Introduction to Teaching Grid Computing

Dr. Clayton Ferner  
University of North Carolina  
Wilmington

Dr. Barry Wilkinson  
University of North Carolina  
Charlotte

OSG&C 2008  
Open Source Grid & Cluster Conference 2008

Thursday May 15, 2008
Outline

- Our Grid Computing course in North Carolina
- What we changed in 2007
- Explanation of Assignments
- Lessons Learned
- Future Improvements
Grid Computing Course

- Taught on North Carolina Research and Education Network (NCREN) that connects all 16 state campuses and also private institutions
  - Fall 2004: 8 sites
  - Fall 2005: 12 sites
  - Spring 2007: 3 sites (Experimental re-designed course.)
Grid Computing Course

- Undergraduate/graduate
- Hands-on with distributed grid infrastructure
- Teleconferencing facilities - students and faculty at many institutions participating
- Expert guest speakers near end of course
- Probably first such course for undergraduate students and so many distributed sites using large-scale teleconferencing facilities and a truly distributed grid infrastructure.
Grid Computing Course
Problem with previous offerings

- In 2004 and 2005, the course was taught from a bottom-up perspective
  - We started with Web services then moved on the Grid services
  - Everything was command-line
  - Only toward the end of the semester did we introduce a workflow editor and schedulers (such as Condor or Sun Grid Engine)
Changes to Grid Course in 2007

- In 2007 we changed the course to be a more top-down perspective
  - Although it might best be described as alternating between high-level and low-level view of Grid Computing
  - We started the course with using tools (i.e. a portal) that a typical Grid User would use
Assignments in 2007

• **Assignment 1** Using Grid computing portal
• **Assignment 2** Using Grid through a command line
• **Assignment 3** Using a scheduler (Condor-G)
• **Assignment 4** Installing GT4 core. Creating, deploying, and testing a GT4 Grid service
Assignments in 2007

- **Assignment 5** Installing and using GridNexus workflow editor to create and execute workflows
- **Assignment 6** Implementing a portlet with OGCSE2/Gridsphere portal
- **Assignment 7** MPI assignment on Grid
- **Mini-project assignment** Developing Grid computing

Assignments 4, 5, and 6 required students to install significant software packages on their computer.
Assignment 1 (Using a Grid Portal)

New User

Fill in form
Provide password and other information

PURSe on-line registration form

Email
- Request Confirmation
- Acknowledgement

CA/System Administrator

Create accounts, set access control, sign certificate, …
Assignment 1 (Using a Grid Portal)

• Students filled out request for account
  - Certificate Request generated and sent to Certificate Authority automatically

• Certificate Authority signed certificate, installed it in the MyProxy server, emailed confirmation to student

• Student logged onto Portal, acquired a proxy, submitted a simple job, created a java program, transferred and ran that program
Course portal (OGCSE2/Gridsphere)

Portal provides single sign-on to all grid resources.
Assignment 2
(Using command-line)

- Students performed the same tasks as in Assignment 1 but used a command-line
  - Install ssh client (e.g. putty) if necessary
  - Set up your credentials (grid-cert-request)
  - Email Certificate Authority to sign certificate
  - Install certificate and create proxy
  - Submit various jobs with globusrun-ws
  - Some jobs submitted with xml job description files
Assignment 3 (Using Condor)

• Create proxy
• Check the status of the Condor pool
• Create a test submit description text file
• Submit jobs and check status
• Using different Condor universes
• Still a command-line interface
• In 2005, we used Sun Grid Engine, which has a GUI interface
### Assignment 4 (Creating and deploying a Grid Service)

<table>
<thead>
<tr>
<th>Install GT 4 core and associated software</th>
</tr>
</thead>
<tbody>
<tr>
<td>- JDK 1.4.2+</td>
</tr>
<tr>
<td>- Ant 1.5.1+</td>
</tr>
<tr>
<td>- Python 2.4+</td>
</tr>
<tr>
<td>- Globus 4.0 core</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Testing installation</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Start container</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Create, deploy, and test simple GT4 Grid Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Deplore prewritten service and test with client</td>
</tr>
<tr>
<td>- Add Functionality to Service</td>
</tr>
<tr>
<td>- Need to handle WSDL (XML) and other files</td>
</tr>
</tbody>
</table>
Assignment 5  
(Using a GridNexus)

- Install GridNexus workflow editor
- Create Web service workflow
- Create Grid service workflow
- Submitting a job to GRAM
Assignment 6 (Implementing portlets with OGCSE2/Gridsphere)

- Install Gridsphere and associated software
  - Java 6 SDK
  - ant
  - tomcat 5.5.20
- Install Gridsphere
  - First start Tomcat

- Create portlets
  - Prewritten odd-even portlet
  - Installation involves handling deployment descriptor files etc.
  - Portlet to add, subtract, multiply and divide two numbers
Assignment 6 (Implementing portlets with OGCSE2/Gridsphere)

This portlet will add, subtract, multiply, or divide INTEGER numbers.

2 + 3 = 5

2 / 5 = 0.4

Add
Subtract
Multiply
Divide

March 23, 2007
Assignment 7 (MPI Program)

- Write simple MPI program (Matrix Multiplication)

- We couldn't do much more than work with “embarrassingly parallel” applications since we only had a few weeks to deal with MPI.

- The students only executed their programs on one cluster (not really using the Grid).
Mini Projects

- Teams of 3 members
- Objective was to create a new Grid assignment
- Assignment had to involve creating a Grid application with a GUI (such as a workflow or portal)
- Assignment had to be written up as though the student would give it to their classmates
  - Focus on dissemination
  - May create future assignments
Mini Projects

- Teams had to provide a written report (with solutions)
- Teams had to give a presentation
Keeping to Assignment Schedules

- Each assignment allocated 1-2 weeks to complete. Posted three dates:
  - Date assignment was set
  - Date that students had to report any system problems that were preventing them from proceeding
  - Date due
- Fall 2004 many system problems (Globus 3.2)
- Fall 2005 much fewer problems (Globus 4.0)
- Spring 2007, no system problems reported
Keeping to Assignment Schedules

- In 2004 and 2005, students created and deployed their Grid services on the same machine.
- An error in a service could make the Globus container inoperable for the other students.
- In 2007, student install the Globus 4 core on an individual PC (lab machine or their own) and used their own container.
- This turned out to be a much better solution!
Lessons Learned

• Students responded positively to using their own computers that were under their direct control.

• Some minor problems with installations but nothing that could not be resolved quickly.

• Using personal computers where possible avoided many problems.
Avoiding Problems

- It requires immense work to prepare for a hands-on Grid computing course.
- Critical that all assignments are fully tested prior to the start of class and that all computer systems are reliable and the software maintained.
- Assignments went much smoother by requiring students to use personal computers when possible.
Future Improvements

- Use GridNexus to create and deploy Grid Service
- Use GUI interface to Grid Scheduling
- Maybe introduce Cloud computing and use Hadoop (MapReduce)
Acknowledgements

Support for the work described here was provided by the National Science Foundation under grants 0410667/0533334, and University of North Carolina Office of the President.
Questions

Dr. Clayton Ferner cferner@uncw.edu
Dr. Barry Wilkinson abw@uncc.edu

Course Webpage:
http://www.cs.uncc.edu/~abw/gridcourse