SGI™ ProPack 1.3 for Linux™ Start Here

Document Number 007-4062-005

SGI™ ProPack 1.3 for Linux™ Start Here
Document Number 007-4062-005

Contents

List of Tables

About This Guide

This document provides information about the SGI ProPack 1.3 for Linux release. It is divided into three chapters:

- Chapter 1, “Release Features,” describes the major features of this release.
- Chapter 2, “Software Overview,” describes the major features of earlier releases that are included in this release.
- Chapter 3, “Quick Configuration and Installation Instructions,” describes the procedures for configuring and installing the SGI ProPack 1.3 for Linux software in various circumstances, including upgrading your system from an earlier SGI Linux Environment release.

The SGI ProPack 1.3 for Linux is an overlay product that adds or enhances features in Linux base distributions from Red Hat (version 6.2), SuSE (version 6.4), or TurboLinux (6.0). Guides for installing and getting started with a base Linux distribution are included with the base Linux distribution release kit.

Reader Comments

If you have comments about the technical accuracy, content, or organization of this document, please tell us. Be sure to include the title and document number of the manual with your comments. (Online, the document number is located in the front matter of the manual. In printed manuals, the document number can be found on the back cover.)

You can contact us in any of the following ways:

- Send e-mail to the following address:

  techpubs@sgi.com

- Use the Feedback option on the Technical Publications Library World Wide Web page:

  http://techpubs.sgi.com

- Contact your customer service representative and ask that an incident be filled in the SGI incident tracking system.
- Send mail to the following address:

  Technical Publications
  SGI
  1600 Amphitheatre Pkwy., M/S 535
  Mountain View, California 94043-1351

- Send a fax to the attention of Technical Publications at the following number:

  +1 650 932 0801

We value your comments and will respond to them promptly.

Release Features

This chapter provides a short introduction to the SGI ProPack 1.3 for Linux and describes the features provided with this release. Chapter 2, “Software Overview,” describes features supported in previous releases that enhance the features of your base Linux distribution.

Note: The information in this file is also available on your SGI ProPack 1.3 for Linux CD as *README.SGI*. Because that file can be updated later in the release cycle, you should check it for the latest information. The file is also installed in */usr/doc/README.SGI*.

The errata that describes issues which arose too late in the release cycle to be included in this document and provides help with all known problems is available at the following URL:

http://support.sgi.com/linux

Manuals for installing your base Linux distribution and getting started in using it are included with your base Linux distribution release kit. Use these manuals as a reference for information not included in this guide. If you need to reinstall your base distribution, you must read “Configuring SGI Monitors on the Base Distribution” on page 13.

Documentation for the SGI ProPack 1.3 software and for Linux in general, including man pages, HOWTO guides, and other relevant documentation from the Linux Documentation Project, is available from the following URL:

http://techpubs.sgi.com

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

http://oss.sgi.com

Note: Packaged with the SGI ProPack 1.3 software is a separate sheet that contains the Software License Agreement. This software is provided to you solely under the terms and conditions of the Software License Agreement. Please take a few moments to review the Agreement.

Feature Overview

The SGI ProPack 1.3 for Linux includes modules that provide manageability and scalability for Internet and high-performance clustering applications. This release also provides SGI enhancements for Linux to international customers.

The SGI ProPack 1.3 for Linux is an overlay product that adds to or enhances features in base Linux distributions from Red Hat (version 6.2), SuSE (version 6.4), or TurboLinux (version 6.0).

The SGI ProPack 1.3 software will most likely come preinstalled on your SGI platform. If you should need to install it, be aware that you must install one of the supported base Linux distributions using the base distributor's installation tools and then install the SGI ProPack software using the ProPack installer. This is also true if you are upgrading from an earlier version of the SGI Linux Environment. Installation and configuration is described in Chapter 3, "Quick Configuration and Installation Instructions."

