sgi®

SGI ProPack™ 4 to SGI ProPack™ 5 for Linux® Migration Guide

Record of Revision

Version	Description
001	July 2006 Original publication. Supports the SGI ProPack 5 for Linux release.

Contents

Migrating from SGI ProPack 4 to SGI ProPack 5 for Linux

This guide contains information about migrating from SGI ProPack 4 for Linux operating system to SGI ProPack 5 for Linux Service operating system.

Note: The SGI ProPack 5 for Linux main download page, caveats page, and information about which configuration files you should save off before migrating to the SGI ProPack 5 release are available at the following location on Supportfolio:http://support.sgi.com/content_request/329516/index.html

---

SGI ProPack contains numerous libraries, utilities, and other extensions designed to supplement a base Linux distribution with additional capabilities specifically for HPC, storage, and visualization environments.

SGI ProPack 5 runs on top of SUSE Linux Enterprise Server 10 (SLES10), a standard, off-the-shelf Linux distribution.

The SGI ProPack 5 for Linux release is certified and available with the 2.6 based Linux distribution of SLES10 only.

About This Guide

This guide provides information about specific differences between SGI ProPack 4 and SGI ProPack 5. For general information about migrating to SUSE Linux Enterprise Server 10 (SLES10), see "SUSE Linux Migration Documentation" on page 2.

This guide contains the following sections:

- "Audience" on page 2
- "SUSE Linux Migration Documentation" on page 2
- "Migration Overview" on page 2
- "Migrating from SGI ProPack 4 for Linux" on page 3
- "Differences between SGI ProPack 4 and SGI ProPack 5" on page 4
- "Major Documentation Changes for SGI ProPack 5" on page 13

Audience

This guide is written for system administrators who are responsible for installing, configuring, and administering the SGI ProPack for Linux operating system running on the SGI Altix family of servers and superclusters.

SUSE Linux Migration Documentation

The SGI ProPack 5 for Linux product includes capabilities and performance improvements ideal for enabling technical and creative users to solve their big compute and data problems by using the Linux operating system and Itanium or Xeon processors. This product adds to or enhances features in SUSE Linux Enterprise Server 10 (SLES10). SGI ProPack 5 for Linux is designed to run on any SGI Altix system.

Novell provides information about migrating to SLES10 at the following location: http://www.novell.com/resourcecenter/ext_item.jsp?itemId=13485.

Novell also provides information about migrating applications to SUSE Linux at the following location: http://www.novell.com/coolsolutions/cooldev/features/a_red_hat_migration_cdev.html.

It is important to read the SLES10 release notes as well for information about changes from prior SLES releases. You will find them in a text file named `RELEASE-NOTES.en.txt` in the `root` directory of the SLES10 DVD. The YaST installation tool can also get the latest version of the release notes directly from Novell. On the installed system, the latest version of the release notes are at the following location:

`/usr/share/doc/release-notes/SUSE_Linux_Enterprise_Server_10/RELEASE-NOTES.en.html`

Migration Overview

This section describes general differences between SGI ProPack 4 for Linux and SGI ProPack 5 for Linux.

SGI ProPack 5 runs layered over SLES 10, with some additions by SGI. SGI does **not** provide its own version of the kernel software.

Kernel fixes and updates are provided by Novell.

Disk devices are quite different in SGI ProPack 5 versus SGI ProPack 3. The SUSE YaST installation tool and other SUSE tools do not recognize XSCSI device names. Device names are of the `/dev/sdX` variety. For most SCSI and Fibre channel disks, persistent naming can be achieved by looking through `/dev/disk/`. Both `/dev/disk` and `/dev/xscsi` are set up using `udev`.

Although an XSCSI compatibility mechanism is set using `udev`, XSCSI paths should not be used in general except for working with Fibre channel device names.

Migrating from SGI ProPack 4 for Linux

Note: SGI ProPack 5 is **not** supported on the Silicon Graphics Prism Visualization System.

