Grid Computing

Mahesh Joshi
joshi031@d.umn.edu

Presentation for Graduate Course in
Advanced Computer Architecture

28th April 2005

Objective

- Overview of the concept and related aspects
- Some practical implications and implementations
- Wherever possible, tying the concepts to what we learned in the course

Overview

- Why Grid Computing
- What is Grid Computing
- Aspects of Grid Computing
- Is this for real?
- References

Why Grid Computing

- Various architectural enhancements exist for increasing computer as well as network speeds and storage capacity
- How effectively is the resulting computing power utilized?
  - Surprisingly low utilization:
    - Mainframes are un-used 40% of the time!
    - UNIX® Servers "serve" less than 10% of the time!
    - Desktops produce "useful work" less than 5% of the time!
- We have fast computers and fast means to connect them – why not utilize them constructively?

Why Grid Computing ..contd.

- Not all problems are solved efficiently by today's super-computers
- Numerically as well as data intensive problems that require a variety of heterogenous resources not availabe on a single computer
- Grid Computing offers a solution

What is Grid Computing

- Simply stated – distributed computing taken to next level
- "*The goal is to create the illusion of a simple yet large and powerful self managing virtual computer out of a large collection of connected heterogeneous systems sharing various combinations of resources.*"
- Driving force is the emerging standardization for sharing resources and higher network bandwidth

What is Grid Computing ..contd.

- Voluntary use of under-utilized resources available over the network
- Balancing out the utilization by exploiting parallelism
- Providing a means for solving highly intensive problems in quasi real time.

What Grid Can Do

- Exploiting underutilized resources
- Parallel CPU capacity
- Virtual resources and virtual organizations for collaboration
- Access to additional resources
- Resource balancing
- Reliability and Management

What Grid Cannot Do

- Not a Silver Bullet!!
- Cannot achieve more without additional hardware or resources
- Cannot introduce parallelism in applications
- Takes more than just one-click installation, configuration is important

Grid Computing compared to..

- Distributed Computing
  - Homogeneous vs. heteregenous
- Cluster Computing
  - Centralized control vs. distributed control
- Peer To Peer (P2P)
  - Merging streams

Towards Grid Computing: a conceptual view

Image Source: Grids and Grid technologies for wide-area distributed computing

Aspects of Grid Computing

- Types of Grids
- General issues in Grid realization
- Fundamental components of a Grid
- Architectural considerations
- Design Features
- Grid topologies

Types of Grids

- Compute Grid / Computational Services Grid
- Information Grid / Data Grid
- Enterprise Grid

Data Grid

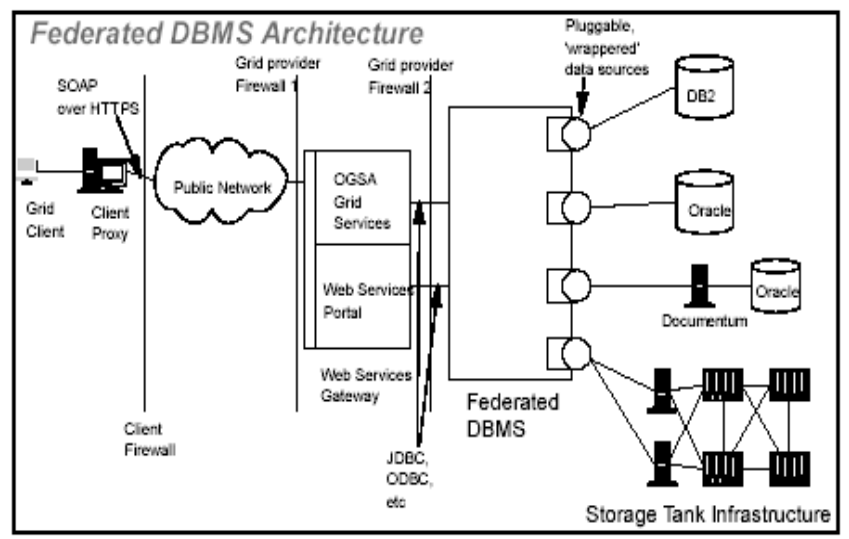

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

General Issues in Grid Realization

- Multiple administrative domains and autonomy
- Heterogeneity
- Scalability
- Dynamicity / Adaptability

