

If the HBA ports for the client node are not listed, see the following:

- "I/O Fencing for AIX" on page 35
- "I/O Fencing for Linux" on page 65

- "I/O Fencing for Mac OS X" on page 93
- "I/O Fencing for Solaris" on page 133
- "I/O Fencing on Windows" on page 193

For further information regarding switch configuration, see the "Define a Switch" and "Update Switch Port Information" sections of the GUI or `cmgr` reference chapters in the *CXFS Administration Guide for SGI InfiniteStorage*

## Verifying Access to XVM Volumes

To verify that a client node has access to all XVM volumes that are required to mount the configured filesystems, log on to a CXFS administration node and run:

```
irix# /usr/cluster/bin/cxfs-config -xvm
```

This will display the list of filesystems and the XVM volume and volume elements used to construct those filesystems. For example:

```
fs stripe1: /mnt/stripel          enabled
  device = /dev/cxvm/stripel
  force = false
  options = []
  servers = cxfsirix5 (0), cxfsirix4 (1)
  clients = cxfsirix4, cxfsirix5, cxfsirix6, cxfsmac4, cxfssun1
xvm:
  vol/stripel                    0 online,open
    subvol/stripel/data          2292668416 online,open
      stripe/stripel            2292668416 online,open
        slice/d9400_0s0         1146334816 online,open
        slice/d9400_1s0         1146334816 online,open

data size: 1.07 TB
```

It is then possible to run the `xvm` command to identify the XVM volumes and disk devices. This provides enough information to identify the device's WWN, LUN, and

controller. In the following example, the `slice/d9400_0s0` from `phys/d9400_0` is LUN 0 located on a RAID controller with WWN 200500a0b80cedb3.

```
irix# xvm show -e -t vol
vol/stripel                0 online,open
  subvol/stripel/data      2292668416 online,open
    stripe/stripel        2292668416 online,open (unit size: 1024)
      slice/d9400_0s0      1146334816 online,open (d9400_0:/dev/rdisk/200500a0b80cedb3/lun0vol/c2p1)
      slice/d9400_1s0      1146334816 online,open (d9400_1:/dev/rdisk/200400a0b80cedb3/lun1vol/c3p1)
```

On all platforms other than Windows, it is then possible to run the `xvm` command on the client to identify the matching disk devices on the client:

```
solaris# /usr/cxfs_cluster/bin/xvm show -e -t vol
vol/stripel                0 online,open
  subvol/stripel/data      2292668416 online,open
    stripe/stripel        2292668416 online,open (unit size: 1024)
      slice/d9400_0s0      1146334816 online,open (d9400_0:pci@9,600000/JNI,FCR@2,1/sd@2,0)
      slice/d9400_1s0      1146334816 online,open (d9400_1:pci@9,600000/JNI,FCR@2/sd@2,1)
```

The process to map device names to identify the target WWN is platform-specific. See:

- "Mapping XVM Volumes to Storage Targets on Mac OS X" on page 98
- "Mapping XVM Volumes to Storage Targets on Solaris" on page 138

---

**Note:** There is no `xvm` command on the Windows platform and therefore no method to map XVM volumes directly to disk devices under Windows.

---

If a disk device has not been found for a particular volume element, the following message will be displayed instead of the device name:

```
no direct attachment on this cell
```

For example:

```
solaris# /usr/cxfs_cluster/bin/xvm show -e -t volvol/stripel
0 online,open,no physical connection
  subvol/stripel/data      2292668416 online,open
    stripe/stripel        2292668416 online,open (unit size: 1024)
      slice/d9400_0s0      1146334816 online,open (d9400_0:no direct attachment on this cell)
      slice/d9400_1s0      1146334816 online,open (d9400_1:no direct attachment on this cell)
```

Using the device information from the CXFS administration node, it should then be possible to determine if the client can see the same devices using the client HBA tools and the RAID configuration tool.

To see the complete list of volumes and devices mappings, especially when XVM failover is configured, run:

```
solaris# /usr/cxfs_cluster/bin/xvm show -v phys
```

For more information about xvm, see the *XVM Volume Manager Administrator's Guide*.



## General Troubleshooting

This chapter contains the following:

- "Identifying Problems" on page 231
- "Typical Problems and Solutions" on page 233
- "Reporting Problems to SGI" on page 238

Also see the following platform-specific sections:

- "Troubleshooting for AIX" on page 39
- "Troubleshooting for Linux" on page 72
- "Troubleshooting for Mac OS X" on page 99
- "Troubleshooting for Solaris" on page 139
- "Troubleshooting for Windows" on page 202

### Identifying Problems

This section provides tips about identifying problems:

- "Is the Client-Only Node Configured Correctly? " on page 231
- "Is the Client-Only Node in Membership?" on page 232
- "Is the Client-Only Node Mounting All Filesystems?" on page 232
- "Can the Client-Only Node Access All Filesystems?" on page 232
- "Are There Error Messages?" on page 233

#### Is the Client-Only Node Configured Correctly?

To determine the current configuration of a node in a cluster, run the following command on a CXFS administration node:

```
/usr/cluster/bin/cxfs-config -all
```

For more information, see "Verifying the Cluster Status" on page 223.

Confirm that the host type, private network, and failure hierarchy are configured correctly, and that no warnings or errors are reported. You should rectify any warnings or errors before proceeding with further troubleshooting.

### **Is the Client-Only Node in Membership?**

To determine if the node is in the cluster membership, use the tools described in "Verifying the Cluster Status" on page 223.

If the client is not in membership, see the following:

- "Verifying the Cluster Configuration" on page 222
- "Verifying Connectivity in a Multicast Environment" on page 222
- "Unable to Achieve Membership" on page 234

### **Is the Client-Only Node Mounting All Filesystems?**

To determine if the node has mounted all configured filesystems, use the tools described in "Verifying the Cluster Status" on page 223.

If the client has not mounted all filesystems, see the following:

- "Verifying the Cluster Configuration" on page 222
- Appendix C, "Mount Options Support" on page 247
- "Determining If a Client-Only Node Is Fenced" on page 236
- "Verifying Access to XVM Volumes" on page 227

### **Can the Client-Only Node Access All Filesystems?**

To determine if the client-only node can access a filesystem, navigate the filesystem and attempt to create a file.

If the filesystem appears to be empty, the mount may have failed or been lost. See "Determining If a Client-Only Node Is Fenced" on page 236 and "Verifying Access to XVM Volumes" on page 227.

If accessing the filesystem hangs the viewing process, see "Filesystem Appears to Be Hung" on page 235.

## Are There Error Messages?

When determining the state of the client-only node, you should check error message logs to help identify any problems.

Appendix A, "Operating System Path Differences" on page 239 lists the location of the `cxfs_client` log file for each platform. This log is also displayed in the Windows version of `cxfs_info`.