Some of the most significant features that SGI ProPack 1.3 for Linux provides are as follows:

- Support for the Advanced Linux Sound Architecture (ALSA) drivers. These drivers make up the newest sound subsystem for Linux and contain many sound card drivers not previously available.
- Performance Co-Pilot (PCP) is a framework and set of services that support system-level performance monitoring and performance management. The PCP open source release provides a unifying abstraction for all of the interesting performance data in a system, and allows client applications to easily retrieve and process any subset of that data using a single API. For more information, see "Performance Co-Pilot" on page 6.
- The device file system (DEVFS) is turned on by default. Sites that expect to connect a large number of devices may find DEVFS very useful in helping to manage them. See Chapter 2, "Software Overview," for more information.

- The “bigmem” kernel support for large memory systems that was available in previous SGI Linux releases has been replaced with the Linux community standard implementation.
- SGI ProPack 1.3 for Linux provides a number of bugfixes for the NFS server and client.
- The SGI ProPack 1.3 for Linux software is designed to run on any x86 Linux system, but only the SGI platforms are guaranteed. SGI hardware platforms supported by this release are documented at the following URL:

  http://support.sgi.com/linux

Qualified Drivers

QLogic 1080/1280 and 2100 drivers have been supplied by QLogic Corporation, and include updated firmware and improved error handling. The Alteon Gigabit Ethernet driver has been modified to recognize and drive the SGI Gigabit Ethernet card.

The SGI ProPack 1.3 for Linux works with other drivers. You can check the following URL for a complete list:

http://support.sgi.com/linux

Patches and Changes to Base Linux Distributions

This section provides an overview of changes the SGI ProPack 1.3 software makes in your base Linux distribution.

The following packages were added by SGI:

- kernprof 1.0 (kernel profiling tool to identify performance bottlenecks)
- ktrace 1.0 (kernel trace tool)
- libdba.so 1.0 (APIs to enhance database performance)
- lockstat 1.0 (spinlock metering analysis)
- sard 0.2 (disk activity statistics/analysis)
- sgi-logos 1.0.1 (SGI logos)

- sgi-fonts 1.0 (SGI fonts)
- sgi-extra-RedHat 1.3 (manipulate system files for SGI value-added features)
- sgi-extra-SuSE 1.3 (manipulate system files for SGI value-added features)
- sgi-extra-TurboLinux 1.3 (manipulate system files for SGI value-added features)
- sgi-initscripts-RedHat 1.0 (manipulate init scripts for SGI value-added features)
- sgi-initscripts-SuSE 1.0 (manipulate init scripts for SGI value-added features)
- sgi-initscripts-TurboLinux 1.2 (manipulate init scripts for SGI value-added features)
- sgi-release (SGI release identification)
- XFree86 4.0 (approximately 40 RPMs for the 230 hardware platform)
- devfsd 1.3.1 (daemon that allows backwards compatibility with old device file system.
- knfsd 1.4.7 (provides kernel NFS server and related tools)
- mkinitrd 2.3 (creates an initial ramdisk image for preloading modules. Only installs on SuSE, because RedHat and TurboLinux distributions already have this RPM.
- mount-2.9u-4_nfsv3 0.3 (provides mount support for NFS version 3**)**
- sgi-propackdocs 1.3 (HTML-formatted documentation for SGI ProPack 1.3 for Linux)
- sgi-propackdocs-print 1.3 (PDF-formatted documentation for SGI ProPack 1.3 for Linux)

Software Overview

This chapter describes features supported in previous releases that enhance the features of your base Linux distribution. For a description of new features, please read Chapter 1, “Release Features.”

The SGI ProPack for Linux provides the Linux kernel version 2.2.15. The ProPack software adds functionality to base Linux distributions that is specific to SGI hardware platforms.

