Grid Computing

Dr. Richard Sinnott

Dr. Colin Perkins

Dr. John Watt

http://csperkins.org/teaching/2004-2005/gc5/

UNIVERSITY
of
GLASGOW

Lecture Outline

- Introduction to the Module/Administrivia
  - Lecturers
  - Aims and Intended Learning Outcomes
  - Prerequisites
  - Module Structure and Timetable
  - Assessment
  - Reading List
- Introduction to Grid Computing
  - What are Computational Grids? Why are they useful?
  - The Architecture of Grid Computing Systems
- Preparation for Tutorial 1

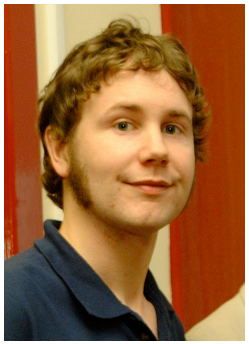
Lecture Outline

- Introduction to the Module/Administrivia
  - Lecturers
  - Aims and Intended Learning Outcomes
  - Prerequisites
  - Module Structure and Timetable
  - Assessment
  - Reading List
- Introduction to Grid Computing
  - What are Computational Grids? Why are they useful?
  - The Architecture of Grid Computing Systems
- Preparation for Tutorial 1

Lecturers

Dr. Richard Sinnott
Kelvin Building, room 234b
Email: ros@dcs.gla.ac.uk
Course Director

Dr. Colin Perkins
Lilybank Gardens, room S154
Email: csp@dcs.gla.ac.uk



Dr. John Watt
Kelvin Building, room 246d
Email: jwatt@dcs.gla.ac.uk

All willing to discuss the module and to answer questions…
But: **please make appointments by email** before coming to see us; we're all busy, and if you make an appointment it allows us to pick a slot when we both have time to talk

Aims of This Module

Grid Computing is a developing area:

- Our understanding of the problem space is still evolving
- Systems and standards change frequently
- Many open research issues
- No-one has all the answers

Accordingly, we aim to:

- Produce Grid-savvy individuals
- Encourage critical thinking about Grid-related technology
- Pique your interest in Grid-related research issues

Intended Learning Outcomes

- Thorough grounding in the architecture of the Grid, and exposure to various implementations of the infrastructure
- Experience in using one particular implementation to construct a Grid-based application
  - Competence in Grid programming
  - Exposure to large-scale cluster computing facilities
- Awareness of current open research issues relating to the Grid architecture and infrastructure
  - How do they differ from distributed systems?
  - What are the challenges in making Grids work?

Prerequisites

- Students are expected to have completed computer science degree-level studies covering the areas of operating systems, distributed algorithms, communications and networks, databases, and internet technologies including web services.
- For Glasgow MSci and Honours students, this largely corresponds to the material covered in OS3, DAS4, CA4 and NCT4.
  - If taking this as a level 4 module, these run in parallel; you may have to read ahead in some modules
  - The material in DBIT4 would also be useful, although it is not a formal co-requirement

- Students are assumed to be experienced at programming in Java and C on Unix-like systems

Prerequisites and Focus

- Students are expected to learn quickly, and to master complex systems, languages and technologies in a self-directed manner
- Focus **will not** be in teaching Grid Computing languages and technologies as such
  - You should be competent programmers who can take software and trial it out yourself
  - Although "some" training given on technologies, languages, etc.
- Focus is on understanding the fundamental computing science topics underlying Grid Computing
  - Why? Grid Computing is a highly dynamic area, where the standards, technologies, and software change all of the time
  - By the end of the module you should understand Grid Computing concepts, and be able to apply them to various scenarios, using a mix of technologies

Timetable

Week beginning…	Monday 12:00-13:00	Thursday 14:00-15:00	Friday 09:00-10:00
27 September		Lecture 1	Lecture 2
4 October	Tutorial 1	Lecture 3	Lecture 4
11 October	Lecture 5	Lecture 6	Tutorial 2
18 October	Lecture 7	Lecture 8	Lecture 9
25 October	Tutorial 3	Lecture 10	Lecture 11
1 November	Tutorial 4	Lecture 12	Tutorial 5
8 November	Lecture 13	Lecture 14	Tutorial 6
15 November	Lecture 15	Lecture 16	Tutorial 7
22 November	Lecture 17	Tutorial 8	Lecture 16
29 November	Lecture 19	Lecture 20	Tutorial 9

