Training a Grid Workforce

Dale R. Thompson*, Amy Apon*, Yuriko Yara*, Jens Mache$, and Russell Deaton*
*University of Arkansas, $Lewis & Clark College

Motivation

The growing capability of Grids as viable compute resource brokers is largely responsible for their acceptance beyond the traditional high performance computing (HPC) research community. As applications become more grid-enabled, the business community is expected to increase significantly its investments in Grid Computing over the next decade. These increasing demands for Grid Computing, coupled with continued advancements in middleware and networking technologies, have raised concerns about the availability of a qualified workforce to build and use the Grid.

Approach

An experiment to connect two institutions on the grid

Ingredients:
- •One undergraduate from Lewis and Clark College
- •One graduate student from the University of Arkansas
- •Funding to support the students from Lewis & Clark College and the University of Arkansas
- •Four eager faculty not on summer support
- •Several PC's and some Fast Ethernet switches
- •One month of time
- •Free software packages
- •Encouragement

Goal – connect the two institutions using Grid protocols, and run a Grid service and an MPI program across the Grid in one month.

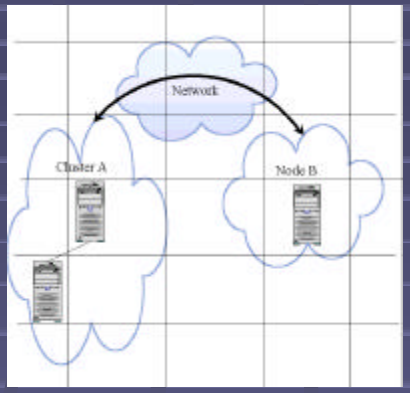
Result – a much better understanding of the knowledge and training required to build and use the Grid!

Software Packages Used

•The Globus 3.0 Toolkit (version 3.0.1)
•Java 2 Platform Standard Edition (j2sdk-1.4.2)
•OSCAR (version 2.2.1)
•MPICH-G2 (version 1.2.5-1)
•RedHat Linux 7.3, with default configurations

At the time of this work, these tools were the latest releases that would interoperate with each other.

Diagram of Experimental System



What is the Grid?

• The Grid allows sharing of computational and data resources across diverse platforms and different organizations.

• The Grid uses the Internet for worldwide communications.

•The Grid provides worldwide computation.

Four levels of Grid

Applications
Application Tool Kits
Grid Services
Grid Fabric

Future Grid workers must have knowledge in several specialized areas

Knowledge of Linux Administration

• the ability to configure the network
• the location of various configuration files
• software installation and setting up permissions
• the use of tools such as ifconfig, iptables/ipchains for firewall configuration, rpm, make, and ant.

Ability and Knowledge for Solving Common Difficulties

•Establishing trust at the host level and the user level
•Synchronizing clocks - clocks should be within 5 minutes or else "handshake failed" is obtained
•Configuring firewall to permit Globus port numbers

User-level Initialization Instructions

Initialize environment
%cd $GLOBUS_LOCATION
% source setenv.sh
%$GLOBUS_LOCATION/etc/globus-user-env.sh
Request user certificate
%grid-cert-request
Sign user certificate (you must be administrator or CA)
%grid-ca-sign -in usercert_request.pem -out usercert.pem
Start a valid proxy
%grid-proxy-init
Delete a proxy
%grid-proxy-destroy

Knowledge of Order for Installing Software

•ANT
•jakarta-oro
•jdk
•bison
•Junit
•Globus 3
•SimpleCA
  –CA and add trust
  –Request Host certificate
  –Sign user certificates

User-level Globus commands

globusrun - submits jobs specified in Resource Specification Language (RSL)
globus-job-run - used for running short "interactive" jobs in the foreground
globus-job-submit - used for running batch programs. It is a batch interface to the GRAM server.
globus-job-status - gets status of a job previously started with globus-job-submit
globus-job-output - collects output of a job started with globus-job-submit
globus-job-clean - stops jobs and cleans up cached output
Using gridFTP
globus-url-copy <sourceURL> <destinationURL>
**%globus-url-copy gsiftp://node1/home/USERNAME/file **
file:/home/USERNAME/file

Activities that Build Knowledge

- Install Linux on a PC
- Install the Globus Toolkit
- Build and use a Grid Service
- Build a cluster using PC's and OSCAR
- Run a simple MPI program across the Grid

A website is being developed for Grid training and access to Grid training activities and materials at http://gotgrid.uark.edu/

D. R. Thompson, A. Apon, Y. Yara, J. Mache, and R. Deaton, “Training a grid workforce,” poster at *Oklahoma Supercomputer Symposium*, Norman, OK, Sep. 25, 2003.