Project no. IST-033576

XtreemOS

Integrated Project
BUILDING AND PROMOTING A LINUX-BASED OPERATING SYSTEM TO SUPPORT VIRTUAL ORGANIZATIONS FOR NEXT GENERATION GRIDS

Initial LinuxSSI Integration and Packaging in OSCAR
D4.1.2

Due date of deliverable: February 28, 2007
Actual submission date: April 3, 2007

Start date of project: June 1^{st} 2006

Type: Deliverable
WP number: WP4.1
Task number: T4.1.2

Responsible institution: INRIA
Editor & and editor's address: Jean Parpaillon
IRISA-INRIA
Campus de Beaulieu
35042 Rennes
France

Version 1.1 / Last edited by Jean Parpaillon / April 2, 2007

Project co-funded by the European Commission within the Sixt Framework Programme		
Dissemination Level		
PU	Public	√
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Revision history:

Version	Date	Authors	Institution	Section affected, comments
1.0	22/02/07	Jean Parpaillon	IRISA	first draft
1.1	07/03/07	Jean Parpaillon	IRISA	Corrections from Christine Morin
1.2	12/03/07	Jean Parpaillon	IRISA	Corrections from David Margery
1.3	29/03/07	Jean Parpaillon	IRISA	Grammar corrections from Nicolas Vigier
1.4	02/04/07	Jean Parpaillon	IRISA	Corrections from Samuel Kortas

Abstract

Linux SSI is the XtreemOS foundation part for clusters. It is based on the Kerrighed technology. The OSCAR project eases the installation, configuration and administration of clusters. This document describes the work done for the integration of Kerrighed in the OSCAR suite of tools. Packaging Kerrighed for the OSCAR supported distributions is the major part of this integration work. Therefore, this document contains description of this packaging. Finally, basic informations are provided to get an Linux-SSI cluster running with OSCAR.

Contents

1 Introduction

Linux-SSI is the core of XtreemOS for clusters. As an SSI (Single System Image) based on the Kerrighed project, it offers illusion that a cluster is a single multi-processor machine. OSCAR is a set of best practices tools to ease the deployment and management of clusters.

Packaging Linux-SSI for OSCAR offers a complete solution for deploying, configuring, managing and using a cluster.

This document presents the Kerrighed and OSCAR projects and the work done to package Kerrighed into OSCAR. Future work is also mentionned.

1.1 Presentation of Kerrighed and Oscar Software

1.1.1 Kerrighed [2]

Linux-SSI is based on the Kerrighed [4] cluster operating system. As a Single System Image, it offers the view of a single machine over a fast network interconnected set of machines. Kerrighed implements a virtual multi-processor SMP-like.

Kerrighed system offers load-balancing through cluster nodes. The load balancing policy is configurable. The high performance communication system can be easily ported on interconnection networks like Gigabit Ethernet, Myrinet, Infiniband, Quadrics, etc. Distributed services implementing global resource management in Kerrighed rely on this communication layer.

Kerrighed can be specialised as a user can selectively enable or disable functionnalities of the SSI through an API or with command-line tools. These functionnalities are:

migration: allow a process to migrate,

distant fork: allow a process to fork on a distant node,

see local /texttt/proc: allow a process to see the local `/proc` filesystem instead of the cluster-wide one,

change capabilities: allow a process to change its capabilities.

The Kerrighed system is implemented with a Linux kernel patch, a kernel module and some optionnal tools.

1.1.2 OSCAR Suite of Tools [5]

OSCAR [3] is a set of best practices tools for clusters. It contains software for deploying, configuring, monitoring and managing clusters. Most of these tools are

distribution independent and OSCAR officially supports 7 distributions: Fedora Core, CentOS, RHEL, Mandriva Linux, SLES, Scientific Linux and Debian. It is implemented as a master node controlling the cluster nodes, as shown on Figure 1.

