



SGI® Performance Suite 1.3 Start Here

007-5680-004

---

## COPYRIGHT

© 2010, 2011, SGI. 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 SGI.

---

## 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

Accelerate, Altix, IRIX, NUMAflex, NUMalink, OpenGL, Performance Co-Pilot, REACT, SGI, the SGI logo, SHMEM, Silicon Graphics, Supportfolio, and XFS are trademarks or registered trademarks of Silicon Graphics International Corp. or its subsidiaries in the United States and other countries.

FLEXlm is a registered trademark of Macrovision Corporation. Intel and Xeon are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Linux is a registered trademark of Linus Torvalds in the U.S. and other countries. Novell and SUSE are registered trademarks of Novell, Inc., in the United States and other countries. QLogic is a registered trademark of QLogic Corporation. Platform Manager is a trademark of Platform Computing, Inc. PostScript is a registered trademark of Adobe Systems, Inc. Red Hat and all Red Hat-based trademarks are trademarks or registered trademarks of Red Hat, Inc. in the United States and other countries. Windows is a registered trademark of Microsoft Corporation in the United States and/or other countries. All other trademarks mentioned herein are the property of their respective owners.

---

## Record of Revision

<b>Version</b>	<b>Description</b>
001	October 2010 Original publication. Supports the SGI Performance Suite 1.0 release.
002	February 2011 Updated to support the SGI Performance Suite 1.1 release.
003	April 2011 Updated to support the SGI Performance Suite 1.2 release.
004	November 2011 Updated to support the SGI Performance Suite 1.3 release.



---

## New Features in This Guide

This update of *SIG Performance Suite 1.3 Start Here* supports the SIG Performance Suite 1.3 release.

### Major Documentation Changes

Changes in this guide for the SIG Performance Suite 1.3 release include the following:

- Updated information in “SIG Performance Suite 1.3 Updates” on page 10.
- Described manual updates in “Documentation Updated for This Release” on page 12.
- Added information about superpages on Altix UV systems in “Superpages” on page 16.



---

# Contents

<b>Record of Revision</b>	iii
<b>Tables</b>	ix
Helpful Online Resources	xii
Related Publications	xiv
Novell Documentation	xiv
Red Hat Documentation	xv
Intel Compiler Documentation	xv
Other Intel Documentation	xv
SGI Documentation	xvi
SGI Performance Suite Manuals	xvi
Additional SGI Software Manuals	xvii
SGI Hardware Manuals	xix
Obtaining SGI Publications	xxii
Conventions	xxii
Reader Comments	xxii
<b>1. Release Features</b>	<b>1</b>
Software Introduction	1
SGI Performance Suite Software Bundles	2
SGI® MPI	2
SGI® REACT™	3
SGI® UPC	4
SGI® Accelerate™	5
Software Package License Information	6
SGI Performance Suite 1.3 CDs	7
Release Information and Release Notes	7
Finding SGI Performance Suite Manuals	7

Getting Updates . . . . .	9
SLES 11 SP1 Updates . . . . .	9
RHEL 6 and RHEL 6.1 Updates . . . . .	10
SGI Performance Suite 1.3 Updates . . . . .	10
SGI Foundation 2.5 Updates . . . . .	11
SGI Open Source Information. . . . .	11
Documentation Updated for This Release . . . . .	12
<b>2. Software Installation . . . . .</b>	<b>15</b>
Systems with Multiple Ethernet Configurations . . . . .	15
XVM Volumes Access in Rescue Mode . . . . .	15
Adding the <code>xfs_quota</code> Variable to the <code>initrd</code> Boot Image . . . . .	16
Superpages . . . . .	16
SLES 11 SP1 or RHEL 6.1 Kernels . . . . .	18
Getting the Latest Firmware/Software for Your SGI System. . . . .	18
License Manager Updates . . . . .	18
Upgrading from an Earlier Release . . . . .	19
<b>3. SGI Performance Suite Products Overview. . . . .</b>	<b>21</b>
SGI Performance Suite 1.3 Products . . . . .	21
Performance Suite Products . . . . .	24
Cpuset Support . . . . .	24
Partitioning. . . . .	25
I/O Subsystems. . . . .	25
XFS Filesystem . . . . .	25
XVM Volume Manager . . . . .	26
HPC Application Tools and Support. . . . .	26
Message Passing Toolkit . . . . .	26
Performance Co-Pilot . . . . .	27
NUMA Data Placement Tools. . . . .	27
<code>dlook</code> Command . . . . .	27
<code>dplace</code> Command . . . . .	28
<code>taskset</code> Command . . . . .	28
<b>Index. . . . .</b>	<b>29</b>



---

# Tables

<b>Table 3-1</b>	SGI Performance Suite 1.3 Products . . . . .	22
------------------	--	----



---

## About This Guide

This guide provides information about the SGI Performance Suite 1.3 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 Installation,” describes the flowcharts of disk partitions and file configurations and layouts that you need to get started and provides instructions for installing SGI Performance Suite software.
- Chapter 3, “SGI Performance Suite Products Overview,” documents the product components that are supported on SGI systems.

Online versions of this document, the release notes, which contain the latest information about software and documentation in this release, the list of RPMs distributed with SGI Performance Suite bundle, can be found in the `/docs` directory on the SGI Performance Suite CD for each product.

This Preface covers the following topics:

- “Helpful Online Resources” on page xii
- “Novell Documentation” on page xiv
- “Red Hat Documentation” on page xv
- “Intel Compiler Documentation” on page xv
- “Other Intel Documentation” on page xv
- “SGI Documentation” on page xvi
- “Obtaining SGI Publications” on page xxii

## Helpful Online Resources

This section describes the electronic support services available for your SGI system running SGI Performance Suite software, as follows:

- Supportfolio

Supportfolio, the SGI support web site, including the SGI Knowledgebase, links for software supports and updates.

<https://support.sgi.com/login>

- Supportfolio SGI Knowledgebase

The SGI Knowledgebase available from Supportfolio has answers to common questions, problem resolutions, how-to documents, and many other useful articles for system administrators.

<https://support.sgi.com/kb/publisher.asp>

- SGI Update Server

The SGI update server (<https://update.sgi.com/>) is available for your convenience. Get updates to the SGI Performance Suite software releases using update tools (for example, YaST, zypper and yum) available in your operating system environment. Refer to the product release notes for more information about specific product details on the SGI update server.

- Supportfolio Software Updates

Check for new updates to your purchased SGI software, and download or order CDs all in one convenient place. Supportfolio Software Updates provides one-stop shopping to check for and get new updates for your SGI software.