- Lectures and tutorials do not always follow the same weekly pattern
  - This is to allow preparation time for tutorials, after the lecture covering the material
  - You **will** be expected to present and discuss papers during the tutorials

Module Outline

Week 1	Lecture 1	Introduction to Grid Computing	Colin Perkins
	Lecture 2	Scalability and Heterogeneity	Colin Perkins
Week 2	Tutorial 1	Discussion of Seminal Grid Papers	Colin Perkins
	Lecture 3	Open Standards and Architectures	Richard Sinnott
	Lecture 4	Implementations of the Grid Architecture	John Watt
Week 3	Lecture 5	Resource Discovery/Information Services	John Watt
	Lecture 6	Web Services	Richard Sinnott
	Tutorial 2	Exploring Web Services Technologies with GT3	John Watt
Week 4	Lecture 7	Grid Security Concepts	Richard Sinnott
	Lecture 8	Virtual Organizations	Richard Sinnott
	Lecture 9	Security in Practice	John Watt
Week 5	Tutorial 3	Lab work and Discussion of Grid Security	Richard Sinnott
	Lecture 10	Job Scheduling and Management - Theory	Colin Perkins
	Lecture 11	Job Scheduling and Management - Practice	Colin Perkins
Week 6	Tutorial 4	Discussion of Job Scheduling Papers	Colin Perkins
	Lecture 12	Workflow Management	John Watt
	Tutorial 5	Q&A on Programming Assignment	All

Module Outline

Week 7	Lecture 13	Data Access, Integration and Management	John Watt
	Lecture 14	Data Provenance and Curation	Seamus Ross*
	Tutorial 6	Discussion of Data Management/Provenance	Richard Sinnott
Week 8	Lecture 15	Data Transfer	Colin Perkins
	Lecture 16	Peer-to-Peer Communication	Colin Perkins
	Tutorial 7	Discussion of Networking Papers	Colin Perkins
Week 9	Lecture 17	Tools for Collaboration	Colin Perkins
	Tutorial 8	Discussion on the Future of Grid Computing	Richard Sinnott
	Lecture 18	The Future of Grid Computing	Richard Sinnott
Week 10	Lecture 19	Sample Applications	Richard Sinnott
	Lecture 20	Review of Major Concepts	All
	Tutorial 9	Q & A	All

*National Digital Curation Centre

- Different subjects taught by different lecturers, representing the broad nature of Grid computing research
- Supported by background reading from the research literature

Tutorials

- Discussion tutorials based on students reviewing & presenting materials to class – **NOT** on sitting there listening to lecturers discuss
  - Groups of students will be selected the week before a tutorial discussion and expected to prepare/present material together
  - Other students will be expected to read the papers, and to ask questions and participate in the discussion
  - The Research Techniques and Advanced Topics in Computing modules will prepare you for this
- Material in the tutorial papers may be in the written examination, so it is to your benefit to participate in the tutorials
  - You **cannot** pass this module just by memorizing facts
  - The examinations will require critical thinking and discussion skills; the tutorials are practice for this

Assessment

- 3 Problem sets (each worth 5% of total mark)
  - Problem set 1: available 7th October… due 14th October
  - Problem set 2: available 28th October… due 4th November
  - Problem set 3: available 15th November… due 26th November
  - All problem sets to be handed in by 5pm on the due date

- Programming assignment (15% of total mark)
  - Available 22nd October, due at 5pm on 29th November

- Written examination (70% of total mark)
  - All material covered in the lectures, tutorials and background reading will be examinable
  - Aim is to test your understanding of the material, not to test your memory of all the details; explain why – don't recite what

Reading List

- There is no set textbook for this module
- Instead, research papers will be made available on the website and in lectures/tutorials
  - Some of these will cover background material, and are provided are a reminder of material you should have covered in your previous studies
    - This may be in the form of the original research literature covering material you've previously seen presented in a textbook
    - The aim of this is to make you familiar with reading research papers
  - Others are new material which you will be expected to understand, and comment upon

Lecture Outline

- Introduction to the Module/Administrivia
  - Lecturers
  - Aims and Intended Learning Outcomes
  - Prerequisites
  - Module Structure and Timetable
  - Assessment
  - Reading List
- Introduction to Grid Computing
  - What are Computational Grids? Why are they useful?
  - The Architecture of Grid Computing Systems
- Preparation for Tutorial 1

What are Computational Grids?

