Introducing Grid Computing to a Masters Course: The Design Rationale and Practice

MAOZHEN LI and MARIOS HADJINICOLAOU

Electronic and Computer Engineering
School of Engineering and Design
Brunel University
Uxbridge
UB8 3PH, UK

Abstract: - The computational grid is rapidly evolving into a large-scale computing infrastructure that facilitates scientists and engineers to solve data and computationally intensive problems by utilizing various resources over the Internet. This paper presents the curriculum design of a one-year taught MSc course in Distributed Computing Systems Engineering currently running at Brunel University in the Unite Kingdom. It reports the design rationale and practice in introducing grid computing to the MSc course.

Key-Words: - Masters Course, Distributed Systems, Grid Computing, Service-Oriented Computing, Curriculum Development, Teaching and Learning

1 Introduction

The Electronic and Computer Engineering (ECE) in the School of Engineering and Design at Brunel University has been successfully running a Masters course in Distributed Computing Systems Engineering for over 10 years. This course currently runs off-campus in Esslingen, Germany. It will be running on-campus at Brunel University in Sept. 2008. Grid computing [1] is emerging as an effective computing paradigm for sharing various resources over the Internet. The computational grid is rapidly evolved into a large-scale computing infrastructure for scientists and engineers to solve data and computationally intensive problems. The past few years have witnessed tremendous development and deployment of grid systems and applications. Representative grid networks and projects being carried out include EGEE (enabling e-Science for Europe, http://public.eu-egee.org/) spanning more than 30 countries with over 150 sites to a myriad of applications, UK e-Science programme facilitating one national e-Science center, and a number of regional e-Science centers, US OSG (open science grid, http://www.opensciencegrid.org/), CNGrid (China National Grid, http://www.cngrid.org).

The ECE in the School of Engineering and Design at Brunel University has strong expertise in grid computing. We have been actively participating in grid computing projects including the EC funded Data Grid (http://eu-datagrid.web.cern.ch/eu-datagrid), GridCC (http://www.gridcc.org), UK PPARC funded GridPP (http://www.gridpp.ac.uk/). The Brunel Information Technology Laboratory (BITLab) in the School of Engineering and Design at Brunel University has various facilities for grid computing, including 3D imaging, a 128-processor computing cluster for science and engineering computations, a 5 Terabyte data store, 3D scanning equipment, a display system for virtual environment visualisation, hardware rendering equipment, a 6 camera motion capture suite, and a high-speed network link to collaborators in the e-Science consortium. BITLab is part of a London based, geographically distributed Grid Tier 2 computing center. BitLab is also a computing node of the EGEE grid network.

Compared with traditional distributed systems, grid systems are larger in sizes crossing over multiple institutions or organizations, more heterogeneous in resources that are dealt with, and more dynamic in computing capacities. In this paper, we present the educational rationale of introducing grid computing
The reminder of the paper is organized as follows. Section 2 introduces the aims and objectives of the course. Section 3 describes the structure of the course. Section 4 presents teaching and learning strategies. Section 5 gives a brief review of current MSc courses in grid computing in the UK, and Section 6 concludes the paper.

2 Course Aims and Objectives

In the process of developing this programme, guidelines and boundaries set by the following references were used:

- Brunel University Learning & Teaching Strategy.
- Brunel University Mission Statement.
- Brunel University Strategic Plan 2002-2007

Specifically, the emphasis on theoretical aspects to very practical problems, and other transferable skills, is clearly consistent with Brunel University’s “Mission Statement” to produce high quality graduates that are of use to the community’ and with the objective of Brunel University’s “Learning and Teaching Strategy” to ensure that our graduates are flexible and able to meet the changing needs of the world of work. The opportunities provided through this programme to tackle significant practical problems, in particular through workshops and implementation projects, is in the spirit of the Strategy’s emphasis on providing student projects that are “problem focussed to simulate or reflect work and community related activity”.

The aim of this programme is to equip high quality and ambitious graduates with the necessary advanced technical and professional skills for an enhanced career either in industry or leading edge research in the areas of distributed systems and Grid computing. Specifically, the main objectives of the programme are:

- To critically appreciate advanced and emerging technologies for developing grid systems;
- To practically examine the development of large-scale grid systems;
- To critically investigate the problems and pitfalls of grid systems in business, commerce, and industry.

These aims and learning outcomes are fully consistent with the Masters Level Descriptor of the QAA Qualifications Framework, in particular their emphasis on “systematic understanding of knowledge”, “critical awareness of current problems”, abilities to “demonstrate self-direction and originality in tackling and solving problems and act autonomously in implementing tasks at a professional or equivalent level” and “continue to advance their knowledge and understanding, and to develop new skills to a high level”.

3 Course Structure

The course comprises 8 taught modules and an MSc dissertation:

- Computer Networks
- Distributed Systems
- Network Security and Data Encryption
- Network Computing
- Grid Middleware Technologies
- Grid System Analysis and Design
- Workshop
- Project Management
- MSc Dissertation

Each of the 8 modules carries 15 credits and the dissertation has 60 credits. Students receiving a Masters degree of the course need to achieve 180 credits. Figure 1 shows the correlation of these modules.

As can be seen in Figure 1, Computer Networks module is the fundamental module for this course. Grid computing demands quality-of-service (QoS) supported networks especially when transmitting large amount of data over the Internet. For this purpose, the focus of Computer Network is on QoS aspects such as ATM networks, IP QoS (integrated service model and differentiated service model), RSVP (resource reservation protocol), MPLS (multi-protocol label switching), IP multicasting.
Distributed Systems module mainly focuses on issues related to traditional distributed systems such as file systems, naming services, replication services.

