

Open Source Cluster Application Resources (OSCAR)

Jaumin Ajdari

Head of Institutional Research Office

South East European University

Tetovo, R. of Macedonia

j.ajdari@seeu.edu.mk

Abstract

OSCAR is free open source software for installing and maintaining a high performance computing (HPC) cluster, especially of the Beowulf type. It contains everything needed to administer and program this type of HPC cluster. Beowulf cluster is scalable performance cluster based on commodity hardware, on a private system network, with open source software infrastructure. It consists of a cluster of PCs or workstations dedicated to running high performance computing tasks.

OSCAR is installed on the top of a standard installation of a supported Linux distribution. It installs and configures all required software for the selected packages according to user input. Then it creates disk images which are used to install nodes in the cluster with all client software and administrative tools needed for immediate use. The default OSCAR setup is generally used for scientific computing, using a message passing interface (MPI) implementation. openMosix and Rocks are similar.

We have used OSCAR to create a cluster of PCs for experimentation of the parallel programming with MPI paradigm.

Keywords: OSCAR, openMosix, Rocks, Linux, Cluster, Beowulf, HPC, MPI.

1. Introduction

Setting up a cluster can involve the installation and configuration of a lot of software as well as reconfiguration of the system and previously installed software. OSCAR (Open Source Cluster Application Resources) is a software package that is designed to simplify cluster installation. A collection of open source cluster software, OSCAR includes everything that you are likely to need for a dedicated, high-performance cluster. OSCAR takes you completely through the installation of your cluster [1].

OSCAR was created and is maintained by the Open Cluster Group (<http://www.openclustergroup.org>), an informal group dedicated to simplify the installation and use of

clusters and broadening their use. Over the years, a number of organizations and companies have supported the Open Cluster Group, including Dell, IBM, Intel, NCSA, and ORNL, to mention only a few.

OSCAR creation started in April 2000 (first OSCAR meeting), it was the first project of the Open Cluster Group (OCG) [2],[4], and the first version of OSCAR appeared in June 2001 as a middleware software for creating, maintaining and programming of High Performance Computing Clusters. The last available version is OSCAR 6.0.3 (released on May 27, 2009, <http://svn.oscar.openclustergroup.org>). There are a lots of middleware software for HPC clusters (<http://www.lcic.org/computational.html>) but OSCAR is one of them that contains all the needed components for one HPC cluster.

2. OSCAR Overview

OSCAR is software (a middleware) for creating, maintaining and programming of clusters especially clusters of workstations. OSCAR is a collection of free open source software needed for creating and using of cluster. Basically, it is designed to be used with an asymmetric cluster. Unless you customize the installation, the computer nodes are meant to be dedicated to the cluster. Typically, you do not log directly to the client nodes but rather work from the headnode. After installing the operating system (Linux, basic installation is enough) on the headnode, you first install OSCAR on the cluster's headnode, and then OSCAR installs the remaining machines (clients), from the headnode. During installation and configuration, OSCAR creates the client image which is a disk image for the client that includes the boot sector, operating system, and other software for the client. Since the headnode is used to build the client image, it is the home for most user services, and is used to administer the cluster. A graphical user interface (GUI) is available to assist with the installation. This GUI based OSCAR wizard consists of several steps that user walks through in sequence to perform necessary operations to build a headnode and the set of compute nodes. The necessary headnode services are configured during this process as well as

the number of nodes and what software to install on the cluster. Prerequisites and install procedure are in [1],[3].

3. OSCAR Components

OSCAR consist a lot of packages for work with cluster [5]. They are:

- **System Installation Suite (SIS)** – is a tool for installing Linux systems over a network. It is used in OSCAR to install the client nodes. SIS also provides the database from which OSCAR obtains its cluster configuration information. It is an image based install tool. Once this image is built, clients are defined and associated with the image. When one of these clients boots using a SIS auto-install environment, the corresponding image is pulled over the network. After installation, SIS also does configuring of the client node.

- **Cluster Command Control (C3)** – C3 tools are a suite of cluster tools that are useful for both administration and application support. The suite includes tools for cluster – wide command execution, file distribution and gathering, process termination, remote shutdown and restart, and system image updates (<http://www.csm.ornl.gov/torc/c3/>).

- **Switcher Environment Manager** – switcher package provides a simple mechanism to allow users to manipulate their environment. The switcher package provides a convenient command-line interface to manipulate the inclusion of packages in a user's environment so users are not required to manually edit their ".dot" files. The OSCAR switcher package provides a two – level set of defaults: a system – level default and a user – level default.

- **The OSCAR Password Installer and User Management (OPIUM)** – The OPIUM package includes facilities which synchronize the cluster's accounts and configures ssh for users. OPIUM configures ssh such that every user can traverse the cluster securely without entering a password, once logged on to the head node.

- **Torque Resource Manager** – is a flexible workload management system based on OpenPBS. The Torque resource manager consists of three major components:

- i. *The Torque server* – This runs on the OSCAR headnode, controls the submission and running of jobs and also tracks the current state of cluster resources.

- ii. *A "mom" daemon on each cluster node* – responsible for actually starting and stopping jobs on the client nodes.

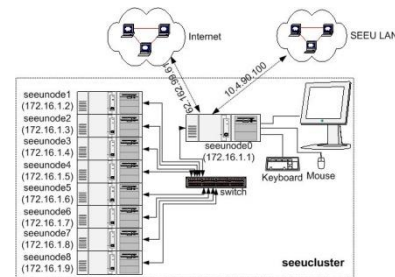
- iii. *The scheduler* – Torque has a first-in-first-out scheduler which it can use. By default OSCAR uses the Maui Scheduler as it is more flexible and powerful.

- **LAM/MPI, MPICH, OpenMPI and PVM** – are the standard interfaces for high performance parallel programming in the message passing paradigm. With those tools, a dedicated cluster or an existing network computing infrastructure can act as a single parallel computer.

- **Ganglia Monitoring System** – Ganglia is a scalable distributed monitoring system for high performance computing systems such as clusters. The base system as configured by OSCAR includes a monitor daemon on each node and a server on the headnode gathering the information. The monitoring system offers a very nice web based GUI to browse historical data and view usage information for the cluster (<http://ganglia.info/>).

4. SEEU cluster

We used OSCAR to build an experimental cluster of type Beowulf named seeucluster. The cluster address is seeucluster.seeu.edu.mk. Our cluster is constructed over the network of 9 PCs and it is defined as an asymmetric cluster.

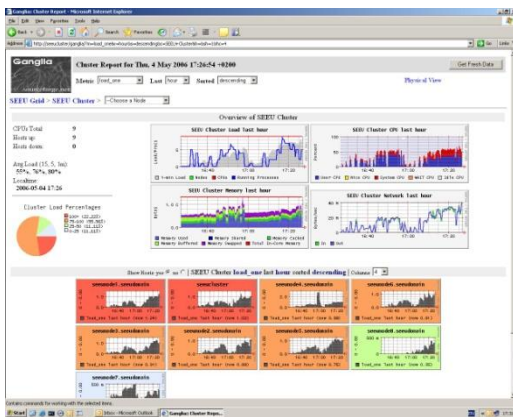


The used hardware and software is:

- i. 9 PC Net Vista IBM A30 (P4 1.5 GHz, RAM 128 MB, HDD 20 GB, Net card Intel Corp.82801BA 100 Mb)
- ii. One switch Cisco System SR224 24 ports 10/100 Mb
- iii. One KVM switch DKVM-8E 8 ports
- iv. Monitor, keyboard and mouse
- v. Linux Red Hat Enterprise 4 AS
- vi. OSCAR 4.2



The seecluster is used for parallel programming using LAM/MPI paradigm. Monitoring of the seecluster is shown with the web interface <http://seecluster.seeu.edu.mk/ganglia>.



5. Conclusion

During installation, testing and use of the seecluster we gained impression that OSCAR is a very powerful package for clustering and especially for cluster of Beowulf type. It is a simple for use and allows anyone who is interested in clustering and/or parallel programming to quickly and easily deploy a medium – size cluster. In the created cluster easily can be added or removed nodes and it can be easily extended or reduced. The modularity of OSCAR allows researchers and companies to easily develop new cluster without any encumbrance in installation, configuration and maintenance.

6. References

- [1] Joseph D. Sloan, High Performance Linux Clusters with OSCAR, Rocks, OpenMosix, and MPI, O'Reilly, November 2004
- [2] Richard Ferri , The OSCAR Revolution, June 1st, 2002, <http://www.linuxjournal.com/article/5559>
- [3] OSCAR Administrator's, November 12, 2006, <http://oscar.openclustergroup.org/public/docs/oscar5.0>
- [4] Benoit des Lignerisa, Stephen Scott, Thomas Naughton, Neil Gorsuch, OSCAR design, implementation and interest for the scientific community. *First OSCAR Symp*, May 2003, <http://hpcs2003.ccs.usherbrooke.ca/papers/>
- [5] Administration Guide, <http://svn.oscar.openclustergroup.org/trac/oscar/wiki/AdminGuide>