Computational Grids

Derived from:
"The Anatomy of the Grid", Foster, Kesselman and Tuecke
"Grids: The Top Ten Questions", Schopf, Nitzberg
"What is The Grid ?", Foster
"Computational Grids", Foster and Kesselman

Nitin Madnani, nmadnani@umiacs.umd.edu

Outline

- Motivation
- Defining a Grid
- Is <X> a grid ?
- Grids in the Wild
- Last Six Years
- Take Home Lessons
- References

Motivation

- Group of participants
- Varying degrees of prior relationship
- Mutually Distrustful
- Need to share resources to perform a task
  - Direct access to software, data, sensors and computers
  - Sharing subject to a set of constraints
    - What, Who, When etc.
  - Dynamic sharing relationships

3

Motivation (contd)

- Biochemists exploit 10,000 computers to screen 100,000 compounds in an hour
- 1,000 physicists worldwide pool resources for petaop analyses of petabytes of data
- Civil engineers collaborate to design, execute & analyze shake-table experiments
- Climate scientists visualize, annotate, & analyze terabyte simulation datasets
- An emergency response team couples real time data, weather model, population data

4

Motivation (contd)

- Regular distributed computing
  - Cannot coordinate across multiple sites
  - Cannot accommodate multiple resource types
  - Offers limited form of sharing
  - Needs centralization
- Not good enough

5

Enter the Grid

- Introduced in mid 90s
- Distributed computing for “Big Science”
- Hardware/Software Infrastructure
- Pools resources to provide computing that is:
  - Dependable
  - Consistent
  - Inexpensive
  - Pervasive

6

What exactly is a grid ?

- Funding Concept/Marketing Slogan
- Need for a clear definition
  - The red-blue cluster is NOT a grid.
- Architecture is not as important
- Is there a checklist ?

7

The Grid Checklist

- Collaboration
  - Computation- and data-rich environments
- Coordination of heterogeneous resources across domains
  - No local management
- Standard/Open Interfaces & Protocols
  - Not application specific
- Guarantees Quality of Service (QoS)
  - Response Time, Throughput, Availability
- Transparent to users

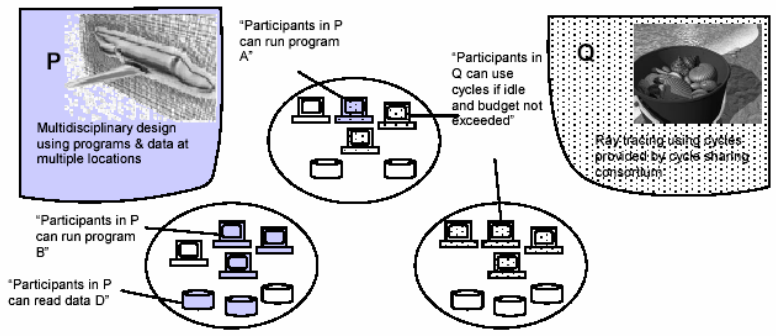
8

Virtual Organizations

- Set of individuals defined by sharing rules
- Dimensions
  - Number & type of participants
  - Resources being shared
  - Types of activities
- Examples:
  - Storage Service Providers
  - SETI@HOME participants
  - Members of a High-Energy Physics Consortium

9

Illustrative Example



10

Is <X> a grid ?

- X = Cluster Management Systems
  - QoS Guarantees [Yes]
  - Distributed Resources [Yes]
  - De-centralized Coordination [No]
  - Standard/Open Protocols [No]
- Verdict: Not a grid. A resource.
- Examples
  - Sun Grid Engine
  - Veridian Portable Batch System

11

Is <X> a grid ? (contd)

- X = The Web
  - QoS Guarantees [No]
  - Distributed Resources [Yes]
  - De-centralized Coordination [No]
  - Standard/Open Protocols [Yes]
- Verdict: Not a grid.

12

Is <X> a grid ? (contd)

- X = Distributed Computing Systems
  - QoS Guarantees [Limited]
  - Distributed Resources [Yes]
  - De-centralized Coordination [Yes]
  - Standard/Open Protocols [Varies]
- Verdict: (Almost) a grid.
- Examples
  - Condor
  - Gnutella

13

Grids in the Wild

- NASA's Information Power Grid (IPG)
  - Computation & Data Grid
  - Distributed heterogeneous resources
    - Computers
    - Databases
    - Instruments
  - Access from any location
  - IPG middleware adding security and control

14

Grids in the Wild (contd)

- NSF's TeraGrid
  - Eight partner sites
    - Including SDSC, NCSA, ORNL, ANL
  - Computation
    - 40 Teraflops, 2 Petabytes, 10-30 Gb/s dedicated network
  - Data
    - Remote Sensing Data
    - Earth Satellite Imagery
    - Genomic Data

15

Last Six Years

- Major Grid Projects underway and in production
- Global Grid Forum created (GGF)
  - Arena for interaction of communities
  - Develop APIs & protocol standards
  - Define best-practice documents
- Globus Toolkit Developed
  - Open-source & general-purpose
  - Negotiate & manage sharing
  - QoS guarantees
- OGSA (Open Grid Services Architecture) Standards
  - Roadmap published
  - Based on Web-services concepts and technologies
  - Supported by IBM, Microsoft, Sun and others.

16

Take Home Lessons

- Grids respect site autonomy
- Grids involve heterogeneity
- Grid resources are not just computers
- Grids focus on the users
  - Maximize performance of my application
  - Disregard effect on system as a whole
- Grid computing research is still in infancy

Grids in the Press

- **"… merely an excuse by computer scientists to milk the political system for more research grants so they can write yet more lines of useless code."**
  - The Economist, 2001
- **"… a solution in search of a problem."**
  - Schopf *et al*, SC'00 Panel
- **"Grid Computing has been more hype than reality."**
  - HP CEO Carly Fiorina, 2003
- **"No one can really define it, everyone wants an app that can do it, and companies that claim to do it are getting a lot of interest."**
  - Slashdot, 2003

"And then one day the grid went down and never came back up."

Questions ?

References

- I. Foster, C. Kesselman, S. Tuecke, *"The Anatomy of the Grid: Enabling Scalable Virtual Organizations"*, International J. Supercomputer Applications, 15(3), 2001
- I. Foster, *"What is the Grid? A Three Point Checklist"*, GRIDToday, July 20, 2002.
- J.M. Schopf and B. Nitzberg, *"Grids: Top Ten Questions*", Scientific Programming, 10(2):103 - 111, August 2002.
- http://www.globus.org/alliance/publications/papers.php