Some of the most significant features that Linux provides are listed below:

- An extensible UNIX-like kernel, supporting symmetric multiprocessing
- Typical commands you would expect to see on a UNIX-like system
- Typical configuration files you would expect to see on a UNIX-like system, along with an optional graphical frontend
- Development tools such as compilers, debuggers, and libraries
- Internet applications such as web servers and browsers, news servers, network utilities, e-mail servers, and clients
- Everything needed for network file sharing with a wide variety of clients
- Desktop environments and graphical applications

The SGI ProPack 1.3 for Linux software provides optimization that enhances performance on database and other workloads. SGI has added a number of features to the Linux kernel and certain packages to provide increased performance and manageability for database workloads (such as Oracle 8i).

The performance enhancements include a kernel-level implementation of POSIX 1003.1-1996 asynchronous I/O, a low-overhead interprocess synchronization mechanism, low overhead and high-volume raw disk I/O, and support for large amounts of physical memory.

The manageability and supportability improvements include kernel spinlock metering (for performance bottleneck analysis), kernel profiling enhancements, kernel memory dump capability with analysis tools, kernel *gdb* hooks. The SGI ProPack 1.3 for Linux also includes version 0.6 of the kernel debugger *kdb*. The features of *kdb* releases are documented at the following URL:

http://oss.sgi.com/projects/kdb

The manageability of the release has been improved by integrating a number of publicly available kernel patches, such as the following:

- Stephen Tweedie's Raw I/O patch, which forms the basis for the SGI raw disk I/O enhancements. This patch is described in "Raw I/O Path Changes."
- The Device File System (CONFIG_DEVFS_FS) patch from Richard Gooch. This patch provides a more consistent naming scheme for hardware and software devices. Sites that expect to connect a large number of devices may find DEVFS very useful in helping to manage them. DEVFS can also provide the traditional Linux names for devices, for backward compatibility, and is otherwise very compatible with the rest of the Linux system.
- The *sard* utility and associated kernel metrics patch for disk traffic analysis. This patch provides additional disk I/O statistics, useful for tuning database layouts and queries.

Performance Co-Pilot

Performance Co-Pilot (PCP) is a framework and services to support system-level performance monitoring and performance management. The PCP open source release provides a unifying abstraction for all of the interesting performance data in a system, and allows client applications to easily retrieve and process any subset of that data using a single API.

A client-server architecture allows multiple clients to monitor the same host, and a single client to monitor multiple hosts (for example, in a Beowulf cluster). This enables centralized monitoring of distributed processing.

Integrated archive logging and replay allow client applications to use the same API to process real-time data from a host or historical data from an archive.

The framework supports APIs and configuration file formats that enable the scope of performance monitoring to be extended at all levels.

The open source release of PCP provides a subset of the features of SGI's Performance Co-Pilot products for RIX (see http://www.sgi.com/software/co-pilot/).

Performance Improvements

The performance enhancements, enumerated in the preceding section, accelerate the performance of I/O intensive applications by streamlining the kernel code and data paths for disk I/O as well as providing larger shared memory segments and a low overhead interprocess synchronization mechanism.

Raw I/O Path Changes

Current file-system-based disk I/O requires fixed size I/O operations (typically 1024 bytes) into kernel buffers, then the data is moved from the kernel buffer to the user program address space. While this allows the file system to cache frequently accessed data, it also consumes excess system bus bandwidth when copying the data from the kernel buffer(s) into the user address space. Both the small size of the I/O (2 sectors) and the copy operation greatly reduce the I/O subsystem throughput for database operations, where transactions and full-table scan operations operate more quickly with no operating system data intervention.

To help alleviate this problem, Stephen Tweedie of Red Hat developed a mechanism that allows disk I/O directly to a buffer in the application address space (historically known as raw (or unprocessed) I/O). This mechanism will lock the required pages of memory to prevent them from being paged out or swapped during the I/O operation. Applications required to perform this type of disk I/O would open the character special device */dev/raw* and bind the disk device to a special raw device using an *ioctl*(2) system call.