To install SGI ProPack 5 on a system running SGI ProPack 4 for Linux Service Pack 3 (or earlier), you must do a complete re-installation on a clean disk. Before attempting the SGI ProPack 5 installation, perform the following:

1. Verify PROM firmware is correct. Systems running SLES10 and SGI ProPack 5 should use PROM 1.3 (or later) for systems, such as, the Altix 4700 with the next generation SHub ASIC, or PROM 4.53 (or later) which ship with SGI ProPack 5. To see the PROM version currently on the system, use the following command:

    `cat /proc/sgi_prominfo/*/version`

    If you find that you do need to update your PROM, you can download the latest PROM directly from Supportfolio Online:https://support.sgi.com/browse_request/suse_patch_browse_results

    Once you have the PROM RPM, follow the steps outlined in the section entitled "Obtaining the Latest SGI Altix System Firmware" in the *SGI ProPack 5 for Linux Start Here* to update the PROM on your system. If a printed version of this manual is not readily available, a PDF version is available in the `/docs` directory on the SGI ProPack 5 CD and online at http://docs.sgi.com.

2. Verify that all bricks or blades on the system are running the same revision of L1 firmware. You can always download the latest L1/L2 system controller software from Supportfolio Online:http://support.sgi.com/linux/

3. Save any configuration files you may have modified. To locate modified configuration files, enter the following command:

```
rpm -V -a | grep "\b c"
```

Save these files on a disk other than the one where SGI ProPack will be installed. These files can serve as a valuable reference after installation if you plan to make similar changes.

4. Save any FLEXlm license files that might be necessary to support licensed applications after the SGI ProPack 5 installation.

5. Save off any other files or data that may exist on the local disk where the SGI ProPack 5 release is to be installed.

6. Disconnect any non-system drives from your system. This will help speed up the installation since the SLES install will not scan those drives. This also helps avoid accidental data deletion.

7. Proceed to the "How To Install SGI ProPack 5 for Linux For the First Time" instructions in the *SGI ProPack 5 for Linux Start Here*.

Differences between SGI ProPack 4 and SGI ProPack 5

This section describes some important differences between SGI ProPack 4 Service Pack 3 and SGI ProPack 5 and covers these topics:

- "Latest PROM Firmware" on page 5

- "YaST Installation Tool" on page 5

- "Linux PAM" on page 5

- "Cpusets" on page 6

- "Comprehensive System Accounting" on page 7

- "Secure Array Services" on page 7

- "REACT Real-Time for Linux" on page 7

- "DAPL for Altix" on page 8

- "Persistent IP Addressing of Ethernet Interfaces" on page 8

- "tmpfs Filesystem" on page 9

- "/dev Directory" on page 9

- "XSCSI Naming" on page 9

- "Faster SCSI Device Booting" on page 9

- "XFS Write Barrier Support" on page 11

- "Unaligned Access Warnings" on page 12

- "Kernels" on page 12

Latest PROM Firmware

Unless you have the most current PROM, the kernel may fail to boot during the installation procedure. The prom RPM is proprietary software on the SGI ProPack 5 CD in SGI/RPMS directory.

Note: Systems running SLES10 and SGI ProPack 5 should use PROM 1.3 (or later) for systems, such as, the Altix 4700 with the next generation SHub ASIC, or PROM 4.53 (or later) which ship with SGI ProPack 5. SGI Altix software support contract customers can download the latest PROM version from the following area on Supportfolio:https://support.sgi.com/browse_request/suse_patch_browse_results

For information on obtaining the latest PROM, see the "Obtaining the Latest SGI Altix System Firmware" section in Chapter 2, "Software Planning and Installation", of the *SGI ProPack 5 for Linux Start Here*.

YaST Installation Tool

The operation of the YaST Installation Tool has changed significantly with the SLES10 release. For more information, see the *SGI ProPack 5 for Linux Start Here*.

Linux PAM

In SLES10, the crond daemon is Pluggable Authentication Modules (PAM) aware. The /etc/pam.d/crond file is for PAM and crond.

You can add the `pam_job.so` to the `crond` file. If you want to control all entry
points at the same time, you can use the `common-session` PAM file. The
`common-session` PAM configuration file is included in most PAM configurations
and includes `login`, `crond`, `sshd`, and so on. For more information, see the
`/usr/share/doc/packages/job/README` file.

Cpusets

The cpuset implementation has changed for SGI ProPack 5, as follows:

- Removes the `page_cache_local`, `page_cache_list`, `slab_cache_local`
  per-cpuset files.

- Adds `memory_spread_slab` and `memory_spread_page` per-cpuset files.

- Adds cpuset control interface to the new page migration facility using
  `memory_migrate` per-cpuset file.

- The `cpu_exclusive` flag, when set, automatically defines scheduler domains.
  The kernel performs automatic load balancing of active threads on available CPUs
  more rapidly within a scheduler domain than it does across scheduler domains.