<https://support.sgi.com/software?cmd=order>

- SGI Customer Support Center

You can open calls to the SGI Customer Support Center online, keep track of all your cases, and receive automatic update information when your case is updated via the Supportfolio Problem Resolution Tools at:

<https://support.sgi.com/caselist>

- SGI Software Licensing

If you need to get your license for SGI software, you can order your software licenses conveniently via e-mail or online.

You can send a blank e-mail to [license@sgi.com](mailto:license@sgi.com), and you will receive a simple template to fill out. You will quickly receive your new license back with complete instructions for installing it back via e-mail.

Or, you can go online to the SGI Key-O-Matic at:

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

## Related Publications

Documents listed in this section contain additional information that might be helpful, as follows:

- “Novell Documentation” on page xiv
- “Red Hat Documentation” on page xv
- “Intel Compiler Documentation” on page xv
- “Other Intel Documentation” on page xv
- “SGI Documentation” on page xvi

## Novell Documentation

This section describes documentation for SUSE Linux Enterprise Server 11 available from Novell.

Documentation for SUSE Linux Enterprise Server 11 is located on your system in the `/docu/en` directory.

---

**Note:** Novell makes SUSE Linux Enterprise Server 11 documentation available online at: <http://www.novell.com/documentation/>

---

The following SLES11 documentation is available:

- *SUSE Linux Enterprise Server 11 Installation Quick Start*  
`art_sles_installquickstart.pdf`
- *SUSE Linux Enterprise Server 11 Deployment Guide*  
`book_sle_deployment.pdf`
- *SUSE Linux Enterprise Server 11 Administration Guide*  
`book_sle_admin.pdf`
- *SUSE Linux Enterprise Server SLES11: Storage Administration Guide*  
`stor_admin.pdf`
- *SUSE Linux Enterprise Server 11 Package Descriptions* available online at:  
<http://www.novell.com/products/linuxenterpriseserver/packages.html>

## Red Hat Documentation

Documentation for Red Hat Enterprise Linux 6 (RHEL 6) or Red Hat Enterprise Linux 6.1 (RHEL 6.1) is located in the `/usr/share/doc` directory. It contains package-specific documentation. If you install the `zISOFS` module, additional documentation in the form of READMEs can be found in `/usr/share/doc/zISOFS-tools-1.x.x`.

On the RHEL6 DVD, `README-{lang}` and `RELEASE-NOTES-{lang}` files are found in the root directory.

---

**Note:** Red Hat makes Red Hat Enterprise Linux 6 (RHEL 6) documentation available online at:

<http://www.redhat.com/docs/manuals/enterprise/RHEL-6-manual/index.html>

---

The following documentation is available:

- *Release Notes for x86\_64 architecture*
- *Red Hat Enterprise Linux Deployment Guide*
- *Red Hat Enterprise Linux Installation Guide*
- *Red Hat Virtualization*

## 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_prg.pdf`).
- *Intel Fortran Libraries Reference* (`for_lib.pdf`).

## Other Intel Documentation

Intel makes Intel server processor information available online at:

<http://www.intel.com/support/processors/sb/CS-029807.htm>

## SGI Documentation

This section describes SGI documentation and covers these topics:

- “SGI Performance Suite Manuals” on page xvi
- “Additional SGI Software Manuals” on page xvii
- “SGI Hardware Manuals” on page xix

---

**Note:** Online and postscript versions of SGI documentation is available at SGI Technical Publications Library at <http://docs.sgi.com>.

---

## SGI Performance Suite Manuals

- *SGI Performance Suite Product Bundles Release Notes*  
Provides the latest information about software and documentation in this release. Release notes that are in a file named `README.TXT` that is available in `/docs` directory on the CD media for each SGI Performance Suite software bundle.
- *Message Passing Toolkit (MPT) User’s Guide*  
Describes industry-standard message passing protocol optimized for SGI computers.
- *MPInside Reference Guide*  
Documents the SGI MPInside MPI profiling tool.
- *REACT Real-Time for Linux Programmer’s Guide*  
Describes the facilities of REACT real-time for Linux that supports real-time programs.
- *Performance Co-Pilot 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.
- *Unified Parallel C (UPC) User’s Guide*  
Documents the SGI implementation of the Unified Parallel C (UPC) parallel extension to the C programming language standard.



- *SGI Altix UV GRU Development Kit Programmer's Guide*  
Documents the SGI Altix UV global reference unit (GRU) development kit. It describes the application program interface (API) that allows direct access to GRU functionality.
- *Linux Resource Administration Guide*  
Provides a reference for people who manage the operation of SGI servers and contains information needed in the administration of various system resource management features such as, Array Services, CPU memory sets and scheduling, and the Cpuset System.
- *Linux Application Tuning Guide for SGI X86-64 Based Systems*  
Provides information about tuning application programs on SGI systems. Application programs include Fortran and C programs written on SGI Linux systems with the compilers provided by Intel.
- *SGI Altix UV Systems Linux Configuration and Operations Guide*  
Provides information on how to perform system configuration and operations under the Linux operating system for SGI Altix UV 100 and SGI UV 1000 services systems.

## Additional SGI Software Manuals

- *SGI Foundation Software 2.5 Start Here*  
Provides information about the SGI Foundation Software 2.5 release that supports SGI differentiated server solutions.
- *SGI Altix UV System Management Node Administrator's Guide*  
Describes the system management node (SMN) installation, configuration and operation for SGI Altix UV 1000 and SGI Altix UV 100 series systems.
- *SGI Management Center for Altix ICE*  
Describes how to perform general system installation, configuration, and operations on the SGI Altix ICE series systems. For other SGI hardware platforms, see the *SGI Management Center System Administrator's Guide* and *SGI Management Center Installation and Configuration*.
- *SGI Management Center Installation and Configuration*  
This guide is intended for system administrators. It describes how to install and configure the SGI Management Center. A companion manual, *SGI Management Center System Administrator's Guide*, describes general cluster administration. For SGI Altix ICE platforms, use the manual *SGI Management Center for Altix ICE*.