Network Security and Encryption module introduces the fundamental theory on network security and encryption with a focus on Public Key Cryptography which is the basis of the Grid Security Infrastructure [4].

Network Computing module covers object oriented programming with Java programming language with a focus on middleware technologies to facilitate one computing node (client) to communicate with another computing node (server). These technologies include RPC (remote procedure call), Java RMI (remote method invocation), CORBA (common object request broker architecture) and XML based Web services. This module appreciates the evolution of these technologies and stresses their applications. The material on Web services provides direct support for the module of Grid Middleware Technologies as the computational grid is rapidly evolved into a service-oriented computing infrastructure [5].

Grid computing is evolved from parallel computing, cluster computing, meta-computing, then service-oriented computing, so Grid Middleware Technologies module mainly covers MPI (Message Passing Interface) and PVM (Parallel Virtual Machine) for parallel applications, Condor (a high throughput batch system, http://www.cs.wisc.edu/condor/) for cluster computing systems, Globus (http://www.globus.org) for grid middleware technologies, OGSA (open grid services architecture) [6], a standard architecture for developing service-oriented grid systems. It is worth noting that this module was designed based on a grid textbook [7] which provides a systematic way in introducing grid core technologies.

Grid System Analysis and Design module is built on the understanding of grid core technologies. This module helps students build knowledge on the way a grid system works, the current practices of grid applications and systems such as EGEE and GridPP, and issues in the design of grid systems such as scalability, reliability, testing, maintenance.

Project Management module helps students understand the critical issues in management of projects and build their knowledge in the application of approaches.

Hands-on laboratory assignments associated with each of the taught modules are organised in the Workshop module to help students build knowledge and capabilities in problem solving.

4. Strategies for Learning and Assessment

During the 1980’s the focus on teaching and learning in higher education was on the teaching part, i.e. how the lecturer organized and managed learning activities [8]. During the 1990’s the focus gradually shifted from teaching to learning [9]. Learning is a process that involves mastering abstract principles, understanding proofs, remembering factual information, acquiring methods, techniques and approaches, recognition, reasoning, debating ideas, or developing behaviour appropriate to specific situations [10].

Extensive researches have been carried out to classify the high quality learning styles and the low quality learning styles, e.g., meaningful learning vs. rote learning [11], logical forming vs. mnemonic concrete [12], generative processing vs. reproductive processing [13], deep learning vs. surface learning [14], transformational learning vs. reproductive learning [15], holistic learning vs. atomistic learning [16].

Learning outcomes in terms of knowledge and understanding are achieved through a mix of lectures, workshops, seminars, self-study, and individual and group project work. In lectures key concepts and ideas are introduced, definitions are stated, results and techniques are explained, and immediate queries discussed. Seminars provide
students with the opportunity to raise at greater length issues arising from the lectures and from private study, for the lecturer to test student understanding through discussion of relevant problems. Workshops and projects are used to foster practical engagement with the taught material. The dissertation plays a key role in deepening understanding, in developing research and literature review skills, and in applying knowledge and skills gained in the programme to plan, execute, and evaluate a significant investigation into a current problem area related to distributed systems and grid computing.

A key part of the strategy for learning has been to provide a solid experiential learning platform based on the Kolb learning cycle [17], and by using small groups [18]. The strength of this approach is clearly the tutorial style with students able to progress at their own pace with a structured work plan to facilitate learning. We also promote student-oriented learning [19]. Collaborative group work and peer review prove effective and useful. In addition, case studies are used in lectures to facilitate students to build pragmatic capabilities in problem solving.

This course employs a range of assessment methods to promote a rounded, diverse and independent approach to learning by the student. Summative assessment (examination) features heavily in the more theoretical modules while formative assessment (e.g. report writing, oral presentation, group project work) features heavily in the workshops.

5. Existing UK Courses in Grid Computing

In 2004, Cranfield University launched an MSc in Grid Computing and e-Engineering, which is thought to have been the first in UK and the second in the world after University of Amsterdam. Six grid-related modules have been designed, which are: Grid Fundamentals, Grid Middleware, Grid Infrastructure, Grid Development and Applications, High-Performance Computing on the Grids, and Nature Inspired Grid Resource Management.

The MSc in Advanced Computing at Imperial College London is a full-time course of 12 months’ duration. The programme offers students the opportunity to study a wide variety of topics in depth and with dedicated experts and aims to prepare students for a rewarding career in computing in particular, and in IT in general. The Grid Computing module includes: Service oriented architectures, Web services, programming models for Grid environments, Grid infrastructures (Globus, Condor, Condor-G), Open Grid Services Architecture, security issues, resource election and job placement, computational economics models.

The e-Science and Grid module has been part of the MSc in Advanced Computer Science at the University of Manchester since 2004. The module is also taken by MSc students in Advanced Computer Science and ICT Management. The module is unique in the UK, being based on the Unicore software (http://www.unicore.eu/).

Compared with these existing courses in grid computing, our course provides a more coherent and systematic way in introducing grid computing technologies with a well-designed structure.

6. Conclusion

In this paper, we have presented the curriculum design of a one-year taught MSc in Distributed Computing Systems Engineering course currently running at Brunel University. The MSc course has been running successfully in Esslingen Germany as an off-campus course for over 10 years. We introduced grid computing to the course in 2007 and 24 students are currently on the course. Most of the students on the course are from industry. The feedback from the students is highly positive. We will be running this course on campus at Brunel University in Sept. 2008.

References:


