

# SGI ProPack™ for Linux® Start Here

007-4558-006

---

## CONTRIBUTORS

Written by Julie Boney and Terry Schultz

Edited by Susan Wilkening

Graphics by Chrystie Danzer

Production by Karen Jacobson

---

## COPYRIGHT

© 2003 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.

---

## LIMITED RIGHTS LEGEND

The electronic (software) version of this document was developed at private expense; if acquired under an agreement with the USA government or any contractor thereto, it is acquired as "commercial computer software" subject to the provisions of its applicable license agreement, as specified in (a) 48 CFR 12.212 of the FAR; or, if acquired for Department of Defense units, (b) 48 CFR 227-7202 of the DoD FAR Supplement; or sections succeeding thereto. Contractor/manufacturer is Silicon Graphics, Inc., 1600 Amphitheatre Pkwy 2E, Mountain View, CA 94043-1351.

---

## TRADEMARKS AND ATTRIBUTIONS

Silicon Graphics, SGI, the SGI logo, IRIX, Origin, Onyx, Onyx2, and XFS are registered trademarks and Altix, NUMAflex, NUMAlink, OpenMP, Performance Co-Pilot, SGI Advanced Linux, SGI ProPack, and SGIconsole are trademarks of Silicon Graphics, Inc., in the United States and/or other countries worldwide.

SGI Advanced Linux Environment 2.1.1 is based on Red Hat Linux Advanced Server 2.1 for the Itanium Processor, but is not sponsored by or endorsed by Red Hat, Inc. in any way.

Cray is a registered trademark of Cray, Inc. FLEXlm is a registered trademark of Macrovision Corporation. Java is a registered trademark of Sun Microsystems, Inc. in the United States and/or other countries. KAP/Pro Toolset and VTune are trademarks and Intel, Itanium, and Pentium are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Linux is a registered trademark of Linus Torvalds, used with permission by Silicon Graphics, Inc. MIPS is a registered trademark of MIPS Technology, Inc. PostScript is a registered trademark of Adobe Systems, Inc. QLogic is a registered trademark of QLogic Corporation. Red Hat is a registered trademark and Red Hat Linux Advanced Server 2.1 and RPM are trademarks of Red Hat, Inc. Windows is a registered trademark of Microsoft Corporation in the United States and/or other countries.

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

---

## Record of Revision

<b>Version</b>	<b>Description</b>
001	February 2003 Original publication
002	February 2003 Updated for SGI ProPack v2.1.1 for Linux release
003	April 2003 Updated for SGI ProPack v2.1.2 for Linux release
004	June 2003 Updated for SGI ProPack v2.2 for Linux release
005	August 2003 Updated for SGI ProPack v2.2.1 for Linux release
006	October 2003 Updated for SGI ProPack v2.3 for Linux release



---

## New Features in This Guide

This update of *SGI ProPack for Linux Start Here* supports the v2.3 release of SGI ProPack for Linux.

### Major Documentation Changes

Changes in this guide for the SGI ProPack 2.3 release include the following:

- Removed information in Chapter 4. Tuning information is now contained in a new document called *Linux Configuration and Operations Guide* available with the SGI ProPack v2.3 release.
- Added information about how to change the location of Linux Kernel Crash dump (LKCD) files and corrected the description of the default location for the crash dump files in “Linux Kernel Crash Dump” on page 6.
- Added information about FLEXlm licenses in “License Manager Updates” on page 10 .
- Added information to the IOC4 serial driver description about serial ports supported on the IOC4 base I/O chipset in Table 3-1 in “SGI ProPack for Linux Products” on page 19.
- Added information about how to connect to an L1 or L2 system controller in “System Controller Firmware” on page 36.



---

# Contents

<b>Record of Revision</b>	iii
<b>Figures</b>	ix
<b>Tables</b>	xi
Related Publications	xiv
Intel Compiler Documentation	xiv
Other Intel Documentation	xiv
SGI Documentation	xiv
Obtaining SGI Publications	xvii
Conventions	xvii
Reader Comments	xvii
<b>1. Release Features</b>	<b>1</b>
Software Introduction	1
CD Contents	2
Installation Overview	4
<b>2. Software Planning and Installation</b>	<b>5</b>
Software Planning	6
Table of Disk Partitions	6
Linux Kernel Crash Dump	6
Installing the SGI Advanced Linux Environment 2.1.1 Base OS	7
License Manager Updates	10
Upgrading from Previous SGI ProPack Releases to SGI ProPack v2.3	11
Installing SGI ProPack for Linux	12
Installing the System Controller Software 1.3 CD	13
Upgrading Your Software	14
Recovering a Damaged Root Filesystem	14
Manually Mounting the Root Filesystem	16

<b>3.</b>	<b>Product Support</b>	. . . . .	. 19
	SGI ProPack for Linux Products	. . . . .	. 19
	Operating System Enhancements.	. . . . .	. 24
	CpuMemSets Support.	. . . . .	. 24
	Cpuset Support	. . . . .	. 25
	Comprehensive System Accounting (CSA).	. . . . .	. 25
	Partitioning.	. . . . .	. 26
	I/O Subsystems.	. . . . .	. 26
	Persistent PCI-X Bus Numbering	. . . . .	. 26
	Persistent Naming of Ethernet Devices .	. . . . .	. 28
	XSCSI Subsystem	. . . . .	. 30
	XSCSI-SCSI Subsystem	. . . . .	. 30
	XFS Filesystem .	. . . . .	. 30
	XVM Volume Manager	. . . . .	. 31
	HPC Application Tools and Support.	. . . . .	. 31
	Message Passing Toolkit	. . . . .	. 31
	Performance Co-Pilot .	. . . . .	. 32
	System Management	. . . . .	. 32
	PROM Chips	. . . . .	. 32
	Extensible Firmware Interface (EFI) .	. . . . .	. 32
	FLEXlm .	. . . . .	. 35
	SGIconsole .	. . . . .	. 35
	System Controller Firmware .	. . . . .	. 36
	NUMA Tools	. . . . .	. 37
	dlook Command .	. . . . .	. 37
	dplace Command	. . . . .	. 38
<b>4.</b>	<b>Performance Tuning</b>	. . . . .	. 39
	<b>Index.</b>	. . . . .	. 41

---

# Figures

<b>Figure 1-1</b>	SGI ProPack for Linux Release CD Contents . . . . .	3
-------------------	---	---



---

# Tables

<b>Table 2-1</b>	devfs Disk Partitions . . . . .	6
<b>Table 3-1</b>	SGI ProPack v2.3 for Linux Products . . . . .	20
<b>Table 3-2</b>	EFI Commands . . . . .	33



---

## About This Guide

This guide provides information about the SGI ProPack for Linux release. It is divided into the following chapters:

- Chapter 1, "Release Features," describes the major features of this release, the CD contents, and documentation information.
- Chapter 2, "Software Planning and Installation," describes the flowcharts of disk partitions and file configurations and layouts that you need to get started, and provides instructions for installing SGI ProPack for Linux.
- Chapter 3, "Product Support," documents the product components that are supported on SGI Altix 3000 systems.

---

**Note:** Performance tuning information for single processor and multiprocessor programs now resides in the *Linux Configuration and Operations Guide*.

---

The information in this guide, other SGI ProPack for Linux documentation, and all other documentation included in the RPMs on the distribution CDs can be found on the CD titled "SGI ProPack v2.3 for Linux - Documentation CD." To access the information on the documentation CD, open the `index.html` file with a web browser. Because this online file can be updated later in the release cycle than this document, you should check it for the latest information.

---

**Note:** The release notes, which contain the latest information about software and documentation in this release, are on the SGI ProPack for Linux Documentation CD in the `root` directory, in a file named `README.TXT`.

---

## Related Publications

Documents listed in this section contain additional information that might be helpful.

### Intel Compiler Documentation

Documentation for the Intel compilers is located on your system in the `/docs` directory of the directory tree where your compilers are installed. If you have installed the Intel compilers, the following documentation is available:

- *Intel C++ Compiler User's Guide* (`c_ug_lnx.pdf`).
- *Intel Fortran Compiler User's Guide* (`for_ug_lnx.pdf`).
- *Intel Fortran Programmer's Reference* (`for_prgr.pdf`).
- *Intel Fortran Libraries Reference* (`for_lib.pdf`).

### Other Intel Documentation

The following references describe the Itanium (previously called "IA-64") architecture and other topics of interest:

- *Intel Itanium 2 Processor Reference Manual for Software Development and Optimization*, available online at <http://developer.intel.com/design/itanium/manuals>.
- *Intel Itanium Architecture Software Developer's Manual*, available online at <http://developer.intel.com/design/itanium/manuals>.
- *Introduction to Itanium Architecture*, available online at <http://shale.intel.com/softwarecollege/CourseDetails.asp?courseID=13> (secure channel required).

### SGI Documentation

The following SGI documentation is available:

- *Linux Device Driver Programmer's Guide*  
Provides information on programming, integrating, and controlling drivers.