- *SGI Management Center System Administrator's Guide*  
This guide describes how you can monitor and control a cluster using the SGI Management Center. A companion manual, *SGI Management Center Installation and Configuration Guide*, describes installing and configuring the SGI Management Center. For SGI Altix ICE platforms, use the manual *SGI Management Center for Altix ICE*.
- *SGI Management Center Quick Start Guide*  
The guide tells you how to get you SGI Management Center up and running.
- *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.
- *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.
- *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.
- *XVM Volume Manager Administrator's Guide*  
Describes the configuration and administration of XVM logical volumes using the XVM Volume Manager.
- *CXFS 6 Administration Guide for SGI InfiniteStorage*  
Documents CXFS running on a storage area network (SAN). CXFS is clustered XFS, a parallel-access shared clustered filesystem for high-performance computing environments.
- *CXFS 6 Client-Only Guide for SGI InfiniteStorage*  
Provides an overview of the installation and configuration procedures for CXFS client-only nodes running SGI CXFS clustered filesystems.
- *Guaranteed-Rate I/O Version 2 for Linux Guide*  
Provides information about GRIO version 2, the second-generation guaranteed-rate I/O product from SGI.

- *High Availability Extension and SGI InfiniteStorage*  
Provides information about creating resources for the high-availability (HA) SGI resource agents that SGI provides for use with the SUSE Linux Enterprise High Availability Extension (HAE) product.
- *DMF 5 Administrator's Guide for SGI InfiniteStorage*  
Describes how to perform administration of the Data Migration Facility (DMF). DMF is a hierarchical storage management system for SGI environments. Its primary purpose is to preserve the economic value of storage media and stored data. This manual describes DMF components, features, and architecture; it focuses on the tasks of configuration and ongoing administrative duties.
- *SGI InfiniteStorage Appliance Manager User's Guide*  
Describes the operation of SGI InfiniteStorage Appliance Manager, which is a web-based interface that lets you configure, manage, and monitor a storage server.
- *SGI Altix UV Systems Controller Software User's Guide*  
Documents the command line interface (CLI) controller commands for the SGI Altix UV 100 and SGI Altix UV 1000 series systems.
- *SGI Altix UV CMC Controller Software User's Guide*  
Documents the chassis manager controller (CMC) command line interface (CLI) software controller commands for the SGI Altix UV 100 and SGI Altix UV 1000 series systems.
- *Windows on SGI Altix UV Software Install Guide*  
Describes how to install Windows Server 2008 R2 Datacenter software on an SGI Altix UV 100 or SGI Altix UV 1000 system. This manual does **not** apply to the SGI Altix UV 10 system.

## SGI Hardware Manuals

- *SGI Altix ICE 8200 Series System Hardware User's Guide*  
Describes the features of the SGI Altix ICE 8200 series systems and provides operating instructions and general troubleshooting information.
- *SGI Altix ICE 8400 Series System Hardware User's Guide*  
Describes the features of the SGI Altix ICE 8400 series systems and provides operating instructions and general troubleshooting information.

- *SGI Altix UV 10 System User's Guide*  
This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Altix UV 10 system. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- *SGI Altix UV 100 System User's Guide*  
This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Altix UV 100 system. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- *SGI Altix UV 1000 System User's Guide*  
This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Altix UV 1000 system. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- *SGI Rackable C1104-2TY9 System User's Guide*  
Describes the features of the SGI Rackable C1104-2TY9 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Rackable C2108-TY11 System User's Guide*  
Describes the features of the SGI Rackable C2108-TY11 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Rackable C3108-TY11 System User's Guide*  
Describes the features of the SGI Rackable C3108-TY11 server system as well as troubleshooting, upgrading, and repairing.
- *SGI C1103-TY12 System User's Guide*  
This guide provides an overview of the architecture and descriptions of the major components that compose the SGI C1103-TY12 system. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- *SGI Octane III Workstation System User's Guide*  
This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Octane III workstation. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- *SGI Octane III Deskside Cluster Server User's Guide*

This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Octane III Deskside Cluster server. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.

- *SGI CloudRack X2 System User's Guide*  
This guide provides an overview of the SGI CloudRack X2 workgroup cluster server's architecture and descriptions of the major components. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- *SGI Origin 400 System User's Guide*  
This manual is written for system technicians who are responsible for troubleshooting, upgrading, and repairing Origin 400 systems. This document provides reference information, feature information, and step-by-step instructions for adding and replacing components in the Origin 400 system.

## 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
- 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  
46600 Landing Parkway  
Fremont, CA 94538

SGI values your comments and will respond to them promptly.





## Release Features

This chapter provides an introduction to SGI Performance Suite 1.3 release, describes its contents, and provides documentation information. It covers the following topics:

- “Software Introduction” on page 1
- “SGI Performance Suite Software Bundles” on page 2
- “Finding SGI Performance Suite Manuals” on page 7
- “Getting Updates” on page 9
- “SGI Open Source Information” on page 11
- “Documentation Updated for This Release” on page 12

## Software Introduction

The SGI Performance Suite 1.3 is software that accelerates the performance of applications running on the Linux operating system. It consists of the components described in the following sections:

- “SGI® MPI” on page 2
- “SGI® REACT™” on page 3
- “SGI® UPC” on page 4
- “SGI® Accelerate™” on page 5

SGI Performance Suite 1.3 is fully supported on the SUSE Linux Enterprise Server 11 Service Pack 1 (SLES 11 SP1) base distribution or the Red Hat Enterprise Server 6 (RHEL 6) or RHEL 6 .1 base distribution. SGI Performance Suite software (with the exception of SGI REACT) also runs on the Community ENTERprise Operating System 6 (CentOS 6) base distribution. While SGI support will answer questions relating to any of these products, it does not provide technical support for the CentOS operating system.

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

<http://www.sgi.com/servers/altix/>

## SGI Performance Suite Software Bundles

This section describes the main contents of the SGI Performance Suite 1.3 release. Manuals associated with the SGI Performance Suite software can be found at the SGI Technical Publications Library located at: <http://docs.sgi.com>.

### SGI® MPI

SGI® MPI is a scalable, high performance message passing interface (MPI) environment including the following software packages:

- Message Passage Toolkit (MPT) software stack (`sgi-mpt`)
- PerfBoost  
SGI PerfBoost uses a wrapper library to run applications compiled against other MPI implementations under the SGI Message Passing Toolkit (MPT) product on SGI platforms
- `memacct`  
Memory accounting utility (see the `memacct(3)` man page)
- `sgi-mpt-shmem`  
Contains the SGI implementation of SHEM which are shared, distributed memory access (SHMEM) routines (see next bullet item)
- `xpmem`  
SGI's NUMALink technology and the XPC and XPMEM kernel modules allows users to create a very low latency, very large, shared-memory cluster for optimized use of Message Passing Interface (MPI) software and logically shared, distributed memory access (SHMEM) routines. The globally addressable, cache coherent, shared memory is exploited by MPI and SHMEM to deliver high performance.
- MPInside

MPInside is a MPI profiling tool which provides valuable information to help MPI application developers to optimize their application.