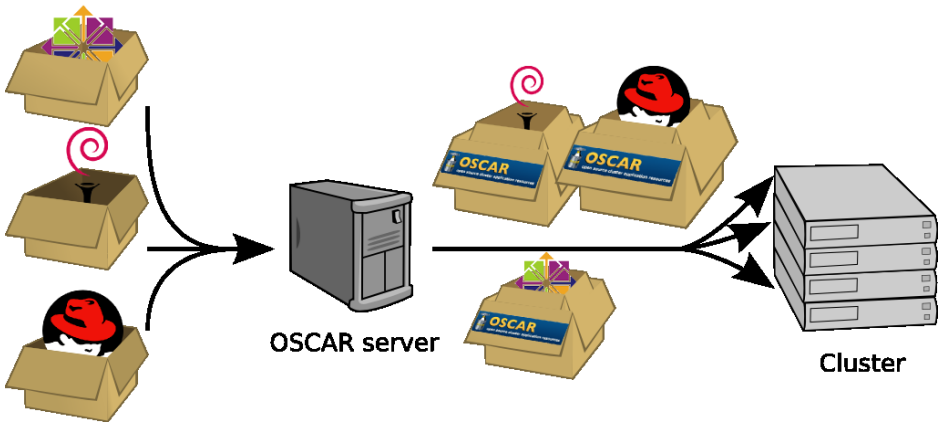

Figure 1: OSCAR architecture

OSCAR defines a packaging format for deploying software on a cluster. This packaging includes distribution native packages, as shown on Figure 2.

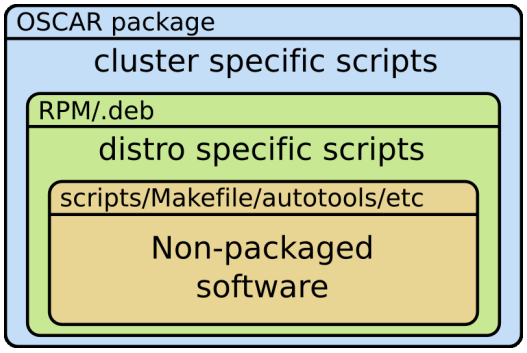

Figure 2: OSCAR Package Architecture

1.1.3 SSI-OSCAR [6]

Since November 2004, Kerrighed has been packaged in the SSI-OSCAR [7] package for the OSCAR versions 4.0 and 4.1. As it includes some specific scripts, it is not an official part of OSCAR, but a third party package. Furthermore, the version of Kerrighed packaged in it is the 1.0.2 which is based on Linux 2.4.29. This

first version of SSI-OSCAR can only be deployed on distributions supported by OSCAR 4.x series.

Current version can be deployed with OSCAR 5.x series and supported distributions include now Debian (in a beta state) [1].

1.2 Kerrighed Deployement Process

The Kerrighed system is made of a modified Linux kernel plus some libraries and tools. When booting, the Kerrighed kernel needs to know the node id. This information is defined at cluster level and not node level, hence a cluster management system, like OSCAR.

To deploy a Kerrighed cluster, the three following steps are required.

Step 1: Install a standard Linux base system on each node.

Step 2: Install the Kerrighed software (kernel, module, tools and librairies).

Step 3: Set up the system to start the cluster.

2 Generic and Kerrighed specific packaging in OSCAR

From the Kerrighed sources [2], we want to build a SSI-OSCAR package. This package can be installed with OSCAR to deploy a Kerrighed cluster .

As seen on Figure 2, the Kerrighed OSCAR package contains several packaging levels.

At the inner-most level, the sources are compiled and installed with generated `Makefiles`. Compilation and installation options can be defined with environment variables or with a `configure` script. The sources are split between several parts, as shown on Table 1.

Compilation and installation of the kernel part is handled with `Kbuild`. This is a `Makefile` targets library specially designed for Linux kernel building. The `Kbuild` system offers the user targets to configure, build and install the Linux kernel and its modules.

Other parts are managed by the `autotools` tools. They configure the sources, build the tools and libraries and install them into the locations specified when configuring.

[1] See http://oscar.openclustergroup.org/supported_distros for supported distributions. Debian support is not official.

