

# *The Lattice Project*

Our Efforts at Grid Computing

Seeded, Grown, and Watered by the Laboratory of  
Michael Cummings

Soon to be a Thriving Tree at UMD

Presented by Adam Bazinet



## *What It Is*

The Lattice Project is an attempt to effectively share computational resources, primarily those at UMD.

We are building a Grid for *computation*, and have not made efforts to enable large-scale data access, storage, or replication.

Though the Grid has been primarily used by scientists and researchers, we eventually expect students and faculty to be using it too.

---

---

# *Grid Software*

Our initial research focused heavily on using the Globus Toolkit. The Globus Toolkit forms the backbone of our Grid system.

We have also done extensive work with BOINC, which enables public participation in the Grid and represents a potentially huge resource.

Finally, Condor has proved itself to be an excellent local resource manager. We are making efforts to deploy this on campus wherever possible.

---

---

# History

Our first Grid system was based on Globus Toolkit 3. We incorporated BOINC, Condor, and Condor-G, and made use of this system well into 2005. In total, Grid projects racked up **over 100 CPU years**.

Starting Summer 2005, we began our efforts to migrate the Grid system to Globus Toolkit 4. This was more like a complete overhaul.

As this new system continues to develop and expand, efforts have been made to involve the rest of campus through the Grid Steering Committee.

---

---

# *Grid Services*

An application can be Grid-enabled and made into a Grid service. These are codes which we trust and 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.

Some desire has been expressed to run arbitrary code on the Grid. Our architecture does not preclude us from making this possible in the future.

---

---

# *User Interfaces*

Our primary interface for job submission and monitoring is command-line based.

We have also provided a Web interface for monitoring job status on the Lattice intranet.

Future work would also see job *submission* take place via a Web portal. In addition, making the command-line interface available throughout campus has been discussed.

---

---

# *Resources*

Much effort has been recently made to federate campus resources using Condor. Each Condor pool is an example of a Grid resource.

Condor pools are functioning in UMIACS, NRSL, and CLFS, as well as on OIT controlled resources.

[see handout for more information]

---

---

# Projects

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

Maile Neel & Joanna Grand – MARXAN

Holly Mortensen and Floyd Reed – MDIV & IM

Fushman Laboratory – CNS

Tree Of Life Project: *Lepidoptera* - GARLI

---

---

# Status

Our current system is a blend between development and production. It is being used to get real work done, and we continue to test and improve it.

The effort to expand resources on and off campus is ever-ongoing. Prospective off-campus resources include Xseed at Bowie State, and a foray into public computing using BOINC.

We will continue to use the system, aggregate resources, recruit sources of work for the Grid, and educate the community.

---

---

# *Current Development*

- Scheduler
- BOINC
- Batch Jobs and MPI Jobs



# 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. Also, not all jobs will run on all resources.

In our GT3-based system, Condor-G, MDS2, and scripts of our own constituted the scheduler.

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 plan to re-launch the Lattice “BOINC” Project. If public trust and interest can be earned, then we could quickly have thousands of new contributors.

This will require setting up a new server running the latest version of the BOINC software. The BOINC web site also needs to be revamped.

The BOINC job manager needs to be updated to be compatible with GT4. If the BOINC API has changed, our compatibility library might also need updating.

---

---

## *Batch Jobs and MPI Jobs*

Our Grid system currently supports serial job submission. Many of our analyses are parameter sweeps, however, so submitting jobs as a batch could save time and disk space. GT4 provides some support for describing a “multi-job”, which is a start.

Some of our resources are MPI-enabled. We would eventually like to be able to send these resources MPI jobs through the Grid.

There is an obvious need to support batch job submission, as well as MPI job submission.

---

---

# *What's the Deal with R?*

R is a free software environment for statistical computing and graphics. It is exceptional in that it must be pre-installed. Most other applications can be transferred to the appropriate resource.

It is conceivable that other applications could fall into this class. Thus, effective procedures for deploying such programs on a large number of machines must be established. These procedures need to be standardized and made available.

---

---

# *Joining the Grid*

First, a resource must be identified.

Condor must be installed on each participating machine. One machine functions as the Condor collector and controls the other nodes.

It has been our practice that this machine also runs Globus. To date, I have been in charge of installing and configuring Globus, though this process will eventually have to be standardized.

---

---

# *Conclusions*

A tremendous amount of progress has been made.

At this point, we are not limited by lack of knowledge or expertise. Progress is only limited by lack of human resources. I see this as the chief problem facing us today.

That said, much of the work is done on a volunteer basis. We appreciate your continued support.

---

---

# *Lattice Project Web Site*

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

---

---