This mechanism, however, is cumbersome to use and suffers from some deficiencies. The primary deficiency with the mechanism comes from its continued use of the file-system buffer-header data structures and associated device queueing routines. While use of the buffer headers is a straightforward mechanism, it implies that I/O operations will still need to be fragmented into 1024-bytes per operation, increasing the kernel overhead significantly. The binding mechanism used to bind an existing block device to a new raw

device is also somewhat cumbersome and counterintuitive to Unix system administrators, who expect to find a relationship in the device namespace between a block device and its corresponding raw device.

To address these concerns, SGI has added additional capabilities to Stephen Tweedie's raw I/O patch that allow large I/O operations directly to the user address space and bypasses the bulk of the kernel I/O queueing code for SCSI and FiberChannel devices.

You can download a *dd* command that is capable of using the raw device features from the following FTP location:

ftp://oss.sgi.com/projects/rawio/download/dd.raw

This feature is off by default, but you can turn it on by setting the **CONFIG_RAW** kernel configuration parameter.

More information about raw I/O is available from the following URL:

http://oss.sgi.com/projects/rawio/faq.html

Large Physical Memory Support

Large amounts of physical memory coupled with the ability to create large (multi-gigabyte) shared memory segments provide a boost in performance to database workloads. SGI includes a configuration parameter to configure the Linux community standard implementation that supports more than 2 GB of physical memory.

You need to run the "bigmem" kernel to use this feature.

Fast Synchronization Mechanism

While the UNIX System V IPC semaphore facility does provide exceptional capability, its performance leaves much to be desired. Many Unix vendors have released a low-overhead inter-application synchronization primitive known as "post/wait."

SGI has included in this release a kernel level implementation of post/wait along with the library containing application API's. The post allows for a process to "wait" for an event. This event can either be a timeout or a "post" from another process. A group of

cooperating processes can use these “post” and “wait” facilities to synchronize among themselves.

In order to use post/wait, the kernel must be compiled with the **CONFIG_PW** configuration variable, and you may optionally set an additional configuration variable, **CONFIG_PW_VMAX**. These variables are described in the configuration help. For a user program to use the post/wait facilities, it must link against *libdba.so*.

For more information on post/wait, please refer to the *postwait*(3) man page.

POSIX Asynchronous I/O

The ability to overlap I/O and processing activities has always been important to high-performance applications. To allow this type of overlap in single-threaded applications, SGI has included a kernel-level implementation of POSIX asynchronous I/O and the associated API library.

The SGI ProPack 1.3 for Linux works with raw devices as well as with file systems including pipes and sockets.

This facility is turned on by setting the **CONFIG_AIO** kernel option. User code can get access to the facility by linking with *libdba.so*. Further information can be found in the */lib/libdba/README* file.

NFS

The following NFS functionality has been added:

- NFS version 3 client and server support
- Network Lock Manager (NLM) version 4 client and server support
- Kernel level NFS and NLM implementation
- A number of bugfixes for the NFS server and client

NFS and NFSD are configured as modules by default, but they can be configured to compile as part of the kernel by setting the **CONFIG_NFS_FS** and **CONFIG_NFSD** configuration parameters. The **CONFIG_NFS_V3** and **CONFIG_NFSD_V3** parameters are set by default and can be turned off if the user wants to use NFS version 2 only. The

CONFIG_NFSD parameter needs to be configured for **LOCKD** to work, so if **CONFIG_LOCKD** is set, **CONFIG_NFSD** should be set also.

Kernel Spinlock Metering

The SGI ProPack 1.3 for Linux includes a feature that allows developers to gather statistical information about the SMP kernel's use of spinlocks and mrlocks (multiple-reader single-writer spinlocks). This functionality is called *spinlock metering*, or *lockmetering*.

Spinlock metering is built into the kernel using the **CONFIG_LOCKMETER** configuration option (in the Kernel Hacking section of **`make xconfig`**). A kernel built with lockmetering will exhibit a small (roughly 1%) performance degradation relative to a kernel that is not configured for lockmetering. See the following URL for additional information:

http://oss.sgi.com/projects/lockmetering

Crash Functionality

The following changes have been made to the Linux crash utility, which are explained briefly below. General information about *lcrash* can be found in the *cmd/lcrash/README* file.