- *Message Passing Toolkit: MPI Programmer's Manual*  
Describes industry-standard message passing protocol optimized for SGI computers.
- *Origin 2000 and Onyx2 Performance Tuning and Optimization Guide*  
Contains information specific to MIPS/IRIX systems, but the general guidelines in the document are hardware and operating system independent.
- *Performance Co-Pilot for IA-64 Linux User's and Administrator's Guide*  
Describes the Performance Co-Pilot (PCP) software package of advanced performance tools for SGI systems running the Linux operating system.
- *Linux Configuration and Operations Guide*  
Provides information on how to perform system configuration and operations for SGI ProPack servers.
- *Linux Resource Administration Guide*  
Provides a reference for people who manage the operation of SGI ProPack servers and contains information needed in the administration of various system resource management features such as Comprehensive System Accounting (CSA), Array Services, CPU memory sets and scheduling, Cpuset System,
- *SGI Altix 3000 User's Guide*  
Provides an overview of the architecture and describes the major components of the SGI Altix 3000 family of servers and superclusters. It also describes the standard procedures for powering up and powering down the system, provides basic troubleshooting information, and includes important safety and regulatory specifications.
- *SGI ProPack v2.3 for Linux Release Notes*  
Provide the latest information about software and documentation in this release. The release notes are on the SGI ProPack for Linux Documentation CD in the `root` directory, in a file named `README.TXT`.
- *SGIconsole 2.0 Start Here*  
Provides an introduction to SGIconsole and information about setting up and configuring SGIconsole hardware and software.
- *Console Manager for SGIconsole Administrator's Guide*  
Provides information about the Console Manager software graphical interface allows you to control multiple SGI servers, SGI partitioned systems, and large single-system image servers.

- *SGI L1 and L2 Controller Software User's Guide*  
Describes how to use the L1 and L2 controller commands at your system console to monitor and manage the SGI Altix 3000 family of servers and superclusters.
- *XFS for Linux Administration*  
Describes XFS, an open-source, fast recovery, journaling filesystem that provides direct I/O support, space preallocation, access control lists, quotas, and other commercial file system features.
- *Event Manager User Guide*  
Provides information about the Event Manger application that collects event information from other applications. This document describes the Event Manager application, the application programming interface that you can use to access it, the procedures that you can use to communicate with it from another application, and the commands that you can use to control it.
- *Embedded Support Partner User Guide*  
Provides information about using the Embedded Support Partner (ESP) software suite to monitor events, set up proactive notification, and generate reports. This revision of the document describes ESP version 3.0, which is the first version of ESP that supports the Linux operating system.

These books will be available in the fall of 2003 from the Technical Publications Library (<http://docs.sgi.com>) and will be included in future ProPack release:

- *Linux Application Tuning Guide*  
Provides information about tuning application programs on SGI Altix 3000 series systems. Application programs include Fortran and C programs written on SGI Linux systems with the compilers provided by Intel.
- *SCSL User's Guide*  
Provides information about the scientific libraries on SGI Altix 3000 series systems and SGI IRIX systems. Topics include dicussions about BLAS, LAPACK, and FFT routines.

## Obtaining SGI Publications

You can obtain SGI documentation in the following ways:

- See the SGI Technical Publications Library at <http://docs.sgi.com>. Various formats are available. This library contains the most recent and most comprehensive set of online books, release notes, man pages, and other information.
- You can view man pages by typing `man <title>` on a command line.

## Conventions

The following conventions are used throughout this publication:

<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.)
<code>manpage(x)</code>	Man page section identifiers appear in parentheses after man page names.
<b>GUI element</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.

## Reader Comments

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

You can contact SGI in any of the following ways:

- Send e-mail to the following address:  
techpubs@sgi.com
- Use the Feedback option on the Technical Publications Library Web page:  
<http://docs.sgi.com>
- Contact your customer service representative and ask that an incident be filed in the SGI incident tracking system.
- Send mail to the following address:  
Technical Publications  
SGI  
1600 Amphitheatre Parkwy, M/S 535  
Mountain View, CA 94043-1351
- Send a fax to the attention of “Technical Publications” at +1 650 932 0801.

SGI values your comments and will respond to them promptly.

## Release Features

This chapter provides an introduction to SGI ProPack for Linux and describes the CD contents, provides documentation information, and provides a brief overview of the installation process.

### Software Introduction

The open-source and collaborative development environment of the Linux community provides a new and different model for computer manufacturers to deliver an operating system for a large server or supercomputer. Using Linux as the operating system for very large computer systems has the benefit of providing better software protection and better integration between a system manufacturer's operating system and high-performance computing (HPC) applications and codes for both independent software vendors and end users. Also, the collaboration on OS improvements in the Linux community as well as among the various computer manufacturers has enabled Linux to evolve and improve in multiple dimensions faster than would otherwise be possible for a single company working on its own operating system.

The SGI ProPack for Linux product includes capabilities and performance improvements ideal for enabling technical and creative users to solve their big compute and data problems by using the Linux operating system and Itanium processors. This product adds to or enhances features in SGI Advanced Linux Environment 2.1.1, the base Linux distribution based on Red Hat Advanced Server 2.1. SGI ProPack is required by and designed to run on any SGI Altix 3000 series system.

SGI hardware platforms and OS configuration settings supported by SGI in this release are documented at the following URL:

<http://support.sgi.com/linux>

## CD Contents

The CDs that you receive with SGI ProPack for Linux are as follows (Figure 1-1 lists the contents of each CD):

- SGI Advanced Linux Environment 2.1.1
  - SGI Advanced Linux Environment 2.1.1 CD1 - Installation CD 1 of 2
  - SGI Advanced Linux Environment 2.1.1 CD2 - Installation CD 2 of 2
  - SGI Advanced Linux Environment 2.1.1 CD3 - Source Code CD 1 of 2
  - SGI Advanced Linux Environment 2.1.1 CD4 - Source Code CD 2 of 2
- SGI ProPack v2.3 for Linux
  - SGI ProPack v2.3 for Linux CD1 - Open/Free Source Software
  - SGI ProPack v2.3 for Linux CD2 - Proprietary Software

---

**Note:** This CD contains software for use on SGI systems only. Redistribution is not permitted. Please see the shrink-wrap license agreement.

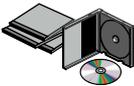
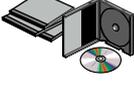
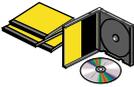
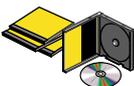
---

- SGI ProPack v2.3 for Linux CD3 - Documentation CD
- System Controller Software 1.3 CD (for IRIX and Linux)

---

**Note:** This software is licensed for use only on Origin and Onyx 3000 series systems (SGI systems based on MIPS processors) and on SGI Altix 3000 series systems (SGI systems based on Intel Itanium processors). This CD contains software provided under restricted or SGI proprietary licensing terms. The licensing terms for items on this CD allow users to install this software only on SGI systems.

---

SGI Advanced Linux Environment 2.1.1		
	SGI Advanced Linux Environment 2.1.1 Installation CD 1 of 2 CD1	Interactive installation procedures
	SGI Advanced Linux Environment 2.1.1 Installation CD 2 of 2 CD2	
	SGI Advanced Linux Environment 2.1.1 Source Code CD 1 of 2 CD3	Source code for base OS and common open source applications
	SGI Advanced Linux Environment 2.1.1 Source Code CD 2 of 2 CD4	
SGI ProPack v2.3 for Linux		
	SGI ProPack v2.3 for Linux Open/Free Source Software CD1	GPL/LGPL and other open source licensed software, latest SGI platform and NUMA support, various kernel performance improvements, CpuMemSets and performance measuring tools, XSCSI infrastructure, QL-SCSI, XFS, CSA, LKCD, kdb
	SGI ProPack v2.3 for Linux Proprietary Software CD2	System PROM, XVM, cpuset and NUMAtools (dlook, dplace, and so on), Array Services, FLEXIm, system partitioning software, MPT
	SGI ProPack v2.3 for Linux Documentation CD CD3	Collection of documentation, FAQs, HOWTOs, and man pages
System Controller Software 1.3 CD		
	System Controller Software 1.3 CD	System Controller software/firmware (for IRIX and Linux). This CD is only included with the ProPack software if you have requested it. You can always download the latest Software Controller software via Supportfolio.

**Figure 1-1** SGI ProPack for Linux Release CD Contents

For a complete list of the RPMs included on the SGI Advanced Linux Environment 2.1.1 and SGI ProPack v2.3 for Linux CDs listed in Figure 1-1 on page 3, see the release notes, which contain the latest information about software and documentation in this release. The release notes are on the SGI ProPack for Linux Documentation CD in the `root` directory, in a file named `README.TXT`.

