

# Kerrighed

Christine Morin, IRISA/INRIA  
PARIS project team  
[cmorin@irisa.fr](mailto:cmorin@irisa.fr)

Procope Meeting, May 3rd 2005

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



# ... Kerrighed V1.0.0 24/02/05 19:03

After many efforts and debugging, Kerrighed 1.0.0 has finally been released !



You can download it at : <http://www.kerrighed.org/download.html>

The next major release (1.2.0) will integrate hot node addition and removal, enhanced distributed file systems fonctionnalités.



## Fixes:

- \* Add permission check on commands "migrate" and "checkpoint".
- \* Avoid process migration on a non existing node !
- \* Fix up cluster wide PIDs problems for threads.
- \* Fix crash during process migration with containers in some corner cases.
- \* Fix a crash in some corner cases when reading a KerFS directory.
- \* Fix problem when migrating an opened but not attached socket.
- \* Fix a problem with migrated programs whose initial shell was opened before Kerrighed was started
- \* Fix numerous problems in checkpoint / restart mechanism.
- \* Fix a crash when reading /proc/<pid>/fd for a process having a large number of open files.
- \* Fix a crash when reading /proc/kerrighed/<pid>/segment for a daemon process
- \* Fix applications seg-fault

## Features:

- \* Support for "mv" in the distributed file system.
- \* krg\_capset now accepts a -k option to change and view capabilities of any process if permitted

# Kerrighed Prototype

Current stable version: Kerrighed 1.0.2 based on Linux 2.4.29 (5 April 05)

Last version based on Linux 2.4

parallel make and three industrial applications successfully executed

## Full SSI

global scheduling

distributed file system

process checkpoint/restart

memory sharing

# RoadMap

4

Kerrighed port to Linux 2.6 (on-going work)