Components of Grid

- Grid Fabric
- Core Grid Middleware
- User-level Grid Middleware
- Grid Applications and Portals

Grid Components – pictorial version

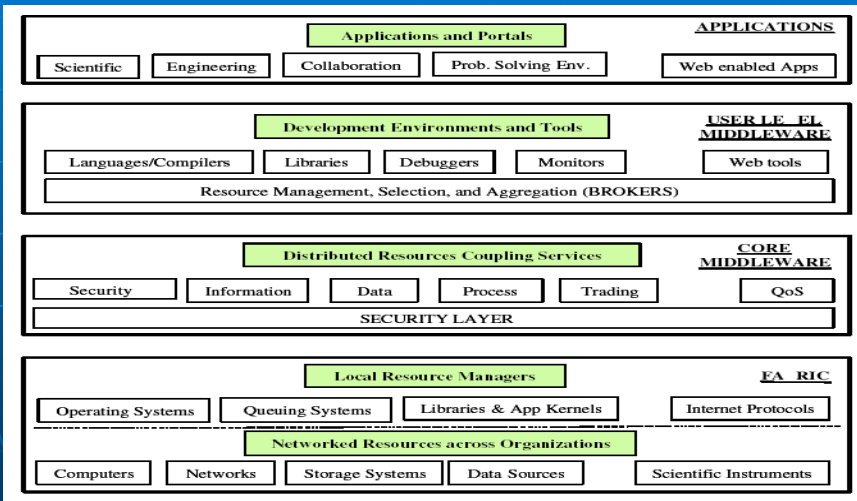

Image Source: Grids and Grid technologies for wide-area distributed computing

A high level view of Grid and interactions within entities

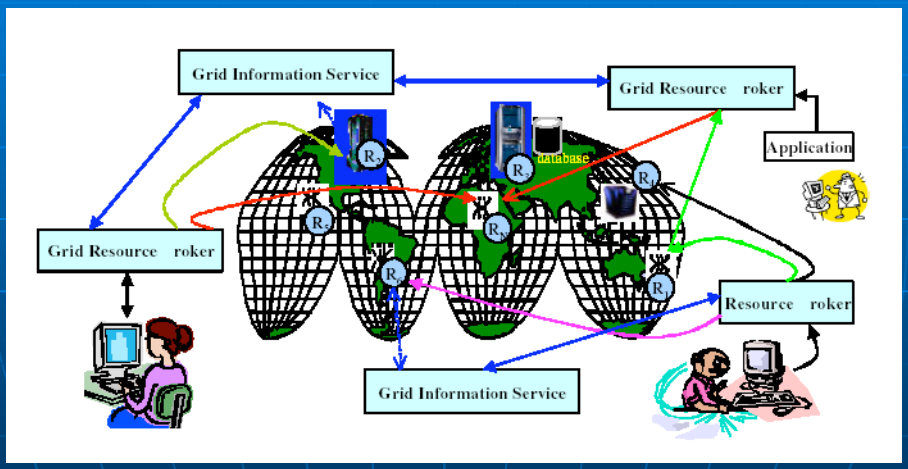

Image Source: Grids and Grid technologies for wide-area distributed computing

Architectural Considerations

- CPU Considerations
  - Parallelizable?
  - Parameter space problem?
- Data considerations
  - Amount of data and time to send it
  - Shared data issues