SGI maintains the following website for open source information that describes projects related to its open source efforts:

<http://oss.sgi.com>

You can also access open source information (such as LKCD) from the following website:

<http://sourceforge.net>

## Installation Overview

SGI ProPack for Linux will most likely come preinstalled on your SGI platform. If you should need to install it, be aware that you must first install the SGI Advanced Linux Environment 2.1.1 CDs, then the SGI ProPack v2.3 Open/Free Source Software (CD1). The SGI ProPack v2.3 Open/Free Source Software CD will then require the SGI ProPack v2.3 Proprietary Software (CD2). Installation is described in Chapter 2, “Software Planning and Installation”.

---

**Note:** There is no Updates CD for the v2.3 release. All updates have been folded back into the SGI Advanced Linux Environment 2.1.1 CDs.

---

## Software Planning and Installation

This chapter describes the planning that you need to do to get started and provides instructions for installing SGI ProPack for Linux.

Your SGI Altix 3000 system comes with a base Linux distribution (SGI Advanced Linux Environment 2.1.1) and the SGI ProPack software preinstalled. This chapter describes how to install the software from the CD if it should ever become necessary to reinstall it.

For security reasons, Linux requires a root password for login. The default password for your preinstalled software is `sgisgi`. After you have logged in, change this root password to a string of your own choice.

The SGI ProPack software works only with the SGI AdvancedLinux Environment 2.1.1. Earlier versions of these distributions, or any other distributions, are not compatible with SGI ProPack software.

Before you install or configure your system, please read Chapter 1, “Release Features” so that you understand the features of SGI ProPack for Linux software and how to configure them. You might also want to familiarize yourself with some or all of the documentation listed in “Related Publications” on page xiv.

This chapter covers the following topics:

- “Software Planning” on page 6
- “Installing the SGI Advanced Linux Environment 2.1.1 Base OS” on page 7
- “License Manager Updates” on page 10
- “Upgrading from Previous SGI ProPack Releases to SGI ProPack v2.3” on page 11
- “Installing SGI ProPack for Linux” on page 12
- “Installing the System Controller Software 1.3 CD” on page 13
- “Upgrading Your Software” on page 14
- “Recovering a Damaged Root Filesystem” on page 14

- “Manually Mounting the Root Filesystem” on page 16

## Software Planning

This section provides disk and file information you need to have about the software that has been shipped to you.

### Table of Disk Partitions

SGI ships systems from the factory with the partition layout shown in Table 2-1.

**Table 2-1** devfs Disk Partitions

Device	Mount Point	Size
/dev/xscsi/pci01.03.0-1/target1/lun0/part1	/boot/efi	500 MB
/dev/xscsi/pci01.03.0-1/target1/lun0/part2	swap	9 GB
/dev/xscsi/pci01.03.0-1/target1/lun0/part3	/	25 GB

For more information on device naming, see “Persistent PCI-X Bus Numbering” on page 26 and “XSCSI Subsystem” on page 30.

### Linux Kernel Crash Dump

Linux Kernel Crash Dump (LKCD) creates files in `/var/log/dump` directory. To save crash dumps to a different location, change the `DUMPDIR` value in `/etc/sysconfig/dump` file.

The `/var/log` directory size will grow as Performance Co-Pilot and Comprehensive System Accounting (CSA) log to their own directories in `/var/log`. These files do not usually exceed 100 MB. For default LKCD configuration details, see the `lkcd_config(1)` man page.

---

**Note:** If there is an `lkcd` dump, LKCD could easily exceed multiple gigabytes in the `/var/log/dump` directory. This is why the default root filesystem is larger. For this reason, some customers may wish to make a separate `/var/log/dump` filesystem or change the configuration of `lkcd`.

---

## Installing the SGI Advanced Linux Environment 2.1.1 Base OS

This section provides steps for installing the SGI Advanced Linux Environment 2.1.1 CD1 and CD2 set. For console setup options and instructions, see the section titled “Monitoring Your Server” in the *SGI Altix 3000 User Guide*.

---

**Note:** The installation procedure has buttons that allow you to go back to the previous screen or to quit the installation. To use these buttons, press the Tab key to highlight the one you want to use and press Enter.

---

1. Insert SGI Advanced Linux Environment 2.1.1 CD1 into the system’s CD-ROM drive and restart the system.
2. While the system is powering up, check the device mapping table. It should look similar to the following:

Device mapping table

```
fs0  : Pci(2|1)/Ata(Primary, Master)/CDROM(Entry1) This is the CD-ROM
fs1  : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part1, Sig00000000)
fs2  : Pci(1|1)/Scsi(Pun0/Lun2)/HD(Part1, Sigg1)
blk0 : Pci(2|1)/Ata(Primary, Master)
blk1 : Pci(2|1)/Ata(Primary, Master)/CDROM(Entry1)
blk2 : Pci(1|1)/Scsi(Pun0/Lun1)
blk3 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part1, Sig00000000)
blk4 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part2, Sig00000000)
blk5 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part3, Sig00000000)
blk6 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part4, Sig00000000)
blk7 : Pci(1|1)/Scsi(Pun0/Lun2)
blk8 : Pci(1|1)/Scsi(Pun0/Lun2)/HD(Part1, Sigg10)
```

3. At the `Shell>` prompt, type the CD-ROM device name, as follows:

```
Shell> fs0:           Type this to change to the CD-ROM device
fs0:\>
```

4. Boot the CD by typing **elilo** and pressing Enter. When the ELILO boot prompt appears, press Enter again.

```
fs0:\> elilo           This is the boot loader
LoadPe: using PE image entry point
LoadPe: using PE image entry point
ELILO boot:           Press ENTER to boot the Linux kernel.
```

5. The CD will then boot and start the installation process.
6. Select the language you would prefer to use for the installation and as the system default. To continue, select **OK**.

---

**Note:** Console support for Asian languages is not available at this time. SGI recommends that customers in Asia set the default system language to English.

---

7. The next screen lets you choose your installation method. Currently, **Local CDROM** is the only choice. To continue, select **OK**.
8. The next screen is the SGI ProPack information screen. To continue, select **OK**.
9. The next screen asks you if you have a mouse installed. For an SGI Altix 3000 system, select **None**. To continue, select **OK**.
10. You will now see a window titled **Installation type**. If you wish to upgrade, select **Upgrade Existing System** and skip now to “Upgrading from Previous SGI ProPack Releases to SGI ProPack v2.3” on page 11 for further instructions. Otherwise, to completely install the system, wiping out any previous data on the disk, choose **Custom** and continue to the next step in this section.
11. In the partition screen, select the **Autopartition** option and press Enter on your keyboard. When warning messages appear, select **IGNORE**.
12. When asked which partitioning scheme to use, select the **Remove all partitions on this system** option. If you have two or more SCSI disks installed in your SGI Altix 3000 system, the installer asks you which drives you wish to use for installation. From the factory, SGI Altix 3000 systems come with the entire operating system installed on one system disk. If you would like the entire operating system installed on one disk, be sure that only one of the disks is selected for installation. Both devices are selected by default. To continue, select **OK**.
13. When the warning message appears about losing all data on the devices, select **Yes** to confirm the partition layout.

14. The next screen allows you to modify the partition layout as needed for your requirements. If you change the partition layout, ensure that the `/boot/efi` filesystem partition type remains set to `vfat`, the swap partition is set to `swap`, and all other file system types are set to `xfs`. To complete the partition layout, select **OK**.
15. If you use `Bootp` or `DHCP` to configure your network interfaces, select **OK**. If you use static IP addresses, fill in the required information and then to continue, select **OK**.

---

**Note:** If you plan to have an Ethernet board in bus 1 slot 2 of the IX-brick, directly to the right of the IO9, see “Persistent Naming of Ethernet Devices” on page 28. (This does not affect most users.)

---

16. If you chose to use a static IP address on the previous screen, you are prompted for a hostname for the system. Enter the hostname and to continue, select **OK**.
17. On the next screen you will be asked if you wish to configure a firewall on this system. Select the appropriate option based on your system’s need and to continue, select **OK**.
18. The next screen allows you to select additional languages to support on the system.

---

**Note:** If you select additional languages, you must select a default language. The default language will be used on your system once installation is complete. If you choose to install other languages, you can change your default language after installation. Console support for Asian languages is not available at this time. SGI recommends that customers in Asia set the default system language to English.

---

19. The next screen allows you to select which time zone this system is in. Select the appropriate information and to continue, select **OK**.
20. The next screen asks you for the root password for this system. Enter the password you would like to use for root access and to continue, select **OK**.
21. The next screen allows you to create a user for the system. Creating a user allows you to log in to the system as a user other than `root`. If you choose to create a user, fill in the required boxes and to continue, select **OK**. Leaving the boxes blank and selecting **OK** continues without creating another user.
22. If you created a user in the previous screen, the next screen allows you to add additional users. Add other users as necessary. To continue, select **OK**.