- The `mem_exclusive` flag, when set, automatically defines constraints for kernel
  internal memory allocations. Allocations of user space memory pages are strictly
  confined by the allocating task's cpuset. Allocations of kernel internal pages are
  only confined by the nearest enclosing cpuset that is marked `mem_exclusive`.

- Adds new `memory_pressure` mechanism that makes it easy to detect when the
  job in a cpuset is running short of memory and needing to page memory out to
  swap.

- The basic cpuset routines are now robust in the face of migration.

- The `cpuset_collides_exclusive` routine detects exclusive cpuset overlap for
  improved error messages.

- The `cpuset_migrate` and `cpuset_migrate_all` routines, which move
  processes from one cpuset to another, are now available.

- New routines to enable building dependent library routines that are immune to
  memory migration are, as follows: `cpuset_get_placement`,
  `cpuset_equal_placement`, and `cpuset_free_placement`.

- Minimal Fortran linking supported for these basic routines, as follows: `cpuset_pin`, `cpuset_size`, `cpuset_where`, and `cpuset_unpin`.

For detailed information, see Chapter 5, "Cpusets on SGI ProPack 5 for Linux" in the *Linux Resource Administration Guide*. You can also find information at the following location:

`/usr/share/doc/packages/cpuset/cpuset.html`

Comprehensive System Accounting

The Comprehensive System Accounting (CSA) software has been updated. The CSA-3.0.0 version is a major cleanup of CSA and removes unsupported CSA record types, simplifies the `/etc/csa.conf` configuration file, and makes changes to header files. As a result of these changes, all user applications need to be recompiled. CSA-3.0.0 still supports accounting data files created in earlier version; however, CSA utilities from earlier releases can not read accounting data files created by CSA-3.0.0 and later.

Secure Array Services

The secure version of Array Services is built to make use of secure sockets layer (SSL) and secure shell (SSH) is new for the SGI ProPack 5 release. The "Secure Array Services" section in Chapter 3, "Array Services" of the *Linux Resource Administration Guide* describes the differences between standard Array Services (`arraysvcs`) and Secure Array Services (`sarraysvcs`).

Standard Array Services is installed by default on an SGI ProPack 5 system. To install Secure Array Services, use the YaST Software Management and use the **Filter->search** function to search for secure array services by name (`sarraysvcs`).

REACT Real-Time for Linux

REACT Real-Time for Linux has changed for SGI ProPack 5, as follows:

- The `reactcfg.pl` and `reactboot.pl` scripts have been replaced by `reactconf.pl`. It includes the ability to dynamically reconfigure real-time CPUs and disable REACT across reboots.

- Adds the SGI Linux Trace feature that generates traces for kernel events such as interrupt handling, scheduling, and system calls

- Adds the ability to shield a CPU from the normally scheduled Linux timer (scheduler) interrupts

- Changes the cpuset structure from /rtcpuN to /rtcpus/rtcpuN. To run a process on a restricted CPU, you must invoke or attach it to a real-time cpuset (that is, a cpuset containing a CPU that does not exist in the bootcpuset.

For more information about changes to REACT Real-Time for Linux, see the *REACT Real-Time for Linux Programmer's Guide*. Information about the new SGI Linux Trace feature is described in Chapter 10, "SGI Linux Trace".

DAPL for Altix

A Direct Access Programming Library (DAPL), new for the SGI ProPack 5 release, is provided for high performance communication between processes on single host or partitioned Altix system. This can be used with Intel Message Passing Interface (MPI) via the remote direct memory access (RDMA) MPI device. See dapl_xpmem(3) for more information.

Note: For the SGI ProPack 5 release, only single host systems are supported. Partitioned systems support is planned for a future Service Pack release.

Persistent IP Addressing of Ethernet Interfaces

Ethernet persistent behavior changes in the SGI ProPack 5 for Linux from prior releases.

The basic change is that the first time a SGI ProPack 5 system is booted after installation, a udev rule defined in the /etc/udev/rules.d/31-net_create_names.rules file is invoked that enumerates all of the Ethernet devices on the system. It then writes another rule file called /etc/udev/rules.d/30-net_persistent_names.rules. This file contains a mapping of the media access control (MAC) addresses to Ethernet IP addresses. A specific physical interface is always mapped to the same Ethernet address. For more information, see the "Persistent Network Interface Names" section in Chapter 1, "Configuring Your System" of the *Linux Configuration and Operations Guide*.

tmpfs Filesystem

The tmpfs filesystem memory allocations have changed for the SGI ProPack 5 release. Prior to the SGI ProPack 5 release, allocations were always done round-robin on all nodes. With SGI ProPack 5, this is now a tmpfs mount option. For more information, see the "Filesystems Changes" section in Chapter 1, "Configuring Your System" of the *Linux Configuration and Operations Guide*.