CPU Considerations – Parallelizing

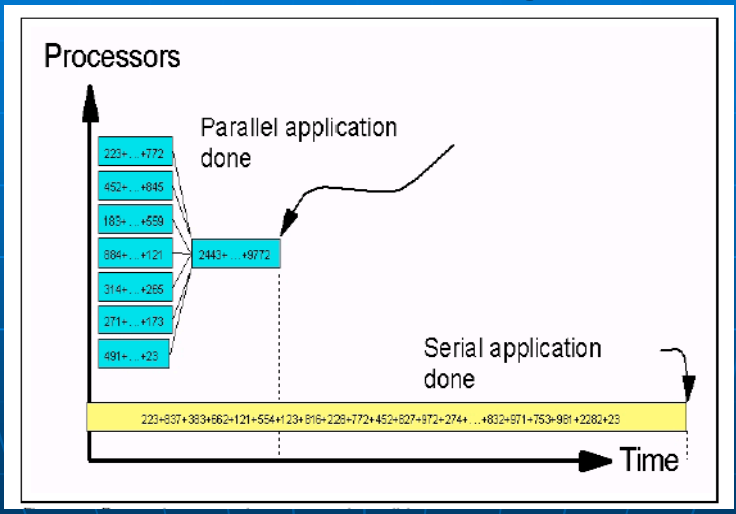

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

CPU Considerations – Parallization not possible

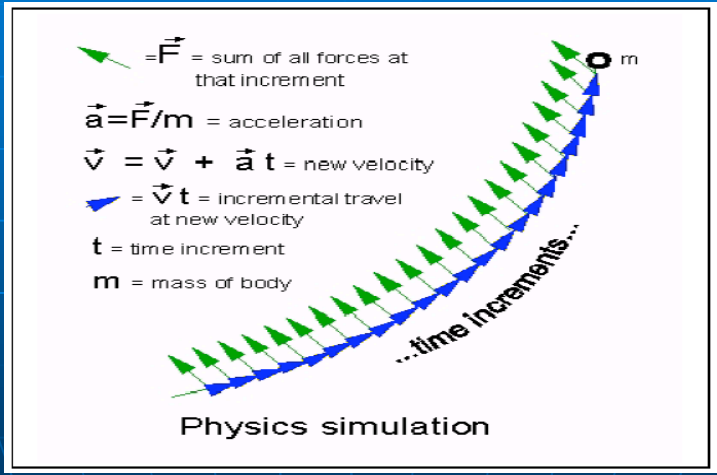

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

Design Features

- Administrative hierarchy
- Communication services
- Information services
- Naming services
- Distributed file systems and caching
- Security and authorization
- System status and fault tolerance

Design Features ..contd.

- Resource management and scheduling
- Computational economy and resource trading
- Programming tools and paradigms
- User and administrative GUI

Grid Topologies

- Intragrids
  - Single organizations
  - No partner integration
  - A single cluster
- Extragrids
  - Multiple organizations
  - Partner integration
  - Multiple clusters

Grid Topologies ..contd.

- Intergrids
  - Many organizations
  - Multiple partners
  - Many multiple clusters