23. The next screen lets you choose your authentication configuration. To continue, select **OK**.
24. The **Package Group Selection** screen appears. All packages should be selected for installation. Select **Everything** (located at the bottom of the list). To continue, press the Tab key to highlight **OK** and press Enter.
25. The **Installation to begin** screen is informational. To continue, select **OK**.
26. The installation process will now begin. During the process, you will be prompted to insert SGI Advanced Linux Environment 2.1.1 CD2. Do this when instructed and to continue, select **OK**.
27. When the installation has finished, select **OK**. The system will reboot. You can remove the CD from the drive.

You are not done yet! Only the base environment has been updated. After the system restarts, to upgrade SGI ProPack, proceed to “Installing SGI ProPack for Linux” on page 12.

## License Manager Updates

This release of ProPack 2.3 includes FLEXlm 9.0. In certain situations, you may need replacement FLEXlm licenses for SGI applications. While the details of this upgrade should be mostly transparent, there is one change that is very visible.

Upgraded applications running on ProPack 2.3 use a different license identifier on SGI Altix series systems than applications running on prior ProPack 2.2.1 or 2.2a releases.

If you have any licensed SGI applications running on your Altix system, please contact SGI to request a new license key or contact your local support center and request a new key at this location:

<http://www.sgi.com/support/supportcenters.html>

Otherwise, these applications may not allow access after you upgrade to ProPack 2.3. New licenses generated for FLEXlm 9.0 will work on SGI ProPack releases running on the SGI Advanced Linux Environment 2.1.1. SGI strongly recommends that you replace the licenses for your SGI applications before upgrading the system to ProPack 2.3 to avoid any possible interruption in service. Third party applications are unaffected by this change.

## Upgrading from Previous SGI ProPack Releases to SGI ProPack v2.3

This section explains how to upgrade from previous SGI ProPack releases, including SGI ProPack v2.1, v2.1.1, or v2.1.2, v2.2.1 to this SGI ProPack v2.3 release. For console setup options and instructions, see the section titled “Monitoring Your Server” in the *SGI Altix 3000 User Guide*.

---

**Note:** In order to upgrade a system to ProPack v2.3, you will need to boot with the SGI Advanced Linux Environment 2.1.1 CD1. All updates to the linux base environment have been included in the SGI Advanced Linux Environment CDs.

---

After you have followed steps 1 through 9 in “Installing the SGI Advanced Linux Environment 2.1.1 Base OS” on page 7, the **Installation Type** window appears. To upgrade, perform the following steps:

1. Select **Upgrade Existing System**, then select **OK** and press **Enter**.
2. Ignore the warning message by selecting **Ignore** and pressing **Enter**. If you have more than one SCSI disk in your system, this warning might be repeated for each SCSI disk.
3. If your system has more than one root filesystem, select the root partition you wish to upgrade then select **OK** and press **Enter**. If you have only one existing root filesystem, you will see a window stating that it will be upgraded. Select **OK** and press **Enter**.
4. A **Customize Packages to Upgrade** window appears. Select **No** and press **Enter**.
5. You will see a window stating that packages to be upgraded are being found. It will take a few minutes to get past this point.
6. An **Upgrade to begin** window appears. Ensure that **OK** is selected and press **Enter**.
7. During the system upgrade process, you might be prompted to insert SGI Advanced Linux Environment 2.1.1 CD2. When instructed, put that CD in the drive and press **Enter**.
8. When the installation has finished, select **OK**. The system will reboot. You can remove the CD from the drive.

You are not done yet! Only the base environment has been updated. After the system restarts, to upgrade SGI ProPack, proceed to “Installing SGI ProPack for Linux” on page 12.

## Installing SGI ProPack for Linux

This section provides steps for installing the SGI ProPack v2.3 for Linux CD1 (Open/Free Source Software) and the SGI ProPack v2.3 for Linux CD 2 (Proprietary Software). For console setup options and instructions, see the section titled “Monitoring Your Server” in the *SGI Altix 3000 User Guide*.

1. Log in to the system as `root`.
2. Insert SGI ProPack v2.3 for Linux CD1 into the system and enter the following command to mount the CD:  

```
mount /dev/cdrom /mnt/cdrom
```
3. Enter `/mnt/cdrom/INSTALL`.
4. The SGI Welcome screen for SGI ProPack will appear. To continue, select **OK**.
5. The next screen lets you choose your installation type. Select **Yes** and select **OK**.
6. The **Package Group Selection** screen appears. This screen allows you to select the type of package group you want to install. You select a package group by using the up and down arrow keys and pressing the space bar to select the one you want. When you select a package group, RPMs for that package group will be installed after you press the Tab key to highlight **OK** and press Enter. All packages are selected by default, except **SGI Licensed**. If you change the selection, ensure that **SGI Proprietary** remains selected. If you will be using SGI licensed software, such as XVM, be sure to select **SGI Licensed**. Otherwise, leave it unselected. To continue, select **OK**.
7. The **Installation to Begin** screen appears. It tells you that a log of the installation will be placed in `/tmp/sgi-install.log`. Press Enter. (**OK** is highlighted by default).
8. The installation begins. You will see the **Package Installation** screen, which tells you which packages are being installed and logs the time it takes to install them.

---

**Note:** If you started the SGI ProPack installation from the system console, you may see a message stating "INIT: *version number* reloading" as the packages are being installed. This message is normal and is not an error.

---

9. During the installation process, you will be asked to insert SGI ProPack v2.3 for Linux CD2. Insert the disk when prompted and press `Enter`.
10. After the installation is complete, the **Complete** screen appears. Press `Enter`. You are returned to the `root` prompt.
11. After the software installation is complete, reboot your system to begin using the newly installed SGI ProPack for Linux software environment. You can reboot by typing `reboot` and pressing `Enter`. You can then remove the CD from the drive.

## Installing the System Controller Software 1.3 CD

This section provides steps for installing the System Controller Software 1.3 (for IRIX and SGI Linux). This software is licensed for SGI Origin and Onyx 3000 series systems based on MIPS processors and for SGI Altix 3000 systems.

1. Log in to the system as `root`, using the root password you used in the previous installation process.
2. Insert the System Controller Software 1.3 CD into the system and enter the following command to mount the CD:  

```
mount /dev/cdrom /mnt/cdrom
```
3. Change directories to `/mnt/cdrom/RPMS/ia64` and enter the following command:  

```
./install
```
4. The installation begins. You will see which packages are being installed. After the installation is complete, you are returned to the `root` prompt.
5. After the software installation is complete, remove the System Controller Software 1.3 CD.

## Upgrading Your Software

When you are ready to upgrade your SGI ProPack software, follow the instructions outlined in the section titled “Installing SGI ProPack for Linux” on page 12, using the newer SGI ProPack CD set.

To reinstall all of the software on your system, follow the entire set of installation instructions listed previously, using the latest SGI ProPack CD set.

## Recovering a Damaged Root Filesystem

If your root filesystem should become damaged, use the following recovery steps:

1. Insert SGI Advanced Linux Environment 2.1.1 CD1 into the system’s CD-ROM drive and restart the system.
2. While the system is powering up, check the device mapping table. It should look similar to the following:

Device mapping table

```
fs0  : Pci(2|1)/Ata(Primary, Master)/CDROM(Entry1) This is the CD-ROM
fs1  : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part1,Sig00000000)
fs2  : Pci(1|1)/Scsi(Pun0/Lun2)/HD(Part1,Siggl)
blk0 : Pci(2|1)/Ata(Primary, Master)
blk1 : Pci(2|1)/Ata(Primary, Master)/CDROM(Entry1)
blk2 : Pci(1|1)/Scsi(Pun0/Lun1)
blk3 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part1,Sig00000000)
blk4 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part2,Sig00000000)
blk5 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part3,Sig00000000)
blk6 : Pci(1|1)/Scsi(Pun0/Lun1)/HD(Part4,Sig00000000)
blk7 : Pci(1|1)/Scsi(Pun0/Lun2)
blk8 : Pci(1|1)/Scsi(Pun0/Lun2)/HD(Part1,Siggl0)
```

3. At the Shell> prompt, type the CD-ROM device name, as follows:

```
Shell> fs0:           Type this to change to the CD-ROM device
fs0:\>
```

4. Boot the CD by typing **elilo linux rescue**, as follows:

```
fs0:\> elilo linux rescue           This starts the boot loader
```

5. The CD will then boot and start the rescue process.

6. Select the language you prefer to use for the rescue process. To continue, select **OK**.

---

**Note:** Console support for Asian languages is not available at this time. SGI recommends that customers in Asia set the language to English.

---

7. The next screen lets you choose the media containing the rescue image. Currently, **Local CDROM** is the only choice. To continue, select **OK**.
8. On the next screen, you have the option of mounting your root drive. If your root filesystem has been damaged, it is important to attempt to mount it before running the `xfstool` command. This ensures that valuable data from the journal is recovered.

If you select **Continue** and the installer hangs or crashes, you will have to restart this process and select **Skip**. Then you can attempt to mount the filesystem manually. See the section titled “Manually Mounting the Root Filesystem”.