- Linux kernel crash dump enhancements. SGI ProPack 1.3 provides a configuration option to allow kernel crash dumps to be available. This option is configured to be on by default, and the default dump space is the first swap partition found when booting. If you are building a new kernel, you can specify *Support kernel crash dump capabilities* in the Kernel Hacking section of **`make xconfig`**.

The crash dump capabilities in the kernel allow the system to create a crash dump when a failure occurs due to a *panic*() call or an exception. For more details on the dump method, compression used, and so on, please read the LKCD FAQ at the following URL:

http://oss.sgi.com/projects/lkcd/faq.html

Information about LKCD is also available in the file *cmd/lcrash/README.lkcd*.

- Boot up process changes. As the system boots up, the */sbin/vmdump* script will be run out of */etc/rc.d/rc.sysinit*. This script saves crash dumps and reads sysconfig variables to open the dump device and configure the system for crash dumps.
- Crash dump configuration options. There are a number of configurable options to save system crash dumps. Please read */etc/sysconfig/vmdump* for more details on the options available. The following list describes what the options allow you to do:
  - Determine if you want to implement crash dumps in the kernel
  - Choose whether to save crash dumps to disk or not
  - Change the location to which the crash dumps are saved
  - Specify any block dump device you want
  - Compress (or not compress) the crash dumps
  - Configure the system to reset (or not reset) after a failure
- The *lcrash* utility now uses the new *librl* library for command line input.

Patches, Configuration Options, Commands, and Libraries

The following list describes patches implemented and enhancements to configuration options, commands, and libraries:

- *librl* library. This new library supplies command line editing and command history functionality. See the */cmd/lcrash/lib/librl/README* file for information on how to use this library. The *lcrash* command uses this library.
- Remote debugging over a serial line. There is a new configuration option, **CONFIG_GDB**, which is used to enable *gdb* debugging. To force a kernel that has been compiled with **CONFIG_GDB** to pause during the boot process and wait for a connection from *gdb*, the parameter **gdb** should be passed to the kernel. This can be done by typing **gdb** after the name of the kernel on the LILO command line. The

patch defaults to use ttyS1 at a baud rate of 38400. These parameters can be changed by using **gdbttyS=***port number* and **gdbbaud=***baud rate* on the command line.

- *rlimits* patch. In the Linux 2.2.15 kernel, faulty rlimit checking will not allow a process to have more than 2 GB total address space, stack size, or locked memory. This release has fixed the rlimit checking, so (subject to other accounting limitations), the kernel honors **RLIM_INFINITY** settings on these resources.
- SMP PTE patch. In stock Linux, the page stealing code that is used under high memory load has a bug that might cause it to steal a page from a process without writing out the contents to swap if the page has been modified by the process. This bug is only present in a multiprocessor machine. SGI ProPack 1.3 for Linux provides a fix for this bug.

Quick Configuration and Installation Instructions

Your SGI machine comes with a base Linux distribution and the SGI ProPack 1.3 for Linux overlay software preinstalled. This chapter describes how to install the software from the CD if at some time you need to reinstall it. The chapter also describes how to upgrade from earlier SGI Linux releases.

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

If you need to install software, be aware that you must install a base Linux distribution using the base distributor's installation tools and then install the SGI ProPack 1.3 for Linux software using its installer as described in this chapter. You should configure the base Linux distribution as described in the installation manual for the base distribution.

Note: The SGI ProPack 1.3 for Linux software only works with Red Hat 6.2, SuSE 6.4, or TurboLinux 6.0. Earlier versions of these distributions, or any other distributions, are not compatible with SGI ProPack 1.3 for Linux.

Before you install or configure your system, please read Chapter 1, "Release Features," and Chapter 2, "Software Overview," so that you understand the features of SGI ProPack 1.3 for Linux and how to configure them.