Kerrighed on Linux 2.6 expected in June

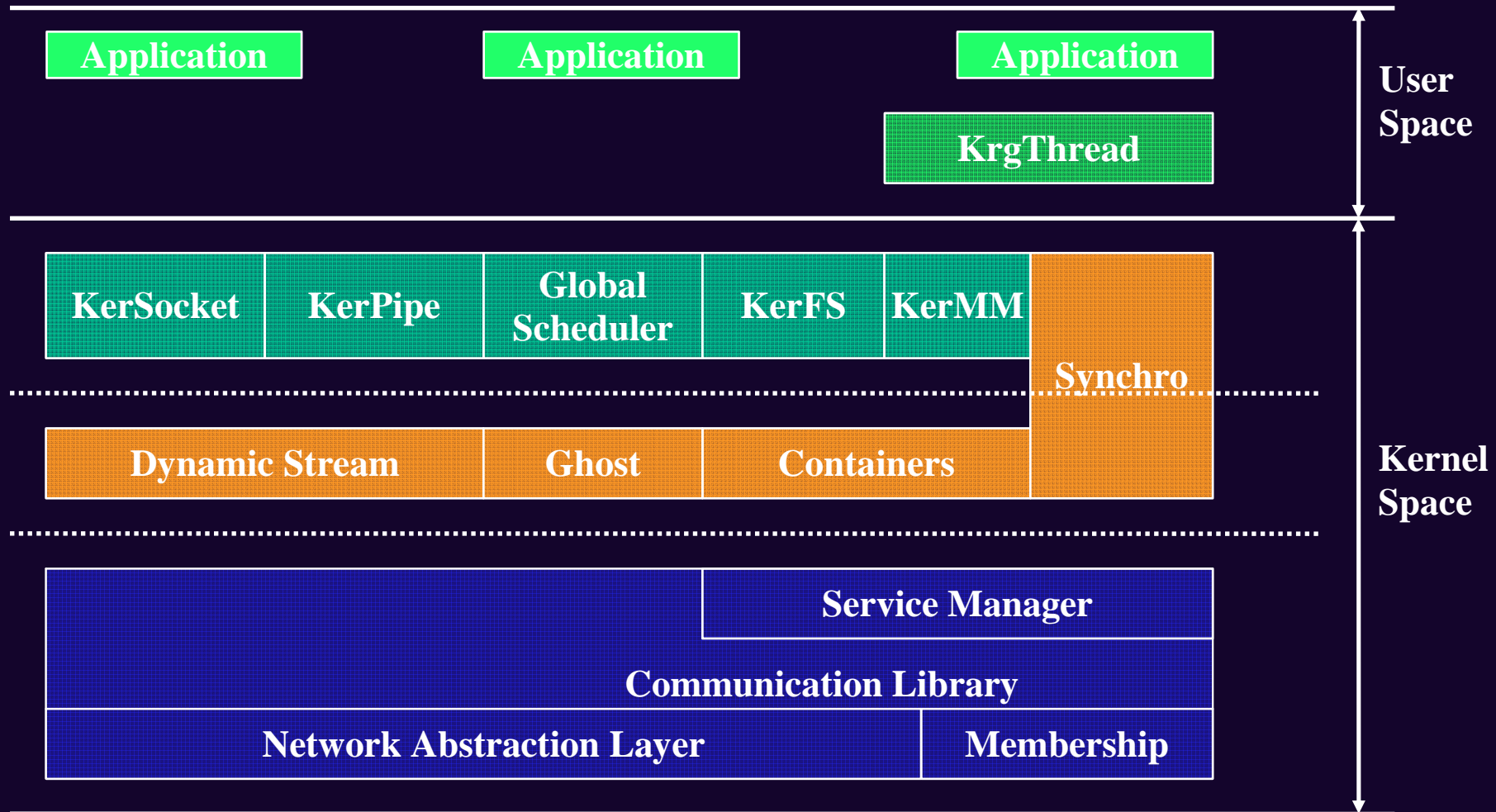
Kerrighed port on Opteron

Expected Summer 2005

Kerrighed for SMP nodes

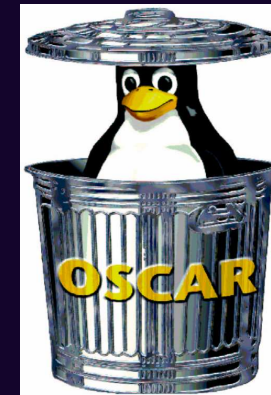
September 2005

# Kerrighed Architecture



# SSI-OSCAR

- OSCAR is a good vector to distribute Kerrighed SSI operating system for clusters
- INRIA industrial post-doc in collaboration with EDF R&D and ORNL (Geoffroy Vallée)
  - Creation of the SSI-OSCAR package
- SSI-OSCAR



<http://ssi-oscar.irisa.fr>

- First version based on Kerrighed V1.0rc8, OSCAR 3.0 for RedHat 9.0 distribution available since November 2004
- Second version based on Kerrighed V1.0.0, OSCAR 4.0 for RedHat 9 and Fedora core 2 (March 2005)
- Full integration in OSCAR to be done in the future (target OSCAR 5.0)

# Checkpoint/Restart for Parallel Applications

7

- Master student (Matthieu Fertré)
  - Integrating checkpointing mechanisms for parallel applications in Kerrighed operating system
- First step: global coordinated strategy
- Work on basic building blocks
  - Global process synchronization
    - identification of objects belonging to a given application
    - set of communicating processes
  - Checkpoint numbering and storage
- First basic prototype by end of June
  - Enhanced prototype for the end of the summer

# Node Addition and Eviction

## Goal

being able to add or stop a node at any time

## Advantages

easier cluster management

## On going work

Expected for the Summer

## Side product

Live CD (expected in August 2005)

# Grid aware OS

- Large scale data sharing (Louis' PhD)
- Resource allocation (Emmanuel's PhD, CIFRE EDF R&D)
  
- Initiative to launch an IP project (XtremOS)
  - IST Call 5

# Future of Kerrighed Open Source Software

10

Kerrighed should become one the reference cluster SSI OS for supporting HPC applications



- Leading Kerrighed evolutions
- Hosting Kerrighed Development Core-team
  - Software robustness
  - Packaging
  - Linux evolution
  - New hardware (processor, network)
  - Integration of research results
- Providing user Support
- Promoting Kerrighed
- Contribution to OSCAR and SSI-OSCAR
- Active to participate to Linux evolutions regarding clustering (osdl)