Each platform also has its own system log for kernel error messages that may also capture CXFS messages. See the following:

- "Log Files on AIX" on page 21
- "Log Files on Linux" on page 49
- "Log Files on Mac OS X" on page 81
- "Log Files on Solaris" on page 106
- "Log Files and Cluster Status for Windows" on page 149

There are various logs also located on the CXFS administration nodes. For more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

## Typical Problems and Solutions

This section contains the following typical problems that apply to any platform:

- "Unable to Achieve Membership" on page 234
- "Filesystem Appears to Be Hung" on page 235
- "Determining If a Client-Only Node Is Fenced" on page 236
- "No HBA WWPNs are Detected" on page 238

## Unable to Achieve Membership

If `cxfs_info` does not report that CMS is UP, do the following:

1. Check that `cxfs_client` is running. See one of the following sections as appropriate for your platform:
  - "Start/Stop `cxfs_client` Daemon for AIX" on page 36
  - "Start/Stop `cxfs_client` for Linux" on page 67
  - "Start/Stop `cxfs_client` for Mac OS X" on page 96
  - "Start/Stop `cxfs_client` for Solaris" on page 135
  - "Start/Stop the CXFS Client Service for Windows" on page 196
2. Verify that the license check has succeeded in the `cxfs_client` log file.
3. Look for other warnings and error messages in the `cxfs_client` log file. See Appendix A, "Operating System Path Differences" on page 239 for the location of the log file on different platforms.
4. Check `cxfs-config` output on the CXFS administration node to ensure that the client is correctly configured and is reachable via the configured CXFS private network. For example:

```
irix# /usr/cluster/bin/cxfs-config -all
```
5. Check that the client is enabled into the cluster by running `clconf_info` on a CXFS administration node.
6. Look in the system log on the CXFS metadata server to ensure the server detected the client that is attempting to join membership and check for any other CXFS warnings or errors.
7. Check that the metadata server has the node correctly configured in its hostname lookup scheme (`/etc/host` file or DNS).
8. If you are still unable to resolve the problem, reboot the client node.
9. If rebooting the client node in step 8 did not resolve the problem, restart the cluster administration daemons (`fs2d`, `cad`, `cmond`, and `crsd`) on the metadata server. This step may result in a temporary delay in access to the filesystem from all nodes.

10. If restarting cluster administration daemons in step 9 did not solve the problem, reboot the metadata server. This step may result in the filesystems being unmounted on all nodes.

## Filesystem Appears to Be Hung

If any CXFS filesystem activity appears to hung in the filesystem, do the following:

1. Check that the client is still in membership and the filesystem is mounted according to `cxfs_info`.
2. Check on the metadata server to see if any messages are more than a few seconds in age (known as a *stuck message*). For example, on IRIX running `icrash` as `root`, the following message was received from cell 4 more than four minutes ago:

```
# icrash
>>>> mesglist
Cell:1
THREAD ADDR          MSG ID TYPE CELL MESSAGE
Time(Secs)
=====
0xa80000004bc86400   10fc  Rcv   4                I_dsxvn_allocate      4:20
```

3. If there is a stuck message, gather information for SGI support:

- Find the stack trace for the stuck thread. For example:

```
>>>> kthread 0xa80000004bc86400

          KTHREAD TYPE          ID          WCHAN NAME
=====
a80000004bc86400    1          100000534  c000000002748008 mtcp_notify
=====
1 kthread struct found

>>>> defkthread 0xa80000004bc86400

Default kthread is 0xa80000004bc86400

>>>> trace
```

```
=====
STACK TRACE FOR XTHREAD 0xa80000004bc86400 (mtcp_notify):

1 istswtch[../os/swtch.c: 1526, 0xc00000000021764c]
2 swtch[../os/swtch.c: 1026, 0xc000000000216de8]
3 thread_block[../os/ksync/mutex.c: 178, 0xc00000000017dc8c]
4 sv_queue[../os/ksync/mutex.c: 1595, 0xc00000000017f36c]
5 sv_timedwait[../os/ksync/mutex.c: 2205, 0xc0000000001800a0]
6 sv_wait[../os/ksync/mutex.c: 1392, 0xc00000000017f038]
7 xlog_state_sync[../fs/xfx/xfx_log.c: 2986, 0xc0000000002a535c]
8 xfs_log_force[../fs/xfx/xfx_log.c: 361, 0xc0000000002a25dc]
9 cxfs_dsxvn_wait_inode_safe[../fs/cxfs/server/cxfs_dsxvn.c: 2011,
0xc00000000046a594]
10 dsvn_getobjects[../fs/cxfs/server/dsvn.c: 3266, 0xc0000000004676fc]
11 I_dsxvn_allocate[../fs/cxfs/server/cxfs_dsxvn.c: 1406, 0xc0000000004699c8]
12 dsxvn_msg_dispatcher[../IP27bootarea/I_dsxvn_stubs.c: 119,
0xc000000000456768]
13 mesg_demux[../cell/mesg/mesg.c: 1130, 0xc000000000408e88]
14 mtc_notify[../cell/mesg/mesg_tcp.c: 1100, 0xc0000000004353d8]
15 tsv_thread[../cell/tsv.c: 303, 0xc000000000437738]
16 xthread_prologue[../os/swtch.c: 1638, 0xc00000000021782c]
17 xtresume[../os/swtch.c: 1686, 0xc0000000002178f8]
=====
```

- Run `cxfsdump` on the metadata server.
  - Run `cxfsdump` on the client that has the stuck message.
  - If possible, force the client that has the stuck message to generate a crash dump.
4. Reboot the client that has the stuck message. This is required for CXFS to recover.

### Determining If a Client-Only Node Is Fenced

To determine if a client-only node is fenced, log in to a CXFS administration node and use the `hafence(1M)` command. A fenced port is displayed as `status=disabled`.

In the following example, all ports that have been registered as CXFS host ports are not fenced:

```
irix# /usr/cluster/bin/hafence -q
Switch[0] "brocade04" has 16 ports
Port 4 type=FABRIC status=enabled hba=210000e08b0042d8 on host o200c
Port 5 type=FABRIC status=enabled hba=210000e08b00908e on host cxfs30
Port 9 type=FABRIC status=enabled hba=2000000173002d3e on host cxfssun3
```

All switch ports can also be shown with hafence:

```
irix# /usr/cluster/bin/hafence -v
Switch[0] "brocade04" has 16 ports
Port 0 type=FABRIC status=enabled hba=2000000173003b5f on host UNKNOWN
Port 1 type=FABRIC status=enabled hba=2000000173003adf on host UNKNOWN
Port 2 type=FABRIC status=enabled hba=210000e08b023649 on host UNKNOWN
Port 3 type=FABRIC status=enabled hba=210000e08b021249 on host UNKNOWN
Port 4 type=FABRIC status=enabled hba=210000e08b0042d8 on host o200c
Port 5 type=FABRIC status=enabled hba=210000e08b00908e on host cxfs30
Port 6 type=FABRIC status=enabled hba=2000000173002d2a on host UNKNOWN
Port 7 type=FABRIC status=enabled hba=2000000173003376 on host UNKNOWN
Port 8 type=FABRIC status=enabled hba=2000000173002c0b on host UNKNOWN
Port 9 type=FABRIC status=enabled hba=2000000173002d3e on host cxfssun3
Port 10 type=FABRIC status=enabled hba=2000000173003430 on host UNKNOWN
Port 11 type=FABRIC status=enabled hba=200900a0b80c13c9 on host UNKNOWN
Port 12 type=FABRIC status=disabled hba=0000000000000000 on host UNKNOWN
Port 13 type=FABRIC status=enabled hba=200d00a0b80c2476 on host UNKNOWN
Port 14 type=FABRIC status=enabled hba=1000006069201e5b on host UNKNOWN
Port 15 type=FABRIC status=enabled hba=1000006069201e5b on host UNKNOWN
```

When the client-only node joins membership, any fences on any switch ports connected to that node should be lowered and the status changed to enabled.

However, if the node still does not have access to the storage, do the following:

- Check that the HBA WWPNs were correctly identified. See "Verifying the I/O Fencing Configuration" on page 226.
- Check the `cxfs_client` log file for warnings or errors while trying to determine the HBA WWPNs. See "No HBA WWPNs are Detected" on page 238.
- Log into the Fibre Channel switch. Check the status of the switch ports and confirm that the WWPNs match those identified by `cxfs_client`.

## No HBA WWPNs are Detected

On most platforms, the `cxfs_client` software automatically detects the world wide port names (WWPNs) of any supported host bus adapters (HBAs) in the system that are connected to a switch that is configured in the cluster database. These HBAs will then be available for fencing.

However, if no WWPNs are detected, there will be messages about loading the HBA/SNIA library.