Configuring SGI Monitors on the Base Distribution

Should you need to install a base Linux distribution, it is safe to allow the X Configuration mechanism to auto-probe your graphics hardware. When you are asked

to provide information for your monitor, you may refer to the following table to get information for various SGI monitors:

Table 3-1 SGI Monitor Configuration Values

Monitor	Horizontal Sync	Vertical Sync
SGI 17-inch 340C	30-95	48-180
SGI 17-inch GDM-17E11	30.0-85	48.0-150.0
SGI 17-inch GDM-2011P	30.0-85.0	48.0-150.0
SGI 17-inch M-7S54SG	30.0-92.0	48.0-160.0
SGI 19-inch CNMB024B	30-100	48-200
SGI 21-inch GDM-5011P	30-107	48-160
SGI 21-inch 420c	30-107	48-160
SGI 21-inch GDM-5411	30-121	48-160
SGI 20-inch GDM-20E21	30.0-96.0	48.0-160.0
SGI 20-inch GDM-4011P	30.0-96.0	48.0-160.0
SGI 21-inch GDM-5011P	30.0-107.0	48.0-160.0
SGI 21-inch GDM-5021PT	30.0-107.0	48.0-160.0
SGI 24-inch GDM-90W11	30.0-96.0	48.0-160.0

Installing SGI ProPack 1.3 for Linux from a CD

Should you need to install SGI ProPack 1.3 for Linux from a CD, first ensure that either Red Hat 6.2, SuSE 6.4, or TurboLinux 6.0 is installed, and then use the procedure in this section.

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

1. Log in as root.
2. Mount the SGI ProPack 1.3 for Linux CD by executing the mount command as you have configured it. A common example is *mount /dev/cdrom /mnt/cdrom*.
3. Change directories to the root directory for the mounted CD. A common example is is */mnt/cdrom*.
4. Execute *./INSTALL*
5. Select the language you want to use for the installation procedure. The default is English. Use the up and down arrow keys to select your language of choice. Press the **Tab** key to highlight *OK* and press **Enter**.
6. The Welcome screen appears. Highlight *OK* and press **Enter**.
7. The Package Group Selection screen appears. This screen allows you to select the type of package you want to install. You select a package by using the up and down arrow keys and pressing the Space bar to select the one you want. When you select a package, RPMs for that package will be installed after you press the **Tab** key to highlight *OK* and press **Enter.**

   You may also choose “Select individual packages,” which lets you choose the specific RPMs that you want to install.
8. You may see the Package Dependencies screen, which tells you if there are additional packages required beyond those you selected. Review the packages. If you want to install them (you should install them unless there is some important reason not to do so), press the **Tab** key to highlight *OK* and press **Enter**.
9. The Installation to Begin screen appears. It tells you that a log of the installation will be placed in */tmp/install.log*. Press the **Tab** key to highlight *OK* and press **Enter**.
10. The installation begins. You will see the Package Installation screen, which tells you which packages are being installed and logs the time it takes to install them.
11. After the installation is complete, the Complete screen appears. Press **Enter**. You are returned to the root prompt.

Recreating Your Preinstalled System

Should you need to recreate your system (returning it to its original state), simply install your base distribution (Red Hat 6.2, SuSE 6.4, or TurboLinux 6.0) as described in the distributor's installation instructions and then install the SGI ProPack 1.3 for Linux from the CD as described in "Installing SGI ProPack 1.3 for Linux from a CD" on page 14.

Upgrading Your SGI Linux Environment

To upgrade your software from earlier SGI Linux releases, you should follow the instructions for upgrading the base distribution. SGI ProPack 1.3 for Linux will only run on Red Hat 6.2, SuSE 6.4, or TurboLinux 6.0.

After you have upgraded your base distribution, install the SGI ProPack 1.3 for Linux as described in "Installing SGI ProPack 1.3 for Linux from a CD" on page 14.

Note: It is not possible to change base distributions when upgrading your SGI ProPack for Linux release. You must install the other base distribution and then install SGI ProPack 1.3 for Linux.