OSCAR surpasses the 100,000 download mark!

Stephen L. Scott [*]; Oak Ridge National Laboratory

Summary

Open Source Cluster Application Resources (OSCAR) – the self-installing software stack for high-performance cluster computing surpasses the 100,000-download milestone. OSCAR is a Linux based cluster computing distribution that does for the high-performance cluster computing community what Linux distributions have done for the workstation community – including: bundled comprehensive software stack, open source, high quality, low (no) cost, best practices, consistency, and community support.

The Open Source Cluster Application Resources (OSCAR), represented by the "penguin-in-the-pail," is a self-installing cluster software stack that provides a complete infrastructure for the installation, maintenance, and operation of high-performance computing (HPC) clusters by leveraging today's open source "best practices" for Beowulf computing while fostering the development of new technologies where needed. During the month of April 2003, the cumulative download count, beginning with the first tracked downloads in June 2001, has surpassed 100,000. (The first public release was April 2001, but download counts did not begin until June.) As the number of downloads attests to, in approximately two years, OSCAR has become one of the (if not the) most popular HPC cluster computing software stack today with a worldwide acceptance.

OSCAR is the flagship offering by the consortium of industry, academia, and research-lab collaborators forming the Open Cluster Group (OCG) unified with the goal of simplifying the installation, operation, and use of HPC clusters. Oak Ridge National Laboratory (ORNL) has been a key player in the success of OSCAR from its inception in January 2000. In addition to ORNL, current consortium core members include: Dell, IBM, Intel, MSC Software, BGS, Indiana University, NCSA, and Sherbrooke University.

The installation of a computing cluster typically involves the installation and configuration of a headnode, which in turn provides services to the compute nodes. Following headnode setup, the compute nodes must be built. This compute node building step is the part of a cluster build that is time consuming and error prone due to the massive replication of operations required for each compute node. OSCAR assists with the headnode configuration and then builds the compute nodes based on a user specified description of the desired software packages to include. This build and configuration process greatly reduces the effort and expertise needed to setup a cluster while providing consistency of build across each node.

The standard OSCAR release targets the typical HPC cluster user. However, the base OSCAR framework is not necessarily tied to only HPC environments. This framework

[*] 865-574-3144, scottsl@ornl.gov

consists of the requisite *core* packages needed to build and maintain a cluster, *included* packages are those commonly used in HPC applications, and the *third-party* designation is used for all others.

The core components enable a user to construct a virtual image of the target machine (cluster node) using System Installation Suite (SIS). There is also an OSCAR database (ODA) that stores cluster configuration information. The final two core packages include a parallel-distributed "shell" tool developed at ORNL called Cluster Command & Control (C3) as part of the ORNL Cluster Power Tools and an environment management facility called Env-Switcher. Using these packages, OSCAR is able to simplify the tasks to install, upgrade, and maintain both the cluster hardware and software infrastructure.

A number of tools are provided within OSCAR to assist with cluster administration, including the previously mentioned SIS and C3 packages. The OSCAR Password Installer and User Management (OPIUM) utility synchronizes user account files across the cluster. The standard user *add* and *delete* commands are wrapped so that they are cluster aware and their effect is delivered cluster wide. A Kernel-Picker tool allows one to substitute a given operating system kernel into the SIS image prior to building nodes. This provides one the flexibility to customize the operating system on a cluster. OSCAR also helps with the configuration of standard cluster services such as the Network File System (NFS) enabling the sharing of files from any node and the Network Time Protocol (NTP) service used to keep a consistent time throughout the cluster as well as synchronizing the headnode to one of the Internet atomic clocks. Another useful configuration is performed by the Loghost package in forwarding the system log (syslog) entries from each compute nodes to the headnode.

HPC services and tools available with configuration for an OSCAR cluster include the Ganglia monitoring system, Hierarchical Data Format (HDF5), the parallel message-passing libraries of Parallel Virtual Machine (PVM) and the MPIs, LAM/MPI and MPICH, as well as the scheduling duo of OpenPBS and MAUI. Pfilter and OpenSSH provide for cluster security. Pfilter is configured on each system to enable one machine, the headnode, to provide packet forwarding while all compute nodes are "behind" it on a private network, allowing the protected machines to access the outside network in a safe manner. OpenSSH enables remote connections to the cluster in a safe manner. The authentication and encryption schemes incorporated can be adjusted for less or more stringent security needs.

Since early 2001 there have been approximately eight major releases. Each of these releases has seen OSCAR evolve and improve. The current development path includes an overhaul of the base installation harness; a new facility to provide a uniform interface to add, delete, and update software; an embracing of the Linux Standards Base (LSB) to provide better multi-distribution support; as well as further exploration into increasing the scalability of the OSCAR distribution and individual components. In the past year, OSCAR has begun to expand into other cluster paradigms including diskless clusters (Thin OSCAR), and high-availability (HA-OSCAR), embracing fault tolerant telco-like capabilities.

For further information on this subject contact:
Dr. Fred Johnson, Program Manager
Mathematical, Information, and Computational Sciences Division
Office of Advanced Scientific Computing Research
Phone: 301-903-3601
fjohnson@er.doe.gov