[2] Available on a subversion repository: svn://scm.gforge.inria.fr/svn/kerrighed/trunk

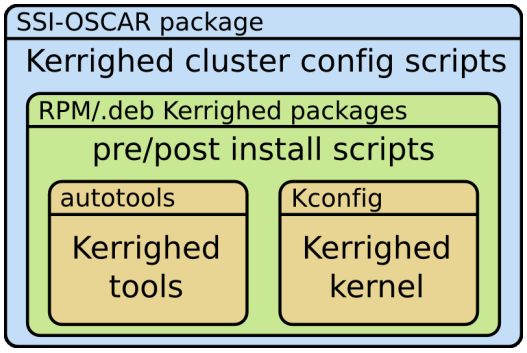

Figure 3: SSI-OSCAR Package Structure

Directory	Description
`kernel`	Linux 2.6.11 + Kerrighed patch
`modules`	Kerrighed module
`libs`	libkerrighed
`tools`	Administration tools
`tests`	Various tests
`doc`	User manual, installation manual

Table 1: Kerrighed sources parts

The `RPM` and `Debian` level generates packages for, respectively, `RPM`-based and `Debian SID` distribution, with specific scripts and tuning for each node independently.

The `OSCAR` package adds scripts to configure nodes cluster wide.

2.1 Autotools Packaging

2.1.1 Autotools Standard Features

Autotools [8] is a suite of tools designed for `Makefile` generation. It allows one to configure, build and install a package with the following commands:

```
> ./configure <options>
> make
> make install
```

Many project configuration options can be given to the `configure` script. They allow one to define installation place of many places of the built software. Running `configure -help` shows all these options.

`Makefiles` generated with `autotools` offer standard targets to the user:

all (default): build the project,

install: install the project,

dist: create a tarball from the project,

clean: clean built objects,

distclean: clean built objects and generated `Makefiles`.

2.1.2 Kerrighed Autotools Features

From `autotools` point of view, Kerrighed is composed of the following packages:

- kernel part,
- libkerrighed,
- tools,
- tests,
- documentation.

Each of them can be enabled or disabled by a `configure` option, with the exception of the kernel part which is the core. When running `configure -help`, one can see the description of these options, as following:

```
--disable-linux-check   Check for the Linux sources to be compatible with
                        Kerrighed (!!! disable at your own risk !!!)
                        [default=enable]
--disable-module        Patch the kernel with Kerrighed and build the module
                        [default=enable]
--disable-libkerrighed  Disable libkerrighed [default=enable]
--enable-libkrgthread   Enable libkrgthread [default=disable]
--disable-tools         Disable tools [default=enable]
--disable-service       Do not install kerrighed service [default: enable if
                        lsb found]
--enable-tests          Enable tests [default=disable]
--disable-pdf-doc       Disable pdf documentation [default=enable]
--enable-html-doc       Enable HTML documentation [default=disable]
--enable-text-doc       Enable fulltext documentation [default=disable]
```

In addition to standard target, the `patch` one patches the Linux kernel sources with Kerrighed.

The `all` target compiles the Kerrighed module and selected packages and documentation. Modified Linux configuration or building is not handled with `autotools`.

The libraries, tools and tests packages contain their own `configure` script which can be run from their respective directories. They contain more options to fine tune the compilation of these packages. Please run `configure -help` to see these options.

Unless specified with `configure` options, the Kerrighed project is installed in the paths as in Table 2.

Path	Description
/usr/local/bin	Tools
/usr/local/lib/libkerrighed.*	Libraries
/usr/local/include/kerrighed/	C headers
/usr/local/share/man/man{1|2|5|7}	Man pages
/usr/local/share/doc/kerrighed/	Manuals
/lib/modules/2.6.11-krg/extra/kerrighed.ko	Kerrighed module
/etc/init.d/kerrighed	Kerrighed service script
/etc/default/kerrighed	Kerrighed service configuration

Table 2: Kerrighed installation paths

2.2 Kbuild Packaging

`Kbuild` has been designed especially for the Linux kernel. Then, it is used for Kerrighed module compilation and installation. `Kbuild` is only used for the Linux kernel project. Updated documentation is available in the Linux source tree [3].

Configuration of the module is done through the Linux configuration interfaces. The Kerrighed module is made of several parts and each one can be disabled or enabled. All these options are available and documented in the "cluster support" section of Linux configuration.

Compilation and installation of the module is done with the `all` and `install` targets of the generated `Makefiles` from autotools (see Section 2.1). These targets launch `Kbuild` targets.

2.3 RPM Packaging

Two packages sets are generated from the Kerrighed project. The first uses as source the Linux kernel and the Kerrighed patch for the kernel. Its content is described in Table 3. The second one deals with all other parts of Kerrighed. The list of its packages is presented in the Table 4.