Infrastructure for Internet-scale Distributed Systems

- "A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities" [1]

- Aim is to make access to computational resources as simple and reliable as access to the electricity grid
  - Readily available
  - Trustworthy and predictable
  - Simple to use
  - Part of the assumed infrastructure

Why Do We Need Grid Computing?

- To make large-scale computing a commodity that anyone can buy, and anyone can use
- To enable "coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations" [2]
- To enable new classes of application, new services, new products

⇒ To make access to remote resources as straight-forward as access to local resources

- Data
- Computation
- Services
- Devices

Applications of Grid Computing

- Supercomputing problems that are too big for single institutions to solve using their own resources
  - Large scale scientific workloads: physics, biochemistry, etc.
  - Large simulations for military use
- Workloads that can effectively utilise "spare" cycles on widely distributed systems, if those can be made available
  - SETI@home
  - distributed.net crypto-busting
- Applications whose peak resource usage far exceeds their average usage, where it would be desirable to rent capacity on demand
- Data intensive applications that need coordinated access to large-scale data repositories that cannot be fully replicated, due to size or security
  - Physics experiments producing terabytes per day
  - Financial modelling using online stock quote databases/history
  - E-commerce: iTunes, Google

Enable *virtual organizations* to form around a problem domain, leveraging assumed computational infrastructure

Building Computational Grids

- Many challenges in building a Computational Grid; to integrate disparate resources into a virtual organization needs:
  - Service and resource management protocols
  - Security solutions
  - Scheduling and workflow management
  - Data management/Provenance services

Today

- All running on a **heterogeneous** network
  - Different hardware and software systems
  - Different administrative domains
  - Different legal domains
  - Different degrees of trust in the users
- **Scaling** beyond traditional distributed systems

Tomorrow

Service and Resource Management

- Need a standard protocol to locate, manage and otherwise interact with resources
- A useful approach is to model resources as **objects**; we require a cross-platform, wide-area, distributed object system
  - Resources maintain state (logical or encapsulating a physical object)
  - Operations can be invoked on resources
  - Resources have types, and can be queried for their type/supported methods

- Typically, implemented using constrained **web services**, although nothing in the architecture requires this particular implementation technique
  - Key point is an architecture neutral protocol with introspection, discovery

[See lectures 5-6]

Security

- To ensure trust in the infrastructure, need to weave services into a coherent, secure, virtual organization
- Implies strong **authentication** is used
  - Of users and the virtual community to which they belong
  - Of services and the organizations running those services
- Proof that interactions are between the intended peers; trust then follows based on reputation, allowing a virtual community to form
  - Stronger than typical web security, which only identifies the service

- Typically implemented using **public key cryptography**

[See lectures 7-9]

Scheduling and Workflow Management

- Distributed scheduling algorithms, resource allocation, workflow management needed to effectively use the resources of the virtual organization
  - Reserving resources, scheduling computation across administrative domains
  - Managing the flow of data, results
  - May have strict quality of service/performance requirements, constraints

- Many approaches; little effective standardization

[See lectures 10-12]