9. If all goes well, the next screen informs you that your system can be found under `/mnt/sysimage`. To enter the rescue shell, select **OK**.
10. At this point, administrators can inspect and attempt to repair or recover any damaged or missing software. What follows is an example of how you might repair a corrupted root filesystem.

While it is important to mount your filesystem to recover any data contained in the journal, you must then unmount the filesystem before running the `xfstool` command. The `mount` command issued without arguments lists currently mounted filesystems, as follows:

```
sh-2.05# mount
```

```
rootfs on / type rootfs (rw)
devfs on /dev type devfs (rw)
/dev/root.old on / type ext2 (rw)
none on /dev type devfs (rw)
/proc on /proc type proc (rw)
/dev/pts on /dev/pts type devpts (rw)
/tmp/cdrom on /mnt/source type iso9660 (ro)
/tmp/loop0 on /mnt/runtime type cramfs (ro)
/dev/xscsi/pci01.03.0-1/target1/lun0/part3 on /mnt/sysimage type xfs (rw)
/dev/xscsi/pci01.03.0-1/target1/lun0/part1 on /mnt/sysimage/boot/efi type vfat (rw)
none on /mnt/sysimage/dev/pts type devpts (rw)
none on /mnt/sysimage/proc type proc (rw)
```

The line ending with `/mnt/sysimage type xfs (rw)` also contains the device corresponding to your root disk. In this case,  
`/dev/xscsi/pci01.03.0-1/target1/lun0/part3`.

11. Enter the following command to unmount the filesystem:

```
umount /mnt/sysimage
```

12. Run `xfs_repair` on the filesystem, as follows:

```
xfs_repair /dev/xscsi/pci01.03.0-1/target1/lun0/part3
```

13. After `xfs_repair` completes, the filesystem has been repaired and you can reset your system.

## Manually Mounting the Root Filesystem

If the installer could not locate your root filesystem, you can locate and mount it manually. This section describes that process. For details, see *XFS for Linux Administration*. In rare cases, mounting a corrupted filesystem could cause the mount command to hang or have trouble. If that happens, see *XFS for Linux Administration* for instructions on how to run `xfs_repair` without mounting first.

1. The following example assumes that your root filesystem is on partition 3 and that it resides on a disk in the first disk bay of the IX-brick. To view the partitions available on the system, issue the following command:

```
sh-2.05# cat /proc/partitions
```

```
major minor #blocks name rio rmerge rsect ruse wio wmerge wsect wuse running use aveq
 5      0   125470 xscsi/pci01.01.0/target0/lun0/disc 218 4524 18968 10952 0 0 0 0 0 10952
10952
 4      0  35843686 xscsi/pci01.03.0-1/target1/lun0/disc 32 101 692 69 2 14 512 10 0 80 80
 4      1   513008 xscsi/pci01.03.0-1/target1/lun0/part1 0 0 0 0 0 0 0 0 0 0 0
 4      2   9438208 xscsi/pci01.03.0-1/target1/lun0/part2 0 0 0 0 0 0 0 0 0 0 0
 4      3  25891840 xscsi/pci01.03.0-1/target1/lun0/part3 28 29 584 50 2 14 512 10 0 61 61
 4     16  35843686 xscsi/pci01.03.0-1/target2/lun0/disc 4 72 108 19 0 0 0 0 0 19 19
 4     17   513008 xscsi/pci01.03.0-1/target2/lun0/part1 0 0 0 0 0 0 0 0 0 0 0
 4     18   9438208 xscsi/pci01.03.0-1/target2/lun0/part2 0 0 0 0 0 0 0 0 0 0 0
 4     19  25891840 xscsi/pci01.03.0-1/target2/lun0/part3 0 0 0 0 0 0 0 0 0 0 0
```

2. When the system is running normally, the XSCSI pathname in the following example represents the system disk in bay 1 of the IX-brick. This example assumes that you put the root filesystem on partition 3.

```
/dev/xscsi/pci01.03.0-1/target1/lun0/part3
```

3. To manually mount the filesystem, enter the following commands:

```
sh-2.06# mkdir /mnt/sysimage
```

```
sh-2.07# mount -t xfs /dev/xscsi/pci01.03.0-1/target1/lun0/part3 /mnt/sysimage
```



## Product Support

This chapter documents the product components that are supported on SGI Altix 3000 systems. (For a list of the products, see Table 3-1 on page 20.)

Descriptions of the product components are grouped in this chapter as follows:

- “SGI ProPack for Linux Products” on page 19
- “Operating System Enhancements” on page 24
- “I/O Subsystems” on page 26
- “HPC Application Tools and Support” on page 31
- “System Management” on page 32
- “NUMA Tools” on page 37

### SGI ProPack for Linux Products

Software provided by SGI for the SGI ProPack for Linux release consists of a kernel RPM for the SGI Altix 3000 product, SGI Advanced Linux Environment 2.1.1 RPMs, and value-added software developed by SGI to run specifically on SGI systems. Table 3-1 provides a description of the SGI ProPack for Linux products.

**Table 3-1** SGI ProPack v2.3 for Linux Products

<b>Product</b>	<b>Description</b>
Application performance measuring tools	<p>The following tools perform program optimization:</p> <p>VTune - This tool, developed and supported by Intel, uses the performance measurement facilities of the Itanium processor to take profiles based on elapsed time or other architected events within the processor. These profiles can be used to measure, tune, and improve application performance.</p> <p>pfmon - This tool, available as open source and licensed under the GPL, provides a command line interface to control the performance measurement facilities of the Itanium processor. Data generated by this tool can be post-processed to produce a variety of reports describing application performance. These reports can be used to measure, tune, and improve application performance. The pfmon package also includes a library interface, <code>libpfmon.a</code>, that can be used to create customized performance measurement tools.</p>
Array Services	Provides a set of tools with kernel support that simplify the management of systems and parallel applications for clusters of SGI systems.
CpuMemSets	Provides the kernel support and infrastructure for implementing processor and memory placement.
Cpuset System	.The Cpuset System is primarily a workload manager tool permitting a system administrator to restrict the number of processors that a process or set of processes may use. A system administrator can use cpusets to create a division of CPUs within a larger system.
CSA	Provides jobs-based accounting of per-task resources and disk usage for specific login accounts on Linux systems.

**Table 3-1** SGI ProPack v2.3 for Linux Products **(continued)**

<b>Product</b>	<b>Description</b>
FLEXlm	Provides a floating license, run-time environment. Includes daemons suitable for serving floating licenses.
IOC4 serial driver	Driver that supports the Internal IDE CD-ROM, NVRAM, and Real-Time Clock. Serial ports are supported on the IOC4 base I/O chipset and the following device nodes are created: /dev/ttyIOC4/0 /dev/ttyIOC4/1 /dev/ttyIOC4/2 /dev/ttyIOC4/3
kdb	Supports kernel debugging of the running system, either directly from the keyboard or over a serial console.
Kernel partitioning support	Provides the software infrastructure necessary to support a partitioned system, including cross-partition communication support.
Kernel performance improvements	Includes community-based patches such as the O(1) scheduler back-ported from the 2.5.x kernel and included in the SGI kernel, several patches that reduce lock contention on the Big Kernel Lock (BKL), and FRlocks (Fast-Reader Locks), which reduce contention on the <code>xtime_lock</code> (used by <code>gettimeofday</code> ).
L1 and L2 System Controller firmware	Provides support for managing and monitoring the power, cooling, and testing functions for a brick and system compute rack.
LKCD	Provides system crash dump analysis tools including <code>lcrash</code> and all user level scripts required for saving and configuring system crash dumps.
MPT	Provides industry-standard message passing libraries optimized for SGI computers.

**Table 3-1** SGI ProPack v2.3 for Linux Products (**continued**)

Product	Description
NUMA tools	Provides a collection of NUMA related tools ( <code>dlook</code> , <code>dplace</code> , and so on).
Performance Co-Pilot collector infrastructure	Provides performance monitoring and performance management services targeted at large, complex systems.
PROM	Allows the CPU to boot and allows you to perform system administration and software installations.
runon	Enables running a command on a particular CPU or set of CPUs.
XFS	Provides a high-performance filesystem for Linux.
XSCSI infrastructure	Provides disk and ATAPI CD-ROM support for the QLogic QLA12160 SCSI Host Bus Adapter. In addition, some QLogic fibre channel host bus adapter cards are supported. For details on which cards are supported, see the <i>SGI Altix 3000 User's Guide</i> . The infrastructure includes support for robust error handling, failover, and SAN. It can be configured with or without the Linux SCSI layers. If configured without the Linux SCSI layer, XSCSI will make the familiar disk device names ( <code>/dev/sda</code> , etc).
XVM	Provides software volume manager functionality such as disk striping and mirroring.