Packages have been tested on Fedora Core, RHEL, CentOS, Mandriva and SuSe distributions.

Each package set is described in a `.spec` file. These files include packages description, package building commands as well as pre-post install script. These files are available in the Kerrighed subversion repository at the following address:

```
svn://scm.gforge.inria.fr/svn/kerrighed/packages/rpm/trunk/specs/
```

More complete informations about RPM packaging are in the deliverable for WP 4.1.1 [1].

[3] Latest version: http://www.kernel.org/git/?p=linux/kernel/git/torvalds/linux-2.6.git;a=tree;f=Documentation/kbuild;hb=HEAD

Package	Depends on	Description
kerrighed-kernel	-	Contains the Kerrighed kernel, which is a Linux kernel modified with Kerrighed patch
kerrighed-source	-	Contains the patched source code of the kernel. Needed for development use.
kerrighed-kernel-doc	-	Contains documentation files from the kernel sources.

Table 3: Kerrighed kernel RPM packages

Package	Depends on	Description
kerrighed	kerrighed-kernel, kerrighed-module, kerrighed-utils, kerrighed-doc, kerrighed-libkerrighed	Main meta-package, install all needed packages for a complete Kerrighed system
kerrighed-module	kerrighed-kernel, kerrighed-utils	Provides the module kerrighed.ko.
kerrighed-utils	kerrighed-libkerrighed	Contains tools to make a fully fonctionnal Kerrighed cluster, like init scripts. It contains the command krgadm.
kerrighed-doc	-	Provides manuals
kerrighed-libkerrighed	kerrighed-module	Provides the kerrighed library to use some advanced features of the Kerrighed OS.
kerrighed-devel	kerrighed-libkerrighed	Provides the kerrighed library development files and static libraries.

Table 4: Kerrighed RPM packages

2.4 Debian Packaging

As for RPM packages, two package sets are generated from the Kerrighed project. For each kernel related package, there are two flavours of the package. Flavours consists of different configurations of the Linux kernel. These two flavours are:

- 686: built with optimization for Pentium Pro and more recent processors.
- 686-dbg: same as before plus debugging symbols and KDB (Kernel DeBugger).

Currently, only the i386 architecture is supported.

The first package set is called `linux-2.6.11-krg` and contains modified Linux packages. The second, called `kerrighed` contains Kerrighed module, tools and libraries. Table 5 presents the list of packages from the `kerrighed` package set and Table 6 presents the list of packages from the `kerrighed-kernel` package set.

All files needed for package generation are available on the Subversion repository of Kerrighed, at the following addresses:

```
svn://scm.gforge.inria.fr/svn/kerrighed/packages/deb/trunk/kerrighed
svn://scm.gforge.inria.fr/svn/kerrighed/packages/deb/trunk/linux-2.6.11-krg
```

2.5 OSCAR Packaging

The OSCAR package provides two more features than the native RPM or Debian packages:

1. It is multi distribution. It contains native packages for several distributions:
   - Debian,
   - RPM packages tested with Fedora Core 5, RHEL/CentOS 4, Mandriva Linux 2006 and SuSe Linux 10.0.
2. Kerrighed is set up on the whole cluster.

The package is named SSI-OSCAR. It contains:

- packages for supported distributions,
- an XML file (`config.xml`) describing the package (name, authors, version, dependencies, license, ...),
- a post-install script setting cluster nodes to boot with Kerrighed.

The post-install configuration is done through a sligtly modified version of the SystemConfigurator tool. This tool configures various system elements of the client nodes, such as network or boot process, from a single file. Description in this file is distribution independent. The modified version can compute node id (required by Kerrighed kernel) and append it to boot parameters.

Like modern Linux distributions, OSCAR integrates a system to download packages from a repository. This piece of software is called OPD (Oscar Package Downloader). The repository is described with a file `oscar_packages.xml`, which contains informations about the available packages and their location.

2.6 Automated Building Scripts to Generate the SSI-OSCAR Packages

As there is not yet a recent release of Kerrighed, packages are generated from subversion snapshot versions. A set of `Makefiles` have been written to automate packages generation. They can be found at the following address:

```
svn://scm.gforge.inria.fr/svn/kerrighed/packages
```

There is a `Makefile` for each packaging system. All these `Makefiles` are included into a main `Makefile`. These files and their main targets are summarized in the Table 7. To see a complete list of targets, run `make help`.

