Course Description

Grid Computing, NGSSC, 2p

Erik Elmroth, Olle Mulmo, and Leif Nixon

February 14, 2003

General information

This course is designed to give a broad overview of the concept of grid computing and state-of-the-art grid software, understanding and practical experience of how computational grids can be used to further research in science and technology, and to review some current and future directions of grid computing research and development.

Prerequisites

Familarity with computer programming and C/C++. Object-oriented scientific programming NGSSC, 3p, Introduction to High Performance Computing NGSSC, 4p, Parallel algorithms with applications to scientific computing, 3p, Scientific visualization NGSSC, 3p, or similar knowledge and experience.

Course contents

Topics covered in the course include the following:

- Introduction to grid computing ($\sim$ 10% of lectures):
  - Overview: Fundamental concepts, grid architecture models, grid middleware functionalities, etc.
  - International overview: Application projects, software development projects, political and administrative aspects on grids, emerging grid standards (OGSA/OGSI), etc.
  - Foundations in algorithm and software development for grids.
- Grid computing middleware and tools ($\sim$ 40% of lectures):
  - Globus: Overview and some specific components, including resource specification language, security, scheduling, file transfer, information services, etc.

  - Grid security: Grid security demands and solutions for, e.g., authentication, authority, assurance, accounting, trust, group communication, for large-scale, dynamic, multi-organization environments.
  - Resource management and scheduling: Characterization of resource management problems based on job requirements, characteristics, and availability of resources. Algorithms, tools and sample resource management systems.
  - Grid portals: Convenient access to grid environments. Functionality and underlying infrastructure for sample general and application specific portals.
  - Data management: Key issues for data management in grids, including file transfer, data replication, data caching issues, catalog issues.
  - Sample grid middleware packages, software tools, and problem solving environments for grids.

- Applications, algorithms, and sample projects ($\sim$ 40% of lectures):
  - NorduGrid: Introduction to the NorduGrid testbed project and software package. Introduction to hands-on tutorial.
  - SweGrid and SweGrid applications: The hardware and software build-up for SweGrid. Overview of planned and ongoing projects, in e.g., high-energy physics, biomedical sciences, earth sciences, space and astro sciences etc. Key issues for their successful use of SweGrid and the grid's implications on future research in these projects.
  - Visualization and interactivity on the grid: Overview of important aspects including relevant aspects of human perception, bandwidth, latency, data storage, package loss, compression, image quantity and quality, interaction technology, integrity and security. Examples from computational steering, data mining, virtual reality and related areas like networked simulators and games.
  - Algorithm development for grids, and e.g., the GrADS project. Focus on the special requirements that grid environments put on algorithm development, e.g., heterogeneity in computers and networks, fault tolerance, possibilities to access special instruments and equipment, etc.
  - Collaborative environments: Grid technologies for large-scale distributed meetings, collaborative work sessions, seminars, lectures, tutorials and training.

- Sample grid research projects ($\sim$ 10% of lectures):
  - One or two invited presentations on current grid research topics.

- Computer projects (preliminary):

- Hands-on introduction to NorduGrid. Sample small ready-to-run exercises.
- Resource management in a Globus environment.
- Hands-on visualization on the grid. Interactive visualization applications, preliminary using Globus and SGI VizServer. Also hands-on experience of one or several active grid projects (e.g. RealityGrid).
- Development of a small MPI-based prototype application software for a large-scale grid.

Literature

The course literature will be based on a collection of articles from journals, conference proceedings and books, and material available via WWW.
Lecture notes including exercises, projects etc will be produced for the course and made available via WWW.

Examination

A number of carefully documented computer assignments.

Preliminary dates

The course 2003 is planned to be given at PDC, KTH, August 18-26, including scheduled and unscheduled work during the weekend, August 23-24. As a guideline, the 80 hours of the two credit points course will roughly be spent on 20 hours with lectures, 45 hours scheduled computer assignments (including introductions and summaries) 15 hours homework (literature studies, unscheduled work on computer assignments).

The following years the course is planned to alternate between HPC2N (Umeå University), NSC (Linköping University), and PDC (KTH).