

SGI® Performance Suite 1.0 Start Here

007-5680-001

COPYRIGHT

© 2010, 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

Altix, IRIX, NUMAflex, NUMAlink, OpenGL, Performance Co-Pilot, 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.

FLEXIm 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 several countries. MIPS is a registered trademark of MIPS Technology, Inc. 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

Version	Description
001	October 2010
	Original publication. Supports the SGI Performance Suite 1.0 release.

Contents

Record of Revision	 			•	. iii
Tables	 				. vii
Helpful Online Resources x
Related Publications	 				. xii
Novell Documentation	 				. xii
Red Hat Documentation	 				. xiii
Intel Compiler Documentation					
Other Intel Documentation	 				. xiv
SGI Documentation	 				. xiv
SGI Performance Suite Manuals.	 				. xiv
Additional SGI Software Manuals					. xv
SGI Hardware Manuals	 				xvii
Obtaining SGI Publications					. xxi
Conventions					
Reader Comments					
Release Features	 				. 1
Software Introduction.	 				. 1
SGI Performance Suite Software Bundles	 				. 2
SGI® MPI	 				. 2
$SGI^{\mathbb{R}} REACT^{TM} $. 3
SGI® UPC					
$SGI^{\mathbb{R}}$ Accelerate TM					
Software Package License Information					
SGI Performance Suite 1.0 CDs					
Release Information, Documentation, and Release No					C

1.

	Getting Updates
	SLES 11 SP1 Updates
	RHEL 6 Updates
	SGI Foundation 2.2 Updates
	SGI Open Source Information.
	Documentation Updated for This Release
2.	Software Installation
	Systems with Multiple Ethernet Configurations
	XVM Volumes Access in Rescue Mode
	Adding the xfs_quota Variable to the initrd Boot Image
	SLES 11 SP1 or RHEL 6 Kernels
	Getting the Latest Firmware/Software for Your SGI System
	License Manager Updates
	Upgrading from an Earlier SLES Base and SGI ProPack Release
3.	SGI Peformance Suite Products Overview.
	SGI Performance Suite 1.0 Products
	Operating System Enhancements.
	Cpuset Support
	Partitioning.
	I/O Subsystems.
	Persistent IP Addressing of Ethernet Interfaces
	XFS Filesystem
	XVM Volume Manager
	HPC Application Tools and Support
	Message Passing Toolkit
	Performance Co-Pilot
	NUMA Data Placement Tools.
	dlook Command
	dplace Command
	taskset Command
	Index

Tables

About This Guide

Note: SGI Performance Suite 1.0 software requires the SGI Foundation Software 2.2 release. For more information, see the *SGI Foundation Software 2.2 Start Here* available at http://docs.sgi.com.

This guide provides information about the SGI Performance Suite 1.0 release. It is divided into the following chapters:

- Chapter 1, "Release Features," describes the major features of this release, the CD contents, SGI Open Source software, SGI system controller software, 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 Peformance Suite Products Overview," documents the product components that are supported on SGI Altix 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 on the SGI Performance Suite CD.

This Preface covers the following topics:

- "Helpful Online Resources" on page x
- "Novell Documentation" on page xii
- "Intel Compiler Documentation" on page xiii
- "Other Intel Documentation" on page xiv
- "SGI Documentation" on page xiv
- "Obtaining SGI Publications" on page xxi

Helpful Online Resources

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

• Supportfolio

Supportfolio, the SGI support web site, including the SGI Knowledgebase, SGI YaST Online Update (YOU) server, SUSE YaST Online Server, links for software supports and updates, as described in this listing.

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 (e.g., 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, 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 xii
- "Red Hat Documentation" on page xiii
- "Intel Compiler Documentation" on page xiii
- "Other Intel Documentation" on page xiv
- "SGI Documentation" on page xiv

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 (installquick.pdf)
- SUSE Linux Enterprise Server 11 Deployment Guide (deployment.pdf)
- SUSE Linux Enterprise Server 11 Administration Guide (admin.pdf)
- SUSE Linux Enterprise Server SLES11: Storage Administration Guide (storage.pdf)
- SUSE Linux Enterprise Server 11 Package Descriptions available online at: http://www.novell.com/products/linuxenterpriseserver/packages.html

Red Hat Documentation

Documentation for RHEL6 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 (RHEL6) 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
- Release Notes *ia64* 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

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

 Intel Xeon processor documentation is available 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 xiv
- "Additional SGI Software Manuals" on page xv
- "SGI Hardware Manuals" on page xvii

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.
- *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 Altix 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 Provides information about tuning application programs on SGI Altix 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 perfrom system configuration and operations under the Linux operating system for SGI Altix UV 100 and SGI UV 1000 servies systems.

Additional SGI Software Manuals

- SGI Foundation Software 2.2 Start Here Provides information about the SGI Foundation Software 2.2 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 Tempo System Administrator's Guide* Describes how to perform general system installation, configuration, and operations on the SGI Altix ICE series systems.
- 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.
- SGI Management Center System Administrator's Guide

This guides 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.

- *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 5 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 5 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 Guide* Provides information about GRIO version 2, the second-generation guaranteed-rate I/O product from SGI.
- SGI InfiniteStorage High Availability Using Linux-HA Heartbeat Provides information about the high-availability (HA) resource agents provided for the Linux-HA Heartbeat product in order to support SGI products
- DMF 4 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.
- *Platform Manager On SGI Altix ICE Systems Quick Reference Guide* Describes how to use Platform Manager management software to perform general system discovery, installation, configuration, and operations on SGI Altix ICE 8200 series systems.

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 XE210 System User's Guide* Describes the features of the Altix XE210 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Altix XE240 System User's Guide* Describes the features of the Altix XE240 server system as well as troubleshooting, upgrading, and repairing.
- SGI Altix XE250 System User's Guide

Describes the features of the Altix XE250 server system as well as troubleshooting, upgrading, and repairing.

- *SGI Altix XE270 System User's Guide* Describes the features of the Altix XE270 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Altix XE310 System User's Guide* Describes the features of the Altix XE310 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Altix XE320 System User's Guide* Describes the features of the Altix XE320 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Altix XE340 System User's Guide* Describes the features of the Altix XE340 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Altix XE500 System User's Guide* Describes the features of the Altix XE500 server system as well as troubleshooting, upgrading, and repairing.
- *SGI Altix XE Clusters Quick Reference Guide* Provides an overview of some of the common operations that system administrators may have to perform to set-up, boot, re-configure (upgrade) or troubleshoot an SGI Altix XE cluster.
- *SGI Altix 330 System User's Guide* Provides an overview of the Altix 330 system components, and it describes how to set up and operate this system
- *SGI Altix 350 System User's Guide* Provides an overview of the Altix 350 system components, and it describes how to set up and operate this system.
- *SGI Altix 350 Quick Start Guide* Guides a knowledgeable user through the installation, setup, and simple configuration of most SGI Altix 350 systems.
- 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 Altix 3700Bx2 User's Guide

•

This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Altix 3700Bx2 family of servers. 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 450 System User's Guide* Provides an overview of the Altix 450 system components, and it describes how to set up and operate this system.
- SGI Altix 4700 User's Guide This guide provides an overview of the architecture and descriptions of the major components that compose the SGI Altix 4700 family of servers. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.
- 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.
- 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.

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:

Convention	Meaning
command	This fixed-space font denotes literal items such as commands, files, routines, path names, signals, messages, and programming language structures.
variable	Italic typeface denotes variable entries and words or concepts being defined.
user input	This bold, fixed-space font denotes literal items that the user enters in interactive sessions. (Output is shown in nonbold, fixed-space font.)
manpage(x)	Man page section identifiers appear in parentheses after man page names.
GUI element	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 1.0 release and describes the software bundle's CD contents and provides documentation information. It covers the following topics:

- "Software Introduction" on page 1
- "SGI Performance Suite Software Bundles" on page 2
- "Getting Updates" on page 6
- "SGI Open Source Information" on page 9
- "Documentation Updated for This Release" on page 9

Note: SGI Performance Suite 1.0 requires the SGI Foundation Software 2.2 release. For more information, see the *SGI Foundation Software 2.2 Start Here* available at http://docs.sgi.com.

Software Introduction

The SGI Performance Suite 1.0 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 Intel processors. This product adds to or enhances features in the SUSE Linux Enterprise Server 11 Service Pack 1 (SLES 11 SP1) base distribution or Red Hat Enterprise Server 6 (RHEL 6) base distribution.

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 1.0 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
- 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

• Unified Parallel C (UPC)

SGI implementation of the Unified Parallel C (UPC) parallel extension to the C programming language standard.

memacct

Memory accounting utility (see the memacct(3) man page)

SHMEM

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.

Unified Parallel C (UPC) User's Guide.

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 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 pthreads 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 hreads 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
- 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
- libgru

SGI Altix UV global reference unit (GRU) libraries

- superpages
- 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

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

SGI[®] Accelerate[™]

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

- cpusets
- numatools
- 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 (hub performance counters)

SGI Performance Co-Pilot Altix UV hub performance counters

- libgru development kit
- SGISolve

The PSLDLT libraries under the SGISOLVE umbrella contains 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

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 1.0 CDs

The following CDs ship with the SGI Performance Suite 1.0 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 Peformance Suite, then you will receive all four CDs.

Release Information, Documentation, and Release Notes

For a complete list of RPMs included in the SGI Performance Suite 1.0 release, see the file called RPMS.txt that is available in the /docs directory on the CD media.

For the latest information about software and documentation in this release, see the release notes that are in a file named README.TXT that is available in /docs directory on the CD media.

Release notes for each of the SGI Performance Suite products are available in the /docs directory of the associated media.

Getting Updates

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

- "SLES 11 SP1 Updates" on page 7
- "RHEL 6 Updates" on page 8

"SGI Foundation 2.2 Updates" on page 8

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.

Customers that received SGI Performance Suite software pre-installed on their new SGI Altix 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.

RHEL 6 Updates

Software updates for RHEL 6 are provided by Red Hat. 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 also use the Package Update (pup), which is a graphical tool for locating and installing updates, as follows:

% pup

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 Foundation 2.2 Updates

Note: See the *SGI Foundation Software 2.2 Start Here* for detailed information on SGI Foundation Software 2.2 release. This section has been included here for your convenience.

Customers that received SGI Foundation Software 2.2 pre-installed on their new SGI Altix system may run the following command as **root user** to add the proper update source for SGI Foundation Software 2.2:

\$ /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.

For additional information on how to configure YaST Online Update to download updates for SGI Foundation 2.2 login to Supportfolio at http://support.sgi.com. Click on the **Browse Collections** link on the left. Click on **SGI ProPack** "SGI ProPack patch downloads and descriptions" on the right. Click on **Supported Major Releases / Service Packs**. Click on **SGI® Foundation Software 2.2** on the right. The **SGI® Foundation Software 2.2** page appears.

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 LKCD) from the following website:

http://sourceforge.net

Documentation Updated for This Release

The following manuals are new or have been updated for this release:

- REACT Real-Time for Linux Programmer's Guide, 007-4746-014
 - Support for SGI REACT 1.0 (a new separate release, and a member of the SGI Performance Suite)
 - Support for running REACT on third-party x86-64 systems. For best performance, run REACT on SGI x86-64 servers. See "REACT Requirements" in Chapter 1.
 - Support for the following new API routines to change the CPU affinity of the given IRQs and change permissions so that non-root users can run REACT: "cpu_sysrt_irq" and "cpu_sysrt_perm" in Chapter 10,
- SGI Tempo System Administrator's Guide, 007-4993-014

Updated for the SGI Tempo v2.2 release.

- Added "Installing RHEL 6 on the Admin Node" in Chapter 2.

- Added information about the discover command generic hardware type used for Ethernet switches in large SGI Altix ICE 8400 configurations in "discover Command" in Chapter 2.
- Added "Configure the Backup DNS Server" in Chapter 2.
- Added information about how to disable InfiniBand switch monitoring in "InfiniBand Configuration" in Chapter 2.
- Updated the information about setting the re_partition_with_slots parameter in "Admin Node Installation Choices Related to Cascading Dual-Boot " in Chapter 2.
- Updated information about rmplists in "crepo Command" in Chapter 3.
- Updated "Using a Custom Repository for Site Packages" in Chapter 3.
- Added "Restoring the grub Boot-loader on a Node" in Chapter 3.
- Added "Configuring the Lightweight Fabric" in Chapter 4.
- SGI Altix UV CMC Controller Software User's Guide, 007-5636-002
 - Update power commands information.
 - Added new instructions for flashing firmware
- SGI Altix UV System Management Node Administrator's Guide, 007-5694-001

New for this release, this manual describes the system management node (SMN) for SGI Altix UV 1000 and SGI Altix UV 100 series systems.

- SGI Altix UV Systems Linux Configuration and Operations Guide, 007-5629-002
 - Updated power command information.
 - Updated installing new firmware information.
- SGI Altix UV GRU Development Kit Programmer's Guide, 007-5668-003
 - Added "Advantages Provided by Directly Programming the GRU" in Chapter 1.
 - Updated "GRU Library Program Example" and added instructions on how to compile and run it in Chapter 1.
- Linux Resource Administration Guide, 007-4413-015
 - Removed Linux kernel jobs chapter.
 - Removed Comprehensive System Accounting (CSA) chapter.

Software Installation

Note: SGI Performance Suite 1.0 software **requires** the SGI Foundation Software 2.2 release. For more information, see the *SGI Foundation Software 2.2 Start Here* available at http://docs.sgi.com.

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 (RHEL 6) and SGI Performance Suite 1.0 software preinstalled.

Installation instructions for SGI Performance Suite 1.0 software are contained in the product release notes in a file named README.TXT that is available in /docs directory on the CD media.

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 SLES 11 SP1 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.

SLES 11 SP1 or RHEL 6 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 has two kernel images, kernel and kernel-xen.

Getting the Latest Firmware/Software for Your SGI System

For information on the latest firmware for your SGI system, see the *SGI Foundation Software 2.2 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, and/or SGI Performance Suite 1.0 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 2.2 Start Here available at http://docs.sgi.com.

Upgrading from an Earlier SLES Base and SGI ProPack Release

Upgrade instructions for SGI Performance Suite 1.0 software are contained in the product release notes in a file named README.TXT that is available in /docs directory on the CD media.

SGI Peformance 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 16.)

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

- "SGI Performance Suite 1.0 Products" on page 15
- "Operating System Enhancements" on page 18
- "I/O Subsystems" on page 19
- "HPC Application Tools and Support" on page 20
- "NUMA Data Placement Tools" on page 21

SGI Performance Suite 1.0 Products

Software provided by SGI for the SGI Performance Suite 1.0 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 (RHEL 6) and value-add software developed by SGI.

For information on how these products are bundle, see "SGI Performance Suite Software Bundles" on page 2.

Product	Description
Array Services	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
	The secure version of Array Services is built to make use of secure sockets layer (SSL) and secure shell (SSH).
	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> .
Cpuset System	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> .
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
Kernel partitioning support	Provides the software infrastructure necessary to support a partitioned system, including cross-partition communication support. For more information on system partitioning, see the SGI Altix UV Linux Configuration and Operations Guide.

 Table 3-1
 SGI Performance Suite 1.0 Products

Product	Description
МРТ	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 (dlook(1), dplace(1), and so on). For more information on NUMA tools, see the <i>Linux Application Tuning Guide</i> .
Performance Co-Pilot collector infrastructure	Provides performance monitoring and performance management services targeted at large, complex systems. For more information on Performance Co-Pilot, see <i>Performance Co-Pilot for IA-64 Linux User's and</i> <i>Administrator's Guide.</i>
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	udev_xsci, is a udev helper for doing XSCSI device names. sgtools, a set of tools for SCSI disks using the Linux SG driver and lsiutil, 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 XVM Volume Manager Administrator's Guide.

 Table 3-1
 SGI Performance Suite 1.0 Products (continued)

SGI does not support the following:

- Base Linux software not released by Novell for SLES 11 SP1 or by Red Hat for RHEL 6 or other software not released by SGI.
- Other releases, updates, or patches not released by Novell for SLES 11 SP1, Red Hat for RHEL 6 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.

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.

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 Altix 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 SGI ProPack 6 for Linux" in the *Linux Resource Administration Guide*.

Partitioning

SGI provides the ability to divide a single SGI Altix 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 Altix to provide high-performance clusters.

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

- "Persistent IP Addressing of Ethernet Interfaces" on page 19
- "XFS Filesystem" on page 20
- "XVM Volume Manager" on page 20

Persistent IP Addressing of Ethernet Interfaces

An Ethernet interface can be given a persistent internet addresses by associating its permanent MAC address, such as 08:00:69:13:f1:aa, with an internet protocol (IP) address, for example 192.168.20.1. An interface with a persistent IP address will be given the same IP address each time the system is booted. For more information, see "Persistent Network Interface Names" in the *Linux Configuration and Operations Guide*.

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.

This release adds a new XVM multi-host failover feature. For more information on this new feature and XVM Volume Manager in general, 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.

For more information on Performance Co-Pilot, see the *Performance Co-Pilot for IA-64 Linux User's and Administrator's Guide.*

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*.

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*.

Index

Α

Array Services, description, 16

С

CD contents, 2 Cpuset system, 18

D

dlook command, 22 dplace command, 22

G

Getting software updates, 6

Н

hardware platforms, 1 HPC support libraries and tools, 20 Message Passing Toolkit, 21 Performance Co-Pilot, 21 I

I/O subsystems for HPC systems, 19 XFS filesystem support, 20 XVM support, 20

L

License manager updates, 12

Μ

Message Passing Toolkit support, 21

Ν

NUMA tools, 21 dllook command, 22 dplace command, 22

Ο

Online Support Services Supportfolio, x Open source website, 9 OS configuration settings, 1 OS enhancements cpuset support, 18 for HPC environments, 18 partitioning, 19

Ρ

Performance Co-Pilot support, 21 Persistent naming Ethernet devices, 19 XSCSI devices, 20 Product list, 15 Product support, 15

R

Release features, 1 RHEL5 updates, 8

S

SGI Foundation 2 updates, 8 SLES11 updates, 7 Software installation, 11 introduction, 1 Software license information, 5 Supportfolio, x

Т

taskset command, 22

U

Unsupported elements, 17

Х

XFS filesystem support, 20 XVM support, 20 XVM Volumes in Rescue Mode, 11