For more information, see these related manuals:

*Message Passing Toolkit (MPT) User's Guide*

*MPInside 3.1 Reference Manual* (ships with the MPInside software module only)

*Linux Application Tuning Guide for SGI X86-64 Based Systems*

For more information on `xpmem` and `SHMEM`, see the “System Partitioning” section of Chapter 2. “Configuring Your System” in the *SGI Altix UV Systems Linux Configuration and Operations Guide*.

## SGI® REACT™

SGI® REACT™ is hard real-time performance solution for Linux and includes the following software packages:

- Linux REACT stack
- External interrupt driver
- REACT/Pro Frame Scheduler

The REACT/Pro Frame Scheduler is an execution manager that schedules `pthread`s or processes on one or more CPUs in a predefined cyclic order (see the `frs(3)` man page).

- Kernel barrier (`kbar`) facility

The kernel barrier (`kbar`) facility provides for the fast wakeup of many blocked user threads. When the barrier is signaled the operating system will use a configurable number of CPUs to quickly wake all the threads blocked on it (see the `kbar(3)` man page).

- `sgi-shield` kernel module

Enables the use of the `cpu_shield` routine that controls timer interrupts on select CPUs

- User level interrupt facility (see the `uli(3)` man page).
- SGI Linux Trace (SLT)

For more information on SGI REACT, see *REACT Real-Time Programmer's Guide*.

## SGI® UPC

SGI® UPC is an implementation of the Unified Parallel C (UPC) parallel extension to the C programming language standard and includes the following software packages:

- `sgi-upc`
- `sgi-upc-devel`
- `xpmem`

Cross process memory mapping. See `xpmem` descriptions in “SGI® MPI” on page 2.

- `sgi-array-services`

SGI Array Services includes administrator commands, libraries, daemons, and kernel extensions that support the execution of parallel applications across a number of hosts in a cluster, or array.

- `sgi-mpt`

SGI MPT runtime environment. See “SGI® MPI” on page 2.

- `sgi-mpt-shmem`

Contains the SGI implementation of SHEM which are shared, distributed memory access (SHMEM) routines. Also see `xpmem` descriptions in “SGI® MPI” on page 2

- `libgru`

SGI Altix UV global reference unit (GRU) libraries

- `superpages`

---

**Note:** The `superpages` module is only supported on SGI Altix UV 100 and UV 1000 series systems. For more information on superpages, see “Superpages” on page 16.

---

- `numatools`

SGI NUMAlink memory placement tools

- libcpuset

The cpuset library provides the mechanisms needed to create, destroy and manage cpusets, and to control the placement of processes and threads within cpusets.

- libbitmask

This Bitmask library supports multi-word bitmask operations for applications programmed in the C programming language. It works in conjunction with Linux kernel support for processor and memory placement on multiprocessor SMP and NUMA systems. The cpuset library depends on this bitmask library.

For more information, see these related manuals:

*Unified Parallel C (UPC) User's Guide*

*Message Passing Toolkit (MPT) User's Guide*

*Linux Resource Administration Guide*

*SGI Altix UV GRU Development Kit Programmer's Guide*

*Linux Application Tuning Guide for SGI X86-64 Based Systems*

## **SGI® Accelerate™**

SGI® Accelerate™ is used to accelerate applications with optimized software libraries and runtime performance tools and includes the following software packages:

- cpusets

The cpuset facility is primarily a workload manager tool permitting a system administrator to restrict the number of processor and memory resources that a process or set of processes may use. A cpuset defines a list of CPUs and memory nodes.

- numatools

See numatools description in “SGI® UPC” on page 4.

- **FFIO**  
Cpusets and NUMA tools provide controls for dedicated resource use which in turn helps optimize performance on a system shared by many users. For I/O intensive applications, FFIO helps improve performance without additional coding or re-tooling of the software logic.
- **sgi-pcp**  
SGI Performance Co-Pilot Altix UV hub performance counters
- **libgru development kit**  
SGI Altix UV Global Reference Unit (GRU), part of the SGI Altix UV Hub application-specific integrated circuit (ASIC), Library development Kit.
- **SGISOLVE**  
The PSLDLT libraries under the SGISOLVE umbrella contain two direct solvers, PSLDLT and PSLDU, and one iterative solver, DIterative (deprecated), for solving symmetric, general unsymmetric, double precision, double complex, 32bit and 64bit integer, sparse linear systems of equations. These solvers are optimized and parallelized for both IN-CORE and OUT-OF-CORE execution on the SGI SMP platforms.

For more information, see these related manuals:

*Linux Resource Administration Guide*

*SGI Altix UV GRU Development Kit Programmer's Guide*

*Performance Co-Pilot Linux User's and Administrator's Guide*

*Linux Application Tuning Guide for SGI X86-64 Based Systems*

## **Software Package License Information**

SGI Performance Suite contains software packages released under a variety of licenses. The licensing terms for all proprietary software packages allow users to install the software only on SGI systems. See the `PACKAGE_LICENSES.txt` file that is available in the `/docs` directory of the appropriate CD media for more details.

## SGI Performance Suite 1.3 CDs

The following CDs ship with the SGI Performance Suite 1.3 release:

- SGI® MPI CD
- SGI® REACT® CD
- SGI® UPC CD
- SGI® Accelerate™ CD

---

**Note:** You only receive the CDs associated with the software bundles you have ordered. If you have ordered the entire SGI Performance Suite, then you will receive all four CDs.

---

## Release Information and Release Notes

For a complete list of RPMs included in the SGI Performance Suite 1.3 release, see the file called `RPMS.txt` that is available in the `/docs` directory on the CD media for each product.

For the latest information about software and documentation in this release, see the release notes that are in a file with the product name and `-readme.txt` suffix that is available in `/docs` directory on the CD media for each product.

## Finding SGI Performance Suite Manuals

The SGI Performance Suite manuals are installed in the following location on your SGI system:

```
/usr/share/doc/packages/sgi-perfsuite-1/
```

Change directory (`cd`) to this location to see a list manual directories. Each manual directory contains a pdf and HTML copy of the manual.

```
sys:/usr/share/doc/packages/sgi-perfsuite-1 # ls
figures      index.html    LX_ConfigOps_AG  MPIinside_UG    PS_Start_Here
GRUDEV_PG   LX_86_AppTune LX_Resource_AG   MPT_UG          UPC_UG
```

For example, `cd` to the `PS_Start_Here` directory and list (`ls`) its contents, as follows:

```
/usr/share/doc/packages/sgi-perfsuite-1 # ls  
pdf sgi_html
```

You can also access the following local URL on your SGI system using the Mozilla Firefox browser, as follows:

```
# firefox -no-remote file:///usr/share/doc/packages/sgi-perfsuite-1/index.html
```

The SGI Performance Suite Documentation page appears, as shown in Figure 1-1.