/dev Directory

In SLES10, the /dev directory is no longer persistent. In SLES9, the /dev directory was part of the root filesystem and although udev managed it, if you created your own device file, it was persistent. In SLES10, /dev is a tmpfs filesystem that starts fresh with each reboot.

XSCSI Naming

XSCSI naming is available on SGI ProPack 5 systems. See the "XSCSI Naming Systems on SGI ProPack Systems" section in Chapter 1, "Configuring Your System" of the *Linux Configuration and Operations Guide* for a discussion of XSCSI names on non-blade systems and domain-based XSCSI names.

Faster SCSI Device Booting

There are three files in the /etc/udev/rules.d directory that can be modified to make systems boot faster that have many logical unit numbers (LUNs) for attached SCSI devices (1000 LUNS or more), as follows:

- `50-udev-default.rules`

- `58-xscsi.rules`

- `60-persistent-storage.rules`

This section describes these rules.

50-udev-default.rules

The rules in the `50-udev-default.rules` file cause SCSI drivers `sd_mod`,`osst`, `st`, `sr_mod`, and `sg` to be loaded automatically when appropriate devices are found. These rules are, as follows:

```
SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="0|7|14", RUN+="/sbin/modprobe sd_mod"
SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="1", SYSFS{vendor}=="On[sS]tream", RUN+="/sbin/modprobe osst"
SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="1", RUN+="/sbin/modprobe st"
SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="[45]", RUN+="/sbin/modprobe sr_mod"
SUBSYSTEM=="scsi_device", ACTION=="add", RUN+="/sbin/modprobe sg"
```

You can comment out all of these rules to save calls to the `modprobe`(8) command for each SCSI device to save boot time, as follows:

```
#SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="0|7|14", RUN+="/sbin/modprobe sd_mod"
#SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="1", SYSFS{vendor}=="On[sS]tream", RUN+="/sbin/modprobe osst"
#SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="1", RUN+="/sbin/modprobe st"
#SUBSYSTEM=="scsi_device", ACTION=="add", SYSFS{type}=="[45]", RUN+="/sbin/modprobe sr_mod"
#SUBSYSTEM=="scsi_device", ACTION=="add", RUN+="/sbin/modprobe sg"
```

Make sure the drivers are loaded by adding them to `INITRD_MODULES` variable in the `/etc/sysconfig/kernel` file and then run the `mkinitrd`(8) command to make sure the changes are picked up.

58-xscsi.rules

The rules in the `58-xscsi.rules` file create the `/dev/xscsi/pci...` persistent device symbolic links. The rules are, as follows:

```
# This rule creates sg symlinks for all SCSI devices (/dev/xscsi/pci..../sg)
KERNEL=="sg*", PROGRAM="/sbin/udev_xscsi %p", SYMLINK+="%c"

# This rule creates symlinks for entire disks/luns (/dev/xscsi/pci..../disc)
KERNEL=="sd*[a-z]", PROGRAM="/sbin/udev_xscsi %p", SYMLINK+="%c"

# This rule creates symlinks for disk partitions (/dev/xscsi/pci..../partX)
KERNEL=="sd*[a-z]*[0-9]", PROGRAM="/sbin/udev_xscsi %p", SYMLINK+="%c"
```

You need the rule for the `disk/luns` symbolic links for XVM in this file, (the middle rule). You can comment out the rule that creates `sg` symbolic links and the rule for disk partition symbolic links (top and bottom rules, respectively).

`60-persistent-storage.rules`

> The rules in the `60-persistent-storage.rules` file are for persistent storage links. They are not necessary and can all be commented out or you can add `GOTO="persistent_storage_end"` at the top of the file to accomplish the same.