3 Using the SSI-OSCAR Package for Deploying a Linux SSI cluster

Here is a brief instructions set to make a Linux SSI cluster running with the help of OSCAR.

To install LinuxSSI with OSCAR, you need a cluster head node and some compute nodes. On the head node, install OSCAR. Then, with the OSCAR wizard, deploy the SSI-OSCAR package on the compute nodes.

3.1 Availability

OSCAR is available at the following address:

`http://oscar.openclustergroup.org/download`

SSI-OSCAR packages are available at the following address:

`http://ssi-oscar.gforge.inria.fr/download/oscar-repository/current/`

3.2 OSCAR wizard

Current version of SSI-OSCAR is designed for OSCAR version 5.0. To install OSCAR 5.0, please refer to the installation manual [4].

The Figure 4 represents the OSCAR wizard. Run Step 0 to enter the address of the SSI-OSCAR repository, as in Section 3.1.

Then, in Step 1, select the `ssi-oscar` package then follows instruction as for a standard OSCAR cluster configuration.

4 Conclusion

Linux SSI eases the use of cluster based grids, as clusters are seen as single machines. OSCAR eases the installation of clusters, providing a wizard based interface to do it.

Integration of Linux SSI in OSCAR suite of tools needs packaging Kerrighed for the distributions OSCAR supports. This has been done for Debian and major RPM-based distributions.

Ongoing work includes testing these packages on a wider bunch of distributions. Currently, the OSCAR interface only manages the deployment of a Kerrighed cluster. The reconfiguration (node id, cluster id) could also be done through the OSCAR interface.

[4] http://oscar.openclustergroup.org/public/docs/oscar5.0/OSCAR5.0_Install_Manual.pdf

Figure 4: OSCAR wizard

References

[1] A. Giniès. D4.1.1: Initial linuxssi integration and packaging in debian, mandriva and redflag distribution. Technical report, XtreemOS, Feb. 2007.

[2] Kerrighed. Kerrighed website (http://www.kerrighed.org/), 2007.

[3] T. Mattson. Oscar: A packaged cluster software stack for high performance computing. Technical report, Open Cluster Group, June 2001.

[4] C. Morin, P. Gallard, R. Lottiaux, and G. Vallée. Towards an efficient single system image cluster operating system. *Future Generation Computer Systems*, 20(2), January 2004.

[5] OSCAR. Oscar website (http://oscar.openclustergroup.org/), 2007.

[6] J. Parpaillon. Ssi-oscar website (http://ssi-oscar.gforge.inria.fr/), 2007.

[7] G. Vallée, S. L. Scott, C. Morin, J.-Y. Berthou, and H. Prisker. Ssi-oscar: a cluster distribution for high performance computing using a single system image. In *The 3rd Annual OSCAR Symposium*, University of Guelph,

Guelph, Ontario, Canada, May 2005. Held in conjunction with the 19th International Symposium on High Performance Computing Systems and Applications (HPCS 2005).

[8] G. V. Vaughan, B. Elliston, T. Tromey, and I. L. Taylor. *GNU Autoconf, Automake, and Libtool*. Sams Publishing, 2000.

