



UMIACS

University of Maryland Institute for Advanced Computer Studies

The Lattice Project

A Computational Grid System

Presented by Adam Bazinet

What It Is

- The Lattice Project is an attempt to effectively share computational resources among departments and institutions, starting with those in the University System of Maryland.
- The Grid is focused on **computation**, and we have not yet made efforts to enable large-scale data access, storage, or replication.
- The Grid has only been used by a handful of researchers thus far. We expect this to change as the system grows and improves.

Grid Software

- We make heavy use of the Globus Toolkit, which forms the backbone of our Grid system. It provides mechanisms for job submission, file transfer, and authentication and authorization of Grid entities, to name a few things.
- We have also done extensive work with BOINC, which enables public participation in the Grid and represents a potentially huge resource. We have developed software that allows Globus, (and hence our Grid system), to submit jobs to a BOINC pool.
- We also work with scheduling software, such as Condor and PBS, that controls local resources. Such software is being deployed where it is most appropriate.

History

- Our previous Grid system was based on Globus Toolkit 3. We also incorporated BOINC, Condor, and Condor-G, and made use of this system well into 2005. Applications run on this system accumulated **over 100 CPU years** of wall clock processing time.
- Starting Summer 2005, we began our efforts to transition the Grid system to Globus Toolkit 4. We took this opportunity to change many different aspects of the system.
- As we continue to develop and expand the system, we have made efforts to involve the community through various Grid Committees, collaborations, and presentations like this one.

Grid Services

- An application can be Grid-enabled and made into a Grid service. These are applications that we trust and then make available to execute on Grid resources.
- To date, over **20** Grid services have been written. Most of these are applications used in life sciences. Of these, only a handful have actually been used a significant amount.
- We have developed a software stack which allows us to create new services quickly and easily. This includes GSBL (Grid Services Base Library) and the GSG (Grid Services Generator).

User Interfaces

- Our primary Grid interface is command-line based. Grid users log on to a specific machine, usually upload whatever input files they need, then submit and monitor jobs using our tools.
- We also provide a Web interface for monitoring job status, which is located on The Lattice Project intranet.
- Future work might see job *submission* take place via a Web interface, which could certainly simplify use of the system. We may also make the command-line interface more widely available.

Projects

Cummings Lab – *gsi* – an application written in R

Maile Neel & Joanna Grand – MARXAN

Holly Mortensen & Floyd Reed – MDIV & IM

Fushman Laboratory – CNS

Tree Of Life Project: *Lepidoptera* - GARLI

Grid Resources

- Much effort has recently been made to federate departmental resources using Condor. Each Condor pool is an example of a Grid resource. We recently integrated a dedicated cluster running PBS into the Grid.
- Condor pools are functioning in UMIACS, PSLA, and CLFS. OIT runs a Condor pool and the aforementioned PBS cluster.
- The effort to expand resources on and off campus is ongoing. At UMCP, we are integrating desktops in the School of Business and a cluster in the Engineering department. We are also working with UMBC to integrate one of their clusters into the Grid.

Current Development

- Scheduler
- BOINC
- Various Job Types

Scheduler

- The scheduler is responsible for matching a job with an appropriate resource. Simply stated, a scheduler must be informed about the state of resources. Furthermore, not all types of jobs will run on all types of resources.
- In our GT3-based system, Condor-G, MDS2, and scripts of our own constituted the scheduler component.
- In the GT4-based system, MDS4 delivers information about remote resources. The matchmaking logic once provided by Condor-G is now provided by GSBL, software on the Grid server.

BOINC

- We are working on a re-launch of the Lattice “BOINC” Project. We hope to eventually have thousands of volunteers contributing their resources to the Grid.
- This requires setting up a new server running the latest version of the BOINC software (version 5). We also create a Web site where public participants sign up for the project.
- The BOINC job manager, our interface between Globus and BOINC, needs to be made compatible with GT4 and BOINC 5.

Various Types of Jobs

- Our Grid system currently supports serial job submission and also has limited support for submitting batches of identical jobs.
- Most of our analyses are parameter sweeps, however, so we would like to add the capability to describe a batch of jobs with varying program input files or varying parameters.
- Our PBS resources are MPI-enabled, and some of our Condor resources support application checkpointing. We would like to give our users the option to run MPI jobs or checkpointing jobs on these resources.

Adding a Grid Resource

- First, a resource must be identified. This could be a computing lab, a cluster, or a number of desktops.
- A local resource manager such as Condor or PBS should be installed to control the local resource.
- We then install Globus software on a machine in the pool (usually a machine dedicated for this purpose), which enables the newly established pool to receive jobs submitted from the Grid.

Lattice Project Web Site

<http://lattice.umiacs.umd.edu>