See the following:

- "I/O Fencing for AIX" on page 35
- "I/O Fencing for Linux" on page 65
- "I/O Fencing for Mac OS X" on page 93
- "I/O Fencing for Solaris" on page 133
- "I/O Fencing on Windows" on page 193

## Reporting Problems to SGI

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 AIX Problems" on page 44
- "Reporting Linux Problems" on page 76
- "Reporting Mac OS X Problems" on page 101
- "Reporting Solaris Problems" on page 142
- "Reporting Windows Problems" on page 211

## Operating System Path Differences

This appendix lists the location of CXFS-specific commands and files. For more information, see the `cxfs_client` man page.

**Table A-1** AIX Paths

Component	Path
CXFS client daemon:	<code>/usr/cxfs_cluster/bin/cxfs_client</code>
Command that normally invokes the client daemon:	<code>/usr/cxfs_cluster/bin/cxfs_cluster</code>
Log file:	<code>/var/tmp/cxfs_client</code>
Options file:	<code>/usr/cxfs_cluster/bin/cxfs_client.options</code>
CXFS status:	<code>/usr/cxfs_cluster/bin/cxfs_info</code>
CXFS license file:	<code>/etc/flexlm/license.dat</code>
Hostname/address information	<code>/etc/hosts</code>
GRIo v2 administration	<code>/usr/cxfs_cluster/bin/grioadmin</code>
GRIo v2 quality of service	<code>/usr/cxfs_cluster/bin/grioqos</code>
XVM query	<code>/usr/cxfs_cluster/bin/xvm</code>

**Table A-2** Linux Paths

---

Component	Path
CXFS client service:	/usr/cluster/bin/cxfs_client
Command that normally invokes the client daemon:	/etc/init.d/cxfs_client
Log file:	/var/log/cxfs_client
Options file:	/etc/cluster/config/cxfs_client.options
CXFS status:	/usr/cluster/bin/cxfs_info
CXFS license file:	/etc/flexlm/license.dat
Hostname/address information	/etc/hosts
GRIo v2 administration	/usr/sbin/grioadmin
GRIo v2 quality of service	/usr/sbin/grioqos
XVM query	/sbin/xvm

---

**Table A-3** Mac OS X Paths

Component	Path
CXFS client daemon:	/usr/cluster/bin/cxfs_client
Command that normally invokes the client daemon:	/Library/StartupItems/cxfs/cxfs
Log file:	/var/log/cxfs_client
Options file:	/usr/cluster/bin/cxfs_client.options
CXFS status:	%ProgramFiles%\CXFS\cxfs_info.exe
CXFS license file:	/etc/flexlm/license.dat
Hostname/address information	/etc/hosts
GRIO v2 administration	/usr/sbin/grioadmin
GRIO v2 quality of service	/usr/sbin/grioqos
XVM query	/usr/cluster/bin/xvm

**Table A-4** Solaris Paths

---

Component	Path
CXFS client daemon:	/usr/cxfs_cluster/bin/cxfs_client
Command that normally invokes the client daemon:	/etc/init.d/cxfs_cluster
Log file:	/var/log/cxfs_client
Options file:	/usr/cxfs_cluster/bin/cxfs_client.options
CXFS status:	/usr/cxfs_cluster/bin/cxfs_info
CXFS license file:	/etc/flexlm/license.dat
Hostname/address information	/etc/hosts
GRIo v2 administration	/usr/sbin/grioadmin
GRIo v2 quality of service	/usr/sbin/grioqos
XVM query	/usr/cxfs_cluster/bin/xvm

---

**Table A-5** Windows Paths

Component	Path
CXFS client service:	%SystemRoot%\system32\cxfs_client.exe
Command that normally invokes the client service:	See "Start/Stop the CXFS Client Service for Windows" on page 196
Log file:	%ProgramFiles%\CXFS\log\cxfs_client.log
Options file:	"Modifying the CXFS Software for Windows" on page 197
CXFS status:	%ProgramFiles%\CXFS\cxfs_info.exe
CXFS license file:	%ProgramFiles%\CXFS\lib\license.dat
Hostname and address information:	%SystemRoot%\system32\drivers\etc\hosts
GRIO v2 administration:	%ProgramFiles%\CXFS\grioadmin.exe
GRIO v2 quality of service:	%ProgramFiles%\CXFS\griooqs.exe
XVM query:	(unsupported)



## **Filesystem and Logical Unit Specifications**

Table B-1 on page 246 summarizes filesystem and logical unit specifications differences among the client-only platforms.

**Table B-1** Filesystem and Logical Unit Specifications

Item	AIX	Linux x86	Linux AMD64/EM64T and Itanium 2	Mac OS X	Solaris	Windows
Maximum filesystem size	2 <sup>64</sup> bytes <sup>1</sup>	16 TB	2 <sup>64</sup> bytes	2 <sup>64</sup> bytes	2 <sup>64</sup> bytes	2 <sup>64</sup> bytes
Maximum file size/offset	16 TB <sup>2</sup>	16 TB <sup>3</sup>	2 <sup>63</sup> -1 bytes	2 <sup>63</sup> -1 bytes	2 <sup>63</sup> -1 bytes	2 <sup>63</sup> -1 bytes
Block size (in bytes) <sup>4</sup>	4096 (XFS default)	512, 1024, 2048, or 4096	512, 1024, 2048, or 4096	4096, 8192, 16384, 32768, or 65536	2048, 4096, 8192, 16384, 32768, or 65536 <sup>5</sup>	512, 1024, 2048, 4096, 8192, 16384, 32768, or 65536
Physical LUN limit	1 TB (AIX 5.1 + MR 5) 2 TB (AIX 5.2 + MR 2)	2 TB	2 TB	2 TB	1 TB	2 TB

<sup>1</sup> About 18 million terabytes

<sup>2</sup> Assumes the default `ulimit` is changed, see "Limitations and Considerations for AIX" on page 22.

<sup>3</sup> Using large file support (`O_LARGEFILE`)

<sup>4</sup> If the filesystem is to be accessible by other platforms in a multiOS cluster, its block size must be supported on all platforms in the cluster.

<sup>5</sup> 8192 is recommended

## Mount Options Support

The table in this appendix list the mount options that are supported by CXFS, depending upon the server platform. Some of these mount options affect only server behavior and are ignored by client-only nodes.

The tables also list those options that are not supported, especially where that support varies from one platform to another. Both the IRIX and the SGI ProPack for Linux mount commands support many additional options, but these options may be silently ignored by the clients, or cause the mount to fail and should be avoided.

For more information, see the IRIX `mount(1M)` and Linux `mount(8)` man pages.

---

**Note:** The following are mandatory, internal CXFS mount options that cannot be modified and are set by `clconfd` and `cxfs_client`:

```
client_timeout
server_list
```

---

The table uses the following abbreviations:

I = IRIX metadata server  
P = ProPack for Linux metadata server  
S = Supported  
n = Not supported

A blank space within the table means that the option has not been verified.

Table C-1 Mount Options Support for Client-Only Platforms

Mount Option	Client-only Platform													
	AIX		IRIX		Linux 64		ProPack for Linux		Mac OS X		Solaris		Windows	
	I	P	I	P	I	P	I	P	I	P	I	P	I	P
biosize	S	S	S	S	S	S	S	S	S	S	S	S	S	
client_timeout	S	S	S	S	S	S	S	S	S	S	S	S	S	S
dmapi	n	n	n	n	n	n	n	n	n	n	n	n	n	n
dmi			S	S			S	S	S	S	S	S	S	
gqnoenforce			S	S			S	S	S	S	S	S	S	
gquota			S	S			S	S	S	S	S	S	S	
grpuid		n	S	n		n	n	n	S	n	S	n	S	n
inode64	S		S		n	n	S	S	n	n	S	S	S	
logbufs			S										S	
noatime			S	S			S	S	S	S	S	S	S	
noauto	n	n	n	n	n	n	n	n	n	n	n	n	n	n
nodev			S	S			S	S	S	S	n	n	S	
noquota			S	S			S	S	S	S	S	S	S	
nosuid			S	S			S	S	S	S	S	S	S	
osyncisdsync			S										S	
pqnoenforce			S										S	
pquota			S										S	
qnoenforce			S	S			S	S	S	S	S	S	S	
quota			S	S			S	S	S	S	S	S	S	
ro			S	S			S	S	S	S	S	S	S	
rw			S	S			S	S	S	S	S	S	S	
server_list	S	S	S	S	S	S	S	S	S	S	S	S	S	S
sunit			S										S	

Mount Option	Client-only Platform													
	AIX		IRIX		Linux 64		ProPack for Linux		Mac OS X		Solaris		Windows	
	I	P	I	P	I	P	I	P	I	P	I	P	I	P
swalloc			S										S	
swidth			S										S	
uqnoenforce			S	S			S	S	S	S	S	S	S	
uquota			S	S			S	S	S	S	S	S	S	
wsync			S										S	



## Error Messages

The following are commonly seen error messages.