XFS Write Barrier Support

> By default, in SLES10, XFS uses write barriers if supported on the target device. Write barriers are used to ensure that the journaled log is physically written to the disk and not kept in the disk's cache; this risks filesystem corruption if power is suddenly lost.
>
> Most modern SCSI and SATA disks will support write barriers. This can be confirmed by checking the `dmesg(8)` command output for the disk driver should show something similar to the following:
>
> `"SCSI device sda: drive cache: write back"`
>
> You can also use the following command:
>
> `% `**`sginfo -c /dev/sda | grep -i 'write cache'`**
>
> Be aware that at this time, XVM does not support write barriers and XFS will fall back to the old behavior when the write barrier request fails.
>
> Write barriers should **not** be enabled on storage that has battery backed and mirrored write caching, such as the SGI TP9300, the SGI TP9700 and InfiniteStorage 4000 RAID. These RAIDs are designed to write all cached data to disk before battery power runs out and write barriers will unnecessarily impact write performance.
>
> Write barriers can be disabled with the `nobarrier` mount option.
>
> The first time you mount the device, check the `syslog` file. If the device does not support barriers, you will see a message similar to the following:
>
> `Jul 12 15:34:01 budgie kernel: Filesystem "xvm-0": Disabling barriers, not  suppor`
>
> Note that some filesystem configurations do not support barriers, either. You should check the `syslog` file, for example, the external log devices:
>
> `Jun  5 22:44:24 budgie kernel: Filesystem "dm-1": Disabling barriers, not supporte`

Unaligned Access Warnings

The section describes unaligned access warnings in SLES10, as follows:

* The kernel generated unaligned access warnings in syslog and on the console, when applications do misaligned loads and stores. This is normally a sign of badly written code and an indication that the application should be fixed.

* Use the prctl(1) command to disable these messages on a per application basis.

* SLES10 offers a new way allowing system administrators to disable these messages on a system wide basis. This is generally discouraged, but useful for the case where a system is used to run third-party applications which cannot be fixed by the system owner.

  In order to disable these messages on a system wide level, do the following as root:

  ```
  echo 1 > /proc/sys/kernel/ignore-unaligned-usertrap
  ```

Kernels

SLES10 contains two kernel variants for ia64 systems, based on the same 2.6.16.x source, called kernel-default and kernel-debug.

For more information, see the "SLES10 Kernels" section in Chapter 2, "Software Planning and Installation" of the *SGI ProPack 5 for Linux Start Here*.

Tools and Library Changes

This section describes changes to tools and libraries between the SGI ProPack 4 and SGI ProPack 5 releases that may impact application developers.

The GCC (GNU Compiler Collection) contains a new application binary interface (ABI) in the GNU C++ compiler 3.4 and higher. GCC 4.1 is included with the SLES10 release. GCC 3.3 was included with the SLES9 release.

To compile and link C++ programs with the Message Passing Interface (MPI library), choose one of the following commands:

```
g++ -o myprog myproc.C -lmpi++abi1002 -lmpi
g++ -o myprog myproc.C -lmpi++abi102 -lmpi
g++ -o myprog myproc.C -lmpi++ -lmpi
```

The `libmpi++abi1002.so` library is compatible with C++ ABI 1002 code generated by GNU g++ compiler version 3.4 and higher. The `libmpi++abi102.so` and `libmpi++.so libraries` are compatible with C++ ABI 102 code generated by GNU g++ compiler versions 3.0 through 3.3 and with Intel C++ 8.0 or later compilers. The `libmpi++.so` library will provide binary compatibility with program binaries previously compiled with prior versions of Message Passing Toolkit (MPT) on systems where the ABI 102 C++ environment is installed (for example on most SLES 9 / SGI ProPack 4 systems). If binary compatibility is desired with future MPT software releases, you should link with `-lmpi++abi1002` or `-lmpi++abi102`. The ABI supported by `libmpi++.so` may change in a future release. To find the C++ ABI generated by the compiler you are using, choose from one of these commands:

```
g++ -E -dM - < /dev/null | grep ABI
icc -E -dM - < /dev/null | grep ABI
```

For more information, see the `mpi`(1) man page.

Major Documentation Changes for SGI ProPack 5

For information on manuals that are new or updated for the SGI ProPack 5 for Linux release, see the "Documentation Updated for This Release" section of Chapter 1, "Release Features" of the *SGI ProPack 5 for Linux Start Here*.

Obtaining Publications

To obtain SGI documentation, go to the SGI Technical Publications Library at http://docs.sgi.com.

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://docs.sgi.com

- Contact your customer service representative and ask that an incident be filed in the SGI incident tracking system.

- Send mail to the following address:

  Technical Publications
  SGI
  1500 Crittenden Lane, M/S 535
  Mountain View, California 94043-1351

We value your comments and will respond to them promptly.