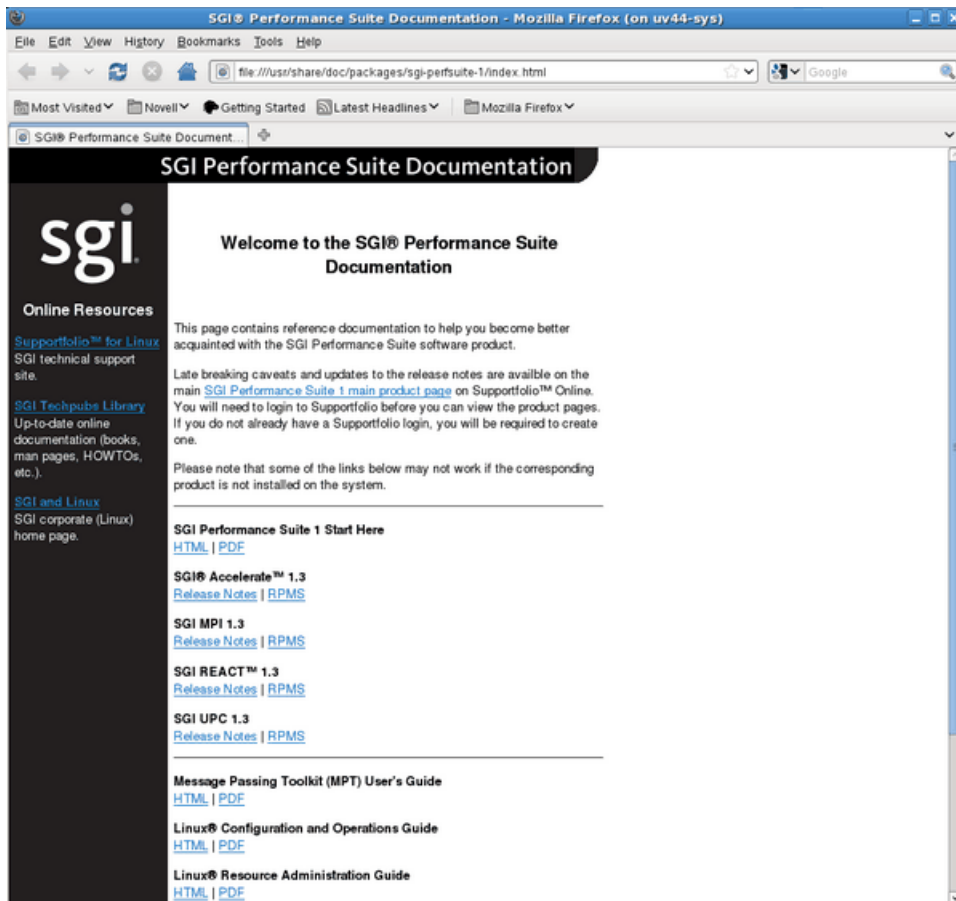


Figure 1-1 SGI Performance Suite Documentation Screen



---

**Note:** Online and postscript versions of SGI documentation is available at SGI Technical Publications Library at <http://docs.sgi.com>.

---

## Getting Updates

This section describes how to get software updates for your system and covers these topics:

- “SLES 11 SP1 Updates” on page 9
- “RHEL 6 and RHEL 6.1 Updates” on page 10
- “SGI Performance Suite 1.3 Updates” on page 10
- “SGI Foundation 2.5 Updates” on page 11

Detailed instructions for downloading and installing updates are available from the SGI product pages on Supportfolio: <https://support.sgi.com/>.

---

**Note:** You must have a Supportfolio account to access this page and the SGI update server.

---

## SLES 11 SP1 Updates

SLES 11 SP1 customers are encouraged to use the YaST Online Update tool to keep the software on their systems updated with the latest maintenance fixes.

Software updates for SLES 11 SP1 are provided by Novell. In order to configure your system to enable online updates, you must register with Novell. For more help with this, you can use the **Novell Customer Center Configuration** tool. The **Novell Customer Center Configuration** module has been moved from the Software to the Support menu in YaST.

## RHEL 6 and RHEL 6.1 Updates

Software updates for RHEL 6 and RHEL 6.1 are provided by Red Hat, Inc. As a Red Hat Enterprise Linux customer, you receive access to the Red Hat Network. This allows you to download updates for your system. For more information on the Red Hat Network and getting updates, see *Red Hat Network Overview* and the **Update** link at <https://www.redhat.com/rhn/>.

You can also use the `yum` software package manager to get updates, as follows:

```
% yum update
```

The `yum` software package manager checks for updates and then updates your system appropriately. You must be registered with RHN, for `yum` to get RHN RHEL 6 updates.

You can configure the `updatesd(8)` notifier daemon to notify you of updates and potentially install the updates for you. Configuration is done via the `yum-updatesd.conf(5)` file.

## SGI Performance Suite 1.3 Updates

Customers that received SGI Performance Suite software pre-installed on their new SGI system may run one of the following commands as **root user** to add the proper update source. Each product has a unique command.

For SGI MPI run:

```
% /usr/sbin/sgi-mpi-update-source
```

For SGI REACT run:

```
% /usr/sbin/sgi-react-update-source
```

For SGI UPC, run:

```
% /usr/sbin/sgi-upc-update-source
```

For SGI Accelerate run:

```
% /usr/sbin/sgi-accelerate-update-source
```

Each script will prompt you for your Supportfolio login and then add the proper SGI update source for the given architecture. These commands only apply to systems running RHEL 6 or RHEL 6 .1 as the base operating system.

For information on where and how to get updates, see the release notes that are in a file with the product name and `-readme.txt` suffix that is available in `/docs` directory on the CD media for each product.

Detailed instructions for downloading and installing SGI Foundation Software 2.5 updates are available from the SGI Foundation Software product pages on Supportfolio:[https://support.sgi.com/content\\_request/786853/index.html](https://support.sgi.com/content_request/786853/index.html)

## SGI Foundation 2.5 Updates

---

**Note:** See the *SGI Foundation Software 2.5 Start Here* for detailed information on SGI Foundation Software 2.5 release. This section has been included here for your convenience.