### Windows Messages

The following are common Windows CXFS messages.

```
cis_driver_init() failed: could not open handle to driver
cis_driver_init() failed: could not close handle to CXFS driver
    The CXFS driver may not have successfully started. Check the system event log for errors.
```

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

Even with `passwd` and `group` warnings above, filesystem mounts should proceed; however, all users will be given `nobody` credentials and will be unable to view or modify files on the CXFS filesystems. For more information about these files, see "Log Files on Solaris" on page 106 and "Log Files and Cluster Status for Windows" on page 149. Also see the log files on the CXFS administration node; for more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

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

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

The registry entries for the location of the `passwd`, `group`, or `fencing.conf` files may be missing, or the path provided on the command line to the CXFS Client service is badly formed. Reset these values by modifying the current installation as described in "Modifying the CXFS Software for Windows" on page 197.

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

Check that the `passwd`, `group` and `fencing.conf` files are in the configured location and are accessible as described in "Checking Permissions on the Password and Group Files for Windows" on page 193.

`no valid users configured in passwd file`

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

`no valid groups configured in group file`

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

`op_failed ERROR: Mount failed for concat0`

A filesystem mount has failed and will be retried.

`unable to create mount point`

`Configured drive letter may already be in use`

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

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

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

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

## CMS Error Messages

The following messages may be logged by CMS.

`CMS excluded cells 0xXXX with incomplete connectivity`

Generated when CMS delivers a membership that excluded some **new** cells that had not established connections with enough cells yet to be admitted. `0xXXX` is a bitmask of excluded cells.

`CMS calculation limited to last membership:configuration change incomplete on cells 0xXXX`

Generated when the leader is attempting to make a configuration change current (that is, actually use the change on all nodes), but some cells in the cluster have not yet received the configuration change staged (uploaded and ready to be made current). `0xXXX` is a bitmask of cells that do not yet have the change in their configuration. Changes make their way through the cluster asynchronously, so this situation is

expected. It can take a few attempts by the CMS leader before all nodes have the change staged. As long as this situation resolves eventually, there is no problem.

```
CMS calculation limited to last membership:recovery incomplete
```

Generated when new members were disallowed due to recovery from the last cell failure that is still being processed.

## Mount Messages

```
cxfs_client: op_failed ERROR : Mount failed for aixdisk0s0
```

A filesystem mount has failed on an AIX node and will be retried

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

A filesystem mount has failed on an Linux 32-bit, Mac OS X, Solaris, or Windows node and will be retried.

## Network Connectivity Messages

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

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



## Summary of New Features from Previous Releases

This appendix contains a summary of the new features for each version of this guide.

### CXFS MultiOS 2.0

Original publication (007-4507-001) supporting Solaris client-only nodes in a multiOS cluster with IRIX metadata servers.

### CXFS MultiOS 2.1

The 007-4507-002 update contains the following:

- Support for Windows NT nodes in a CXFS multiOS cluster. Platform-specific information is grouped into separate chapters.
- Support for up to four JNI HBAs in each CXFS Solaris node.

---

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

---

### CXFS MultiOS 2.1.1

The 007-4507-003 update contains the following:

- References to using the latest software from the JNI website (<http://www.jni.com/Drivers>).
- Information about ensuring that appropriate software is installed on the IRIX nodes that are potential metadata servers.
- Clarifications to the use of I/O fencing and serial reset.
- Corrections to the procedure in the “Solaris Installation Overview” section and other editorial corrections.

## CXFS MultiOS 2.2

The 007-4507-004 update contains the following:

- Support for Microsoft Windows 2000 nodes in a CXFS MultiOS cluster. This guide uses *Windows* to refer to both Microsoft Windows NT and Microsoft Windows 2000 systems.
- Support for SGI TP9100s. For additional details, see the release notes.
- A new section about configuring two HBAs for failover operation.
- Support for the JNI 5.1.1 and later driver on Solaris clients, which simplifies the installation steps.
- DMAPI support for all platforms.
- Removal of the Solaris limitation requiring more kernel threads.

## CXFS MultiOS 2.3

The 007-4507-005 update contains the following:

- Updated Brocade Fibre Channel switch firmware levels.
- Filename corrections the chapters about FLEXlm licensing for Windows and modifying CXFS software on a Solaris system.

## CXFS MultiOS 2.4

The 007-4507-006 update contains the following:

- Support for Sun Microsystems Solaris 9 and specific Sun Fire systems.
- Support for the JNI EZ Fibre release 2.2.1 or later.
- A cluster of as many as 32 nodes, of which as many as 16 can be CXFS administration nodes; the rest will be client-only nodes.
- Information about the **Node Function** field, which replaces node weight. For Solaris and Windows nodes, **Client-Only** is automatically selected for you. Similar fields are provided for the `cmgr` command. For more information, see the *CXFS Administration Guide for SGI InfiniteStorage*.

- Clarification that if the primary HBA path is at fault during the Windows boot up (for example, if the Fibre Channel cable is disconnected), no failover to the secondary HBA path will occur. This is a limitation of the QLogic driver.
- Reference to the availability of cluster information on Windows nodes.
- Information about enabling Brocade Fibre Channel switch ports.
- Additional information about functional limitations specific to Windows, and performance considerations, and access controls.

## CXFS MultiOS 2.5

The 007-4507-007 update contains the following:

- Support for the IBM AIX platform, Linux on supported 32-bit platforms, SGI ProPack for Linux on Altix servers.
- Support for a cluster of up to 48 nodes, 16 of which can be CXFS administration nodes; the rest must be client-only nodes.
- For Windows nodes, user identification with lightweight directory access protocol (LDAP).
- Support of forced unmount of filesystems on Windows nodes.
- Information about protecting data integrity if JNI Fibre Channel cables are disconnected or fail.
- Support for the SGI TP9500 RAID.
- Support for the QLogic 2342 host bus adapter.
- Information about new `cxfs-reprobe` scripts on AIX, IRIX, Linux, and Solaris nodes. These scripts are run by either `clconfd` or `cxfs_client` when they need to reprobe the Fibre Channel controllers. The administrator may modify these scripts if needed.
- Information about setting the `ntcp_nodelay` system tunable parameter in order to provide adequate performance on file deletes.
- Automatic detection of HBAs is provided for Linux, Solaris, and Windows nodes.

## CXFS MultiOS 3.0

The 007-4507-008 update contains the following:

- Support for the Microsoft Windows XP client.

---

**Note:** The CXFS multiOS 3.0 release is the last release that will support the Microsoft Windows NT 4.0 platform. The 3.1 release will not include software for Windows NT 4.0.

---

- Clarifications to the terminology and installation information for Linux 32-bit clients.
- Information about Linux 64-bit clients running SGI ProPack for Linux on SGI Altix 3000 systems has been removed and will appear in the *CXFS Administration Guide for SGI InfiniteStorage* that support CXFS 3.0 for SGI ProPack 2.3 for Linux.