SGI does not support the following:

- Software obtained from other places (that is, not released by SGI).
- Other releases, updates, or patches from Red Hat.
- Software patches, drivers, or other changes obtained from the Linux community or other vendors.
- Kernels recompiled or reconfigured to run with parameter settings or other modules as not specified by SGI. In particular, the following kernel components are

not supported due to quality, functionality, or performance issues: ext3, ReiserFS, LVM, and md. You should use XFS and XVM instead.

- Unsupported hardware configurations and devices.

## Operating System Enhancements

Building on the Linux operating system's rapid expansion and improvements for general commercial and enterprise environments, SGI has focused on improving Linux capabilities and performance specifically for high performance computing's (HPC's) big compute and big data environments. Thus, SGI has leveraged its experience with NUMAflex and HPC from its IRIX operating systems and MIPS processor-based systems and concentrated on the Linux kernel improvements specifically important to HPC environments.

### CpuMemSets Support

CpuMemSets provides a Linux kernel facility that enables system services and applications to specify on which CPUs they can be scheduled, and from which nodes they can allocate memory. The SGI ProPack kernel and library installation automatically include support for CpuMemSets.

The default configuration makes all CPUs and all system memory available to all applications. You can use the CpuMemSets facility to restrict any process, process family, or process virtual memory region to a specified subset of the system CPUs and memory.

The `runon` command relies on CpuMemSets to enable a user to run a specified command on a specified list of CPUs. Both a C shared library and Python language module are provided to access the CpuMemSets system interface.

In future releases of SGI ProPack for Linux, SGI anticipates providing additional facilities that provide other system services that provide convenient access to the CpuMemSets facility.

For further documentation and details on CpuMemSets support, see the chapter titled "CPU Memory Sets and Scheduling" in the *Resource Administration Guide for Linux* or browse the files that are installed as part of the CpuMemSets RPM, as listed by the following command:

```
rpm -ql CpuMemSets
```

In particular, see the man pages for `cpumemsets` and `runon`.

CpuMemSets is an SGI open source project, also available from the following location:

```
http://oss.sgi.com/projects/cpumemsets
```

## Cpuset Support

The cpuset system is primarily a workload manager tool that permits a system administrator to restrict the number of processors that a process or set of processes can use.

A cpuset is a named set of CPUs, which can be defined to be restricted or open. A restricted cpuset allows only processes that are members of the cpuset to run on the set of CPUs. An open cpuset allows any process to run on its CPUs, but a process that is a member of the cpuset can run only on the CPUs belonging to the cpuset. A cpuset is defined by a cpuset configuration file and a name.

A system administrator can use cpusets to create a division of CPUs within a larger system. Such a divided system allows a set of processes to be contained to specific CPUs, reducing the amount of interaction and contention those processes have with other work on the system. In the case of a restricted cpuset, the processes that are attached to that cpuset will not be affected by other work on the system; only those processes attached to the cpuset can be scheduled to run on the CPUs assigned to the cpuset. An open cpuset can be used to restrict processes to a set of CPUs so that the effect these processes have on the rest of the system is minimized.

## Comprehensive System Accounting (CSA)

The port of Comprehensive System Accounting (CSA) software packages from IRIX to Linux is the result of an open source collaboration between SGI and Los Alamos National Laboratory (LANL) to provide jobs-based accounting of per-task resources and disk usage for specific login accounts on Linux systems.

Providing extensive system accounting capabilities is often important for very large systems, especially when the system will be shared or made available for other organizations to use. CSA uses a Job Containers feature, which provides on Linux the notion of a *job*. A job is an inescapable container and a collection of processes that enables CSA to track resources for any point of entry to a machine (for example, interactive login, cron job, remote login, batched workload, and so on).

CSA on Linux is an SGI open source project, also available from the following location:

<http://oss.sgi.com/projects/csa>

For further documentation and details on CSA support, see the chapter titled “Comprehensive System Accounting” in the *Resource Administration Guide for Linux*.

## Partitioning

SGI provides the ability to divide a single SGI Altix 3000 system into a collection of smaller system partitions. Each partition runs its own copy of the operating system kernel and has its own system console, root filesystem, IP network address, and physical memory. All partitions in the system are connected via the SGI high-performance NUMAlink interconnect, just as they are when the system is not partitioned. Thus, a partitioned system can also be viewed as a cluster of nodes connected via NUMAlink.

Benefits of partitioning include fault containment and the ability to use the NUMAlink interconnect and global shared memory features of the Altix 3000 to provide high-performance clusters.

For further documentation and details on partitioning, see the *SGI ProPack v2.3 for Linux Release Notes*.

## I/O Subsystems

Although some HPC workloads might be mostly CPU bound, others involve processing large amounts of data and require an I/O subsystem capable of moving data between memory and storage quickly, as well as having the ability to manage large storage farms effectively. The SCSI subsystem, XFS filesystem, XVM volume manager, and data migration facilities were leveraged from IRIX and ported to provide a robust, high-performance, and stable storage I/O subsystem on Linux.

The following sections describe persistent PCI-X bus numbering, persistent naming of Ethernet devices, the XSCSI subsystem, the XSCSI-SCSI subsystem, the XFS filesystem, and the XVM Volume Manager.

### Persistent PCI-X Bus Numbering

Persistent PCI-X bus numbering ensures that bus numbers can remain the same across reboots in case of faulty hardware or reconfiguration. During platform initialization, as buses are discovered, they are assigned a logical bus number. Each logical bus number is

unique, systemwide. The default number of buses supported by SGI Altix 3000 systems is 256 (numbered 0 to 255).

By default, bus numbers are allocated starting from the lowest C-brick module ID to which the I/O brick is connected. An I/O brick is either an IX-brick or a PX-brick on SGI Altix 3000 systems. Each I/O brick is allocated 0x10 buses, although the current I/O bricks support only six buses each. Therefore, bus numbers are not contiguous across the system. Bus numbers are sparse and have holes in them.

An I/O brick has six physical buses. These buses are numbered 0x1 through 0x6, left to right, looking at the back of the I/O brick. If there is more than one I/O brick on the system, the buses on the next I/O brick are numbered 0x11 through 0x16. The rationale for this numbering is that the bus numbers of each I/O brick are stamped on the back of the brick and they start from 1. Therefore, the rightmost digit of a bus number corresponds to the actual stamped number on the I/O brick.

If you have only one I/O brick, you do not need persistent bus numbering. However, if you have more than one I/O brick, persistent bus numbering is strongly recommended, so that if an I/O brick fails to boot, your bus numbers are still the same.

To make use of persistent bus numbering, you must supply the `ioconfig=` parameter to the kernel.

The `ioconfig=` parameter takes a comma-separated list of I/O brick numbers as an argument. The following manual boot example uses `elilo`. It tells the kernel that the I/O brick represented by `101.01` is assigned bus numbers 0x1 through 0x6 and the I/O brick represented by `101.02` is assigned bus numbers 0x11 through 0x16.

```
Shell> elilo vmlinux ioconfig="101.01,101.02" root=/dev/xscsi/pci01.03.0-1/target1/lun0/part3
```

You can find the numbers to use with `ioconfig` by looking at the components of your system from your L2 controller, as in the following example:

```
l2-pumpkin-001-L2>pwr
001c11:
power appears on
001c27:
power appears on
101i01:
power appears on
101p02:
power appears on
```

In this example, 101i01 is an IX-brick and represents the 101.01 in the previous `elilo` boot example. The 101p02 notation is a PX-brick and represents 101.02.

Most customers will have their system set up to boot automatically. This means that you should update your `elilo.conf` file with the `ioconfig` parameter. The `elilo.conf` file is available in the `/boot/efi/EFI/sgi` directory on your Linux system. The following is an example `elilo.conf` file:

```
prompt
timeout=50
relocatable
default=sgilinux
append="ioconfig=101.01,101.02"

image=vmlinuz-2.4.19-sgi21r4
    label=sgilinux
    read-only
    root=/dev/xscsi/pci01.03.0-1/target1/lun0/part3
```

In the previous example, pay special attention to the `append=` line. That is, notice where the `ioconfig` information goes when using `elilo.conf`.

## Persistent Naming of Ethernet Devices

Persistent naming of Ethernet devices is an SGI proprietary mechanism and is supported on SGI Altix 3000 systems. Persistent naming refers to the mechanism that ensures that the Gigabit Ethernet card on the IO9 interface of an SGI Altix 3000 system is set up to always be `eth0`. It guarantees that the base Ethernet device number is assigned to the correct MAC address on an SGI Altix 3000 system even when multiple Ethernet devices are present in the system.

The `/etc/sysconfig/networking/eth0_persist` file contains the mapping of Ethernet device numbers to MAC addresses. If the file does not exist, it is created by the `/etc/rc.d/init.d/eth_persist` script, which is run at boot time. To ensure that `eth0` is indeed assigned to the MAC address of the IO9 Ethernet card, it might be necessary to edit the file after the first time an SGI Altix 3000 system has been brought up after a clean install.