---

Customers that received SGI Foundation Software 2.5 pre-installed on their new SGI system may run the following command as **root user** to add the proper update source for SGI Foundation Software 2.5:

```
$ /usr/sbin/sgi-foundation-update-source
```

This script will prompt you for your Supportfolio login and then add the proper SGI update source for the given architecture.

Detailed instructions for downloading and installing SGI Foundation Software 2.5 updates are available from the SGI Foundation Software product pages on Supportfolio:[https://support.sgi.com/content\\_request/194480/index.html](https://support.sgi.com/content_request/194480/index.html).

## SGI Open Source Information

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 Open SpeedShop) from the following website:

<http://sourceforge.net>

## Documentation Updated for This Release

The following manuals have been updated for this release:

- *Unified Parallel C (UPC) User's Guide, 007-5604-003*
  - Updated "UPC Quick Start on SGI Altix UV or SGI Altix ICE Systems" in Chapter 2.
  - Added `UPC_CAUTIOUS_STRICT`, `UPC_GRU_DOMAIN_SIZE`, and `UPC_IB_BUFFER_SIZE` variable descriptions in "UPC Runtime Library Environment Variables" in Chapter 2.
- *Linux Application Tuning Guide for SGI X86-64 Based Systems, 007-5646-003*
  - Updated compiler information in "Compiler Overview" in Chapter 2.
  - Updated compiler information in "Other Compiling Environment Features" in Chapter 2.
  - Updated VTune information in "Using VTune for Remote Sampling" in Chapter 3.
  - Added information about the MPInside Reference Guide in "MPInside Profiling Tool" in Chapter 6
- *REACT Real-Time for Linux Programmer's Guide, 007-4746-016*
  - Added support for the `-c` and `-I` options, which let you invoke a subshell or specified command line on the specified real-time CPU. See "react Command-Line Summary" and "Executing Commands on a Real-Time CPU" in Chapter 9, respectively.
  - Added a new `cpu_sysrt_move` library routine to assign the specified process so that it will run only on the specified CPU, which can optionally be real-time or not. See "cpu\_sysrt\_move" in Chapter 10.
  - Added support for the Linux Trace Tool Next Generation (LTTng) and described in Chapter 11, "Linux Trace Toolkit Next Generation" Note that SGI Linux Trace is deprecated and will be removed in a future release.

- Added a new "libreact API Example" in Appendix A
- Removed the `set_affinity` code example, whose functionality has been replaced by the `cpu_sysrt_move` library routine.
- *Message Passing Toolkit (MPT) User's Guide, 007-3773-019*
  - Updated information about SHMEM™ programming model in Chapter 1, "Introduction".
  - Updated release information in Chapter 2, "Administrating MPT".
  - Added information about using huge pages to optimize the performance of MPI applications in "Using Huge Pages" in Chapter 3. Corrected URL to Berkeley Lab Checkpoint/Restart (BLCR) site in Chapter 7, "Checkpoint/Restart".
- *Linux Resource Administration Guide, 007-4413-016*
  - Added "Configuring Array Services for Use with MPI" in Chapter 1.
  - Updated information in "Boot Cpuset" in Chapter 2.
  - Updated "Configuring a User Cpuset for Interactive Sessions" in Chapter 2.
- *SGI Altix UV GRU Development Kit Programmer's Guide, 007-5668-004*
  - Added additional information about the global reference unit (GRU) in "Introduction" in Chapter 1.
  - Updated "Advantages Provided by Directly Programming the GRU" in Chapter 1.
  - Updated "GRU Access Functions" in Chapter 1.
- *SGI Simple Linux Utility for Resource Management Install Guide , 007-5718-001*

New for this release, documents the SGI implementation of the Simple Linux Utility for Resource Management (SLURM) and how to install and configure it.



## Software Installation

Your SGI system comes with a base Linux distribution SUSE Linux Enterprise Server 11 Service Pack 1 (SLES 11 SP1) or Red Hat Enterprise Linux 6.1 and SGI Performance Suite 1.3 software preinstalled.

Installation instructions for SGI Performance Suite 1.3 software are contained in the product release notes in a file with the product name and `-readme.txt` appended to it that is available in `/docs` directory on the CD media. For example, for the SGI Accelerate 1.3 product this file is `SGI-Accelerate-1.3-readme.txt`.

### Systems with Multiple Ethernet Configurations

When doing a network installation on a system with multiple Ethernet interfaces, there is no guarantee about initial interface names. The first interface discovered will be named "eth0", and so on. This can make network installations confusing. If this is a problem, a workaround is to remove the Ethernet cards that are **not** being used to install software over the network. Once the installation is complete, the Ethernet cards can be reinstalled. A persistent naming mechanism ensures that an interface, once named, retains its name across system reboots.

### XVM Volumes Access in Rescue Mode

You cannot access XVM Volumes in Rescue Mode. It is possible to boot into rescue mode from the operating system DVD, but that DVD does not contain the `xvm` kernel module. This makes it impossible to perform maintenance on XVM volumes from rescue mode at this time. SGI recommends that you do not migrate any system directories to XVM since there is no way to access them from rescue mode.

## Adding the `xfs_quota` Variable to the `initrd` Boot Image

You need to add `xfs_quota` to the `INITRD_MODULES` variable in the `/etc/sysconfig/kernel` file when `rootflags=quota` is set or the `initrd` boot loader will panic your system at boot time. This technique is SLES specific.

## Superpages

The superpages mechanism on SGI Altix UV systems allows you to map large areas of reserved memory into your application. `libsuperpages` is a library for allocating and mapping Altix UV superpages (see the `libsuperpages(3)` man page).

Superpages are defined in an EFI firmware variable, as a kernel boot option that can only be adjusted by the superuser. When enabled, the Altix UV will reserve at boot time the amount of specified memory to be used via superpages, so that it is not available to the operating system for its regular activities. The UV BIOS passes superpages addresses to the superpages driver as an Advanced Configuration and Power Interface (ACPI) device.

The `superpages_define -w` command creates the EFI variable to define superpages. You must be a superuser (**root user**) to execute this command. After execution, to enable superpages, you need to reboot the system. The system BIOS then reserves the system memory for superpages rather than providing it to the operating system. The `superpages_define` command is described in more detail in the examples that follow.

The `superpages` command allows you to test the functionality of an Altix UV system's superpages. For a usage statement, enter `superpages -h`.

The `superpages_define` command is used to configure superpages. For a usage statement, enter `superpages_define -h`.

You can use the `superpages_define -L` option to display valid superpage sizes, as follows:

```
uv44-sys:~ # superpages_define -L
Valid superpage sizes:
[1] 0x200000 2M (2097152)
[2] 0x400000 4M (4194304)
[3] 0x1000000 16M (16777216)
[4] 0x2000000 32M (33554432)
[5] 0x8000000 128M (134217728)
```



---

```
[6] 0x20000000 512M (536870912)
[7] 0x40000000 1G (1073741824)
[8] 0x100000000 4G (4294967296)
[9] 0x400000000 16G (17179869184)
[10] 0x1000000000 64G (68719476736)
[11] 0x4000000000 256G (274877906944)
[12] 0x10000000000 1T (1099511627776)
```

You can use the `superpages_define -c` option to see the available memory ranges for superpage allocation, as follows:

```
uv44-sys:~ # superpages_define -c
available memory ranges for superpage allocation:
0x100000000-0x87fffffff 0x780000000 30G
0x1000000000-0x17feffffff      0x7ff000000 31G
available memory: 61G (63472M)
```

To show what four Gigabyte (4G) areas of memory could be used as superpages, perform the following:

```
uv44-sys:~ # superpages_define -p 4G

in /etc/superpages.conf format:
# for pagesize: 0x100000000 4G
size_4G 4G #node0 0x100000000:3 #node1 0x400000000:4 #node2
0x1000000000:4 #node          3 0x1400000000:3
# 14 4G superpages
```

To show what 4G areas of memory could be used as superpages, but will limit such areas to two per node, perform the following:

```
uv44-sys:~ # superpages_define -p 4G -n 2

in /etc/superpages.conf format:
# for pagesize: 0x100000000 4G
size_4G 4G #node0 0x100000000:2 #node1 0x400000000:2 #node2
0x1000000000:2 #node3 0x1400000000:2
# 8 4G superpages
```

Once you are satisfied with your memory selection, perform commands similar to the following to define superpages:

```
uv44-sys:~ # superpages_define -p 4G -n 2 > /etc/superpages.conf
uv44-sys:~ # superpages_define -w
```

When you reboot your Altix UV system, those chosen areas will be superpages.

---

**Note:** You can use huge pages to optimize the performance of your MPI application. For more information, see “Using Huge Pages” section in Chapter 3, “Getting Started” in the *Message Passing Toolkit (MPT) User’s Guide*, PN 007-3773-019.

---

## SLES 11 SP1 or RHEL 6.1 Kernels

For x86\_64 systems, SLES 11 SP1 contains `kernel-default` and `kernel-xen`. The `kernel-default` variant provides multi-processor support.

For x86\_64 systems, RHEL 6 or RHEL 6.1 has one kernel image, `kernel` and `kvm` is the only supported virtualization.

## Getting the Latest Firmware/Software for Your SGI System

For information on the latest firmware for your SGI system, see the *SGI Foundation Software 2.5 Start Here* available at <http://docs.sgi.com>.

## License Manager Updates

---

**Important:** Please backup your license key files prior to installing SLES11 SP1, RHEL 6.1, and/or SGI Performance Suite 1.3 software. A license key file is **not** included in these RPMs.

---

Your original license key files will need to be restored after you complete your installation.

In terms of software licensing support, the operating system distributors do not ship any licensing code. Independent software vendors (ISVs) are responsible for obtaining and including software licensing components into their products. SGI products ship with SGI licensing software.

---

For information on the SGI software licensing mechanism called LK, see the *SGI Foundation Software 2.5 Start Here* available at <http://docs.sgi.com>.

## **Upgrading from an Earlier Release**

Upgrade instructions for SGI Performance Suite 1.3 software are contained in the product release notes for each product in a file with the product name and `-readme.txt` suffix that is available in `/docs` directory on the CD media.



## **SGI Performance Suite Products Overview**

This chapter documents the product components that are supported on the SGI computer systems. (For a list of the products, see Table 3-1 on page 22.)

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

- “SGI Performance Suite 1.3 Products” on page 21
- “Performance Suite Products” on page 24
- “I/O Subsystems” on page 25
- “HPC Application Tools and Support” on page 26
- “NUMA Data Placement Tools” on page 27

### **SGI Performance Suite 1.3 Products**

Software provided by SGI for the SGI Performance Suite 1.3 release consists of kernel modules for SGI software built against the kernels in SUSE Linux Enterprise Server 11 SP1 or Red Hat Enterprise Server 6.1 and value-add software developed by SGI. For

information on how these products are bundled, see “SGI Performance Suite Software Bundles” on page 2.

**Table 3-1** SGI Performance Suite 1.3 Products

Product	Description
Array Services	<p>Array Services includes administrator commands, libraries, daemons, and kernel extensions that support the execution of parallel applications across a number of hosts in a cluster, or array. The Message Passing Interface (MPI) uses Array Services to launch parallel applications. For information on MPI, see the Message Passing Toolkit (MPT) User’s Guide</p> <p>The secure version of Array Services is built to make use of secure sockets layer (SSL) and secure shell (SSH).</p> <p>For more information on standard Array Services or Secure Array Services (SAS), see the Array Services chapter in the <i>Linux Resource Administration Guide</i>.</p>
Cpuset System	<p>The Cpuset System is primarily a workload manager tool permitting a system administrator to restrict the number of processors and memory resources that a process or set of processes may use. A system administrator can use cpusets to create a division of CPUs and memory resources within a larger system. For more information, see the “Cpusets” chapter in the <i>Linux Resource Administration Guide</i>.</p>
IOC4 serial driver	<p>Driver that supports the Internal IDE CD-ROM, NVRAM, and Real-Time Clock.</p> <p>Serial ports are supported on the IOC4 base I/O chipset and the following device nodes are created:</p> <pre>/dev/ttyIOC4/0 /dev/ttyIOC4/1 /dev/ttyIOC4/2 /dev/ttyIOC4/3</pre>
Kernel partitioning support	<p>Provides the software infrastructure necessary to support a partitioned system, including cross-partition communication support. Partitioning is supported on SGI Altix UV 100 and UV 1000 systems only. For more information on system partitioning, see the <i>SGI Altix UV Linux Configuration and Operations Guide</i>.</p>

**Table 3-1** SGI Performance Suite 1.3 Products (continued)

Product	Description
MPT	Provides industry-standard message passing libraries optimized for SGI computers. For more information on MPT, see the <i>Message Passing Toolkit (MPT) User's Guide</i> .
NUMA tools	Provides a collection of NUMA related tools ( <code>dllook(1)</code> , <code>dplace(1)</code> , and so on). For more information on NUMA tools, see the <i>Linux Application Tuning Guide for SGI X86-64 Based Systems</i> .
Performance Co-Pilot collector infrastructure	Provides performance monitoring and performance management services targeted at large, complex systems.
REACT real-time for Linux	Support for real-time programs. For more information, see the <i>REACT Real-Time for Linux Programmer's Guide</i> .
Utilities	<code>udev_xscsi</code> , is a <code>udev</code> helper for doing XSCSI device names. <code>sgtools</code> , a set of tools for SCSI disks using the Linux SG driver and <code>lsiutil</code> , the LSI Fusion-MPT host adapter management utility.
XVM	Provides software volume manager functionality such as disk striping and mirroring. For more information on XVM, see the <i>XVM Volume Manager Administrator's Guide</i> .

