High Performance Computing Rocks at NYU

A Look at Rocks Cluster Distribution for HPC Researchers

By Will Wilson, Yingkai Zhang, and David Ackerman
will.wilson@nyu.edu; yingkai.zhang@nyu.edu; david.ackerman@nyu.edu

A Linux cluster is composed of two or more computers, each running the Linux open source operating system and connected in such a way that they collectively perform like a more powerful computer. This flexible, scalable, and cost-effective configuration is rapidly becoming the dominant system employed by researchers requiring high performance computing (HPC). In an article entitled "Linux Clusters for the Mainstream Manager,"[1] Sean Dague of the IBM Linux Technology Center says Linux clusters are "like 1,000,000 ants vs. one elephant." But keeping a million ants in lock step can be a trying task if not done properly. To simplify the process, the NPACI Rocks Cluster Distribution, based on Red Hat Linux, pulls together the best of open source software to make clusters easy to "deploy, manage, upgrade and scale."[2]

The Rocks Cluster Distribution delivers a stable HPC platform by uniting Linux with low-cost commodity hardware. The growing benefit of using such hardware to tackle HPC tasks is due to a price/performance advantage over more expensive shared memory machines. That advantage, however, can disappear quickly if system administrators get bogged down with maintaining a large number of nodes. Rocks employs a clever technique to avoid this situation by making a complete operating system installation the basic management tool. An automated installation process is far more efficient and effective than an alternative process that, for example, involves tracking down nodes that are out of synch and require patching. Rocks leverages the automated installation methods of Red Hat's kickstart to install nodes and allows systems administrators to bring up a cluster in a relatively short time.

The physical assembly of a Rocks cluster requires network connectivity, two or more computers—a front-end node and at least one dedicated compute node—and a sturdy rack (or racks) to house the cluster. NYU's Information Technology Services is currently running two Rocks cluster installations, one built around 16 Dell Xeon servers with GB Ethernet (see inset on next page), and the other running on eight dual processor AMD Opteron servers with both GB Ethernet and low latency Myrinet. Each of our Rocks clusters is configured with a single front-end node (where cluster users login, submit, and monitor their jobs) and several compute nodes.

The many services required to manage a Linux cluster—NFS, NIS/411, DHCP, NTP, MySQL, HTTP, to name a few—are run on the front-end node. This node is also responsible for kickstarting or automatically installing the compute nodes. By default, the front-end also acts as the gateway to the outside, since it is the only node with an active external interface. The front-end node requires an experienced systems administrator to maintain the required services and to perform the administrative tasks that multi-user systems typically require—e.g., assigning accounts, performing software installs and configurations, and so on.

The compute nodes are the "workhorses" of the cluster.

1. Jacqueline Emigh, *EarthWeb*, September 25, 2003, *http://networking.earthweb.com/netsysm/article.php/3083551.*
2. *http://rocks.npaci.edu/papers/rocks-documentation/preface.html*

The CPU-intensive calculations researchers submit are run on the compute nodes. The data from compute node calculations is collected on the front-end by way of an NFS auto-mounted file system.

Rocks maintains a MySQL database for the cluster configuration files. Changes made to the database are used to generate Linux configuration files, and these files are pushed out to the compute nodes during the kickstart process. An Apache server on the front-end gives a system administrator easy access to the MySQL database and the Ganglia cluster monitoring software. To access the management or monitoring services, a system administrator can simply start a Mozilla browser on the front-end node.

The key to the Rocks Cluster Distribution is its ability to rapidly deploy numerous nodes with quick, automated installations (less than ten minutes per node). This method helps maintain stability among the nodes, and scales very well when expanding the cluster. See the inset below to learn more about how ITS can help you use a Linux cluster and Rocks to facilitate your research. For more information about Rocks, see *http://rocksclusters.org/.*

Will Wilson is a Senior Systems Administrator in ITS e*Services; Yingkai Zhang is a professor of Chemistry in NYU's Faculty of Arts and Sciences; David Ackerman is Executive Director for ITS* e*Services and Digital Library Initiatives.*

ITS Offers HPC Services to NYU Researchers

If your research requires High Performance Computing, Information Technology Services can help you. If you are writing a grant proposal, we can advise you on the infrastructure aspects of the grant. ITS can help you architect, select, and negotiate the price for your computational platform and network.

We have extensive experience with both clusters and large memory systems and can host, set up the hardware, and configure your software and application. We handle the day-to-day management of the system, including daily backups and security monitoring, and will perform maintenance during scheduled off-hour periods that will be announced to you well in advance.

We work 24x7x365 to ensure the highest possible levels of system availability. ITS will supply you with a service level agreement (SLA) that will spell out the services we provide, including our system availability guarantee.

We think you will enjoy working with ITS' highly professional and knowledgeable staff. By availing yourself of ITS HPC services, you will be free to spend more time on your research!

For more information, please contact ***hpc@nyu.edu***.

Case Study

In December 2003, ITS successfully used the Rocks Cluster Distribution to collaboratively set up a Linux Xeon cluster (16 nodes, 32 processors) with NYU Chemistry Professor Yingkai Zhang, whose research involves the computer simulation of enzyme reactions.

Inadequate computational power has been the major bottleneck for his group's research productivity. With the Rocks Distribution, this cluster becomes an attractive option because of its excellent price/performance ratio.

The cluster is now managed by ITS with the Rocks Distribution, and is stable and productive.