Besides ensuring that the mapping of Ethernet device numbers to MAC addresses persists as cards are added to a system, persistent naming also allows system administrators to control the way in which Ethernet devices are numbered. For example,

if the Ethernet card with device number ethX is lost and the system administrator tries to recover by using the Ethernet card with device number ethY, it is possible to force the latter card to take on Ethernet device number ethX by editing the `/etc/sysconfig/networking/eth0_persist` file accordingly.

Following is a sample `/etc/sysconfig/networking/eth0_persist` file:

```
eth0 08:00:69:13:dc:ec
eth1 08:00:69:13:72:e8
```

The content of this file results in the following configuration:

```
[root]# ifconfig -a
eth0 Link encap:Ethernet HWaddr 08:00:69:13:DC:EC
      inet addr:128.162.246.125 Bcast:128.162.246.255 Mask:255.255.255.0
      UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
      RX packets:843 errors:0 dropped:0 overruns:0 frame:0
      TX packets:1245 errors:0 dropped:0 overruns:0 carrier:0
      collisions:0 txqueuelen:100
      RX bytes:386044 (376.9 Kb) TX bytes:126741 (123.7 Kb)
      Interrupt:59

eth1 Link encap:Ethernet HWaddr 08:00:69:13:72:E8
      BROADCAST MULTICAST MTU:1500 Metric:1
      RX packets:136 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:8850 (8.6 Kb) TX bytes:0 (0.0 b)
      Interrupt:63
```

If the slot directly to the right of the IO9 is populated by an Ethernet board, that is, bus 1, slot 2, and if you re-install Linux after it has been placed there or if you remove the `eth0_persist` persistent naming file, the IO9 could become eth1 instead of eth0. This is not recommended because it could affect your product licenses. If the Ethernet board is in any other IX-brick slot, you will not encounter this problem.

If it is necessary to have an Ethernet board in IO9 bus 1 slot 2, install the OS, including SGI ProPack, with the Ethernet board removed. After SGI ProPack is installed, you can replace the Ethernet board in bus 1 slot 2.

## XSCSI Subsystem

The SGI XSCSI subsystem on Linux leverages from IRIX functionality to provide more robust error handling, failover, and storage area network (SAN) infrastructure support as well as long-term large system performance tuning. XSCSI takes advantage of specific features of SGI architecture that standard open source drivers cannot without rewriting.

The naming convention of XSCSI device names is shown in the following example:

```
/dev/xscsi/pci01.03.0-1/target1/lun0/part1
```

Components of the XSCSI device name are as follows:

pci01	System bus number 0x1
03	Device number 3 on that bus
0-1	Logical unit 0 port 1

Notice that the device number (slot number), logical unit, and port number are fixed. These will never change. However, the system bus number could change because of a hardware problem (such as the I/O brick not booting) or a reconfiguration.

Persistent PCI-X bus numbering (see “Persistent PCI-X Bus Numbering” on page 26), if enabled, provides persistent naming to prevent bus number changes even when the hardware fails or is reconfigured. If you use XSCSI names for mounting or locating devices and you also use persistent bus numbering, your XSCSI device names will always be persistent across reboots.

## XSCSI-SCSI Subsystem

The XSCSI-SCSI subsystem provides SCSI device emulation of XSCSI devices. It allows programs written for the SCSI disk driver, the SCSI tape driver, and the SCSI generic driver to use devices controlled by the XSCSI drivers.

## XFS Filesystem

The SGI XFS filesystem provides a high-performance filesystem for Linux. XFS is an open-source, fast recovery, journaling filesystem that provides direct I/O support, space preallocation, access control lists, quotas, and other commercial file system features. Although other filesystems are available on Linux, performance tuning and

improvements leveraged from IRIX make XFS particularly well suited for large data and I/O workloads commonly found in HPC environments.

For more information on the XFS filesystem, see *XFS for Linux Administration*.

## **XVM Volume Manager**

The SGI XVM Volume Manager provides a logical organization to disk storage that enables an administrator to combine underlying physical disk storage into a single logical unit, known as a logical volume. Logical volumes behave like standard disk partitions and can be used as arguments anywhere a partition can be specified.

A logical volume allows a filesystem or raw device to be larger than the size of a physical disk. Using logical volumes can also increase disk I/O performance because a volume can be striped across more than one disk. Logical volumes can also be used to mirror data on different disks.

## **HPC Application Tools and Support**

SGI has ported HPC libraries, tools, and software packages from IRIX to Linux to provide a powerful, standards-based system using Linux and Itanium 2-based solutions for HPC environments. The following sections describe some of these tools, libraries, and software.

### **Message Passing Toolkit**

The SGI Message Passing Toolkit (MPT) provides industry-standard message passing libraries optimized for SGI computers. On Linux, MPT contains MPI and SHMEM APIs, which transparently utilize and exploit the low-level capabilities within SGI hardware, such as its block transfer engine (BTE) for fast memory-to-memory transfers and the hardware memory controller's fetch operation (fetchop) support. Fetchops enable direct communication and synchronization among multiple MPI processes while eliminating the overhead associated with system calls to the operating system.

Parallel workloads, such as MPI jobs, can be launched, monitored, and controlled across a cluster or partitioned system using the SGI Array Services software. Array Services provides the notion of an array session, which is a set of processes that can be running on

different cluster nodes or system partitions. Array Services is implemented using Process Aggregates (PAGGs), which is a kernel module that provides process containers. PAGGs has been open-sourced by SGI for Linux.

For more information on the Message Passing Toolkit, see the *Message Passing Toolkit: MPI Programmer's Manual*.

## Performance Co-Pilot

The SGI Performance Co-Pilot software was ported from IRIX to Linux to provide a collection of performance monitoring and performance management services targeted at large, complex systems. Integrated with the low-level performance hardware counters and with MPT, Performance Co-Pilot provides such services as CPU, I/O, and networking statistics; visualization tools; and monitoring tools.

## System Management

This section describes system management tools. These include the hardware and software environment required to boot the system, license management, system console, and system controllers.

## PROM Chips

Programmable read-only memory (PROM) chips are placed in your computer at the factory with software programmed into them that allows the CPU to boot and allows you to perform system administration and software installations. The PROM chips are not part of your disk or your operating system; they are the lowest level of access available for your system. You cannot erase them or bypass them. For more information on PROM and the L1 controller, see the *SGI L1 and L2 Controller Software User's Guide*.

## Extensible Firmware Interface (EFI)

SGI Altix 3000 systems provide the Extensible Firmware Interface (EFI), a supporting platform to provide input to the CPU and to handle its output. In addition, the EFI controls the server's boot configuration, maintaining the boot menu in durable, nonvolatile memory.

SGI ProPack v2.3 uses the `elilo-3.3` package, which is fully compliant with EFI specification 1.10 with regard to where the bootloader (`elilo.efi`) should be located in the EFI system partition. According to that specification, the bootloader must be located in a dedicated vendor directory. On SGI Altix systems, that directory is `/boot/efi/EFI/sgi/`. As of version 3.3, `elilo` looks for its configuration file **only** in the directory from which it is loaded. Because this also applies to the kernel, the kernel image, which was previously installed in `/boot/efi/`, is now installed in `/boot/efi/EFI/sgi/`. For further documentation and details about this change, see the `elilo` documentation in `/usr/share/docs/`.

From the EFI prompt, you can perform basic file-management tasks (including text editing), make configuration changes, or write scripts that execute at boot time. For a summary of EFI commands, see Table 3-2.

**Table 3-2** EFI Commands

EFI Command	Description
<code>alias [-bdv] [sname] [value]</code>	Sets or gets alias settings
<code>attrib [-b] [+/- rhs] [file]</code>	Views or sets file attributes
<code>bcfg</code>	Configures boot driver and load options
<code>cd [path]</code>	Updates the current directory
<code>cls [background color]</code>	Clears screen
<code>comp file1 file2</code>	Compares two files
<code>cp file [file] ... [dest]</code>	Copies files or directories
<code>date [mm/dd/yyyy]</code>	Gets or sets date
<code>dblk device [Lba] [blocks]</code>	Performs hex dump of block I/O devices
<code>dh [-b] [-p prot_id]   [handle]</code>	Dumps handle information
<code>dmpstore</code>	Dumps variable store
<code>echo [-on   -off]   [text]</code>	Echoes text to stdout or toggles script echo
<code>edit [file name]</code>	Edits a file
<code>endfor</code>	Script only: Delimits loop construct

