Introduction to Grid Computing

These slides are borrowed from course materials for a Grid Computing course developed by Barry Wilkinson at UNC Charlotte
Grid Computing

• Using geographically distributed and interconnected computers together for computing and for resource sharing.

“The grid virtualizes heterogeneous geographically disperse resources” from "Introduction to Grid Computing with Globus," IBM Redbooks
History

• Began in mid 1990’s with experiments using computers at geographically dispersed sites.

• Seminal experiment – “I-way” experiment at 1995 Supercomputing conference (SC’95), using 17 sites across the US running:
  – 60+ applications.
  – Existing networks (10 networks).
• Grid computing is about collaborating and resource sharing as much as it is about high performance computing.
Interconnections and Protocols

Focus now on:

• using standard Internet protocols and technology, i.e. HTTP, SOAP, web services, etc.,
NSF Network for Earthquake Engineering Simulation (NEES)

Transform our ability to carry out research vital to reducing vulnerability to catastrophic earthquakes

from I. Foster
CERN grid

Large Hadron Collider experimental facility for complex particle experiments at CERN (European Center for Nuclear Research, near Geneva Switzerland).
Campus Grids

Several examples of grids within one university and across campuses

Example

University of Virginia Campus Grid

Users:
- BioMed
- Physics
- BioChem
- UVA CS
- Campus workshops
Globus

- A “toolkit” of services and packages for creating the basic grid computing infrastructure
- Higher level tools added to this infrastructure
- Version 4 is web-services based
- Some non-web services code exists from earlier versions (legacy) or where not appropriate (for efficiency, etc.).
Globus Toolkit

Five parts:

• Common Runtime
  – GT Core for building new services

• Security
  – To provide secure access. Based upon Grid Security Infrastructure (GSI)

• Execution management
  – Initiation, monitoring, management, scheduling and coordination of executable programs (jobs)

• Data management
  – Discover, transfer, and access large data

• Information services
  – Discover & monitor dynamic services
1. secure environment
   - GridFTP
   - GSI

2. discover resource
   - MDS

3. submit job
   - GRAM

4. transfer data

Web Services-Based Grid Computing

• Grid Computing is now strongly based upon web services.

• Large number of newly proposed grid computing standards:
  – WS-Resource Framework
  – WS-Addressing
  – etc., etc. .....
Other software

Meta-schedulers – to allow job to be scheduled across grid resources.
Taken some time to develop meta-schedulers.

Example

Gridway

Pre-existing local schedulers schedule jobs once at a local cluster (we used Condor)
Security

• A big issue.

• Has to cross administrative domains.

• Agreed mechanisms.

• Focus is on Internet security mechanisms, modified to handle the special needs of grid computing.