Package	Depends on	Description
kerrighed	linux-image-2.6.11 +krg<version>-686 kerrighed-module-2.6.11 +krg<version>-686, kerrighed-tools, libkerrighed	Meta package depending on all needed packages for a complete Kerrighed system
kerrighed-dev	linux-image-2.6.11 +krg<version>-686-dbg, linux-headers-2.6.11 +krg<version>-686-dbg, kerrighed-module-2.6.11 +krg<version>-686-dbg, kerrighed-tools, libkerrighed, libkerrighed-dev, kerrighed-doc	Meta package depending on all needed packages for using and developping Kerrighed
kerrighed-doc	-	Documentation files for Kerrighed
kerrighed-tools	libkerrighed	Contains tools to make a fully fonctionnal Kerrighed cluster, like init scripts. It contains the command krgadm.
kerrighed-module-source	-	Contains sources of the Kerrighed module and script to build it with the kernel-package on Debian.
kerrighed-module-686	kerrighed-module-2.6.11+krg<version>-686	Virtual package depending on the latest available Kerrighed module for 686 flavour.
kerrighed-module-2.6.11+krg<version>-686	linux-image-2.6.11+krg<version>-686	Contains the Kerrighed module for 686 flavour.
kerrighed-module-<arch>-dbg	kerrighed-module-2.6.11 +krg<version>-686-dbg	Virtual package depending on the latest available Kerrighed for 686-dbg flavour.
kerrighed-module-2.6.11 +krg<version>-686-dbg	linux-image-2.6.11 +krg<version>-686-dbg	Contains the Kerrighed module for 686-dbg flavour.
libkerrighed	kerrighed-module-2.6.11+krg<version>-686 or kerrighed-module-2.6.11+krg<version>-686-dbg, linux-image-2.6.11+krg<version>-686 or linux-image-2.6.11+krg<version>-686-dbg	Contains the Kerrighed library to use advanced features of the OS.
libkerrighed-dev	libkerrighed	Contains headers and static version of Kerrighed library.

Table 5: Kerrighed Debian packages

Package	Depends on	Description
linux-source-2.6.11	-	Provides Linux sources patched with Kerrighed. For use with the kernel-package tools from Debian.
linux-doc-2.6.11	-	Documentation from Linux source tree.
linux-manual-2.6.11	-	Linux kernel API manual pages.
linux-patch-debian-2.6.11	-	Debian patches to Linux 2.6.11 vanilla.
linux-tree-2.6.11	linux-patch-debian-2.6.11, linux-source-2.6.11	Linux kernel source tree for building Debian kernel images. This is a virtual package and is used for building a new kernel from Linux with Kerrighed sources.
linux-support-2.6.11 +krg<version>	-	Contains Kbuild system and other files for Linux kernel building.
linux-headers-2.6.11 +krg<version>-all	linux-headers-2.6.11 +krg<version>-all-386	Depends on Linux headers for all flavours of all architectures.
linux-headers-2.6.11 +krg0r1356-all-386	linux-headers-2.6.11 +krg<version>-686, linux-headers-2.6.11 +krg<version>-686-dbg	Depends on Linux headers for all flavours of i386 architecture.
linux-headers-2.6.11 +krg<version>	-	Provides common (non architecture or flavour specific) headers files from Linux source tree.
linux-headers-2.6.11 +krg<version>-686	linux-headers-2.6.11 +krg<version>	Provides headers file from Linux kernel for 686 flavour.
linux-headers-2.6.11 +krg<version>-686-dbg	linux-headers-2.6.11 +krg<version>	Provides headers file from Linux kernel for 686-dbg flavour.
linux-image-2.6.11 +krg<version>-686	-	Contains Linux kernel patched with Kerrighed in 686 flavour plus all modules in a standard Debian Linux kernel.
linux-image-2.6.11 +krg<version>-686-dbg	-	Contains Linux kernel patched with Kerrighed in 686-dbg flavour plus all modules in a standard Debian Linux kernel.

Table 6: Kerrighed kernel Debian packages

Target	Description
Makefile	
`clean`	Clean generated files.
`help`	Prints available targets and environment variables.
common.mk	
`tarball_files`	Build a tarball `kerrighed-<version>.tar.gz` from Kerrighed SVN repository.
`common_clean`	Cancel generated tarballs and temporary files to build it.
debian.mk	
`deb`	Generates all Debian packages, as described in section 2.4.
`deb_clean`	Cancel Debian packages and temporary files to build them.
apt.mk	
`apt`	Build and update the official Debian Kerrighed repository.
`apt_build`	Generates meta datas for apt repository.
`apt_sync`	Update packages and repositories informations in the official repository.
`apt_clean`	Cancel repositories meta datas.
rpm.mk	
`rpm`	Generates RPM packages as described in section 2.3.
`rpm_kernel`	Generates kernel related RPM packages.
`rpm_kerrighed`	Generates Kerrighed RPM packages.
`rpm_clean`	Cancel RPM files and temporary files to build them.
oscar.mk	
`oscar`	Generates SSI-OSCAR package and update the official repository.

Table 7: Building scripts