**Table 3-2** EFI Commands (**continued**)

EFI Command	Description
endif	Script-only: Delimits IF THEN construct
err [ <i>level</i> ]	Sets or displays error level
exit	Exits
flash <i>filename</i>	Flashes PROM on C-brick
for var in <i>set</i>	Script-only: Indicates loop construct
getmtc	Gets next monotonic count
goto <i>label</i>	Script-only: Jumps to label location in script
guid [-b] [ <i>sname</i> ]	Dumps known guid IDs
help [-b] [ <i>internal command</i> ]	Displays this help
if [not] <i>condition</i> then	Script-only: Indicates IF THEN construct
load <i>driver_name</i>	Loads a driver
ls [-b] [ <i>dir</i> ] [ <i>dir</i> ] ...	Obtains directory listing
map [-bdvr] [ <i>sname[:]</i> ] [ <i>handle</i> ]	Maps shortname to device path
mem [ <i>address</i> ] [ <i>size</i> ] [;MMIO]	Dumps memory or memory mapped I/O
memmap [-b]	Dumps memory map
mkdir <i>dir</i> [ <i>dir</i> ] ...	Makes directory
mm <i>address</i> [ <i>width</i> ] [; <i>type</i> ]	Modifies memory: Mem, MMIO, IO, PCI
mode [ <i>col row</i> ]	Sets or gets current text mode
mount <i>BlkDevice</i> [ <i>sname[:]</i> ]	Mounts a filesystem on a block device
mv <i>sfile dfile</i>	Moves files
pause	Script-only: Prompts to quit or continue
pci [ <i>bus dev</i> ] [ <i>func</i> ]	Displays PCI device(s) info
reset [cold/warm] [ <i>reset string</i> ]	Indicates cold or warm reset

**Table 3-2** EFI Commands (**continued**)

EFI Command	Description
<code>rm file/dir [file/dir]</code>	Removes file or directories
<code>set [-bdv] [sname] [value]</code>	Sets or gets environment variable
<code>setsize newsize fname</code>	Sets the files size
<code>stall microseconds</code>	Delays for <i>x</i> microseconds
<code>time [hh:mm:ss]</code>	Gets or sets time
<code>touch [filename]</code>	Views or sets file attributes
<code>type [-a] [-u] [-b] file</code>	Types file
<code>ver</code>	Displays version information
<code>vol fs [volume label]</code>	Sets or displays volume label

## FLEXlm

FLEXlm is a flexible license management system from Macrovision that lets independent software vendors (ISVs) license their products and helps system administrators install and manage licenses with minimal overhead. It supports a wide range of licensing options, including simple node-locked licenses and floating licenses with redundant servers.

To build licensed software, ISVs must purchase a set of keys from Macrovision. System administrators can install license servers anywhere. Products purchased from SGI are typically licensed using FLEXlm.

For more information, visit

<http://www.macrovision.com/solutions/esd/flexlm/flexlm.shtml>

## SGIconsole

SGIconsole is a combination of hardware and software that provides console management and allows monitoring of multiple SGI servers running the IRIX operating

system and SGI ProPack for Linux. These servers include SGI partitioned systems and large, single-system-image servers, including SGI Altix 3000 servers.

SGIconsole consists of an 1U rackmountable SGI server based on the Intel Pentium processor, a serial multiplexer or Ethernet hub, and a software suite that includes the Console Manager package and Performance Co-Pilot, which provides access to common remote management tools for hardware and software.

Console Manager is a graphical user interface for the SGIconsole management and monitoring tool used to control multiple SGI servers. SGIconsole also has a command line interface. For more information on SGIconsole, see the *SGIconsole Start Here*.

## System Controller Firmware

The L1 and L2 controllers are system controller firmware used in SGI systems.

The L1 controller is embedded in each brick in SGI Origin and Onyx 3000 series systems and in SGI Altix 3000 systems. It provides power and control sequencing, along with temperature and power monitoring for each brick.

The L2 controller is a rack-level controller that monitors and controls the bricks in its rack. All L2 controllers in a system are networked together and they consolidate the control and monitoring information from each brick to provide system-level control and monitoring.

The following references describe how to install and connect to an L1 or L2 controller:

- Procedure 3-10, “Determining the Hardware Address of the L2 System Controller” in chapter 3 of the *SGIconsole 2.0 Start Here* describes how the L2 system controllers come online after a system boots and how to set the eth1 port on the SGI 1100 server to match when a DHCPD server is running on a subnet.
- Procedure 3-2, “Determining the L2 Address of a New L2 Controller” in chapter 3 of the *Console Manager for SGIconsole Administrator’s Guide* describes how to find the L2 address of a new L2 controller.
- Procedure 4-1, “Connecting to a System Console” in chapter 4 of the *Console Manager for SGIconsole Administrator’s Guide* describes how to connect to a system console.

- Procedure 2-2, “Connecting to the L1 Controller” in chapter 2 of the *Linux Configuration and Operations Guide* describes how to connect to an L1 controller and from there an L2 controller with or without SGIconsole installed.
- Procedure 2-3, “Connecting to the L2 Controller” in chapter 2 of the *Linux Configuration and Operations Guide* describes how to connect to the L2 controller with or without SGIconsole installed.

For more information on the L1 and L2 system controller firmware, see the *SGI L1 and L2 Controller Software User’s Guide*.

## NUMA Tools

This section describes the commands that are currently provided with the collection of NUMA related tools.

### d1ook Command

The `d1ook` command displays the memory map and CPU use for a specified process. The following information is printed for each page in the virtual address space of the process:

- The object that owns the page (file, SYSV shared memory, device driver, and so on)
- Type of page (RAM, FETCHOP, IOSPACE, and so on)
- If RAM memory, the following information is supplied:
  - Memory attributes (SHARED, DIRTY, and so on)
  - Node on which that the page is located
  - Physical address of page (optional)

Optionally, the amount of elapsed CPU time that the process has executed on each physical CPU in the system is also printed.

## **dplace Command**

The `dplace` command binds a related set of processes to specific CPUs or nodes to prevent process migrations. In some cases, this tool improves performance because of the occurrence of a higher percentage of memory accesses to the local node.

## Performance Tuning

---

**Note:** The performance tuning information in this chapter is now available in Chapter 3, “Performance Tuning” in the new *Linux Configuration and Operations Guide* available with the SGI ProPack v2.3 release.

---



---

# Index

## A

Array Services, description, 20  
Autopartition option, 8

## B

Booting the CD, 8, 14  
Bus numbering, 26  
Buttons for returning, 7

## C

CD contents, 2  
Comprehensive System Accounting (CSA) support,  
25  
CpuMemSet support, 24  
Cpuset system, 25  
CSA support, 25

## D

Disk partition table, 6  
dlook command, 37  
dplace command, 38

## E

EFI commands, 33  
EFI support, 32  
Ethernet device naming, 28  
Extensible Firmware Interface (EFI) support, 32

## F

Filesystem recovery, 14  
Firewall configuration, 9  
FLEXlm support, 35

## H

hardware platforms, 1  
HPC support  
libraries and tools, 31  
Message Passing Toolkit, 31  
Performance Co-Pilot (PCP), 32

## I

Installation  
additional users, 9  
CDs, 12  
log, 12  
method selection, 8, 15  
open/free source software, 12

- overview, 4
- package group selection, 10, 12
- password for access, 9
- proprietary software, 12
- reboot, 13
- SGI Linux Environment 7.2 base OS, 7
- SGI ProPack for Linux, 12
- system controller software, 13
- time zone, 9
- upgrading, 14
- user creation, 9

#### I/O subsystems

- for HPC systems, 26
- XFS filesystem support, 30
- XSCSI support, 30
- XSCSI-SCSI support, 30
- XVM support, 31

- IP address selection, 9

#### L

- L1 and L2 controllers, 36
- Language selection, 8, 9, 15
- Linux Kernel Crash Dump (LKCD), 6
- LKCD for file configuration, 6

#### M

- Message Passing Toolkit support, 31

#### N

- NUMA tools, 37
  - dllook command, 37
  - dplace command, 38

#### O

- Open source website, 4
- OS configuration settings, 1
- OS enhancements
  - Comprehensive System Accounting (CSA), 25
  - CpuMemSets, 24
  - cpuset support, 25
  - for HPC environments, 24
  - partitioning, 26

#### P

- Package group selection, 10, 12
- Partition selection, 8
- Password requirement, 5
- Performance
  - measuring, 20
- Performance Co-Pilot support, 32
- Persistent bus numbering, 26
- Persistent naming
  - bus numbers, 26
  - Ethernet devices, 28
  - XSCSI devices, 30
- Product list, 19
- Product support, 19
- PROM support, 32

#### R

- Release features, 1
- Remove partitions, 8
- Rescue mode, 14
- Root access password, 9

**S**

## Screen

- Complete, 13
- Installation to Begin, 12
- Package Group Selection, 10, 12
- Package Installation, 12
- Welcome, 12

SGIconsole support, 35

## Software

- installation, 5
- introduction, 1
- planning, 6

System controller firmware support, 36

## System Controller Software

- CD, 13

## System management

- EFI, 32
- FLEXlm, 35
- PROM chips, 32
- SGIconsole, 35
- system controller firmware, 36
- tools, 32

**U**

Unsupported elements, 22

Upgrading from previous releases, 11

Upgrading software, 14

**X**

XFS filesystem support, 30

XSCSI support, 30

XSCSI-SCSI support, 30

XVM support, 31

