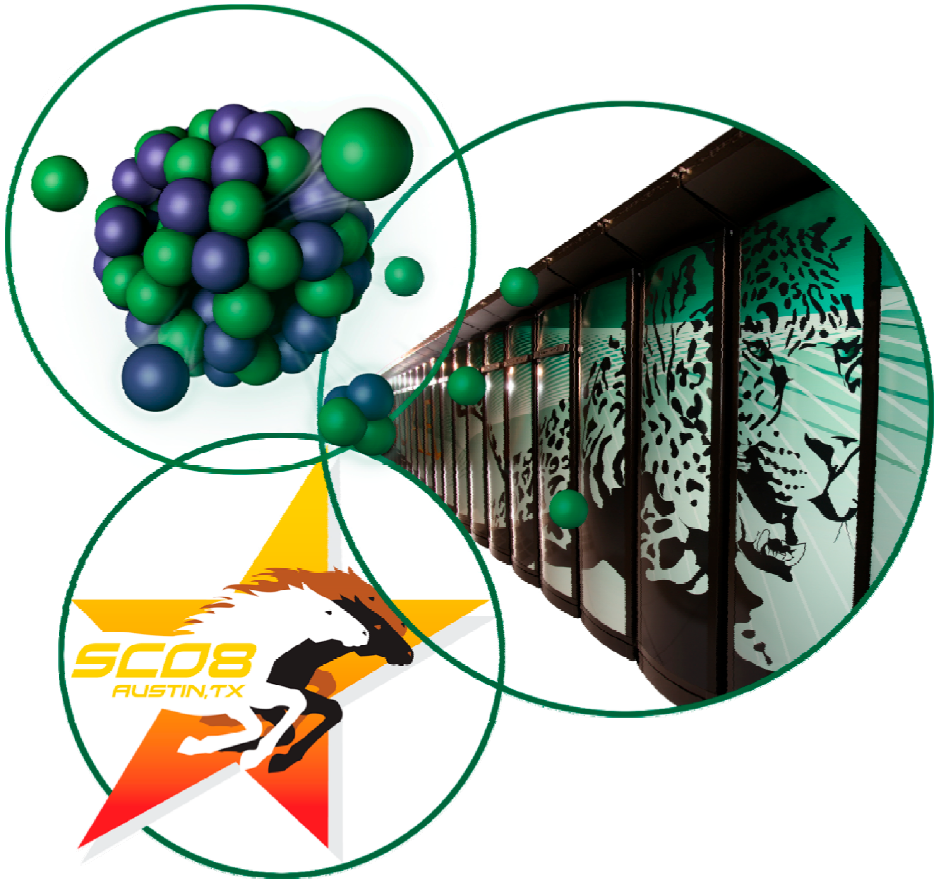
Open Source Cluster Application Resources (OSCAR)



Presented by

Stephen L. Scott
Thomas Naughton
Geoffroy Vallée

Computer Science Research Group
Computer Science and Mathematics Division

Open Source Cluster Application Resources

- Snapshot of best known methods for building, programming, and using clusters
- International consortium of academic, research, and industry members

OSCAR

INRIA

OAK RIDGE National Laboratory

Bald Guy SOFTWARE

KERLABS

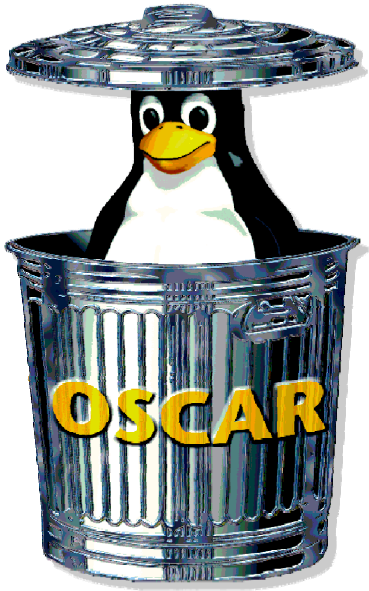
intel

NEC

LOUISIANA TECH UNIVERSITY

pervasive technology labs AT INDIANA UNIVERSITY

OSCAR background



- Concept first discussed in January 2000
  - First organizational meeting in April 2000
  - Cluster assembly is time consuming and repetitive
  - Nice to offer a toolkit to automate
- Leverage wealth of open source components
- First public release in April 2001
- Over 7 years of project development and 6 specialized versions
- Current stable: oscar-5.1; development: oscar-5.2

What does OSCAR do?

- Wizard-based cluster software installation
  - Operating system
  - Cluster environment
- Automatically configures cluster components
- Increases consistency among cluster builds
- Reduces time to build/install a cluster
- Reduces need for expertise

OSCAR design goals

Reduce overhead for cluster management	Leverage "best practices" whenever possible	Extensibility for new software and projects
• Keep the interface simple • Provide basic operations of cluster software and node administration • Enable others to reuse and extend system—deployment tool	• Native package systems • Existing distributions • Management, system, and applications	• Modular metapackage system/API—"OSCAR Packages" • Keep it simple for package authors • Open source to foster reuse and community participation • Fosters "spin-offs" to reuse OSCAR framework

OSCAR overview

Framework for cluster management

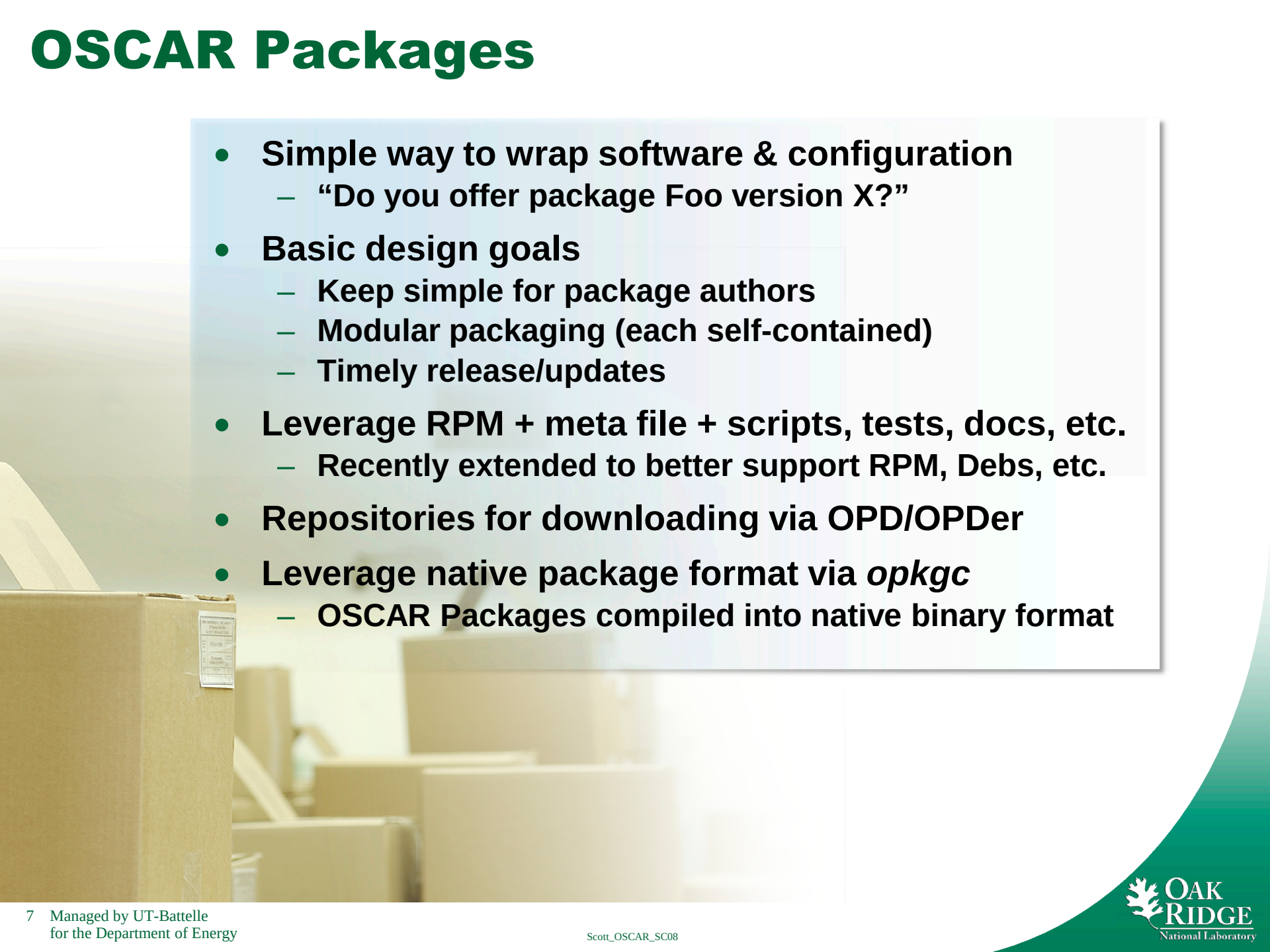
- Simplifies installation, configuration, and operation
- Reduces time/learning curve for cluster build
  - Requires preinstalled head node with supported Linux distribution
  - Thereafter, wizard guides user through setup/install of entire cluster

Package-based framework

- Content: Software + configuration, tests, docs
- Types:
  - Core: SIS, C3, Switcher, ODA, OPD, APItest, Support Libs
  - Non-core: Selected and third-party (PVM, LAM/MPI, OpenMPI, Toque/Maui, etc.)
- Access: Repositories accessible via standard binary package management tools
  - Ex. YUM, APT

OSCAR Packages

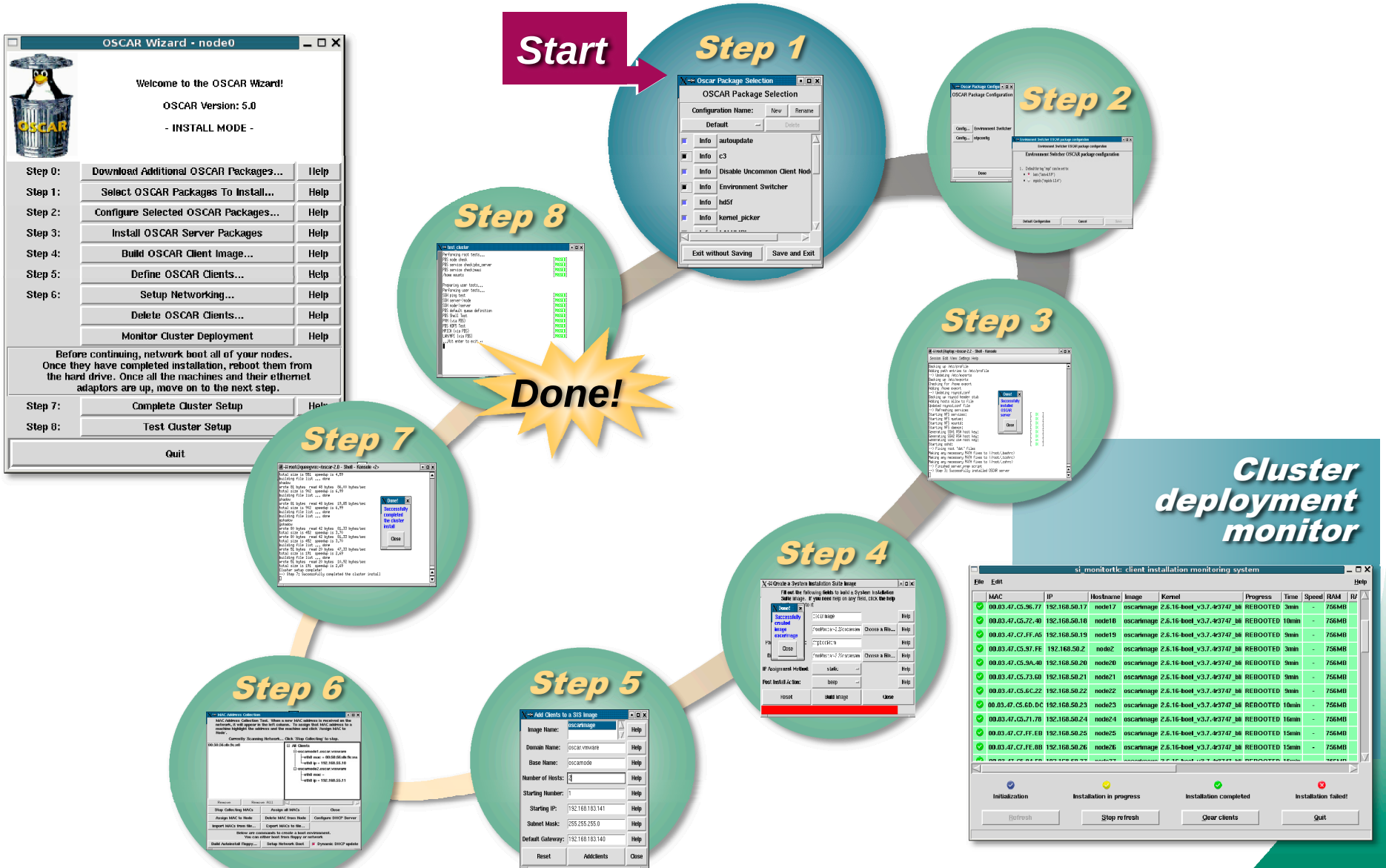
- Simple way to wrap software & configuration
  - "Do you offer package Foo version X?"
- Basic design goals
  - Keep simple for package authors
  - Modular packaging (each self-contained)
  - Timely release/updates
- Leverage RPM + meta file + scripts, tests, docs, etc.
  - Recently extended to better support RPM, Debs, etc.
- Repositories for downloading via OPD/OPDer
- Leverage native package format via *opkgc*
  - OSCAR Packages compiled into native binary format



OSCAR Packages (latest enhancements)

- Maintain versatilty and improve manageability
  - High-level opkg description
  - Use ‘opkgc’ to convert to lower-level native binary pkg(s)
  - Manage binary opkgs via standard tools (rpm/yum, dpkg/apt)
- Package repositories
  - Local repos for restricted access (all via tarball)
  - Online repos for simplified access (opkgs via yum/apt)
- Basis for future work
  - Easier upgrades
  - Specialized OSCAR releases (reuse oscar-core with custom opkgs)

OSCAR—cluster installation wizard



OSCAR components

Administration/ configuration

- System Installation Suite (SIS), Cluster Command and Control (C3), OPIUM, and cluster services (dhcp, nfs, ntp, ssh, etc.)
- OSCAR tools: oscar, oscar-config, yume/rapt, etc.

HPC services/ tools

- Parallel libs: MPICH, LAM/MPI, PVM, Open MPI
- OpenPBS/MAUI, Torque, SGE
- Ganglia, JobMonarch
- Other third-party OSCAR Packages

Core infrastructure/ management

- SIS, C3, Env-Switcher/Modules
- OSCAR DAtabase (ODA), OSCAR Package Downloader (OPD)
- OSCAR Package Compiler (OPKGC)

OSCAR: C3 power tools

- Command-line interface for cluster-system administration and parallel-user tools
- Parallel execution *cexec*
  - Execute across a single cluster or multiple clusters at same time
- Scatter/gather operations *cpush/cget*
  - Distribute or fetch files for all node(s)/cluster(s)
- Used throughout OSCAR
  - Mechanism for clusterwide operations

OAK RIDGE National Laboratory

OSCAR highlights

OSCAR Stable (v5.1)

- OSCAR Package Compiler (opkgc) in full use
- New database schema
- Supported platforms:
  - Fedora Core 7, 8, 9 (i386, x86_64)
  - Red Hat EL 4, 5 (i386, x86_64)
  - SuSE 10.2, 10.3 (x86_64)
  - Yellow Dog Linux 5 (ppc64)

In progress

- Merge of oscar-5.1 development branch with trunk
- Online repositories support for OSCAR installs
  - `ex. aptitude install oscar`
- New tools/changes to support online repositories
- New GUI
- Virtualization support

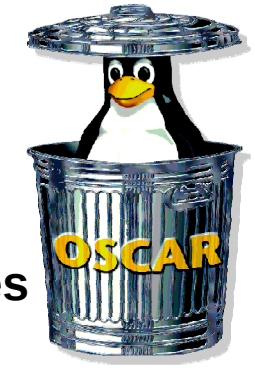
More OSCAR information...

Home page	oscar.OpenClusterGroup.org
Development page	svn.oscar.openclustergroup.org/trac/oscar
Mailing lists	oscar-users@lists.sourceforge.net oscar-devel@lists.sourceforge.net
Open cluster group	www.OpenClusterGroup.org
OSCAR symposium	www.csm.ornl.gov/srt/oscar09

OSCAR research supported by the

Office of Advanced Scientific Computing Research, Office of Science,
U. S. Department of Energy, under contract no. DE-AC05-00OR22725 with UT-Battelle, LLC.

Specialized OSCAR versions



- **OSCAR-V**
  - **Virtualization: Management of clusters containing virtual machines**
- **HA-OSCAR**
  - **High-availability: RAS management for HPC clustering**
- **OSCAR-Pro**
  - **Commercial: NEC's supported version for its HPC systems**
- **SSI-OSCAR**
  - **Single System Image: Management for Kerrighed SSI clusters**
- **SSS-OSCAR**
  - **Scalable System Software: Deployment of research components/prototype**
- **Thin-OSCAR**
  - **Diskless: Support for diskless HPC clusters**

Contacts

Stephen L. Scott

Computer Science Research Group
Computer Science and Mathematics Division
(865) 574-3144
scottsl@ornl.gov

Thomas Naughton

Computer Science Research Group
Computer Science and Mathematics Division
(865) 576-4184
naughtont@ornl.gov

Geoffroy Vallée

Computer Science Research Group
Computer Science and Mathematics Division
(865) 574-3152
valleegr@ornl.gov