Grid Topologies - pictorial

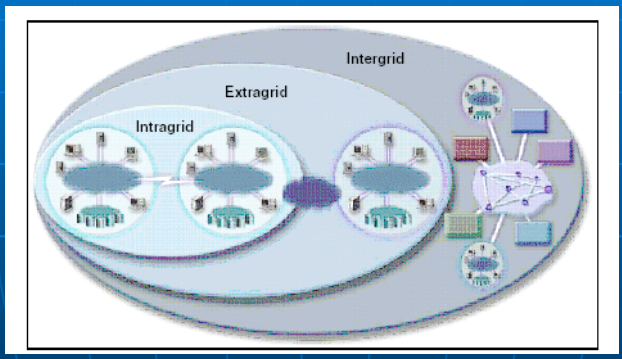

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

An Intragrid

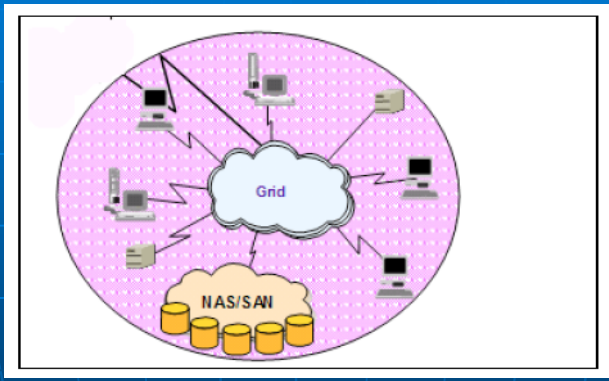

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

An Extragrid

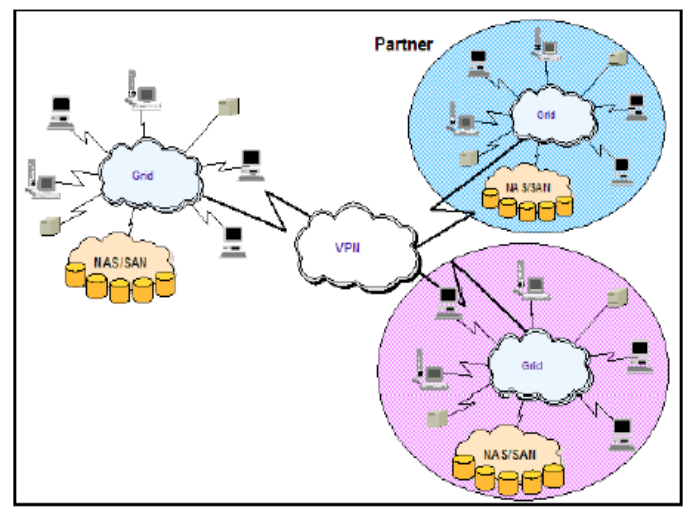

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

An Intergrid

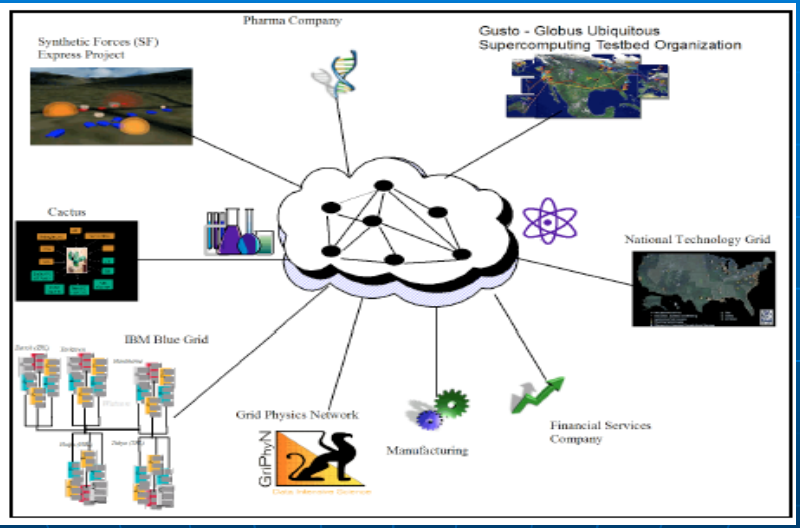

Image Source: Introduction to Grid Computing with Globus – IBM Redbook

Is This For REAL?

- Very Much!!
- Real Grids
  - NASA IPG, the World Wide Grid, and the NSF TeraGrid
- An intragrid example: developed at University of Wisconsin and available at UMD
  - Condor - http://www.cs.wisc.edu/condor/

Is This For REAL? ..contd.

- Grid Forums
  - Global Grid Forum (http://www.ggf.org/)
- Grid Standards
  - Open Grid Services Architecture, an industry blueprint for standards based Grid Computing
  - Implemented by: the IBM Globus Toolkit. Provides a set of tools for application programming (APIs) and system development kits (SDKs)

References

- Introduction to Grid Computing with Globus – IBM Redbook
- Grid Explained, IBM Whitepaper, Jan. 2005
- http://www.gridcomputing.com/
- Grids and Grid technologies for wide-area distributed computing, Mark Baker, Rajkumar Buyya and Domenico Laforenza, Appeared in Journal of Software Practice and Experience Softw. Pract. Exper. 2002; (in press) (DOI: 10.1002/spe.488)
- The Grid: A New Infrastructure for 21st Century Science, Ian Foster, Feature Arcticle, Physics Today, VOL 55, ISSUE 2, Pg 42
- Scooped, Again, Jonathan Ledlie, Jeff Shneidman, Margo Seltzer, John Huth. Harvard University

Any Questions?

Thank You!!