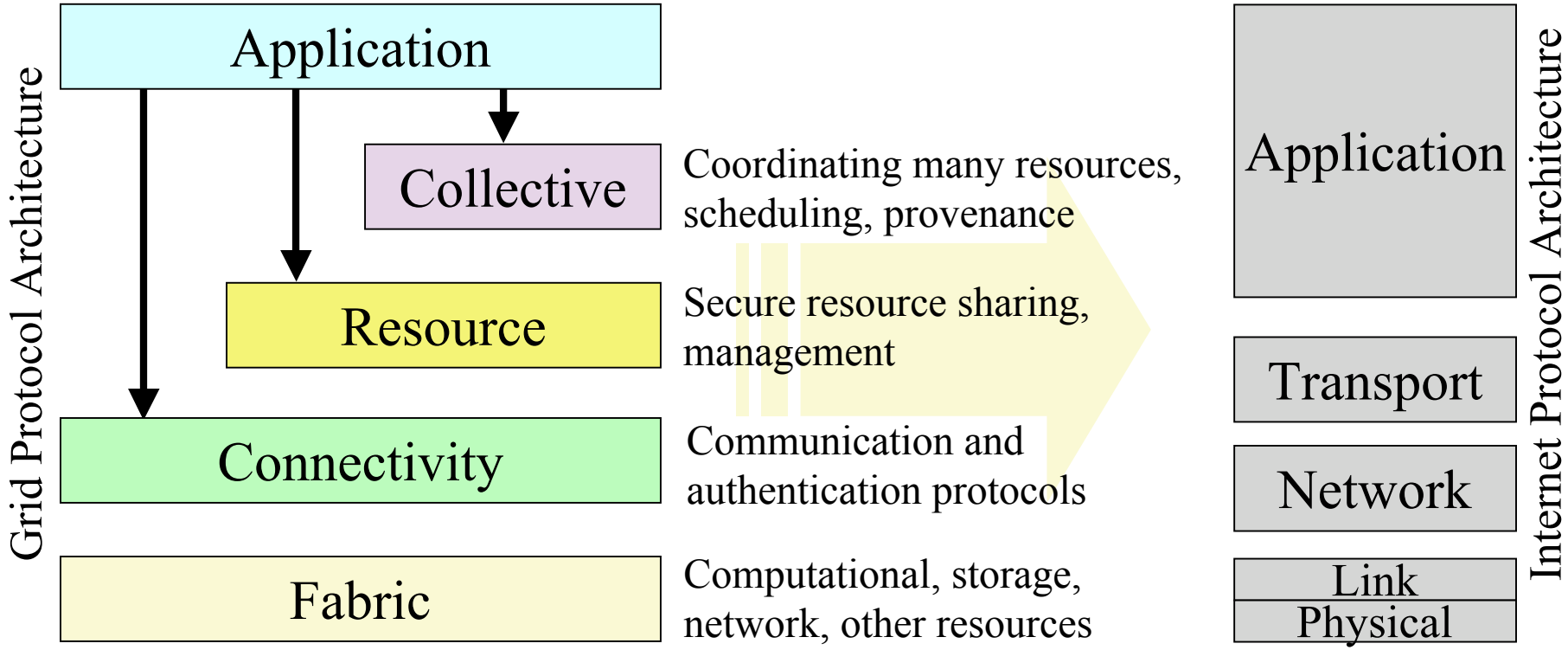
Data Management and Provenance

- The supporting infrastructure/supply chain
  - Distribute data and results to the appropriate places
  - Replication management
  - Change control
  - Tracking transformations/operations on the data
- Ensure **data provenance** as workflow progresses; represent in a standard form

- Basic data distribution using Internet protocols well understood, although hard to ensure quality of service
- Effective replication management, change control, provenance open research issues

[See lectures 13-16]

An Architecture For Grid Computing



- Architecture defines **middleware** to provide abstract services for applications, hiding details of lower layers
- May be recursively applied; *n*-tier architecture

[See lectures 3-4]

Lecture Outline

- Introduction to the Module/Administrivia
  - Lecturers
  - Aims and Intended Learning Outcomes
  - Prerequisites
  - Module Structure and Timetable
  - Assessment
  - Reading List
- Introduction to Grid Computing
  - What are Computational Grids? Why are they useful?
  - The Architecture of Grid Computing Systems
- Preparation for Tutorial 1

Tutorial 1: Discussion of Seminal Grid Papers

The first tutorial is on Monday at 12:00pm

We will be discussing the following papers:

1. I. Foster and C. Kesselman, "*Computational Grids*", in *The Grid: Blueprint for a Future Computing Infrastructure*, I. Foster and C. Kesselman (Eds), Morgan Kaufmann, 1998.
2. I. Foster, C. Kesselman, and S. Tuecke, "*The Anatomy of the Grid*", International Journal of Supercomputer Applications, Volume 15, Number 3, 2001.

Need to split into groups; prepare short presentations on the papers…