## CXFS MultiOS 3.1

The 007-4507-009 update contains the following:

- Support for the Apple Computer, Inc. Mac OS X operating system on client-only nodes.
- Support for a cluster of up to 64 nodes.
- Information about the SGI TP9300, SGI TP9300S, and SGI TP9500S.
- Information about setting the LUN discovery method for Solaris systems using the SGI TP9100 1-Gbit controller
- Additional AIX troubleshooting information.

## CXFS MultiOS 3.2

The 007-4507-010 update contains the following:

- Support for Mac OS X 10.3.5 and Apple host bus adapters (HBAs).

---

**Note:** Mac OS X 10.2.x and the Astera HBA are not supported with the CXFS 3.2 release.

---

- Support for Red Hat Enterprise Linux 3. If you are running a Red Hat Enterprise Linux 3 kernel and you want to use quotas on a CXFS filesystem, you must install the quota package.
- Support for the Sun Fire V210 server as a multiOS client platform.
- A summary of the maximum filesystem size, file size, and block size for each platform.
- Information about the environment variables you must define in the `/etc/cluster/config/cxfs_client.options` file in order for the `/etc/cluster/config/cxfs-reprobe` script to appropriately probe all of the targets on the SCSI bus for the Linux platform on third-party hardware.
- Availability of the new `xvm_maxdma` attribute to the AIX `chdev` command, used to change the maximum XVM direct memory access (DMA) size to improve direct I/O performance.
- Information about ensuring proper hostname configuration for a Windows node.
- XVM volume names are limited to 31 characters and subvolumes are limited to 26 characters.
- Information about mount options.
- Updates to the procedure for installing the AMCC JNI HBA.
- Clarification that the AMCC JNI HBA that is provided by Sun Microsystems **does not function with CXFS** and cannot be configured to do so. You must purchase the JNI HBA directly from AMCC.



---

## Glossary

### **active metadata server**

A server-capable administration node chosen from the list of potential metadata servers. There can be only one active metadata server for any one filesystem.

### **administration node**

A node in the pool that is installed with the `cluster_admin.sw.base` software product, allowing the node to perform cluster administration tasks and contain a copy of the cluster database. There are two types of administration nodes: server-capable administration nodes and client administration nodes.

### **administrative stop**

See *forced CXFS shutdown*

### **cell ID**

A number associated with a node that is used by the CXFS software and appears in messages.

### **CLI**

Underlying command line interface commands used by the CXFS Manager graphical user interface (GUI) and the `cmgr` command.

### **client**

See *CXFS client node*, *CXFS client-only node* and *administration node*.

### **client administration node**

A node that is installed with the `cluster_admin` software product, allowing the node to perform cluster administration tasks and contain a copy of the cluster database, but is not capable of coordinating CXFS metadata. Only to be used with FailSafe.

**client-only node**

A node that is installed with the `cxfs_client.sw.base` software product; it does not run cluster administration daemons and is not capable of coordinating CXFS metadata. Any node can be client-only node. See also *CXFS administration node* and *client administration node*.

**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. Cluster names must be unique.

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 administration daemons**

The set of daemons on a CXFS administration node that provide the cluster infrastructure: `fs2d`, `cad`, `cmond`, `crsd`.

**cluster administrator**

The person responsible for managing and maintaining a cluster.

**cluster database**

Contains configuration information about all nodes and the cluster. The database is managed by the cluster administration daemons.

**cluster domain**

XVM concept in which a filesystem applies to the entire cluster, not just to the local node. See also *local domain*.

**cluster database membership**

The group of administration nodes in the **pool** that are accessible to cluster administration daemons and therefore are able to receive cluster database updates; this may be a subset of the nodes defined in the pool. The cluster administration daemons manage the distribution of the cluster database (CDB) across the administration nodes in the pool. (Also known as *user-space membership* and *fs2d database membership*.)

**cluster ID**

A unique number within your network in the range 1 through 128. The cluster ID is used by the operating system kernel to make sure that it does not accept cluster information from any other cluster that may be on the network. The kernel does not use the database for communication, so it requires the cluster ID in order to verify cluster communications. This information in the kernel cannot be changed after it has been initialized; therefore, you must not change a cluster ID after the cluster has been defined. Clusters IDs must be unique.

**cluster mode**

One of two methods of CXFS cluster operation, `Normal` or `Experimental`. In `Normal` mode, CXFS resets any node for which it detects heartbeat failure; in `Experimental` mode, CXFS ignores heartbeat failure. `Experimental` mode allows you to use the kernel debugger (which stops heartbeat) without causing node failures. You should only use `Experimental` mode during debugging.

**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.

**cluster node**

A node that is defined as part of the cluster. See also *node*.

**coexecution**

The ability to run CXFS and IRIS FailSafe together.

**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 daemon**

The daemon (`cxfs_client`) that controls CXFS services on a client-only node.

**CXFS control daemon**

The daemon (`clconfd`) that controls CXFS services on an administration node.

**CXFS database**

See *cluster database*.

**CXFS kernel 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 kernel membership. Once accepted, the node can share the filesystems of the cluster. (Also known as *kernel-space membership*.) CXFS kernel membership differs from *cluster database membership* and FailSafe membership. For more information about FailSafe, see *FailSafe Administrator's Guide for SGI InfiniteStorage*.

**CXFS services**

The enabling/disabling of a node, which changes a flag in the cluster database. This disabling/enabling does not affect the daemons involved. The daemons that control CXFS services are `clconfd` on an administration node and `cxfs_client` on a client-only node.

**CXFS services start**

To enable a node, which changes a flag in the cluster database, by using an administrative task in the CXFS GUI or the `cmgr` command.

**CXFS services stop**

To disable a node, which changes a flag in the cluster database, by using a CXFS graphical user interface (GUI) or the `cmgr` command. See also *forced CXFS shutdown*.

**CXFS shutdown**

See *forced CXFS shutdown* and *shutdown*

**CXFS tiebreaker node**

A node identified as a tiebreaker for CXFS to use in the process of computing CXFS kernel membership for the cluster, when exactly half the nodes in the cluster are up and can communicate with each other. There is no default CXFS tiebreaker. SGI recommends that the tiebreaker node be a client-only node. The CXFS tiebreaker differs from the FailSafe tiebreaker; see *FailSafe Administrator's Guide for SGI InfiniteStorage*.

**database**

See *cluster database*.

**database membership**

See *cluster database membership*.

**details area**

The portion of the GUI window that displays details about a selected component in the view area. See also *view area*.

**domain**

See *cluster domain* and *local domain*.

**FailSafe Membership**

The group of nodes that are actively sharing resources in the cluster, which may be a subset of the nodes defined in a cluster. FailSafe membership differs from *CXFS kernel membership* and *cluster database membership*. For more information about FailSafe, see *FailSafe Administrator's Guide for SGI InfiniteStorage*.

**failure action hierarchy**

The set of instructions that determine what happens to a failed node; the second instruction will be followed only if the first instruction fails; the third instruction will be followed only if the first and second fail. The available actions are: *I/O fencing*, *reset*, and *shutdown*.

**fencing**

See *I/O fencing*.

**fencing recovery**

The process of recovery from fencing, in which the affected node automatically withdraws from the CXFS kernel membership, unmounts all file systems that are using an I/O path via fenced HBA(s), and then rejoins the cluster.

**forced CXFS shutdown**

The withdrawal of a node from the CXFS kernel membership, either due to the fact that the node has failed somehow or by issuing an `admin cxfs_stop` command. This disables filesystem and cluster volume access for the node. The node remains enabled in the cluster database. See also *CXFS services stop* and *shutdown*.

**fs2d database membership**

See *cluster database membership*.

**heartbeat messages**

Messages that cluster software sends between the nodes that indicate a node is up and running. Heartbeat messages and *control messages* are sent through the node's network interfaces that have been attached to a control network.

**heartbeat interval**

The time between heartbeat messages. The node timeout value must be at least 10 times the heartbeat interval for proper CXFS operation. The higher the number of heartbeats (smaller heartbeat interval), the greater the potential for slowing down the 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.

**kernel-space membership**

See *CXFS kernel membership*.

**local domain**

XVM concept in which a filesystem applies only to the local node, not to the cluster. See also *cluster domain*.

**log configuration**

A log configuration has two parts: a *log level* and a *log file*, both associated with a *log group*. The cluster administrator can customize the location and amount of log output, and can specify a log configuration for all nodes or for only one node. For example, the *crsd* log group can be configured to log detailed level-10 messages to the *crsd-foo* log only on the node *foo* and to write only minimal level-1 messages to the *crsd* log on all other nodes.

**log file**

A file containing notifications for a particular *log group*. A log file is part of the *log configuration* for a log group.

**log group**

A set of one or more CXFS processes that use the same log configuration. A log group usually corresponds to one daemon, such as *gcd*.

**log level**

A number controlling the number of log messages that CXFS will write into an associated log group's log file. A log level is part of the log configuration for a log group.

**membership**

See *cluster database membership* and *CXFS kernel membership*.

**membership version**

A number associated with a node's cell ID that indicates the number of times the CXFS kernel membership has changed since a node joined the membership.

**metadata**

Information that describes a file, such as the file's name, size, location, and permissions.

**metadata server**

The administration node that coordinates updating of meta data 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.

**metadata server recovery**

The process by which the metadata server moves from one node to another due to an interruption in CXFS services on the first node. See also *recovery*

**multiOS**

A cluster that is running multiple operating systems, such as IRIX and Solaris.

**multiport serial adapter cable**

A device that provides four DB9 serial ports from a 36-pin connector.

**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.

See also *CXFS administration node*, *client administration node*, *client-only node*, *server-capable administration node*, and *standby node*,

**node ID**

An integer in the range 1 through 32767 that is unique among the nodes in the pool. If you do not specify a number, CXFS will calculate an ID for you. You must not change the node ID number after the node has been defined.

**node membership**

The list of nodes that are active (have CXFS kernel membership) in a cluster.

**node timeout**

If no heartbeat is received from a node in this period of time, the node is considered to be dead. The node timeout value must be at least 10 times the heartbeat interval for proper CXFS operation.

**notification command**

The command used to notify the cluster administrator of changes or failures in the cluster and nodes. The command must exist on every node in the cluster.

**owner host**

A system that can control a node remotely, such as power-cycling the node. At run time, the owner host must be defined as a node in the pool.

**owner TTY name**

The device file name of the terminal port (TTY) on the *owner host* to which the system controller is connected. The other end of the cable connects to the node with the system controller port, so the node can be controlled remotely by the owner host.

**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` 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).

**port password**

The password for the system controller port, usually set once in firmware or by setting jumper wires. (This is not the same as the node's root password.)

**potential metadata server**

A server-capable administration node that is listed in the metadata server list when defining a filesystem; only one node in the list will be chosen as the active metadata server.

**quorum**

The number of nodes required to form a cluster, which differs according to membership:

- For CXFS kernel membership:
  - A majority (>50%) of the server-capable nodes in the cluster are required to **form** an initial membership
  - Half (50%) of the server-capable nodes in the cluster are required to **maintain** an existing membership
- For cluster database membership, 50% of the **nodes in the pool** are required to form and maintain a cluster.

**recovery**

The process by which a node is removed from the CXFS kernel membership due to an interruption in CXFS services. It is during this process that the remaining nodes in the CXFS kernel membership resolve their state for cluster resources owned or shared with the removed node. See also *metadata server recovery*

**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.

**reset**

The failure action that performs a system reset via a serial line connected to the system controller. The reset may be a powercycle, serial reset, or NMI (nonmaskable

interrupt). This failure action hierarchy choice is recommended for all potential metadata servers.

**server-capable administration node**

A node that is installed with the `cluster_admin` product and is also capable of coordinating CXFS metadata.

**shutdown**

The fail action hierarchy selection that tells the other nodes in the cluster to wait before reforming the CXFS kernel membership. The surviving cluster delays the beginning of recovery to allow the node time to complete the shutdown. See also *forced CXFS shutdown*.

**snooping**

A security breach involving illicit viewing.

**split-brain syndrome**

A situation in which multiple clusters are formed due to a network partition and the lack of reset and/or CXFS tiebreaker capability.

**spoofing**

A security breach in which one machine on the network masquerades as another.

**standby node**

A server-capable administration node that is configured as a potential metadata server for a given filesystem, but does not currently run any applications that will use that filesystem.

**storage area network (SAN)**

A dedicated, high-speed, scalable network of servers and storage devices designed to enhance the storage, retrieval, and management of data

**system controller port**

A port sitting on a node that provides a way to power-cycle the node remotely. Enabling or disabling a system controller port in the cluster database tells CXFS whether it can perform operations on the system controller port.

**system log file**

Log files in which system messages are stored

**tiebreaker node**

See *CXFS tiebreaker node*.

**user-space membership**

See *cluster database membership*.

**view area**

The portion of the GUI window that displays components graphically. See also *details area*.

---

## Index

100baseT TCP/IP network, 7

### A

ACL problem and AIX, 43  
acledit, 22  
aclget, 22  
aclput, 22  
ACLs  
    AIX, 22, 24  
    Linux, 51  
    Mac OS X, 84  
    Solaris, 107  
    Windows, 157  
Active Directory user ID mapping method, 187  
adapter parameters, 119, 122  
admin cxfs\_unmount, 220  
administrative tasks, 4  
AIX  
    ACLs, 22  
    client software installation, 31  
    commands installed by CXFS, 21  
    common problems, 42  
    FLEXlm license verification, 33  
    hardware, 20  
    HBA installation, 28  
    identifying problems, 231  
    ifconfig, 29  
    installation overview, 25  
    kernel extensions, 45  
    limitations, 22  
    log files, 21  
    manual CXFS startup/shutdown, 36  
    modify the CXFS software, 37  
    NFS export scripts, 6  
    operating system version, 20

    postinstallation steps, 35  
    preinstallation steps, 28  
    problem reporting, 44  
    requirements, 20  
    software  
        maintenance, 37  
        upgrades, 37  
    space requirements, 32  
alog, 44  
AMCC JNI HBA, 104

### B

bandwidth, 2, 5  
/bin/hostname  
    Linux, 52  
    Mac OS X, 86  
BIOS version, 147  
block size, 246  
boot command, 112  
buffered I/O, 5

### C

\$c or \$C, 142  
client software installation  
    AIX, 31  
    Linux, 61  
    Mac OS X, 92  
    Solaris, 131  
    Windows, 185  
client-only node configuration  
    add to the cluster, 217  
    define the node, 216  
    define the switch, 218

- modify the cluster, 217
- mount filesystems, 220
- permit fencing, 216
- start CXFS services, 220
- verify the cluster, 223
- client-only nodes added to cluster, 217
- cluster
  - configuration, 215
  - size, 10
  - verification, 223
- cluster administration, 4
- cmgr command, 215
- commands installed
  - AIX, 21
  - Linux, 49, 50
  - Mac OS X, 80
  - Solaris, 105
  - Windows, 148
- common problems, 233
- concepts, 2
- console log, 44
- core files, 142
- CPU types for Linux, 48
- cpuinfo, 53, 76
- crash dumps
  - Solaris, 142
  - Windows, 212
- crash utility and gathering output, 142
- crontab, 10
- CXFS
  - GUI and cmgr, 215
  - software removal on Windows, 199
  - startup/shutdown
    - Windows, 196
- CXFS Client log color meanings, 151
- CXFS Client service command line arguments, 187
- CXFS Info icon color meanings, 152
- CXFS startup/shutdown
  - AIX, 36
  - Linux, 67
  - Mac OS X, 96
  - Solaris, 135

- Windows, 196
- cxfs\_client
  - daemon is not started
    - Linux, 74
    - Mac OS X, 100
    - Solaris, 140
  - service is not started
    - AIX, 42
- cxfs\_cluster, 36, 67
- cxfs\_cluster command, 135
- cxfs\_info, 148
  - state information, 224
- cxfsdump, 76, 143
- cxfslicense, 148

## D

- define a client-only node, 216
- defragmentation software, 10
- devfs, 74
- dflt\_local\_status, 220
- direct-access I/O, 2
- disk device verification for Solaris, 124
- display LUNs for QLogic HBA, 178
- distributed applications, 4
- dmesg command, 128
- DNS
  - AIX, 29
  - Linux, 58
  - Mac OS X, 90
  - Solaris, 126
  - Windows, 183
- DOS command shell, 183, 203
- dumps and output to gather, 142

## E

- entitlement ID, 13, 176
- Entitlement Sheet, 7

- error messages, 233
  - /etc/fencing.conf and AIX, 35
  - /etc/hostname.<interface>, 128
  - /etc/hosts
    - AIX, 28
    - Linux, 57
    - Mac OS X, 83
  - /etc/inet/ipnodes, 125
  - /etc/init.d/cxfs\_client, 67
  - /etc/init.d/cxfs\_cluster command, 135
  - /etc/netmasks, 129
  - /etc/nodename file, 128
  - /etc/nsswitch.conf, 126
  - /etc/nsswitch.conf file, 14
  - /etc/redhat-release, 62
  - /etc/sys\_id, 128
  - examples
    - add a client-only node to the cluster, 217
    - CXFS software installation
      - AIX, 32
      - Linux, 62
      - Solaris, 132
      - Windows, 186
    - define a node, 216
    - define a switch, 218
    - /etc/hosts file
      - Linux, 58
    - /etc/inet/hosts file
      - Linux, 58
    - /etc/inet/ipnodes file
      - Solaris, 126
    - fabric enable, 113
    - ifconfig
      - AIX, 29, 31
      - Linux, 58, 60
      - Mac OS X, 90
      - Solaris, 126, 130
    - JNI GUI screens, 116
    - modify the cluster, 217
    - modifying the CXFS software
      - AIX, 37
      - Solaris, 136
      - Windows, 197
    - mount filesystems, 220
    - name services
      - Linux, 58
      - Solaris, 126
    - ping
      - AIX, 31
      - Linux, 60
      - Mac OS X, 91
      - Solaris, 130
    - ping output for Solaris, 130
    - private network interface test
      - AIX, 31
      - Linux, 60
      - Mac OS X, 91
      - Solaris, 130
    - private network interface test for Solaris, 130
    - .rhosts, 129
    - start CXFS services, 220
    - verify the cluster configuration, 223
    - Windows Client service command line
      - options, 198
  - EZ Fibre GUI, 114
  - ezf, 115
- ## F
- fabric mode, 113
  - fail action hierarchy, 216
  - FailSafe coexecution, 9
  - failure on restart, 210
  - FcFabricEnabled, 113
  - FcLoopEnabled, 113
  - fence specification in node definition, 216
  - fencing.conf and AIX, 35
  - fencing.conf file, 65, 93, 133, 193
  - Fibre Channel HBA
    - See "host bus adapter", 55
  - Fibre Channel requirements
    - AIX, 20

- Solaris, 104
- file size and CXFS, 5
- file size/offset maximum, 246
- filesystem does not mount
  - AIX, 43
  - Solaris, 141
  - Windows, 209
- filesystem network access, 3
- filesystem specifications, 246
- find and crontab, 10
- FLEXlm
  - license key, 8
- FLEXlm license, 12
  - installation, 13
  - SGI webpage, 12
  - verification
    - Windows, 177
- forced unmount, 11
- format command, 124
- free disk space required, 147
- fuser, 12

## G

- G5 Xserve, 80
- genkex, 45
- guided configuration, 215

## H

- hangs and output to gather, 142
- hardware installed, 76
- hardware requirements
  - AIX, 20
  - all platforms, 7
  - Linux, 48
  - Mac OS X, 80
  - Solaris, 104
  - Windows, 147
- HBA, 113

- AIX, 20, 28
- Linux, 48, 55
- Mac OS X, 87
- problems, 142
- Solaris, 104, 111
- Windows, 147, 177

hierarchy of fail actions, 216

host bus adapter

- See "HBA", 177

host ID

- AIX, 26
- Linux, 52
- Mac OS X, 86
- Solaris, 110

hostname

- AIX, 26
- Linux, 52
- Mac OS X, 82, 86
- Solaris, 110

hostname resolution, 13

hostname.<interface>, 128

hosts file

- Linux, 57

## I

- I/O fencing, 8
- I/O operations, 2
- identifying problems, 233
- ifconfig, 52
  - AIX, 29, 31
  - Linux, 58, 60
  - Mac OS X, 90
  - Solaris, 126, 130
- initial setup services, 2
- inode64 mount option
  - Mac OS X, 81
- install.sh script, 115
- installation overview, 25
- installed packages, 143

- installed patches, 143
- installp, 32
- integrated Ethernet, 127
- Intel Pentium processor, 147
- interface for the private network, 128
- internode communication, 14
- introduction, 1
- IP address, changing, 14
- ipconfig, 183
- ipnodes, 125
- IRIX
  - labels in warning messages, 124

## J

- JBOD, 7
- JNI HBA, 104
- JNIC146x, 113
- jnic146x.conf, 113
- JNIsnia, 113

## K

- kdb, 44, 77
- kernel modules and versions, 143
- kernel running on Linux, 76
- /kernel/drv/jnic146x.conf, 113, 114

## L

- large file support, 246
- large files, 3
- LDAP generic user ID mapping method, 188
- license, 12
  - See also "FLEXlm license", 12
  - verification on AIX, 33
  - verification on Linux, 64
  - verification on Mac OS X, 93
  - verification on Solaris, 132

- verification on Windows, 191
- licensing, 8
- Linux
  - client software installation, 61
  - commands installed by CXFS, 49, 50
  - common problems, 74
  - HBA installation, 55
  - identifying problems, 231
  - ifconfig, 60
  - installation overview, 51
  - limitations, 50
  - log files, 49
  - manual CXFS startup/shutdown, 67
  - NFS export scripts, 50
  - preinstallation steps, 57
  - problem reporting, 76
  - requirements, 48
  - software
    - maintenance, 69
    - software maintenance, 69
    - space requirements, 62
- linux
  - FLEXlm license verification, 64
- log files
  - AIX, 21
  - Linux, 49
  - Mac OS X, 81
  - Solaris, 106
  - Windows, 149, 210
- lslpp, 21, 34, 45
- lsmod, 76
- lspci, 76
- LUN
  - logical unit, 122
  - LUN 31, 122
  - mapping, 122
  - zoning, 121
- LUN limit, 246

**M**

- MAC address, 176
- Mac OS X
  - access control lists, 84
  - client software installation, 92
  - commands installed, 80
  - common problems, 100
  - FLEXlm license verification, 93
  - hardware platforms, 80
  - HBA, 87
  - hostname, 82
  - identifying problems, 231
  - ifconfig, 90
  - installation overview, 84
  - limitations and considerations, 81
  - log files, 81
  - manual CXFS startup/shutdown, 96
  - modifying CXFS software, 97
  - NetInfo Manager, 83
  - power-save mode disabling, 91
  - preinstallation steps, 89
  - private network, 89
  - problem reporting, 101
  - removing CXFS software, 97
  - requirements, 80
  - software maintenance, 96
  - UID and GID mapping, 83
  - upgrading CXFS software, 96
- maintenance and CXFS services, 10
- manual CXFS startup/shutdown
  - AIX, 36
  - Linux, 67
  - Windows, 196
- mapping LUNs, 122
- mdb, 142
- memory error and AIX, 43
- memory map maximum
  - Mac OS X, 81
- memory mapped shared files, 4
- metadata, 3, 5
- metadata server, 4