SGI does not support the following:

- Base Linux software not released by Novell for SLES 11 SP1 or by Red Hat for RHEL 6.1 or other software not released by SGI.
- SGI does not provide technical support for the CentOS 6 operating system.
- Other releases, updates, or patches not released by Novell for SLES 11 SP1, Red Hat for RHEL 6.1 or by SGI for SGI Performance Suite software.
- Software patches, drivers, or other changes obtained from the Linux community or vendors other than Novell, Red Hat, and SGI.
- Kernels recompiled or reconfigured to run with parameter settings or other modules as not specified by Novell or Red Hat and SGI.
- Unsupported hardware configurations and devices.

## Performance Suite Products

SGI Performance Suite provides application acceleration components for software developers and end-users. SGI Accelerate, SGI MPI, SGI REACT, and SGI UPC contain libraries and tools that enable software developers to develop, profile, and tune applications for faster performance. End users benefit from running their applications with the runtime acceleration tools supplied in SGI Accelerate and SGI MPI. This section describes some key components.

### Cpuset Support

The cpuset facility is primarily a workload manager tool permitting a system administrator to restrict the number of processors and memory resources that a process or set of processes may use. A cpuset defines a list of CPUs and memory nodes. A process contained in a cpuset may only execute on the CPUs in that cpuset and may only allocate memory on the memory nodes in that cpuset. Essentially, cpusets provide you with a CPU and memory containers or "soft partitions" within which you can run sets of related tasks. Using cpusets on an SGI system improves cache locality and memory access times and can substantially improve an applications performance and runtime repeatability. Restraining all other jobs from using any of the CPUs or memory resources assigned to a critical job minimizes interference from other jobs on the system. For example, Message Passing Interface (MPI) jobs frequently consist of a number of threads that communicate using message passing interfaces. All threads need to be executing at the same time. If a single thread loses a CPU, all threads stop making forward progress and spin at a barrier. Cpusets can eliminate the need for a gang scheduler.

Cpusets are represented in a hierarchical virtual file system. Cpusets can be nested and they have file-like permissions.

In addition to their traditional use to control the placement of jobs on the CPUs and memory nodes of a system, cpusets also provide a convenient mechanism to control the use of Hyper-Threading Technology.

For detailed information on cpusets, see Chapter 6, "Cpusets on Linux" in the *Linux Resource Administration Guide*. Information about cpusets is also available in Chapter 5, "Data Placement Tools" in the *Linux Application Tuning Guide for SGI X86-64 Based Systems*.



## Partitioning

SGI provides the ability to divide a single SGI Altix UV 100 or Altix UV 1000 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 SGI systems to provide high-performance clusters.

For further documentation and details on partitioning, see the *SGI Altix UV Systems Linux Configuration and Operations Guide*.

## 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 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. This section covers the following topics:

- “XFS Filesystem” on page 25
- “XVM Volume Manager” on page 26

## 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.

For more information, see the *XVM Volume Manager Administrator's Guide*.

## **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 Xeon-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 memory mapping within and between partitions for fast memory-to-memory transfers and the hardware memory controller's fetch operation (fetchop) support. Fetchops and other shared memory techniques enable ultra fast communication and synchronization between MPI processes in a parallel application.

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 (MPT) User's Guide*.

## 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.

## NUMA Data Placement Tools

This section describes the commands that are currently provided with the collection of NUMA related data placement tools that can help you with tuning applications on your system.

---

**Note:** Performance tuning information for single processor and multiprocessor programs resides in *Linux Application Tuning Guide for SGI X86-64 Based Systems*.

---

### dlook Command

The `dlook(1)` 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(1)` 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.

## **taskset Command**

The `taskset(1)` command is used to set or retrieve the CPU affinity of a running process given its PID or to launch a new command with a given CPU affinity. CPU affinity is a scheduler property that "bonds" a process to a given set of CPUs on the system. The Linux scheduler will honor the given CPU affinity and the process will not run on any other CPUs. Note that the Linux scheduler also supports natural CPU affinity; the scheduler attempts to keep processes on the same CPU as long as practical for performance reasons. Therefore, forcing a specific CPU affinity is useful only in certain applications.

For more information on NUMA tools, see Chapter 5, "Data Placement Tools" in the *Linux Application Tuning Guide for SGI X86-64 Based Systems*.

---

# Index

## A

Array Services, description, 22

## C

CD contents, 2

Cpuset system, 24

## D

dlook command, 27

dplace command, 28

## G

Getting software updates, 9

## H

hardware platforms, 2

HPC support

libraries and tools, 26

Message Passing Toolkit, 26

Performance Co-Pilot, 27

## I

I/O subsystems

for HPC systems, 25

XFS filesystem support, 25

XVM support, 26

## L

License manager updates, 18

## M

Message Passing Toolkit  
support, 26

## N

NUMA tools, 27

dlook command, 27

dplace command, 28

## O

Online Support Services  
Supportfolio, xii

Open source website, 11

OS configuration settings, 2

OS enhancements  
  cpuset support, 24  
  for HPC environments, 24  
  partitioning, 25

## P

Performance Co-Pilot support, 27  
Persistent naming  
  XSCSI devices, 25  
Product list, 21  
Product support, 21

## R

Release features, 1  
RHEL5 updates, 10

## S

SGI Foundation 2 updates, 11  
SLES11 updates, 9  
Software  
  installation, 15  
  introduction, 1  
Software license information, 6  
Supportfolio, xii

## T

taskset command, 28

## U

Unsupported elements, 23

## X

XFS filesystem support, 25  
XVM support, 26  
XVM Volumes in Rescue Mode, 15