- mirroring feature and license, 12
- modify cluster command, 217
- modinfo, 143
- modules and versions, 143
- modules loaded on Linux, 76
- mount filesystems, 220
- mount options support, 247
- mount-point nesting on Solaris, 106
- msgbuf, 142
- \$(msgbuf, 142
- multiOS cluster, 1

**N**

- name restrictions, 13
- name service daemon, 126
- nested mount points on Solaris, 106
- NetInfo Manager, 83
- netmasks, 129
- network
  - information service, 126
  - interface configuration, 14
  - requirements, 7
- NFS, 5
- NFS and CXFS, 106
- NFS export scripts
  - AIX, 6
  - Linux, 50
- NIS, 126
  - Linux, 58
  - Solaris, 126
- nsd, 126
- nsswitch.conf, 126
- number of nodes supported, 9

**O**

- O2, 7
- O\_LARGEFILE, 246

oslevel, 45

## P

packages installed

AIX, 45

Linux, 76

Solaris, 143

panic and AIX, 43

partitioned system licensing, 8

passwd and group files user ID mapping

method, 159

patches installed, 143

performance considerations, 4

permissions for Windows passwd and group

files, 193

physical CPU count, 53

physical LUN limit, 246

ping, 31, 60, 91, 130

pkgadd command, 106, 131

pkginfo, 143

pkginfo command, 132

plug and play, 207

plumb, 128

postinstallation steps

AIX, 35

Windows, 191

Power Mac, 80

power management software, 11

power-save mode for Mac OS X, 91

preinstallation steps

AIX, 28

Linux, 57

Mac OS X, 89

Solaris, 125

Windows, 181

premount and postmount scripts, 6, 50

primary hostname

Solaris, 125

Windows, 183

private network

AIX, 28

heartbeat and control, 14

interface test

AIX, 31

Linux, 60

Mac OS X, 91

Solaris, 130

Linux, 57

Mac OS X, 89

required, 7

Solaris, 125

windows, 181

problem reporting

AIX, 44

Linux, 76

Mac OS X, 101

Solaris, 142

Windows, 211

/proc/cpuinfo, 53

processor type on Linux, 76

%ProgramFiles%\CXFS directory, 185

%ProgramFiles%\CXFS\log\cxfs\_client.log

file, 210

pSeries systems, 20

public network

Solaris, 127

## Q

Qlogic HBA installation, 177

QLogic HBA model numbers and driver

versions, 147

## R

\$r, 142

READ\_CAPACITY SCSI command, 107

recommendations, 9

remove CXFS software

- Windows, 199
- removing CXFS software
  - Mac OS X, 97
- requirements
  - AIX, 20
  - all platforms, 7
  - Linux, 48
  - Mac OS X, 80
  - Solaris, 104
  - Windows, 147
- reset lines, 8
- /.rhosts, 129
- rpm, 76

## S

- /sbin/ifconfig, 86
- serial reset lines, 8
- service pack, 148
- set dflt\_local\_status, 220
- setup program for Windows, 185
- setup services, 2
- Setup.exe, 198
- SGIcxfx package, 131
- showrev, 143
- Silicon Graphics O2, 7
- single-user mode in Solaris, 125
- size of the cluster, 9
- small files, 5
- snap, 44
- SNIA API package, 113
- software maintenance
  - AIX, 37
  - Linux, 69
  - Mac OS X, 96
  - Solaris, 135
  - Windows, 196
- software requirements
  - AIX, 20
  - all platforms, 7
  - Linux, 48

- Mac OS X, 80
- Solaris, 104
- Windows, 147

- software upgrades
  - AIX, 37
  - Mac OS X, 96
  - Solaris, 136
  - Windows, 198
- Solaris
  - AMCC JNI HBA installation, 111
  - client software installation, 131
  - commands installed by CXFS, 105
  - common problems, 140, 141
  - FLEXlm license verification, 132
  - identifying problems, 231
  - ifconfig, 130
  - installation overview, 109
  - kernel
    - modules and versions, 143
  - limitations, 106
  - log files, 106
  - modify the CXFS software, 136
  - non-disk devices and AMCC JNI controllers, 107
  - operating system version, 104
  - preinstallation steps, 125
  - problem reporting, 142
  - requirements, 7, 104
  - single-user mode, 125
  - software
    - maintenance, 135
    - software upgrade
      - upgrades, 136
    - space requirements, 131
    - start/stop cxfx\_client, 135
- space requirements
  - AIX, 32
  - Linux, 62
  - Solaris, 131
- standby node, 10
- start
  - AIX, 36

- CXFS client service
  - Windows, 196
- CXFS processes
  - Mac OS X, 96
- CXFS services, 196, 220
- cxfs\_client
  - Linux, 67
  - Solaris, 135
- startup/shutdown of CXFS
  - Mac OS X, 96
- stop CXFS client service
  - Windows, 196
- stop CXFS processes
  - Mac OS X, 96
- stop cxfs\_client
  - AIX, 36
  - Linux, 67
  - Solaris, 135
- Storage Networking Industry Association
  - application programming interface
    - package, 113
- subnet, 8
- Sun hardware, 104
- switch definition, 218
- switchshow, 35, 66, 94, 134, 194
- sys\_id, 128
- sysctl, 77
- system core files, 142
- system device location problems, 74

## T

- TCP/IP network requirements, 7
- telnet port and I/O fencing, 8
- \$(threadlist, 142
- TRIX and Solaris nodes, 9
- troubleshooting, 231
- Trusted IRIX and Solaris nodes, 9

## U

- UFS and CXFS, 106
- Ultra Enterprise platforms, 104
- umount, 12
- uname, 32, 65, 76, 131
- upgrade CXFS software
  - AIX, 37
  - Mac OS X, 96
  - Solaris, 136
  - Windows, 198
- user administration, 4
- User ID mapping methods, 159
  - Active Directory, 187
  - Generic LDAP, 188
- user mapping problems on Windows, 208
- /usr/bin/hostid
  - AIX, 26
  - Solaris, 110
- /usr/bin/hostname
  - AIX, 26
  - Solaris, 110
- /usr/bin/showrev, 143
- /usr/cxfs\_cluster/bin/cxfs\_cluster, 36

## V

- /var/crash/<hostname>, 142
- /var/log/cxfs\_client, 21, 49
- verify
  - cluster, 223
- versions of modules installed, 143

## W

- warning message and IRIX labels, 124
- Windows
  - ACLs, 157
  - client software installation steps, 185

- common problems, 202
- crash dumps, 212
- CXFS commands installed, 148
- CXFS software removal, 199
- debugging information, 212
- failure on restart, 210
- filesystems not displayed, 209
- FLEXlm license verification, 177, 191
- identifying problems, 231
- installation overview, 174
- ipconfig, 183
- large log files, 210
- log files, 149
- LUNs, 178
- manual CXFS startup/shutdown, 196
- memory configuration, 210
- modify the CXFS software, 197
- postinstallation steps, 191
- preinstallation steps, 181
- problem reporting, 211
- Qlogic HBA installation, 177
- requirements, 7, 147
- software maintenance, 196
- software upgrades, 198
- verify networks, 183
- Windows/Setup.exe, 185
- worldwide node name, 114

- worldwide number, 35
- worldwide port name, 65, 93, 114, 133, 194
  - Linux, 65, 238
  - Mac OS X, 93
  - Solaris, 133
  - Windows, 193
- WWNN, 114
- WWPN, 35, 65, 93, 114, 133, 194
  - Linux, 65, 238
  - Mac OS X, 93
  - Solaris, 133
  - Windows, 193

## X

- xfp\_repair, 10
- Xserve, 80
- XVM mirroring license, 12
- XVM version 1 failover
  - supported on IRIX and SGI ProPack for Linux only, 17
- XVM version 2 failover
  - general information, 17
  - Linux, 72
  - Mac OS X, 98