Grid Computing

By
Rajeev Wankar

wankarcs@uohyd.ernet.in

What to learn?

• Distributed Computing Concepts
• Web Services
• Grid Services using Globus Toolkit
• Cloud Computing Introduction using Aneka
What to know?

- **Syllabus**
  - Download from the course web site
  - Will be flexible

- **Books (Grid)**
  - Foster, et. al. “Anatomy of the Grid” (globus.org)
  - Foster et. al. “Physiology of the Grid” (globus.org)
  - Start reading about Globus Toolkit GT 4.0
  - Anirban Chakrabarti, Grid Computing Security
  - The Globus Toolkit 4 Programmer’s Tutorial

- **Books (SOA)**
  - Beginning Java web services, Henry Bequet
    *et. al.*, a! press
  - Programming web services with SOAP,
    James Snell *et. al.*, O’Reilly publisher

- **Course Material**
**Announcements**

- At the end of this month, you need to decide what tools you wish to work with:
  - **Languages:**
    - Java, Python, Perl, C# and .NET are acceptable
    - I will only use Java examples
  - **Operating Systems:**
    - Much is based on Apache, so Linux is best for deployment.
    - XP is OK for development and for Aneka
    - Any tool is well with me as long as I can read the software
  - **Toolkits and programs – begin with:**
    - JDK and Ant (compiler)

---

**Announcements: Course Roadmap**

- The course will cover topics in the following order:
  - Intro to distributed systems
  - Intro to XML technologies
  - Web Services (standard)
  - Intro to Grid Computing
  - Globus Toolkit, Security Aspects
  - Other Grid Middleware
  - Web Services Resource Framework (WSRF)
  - Grid Portals
  - Cloud Computing with Aneka
What is the Grid?

“It is an infrastructure that enables the integrated collaborative use of high-end computers, networks, databases and scientific instruments owned and managed by multiple organizations”

- Grid applications often involve large amount of data and/or computing and often require secure resources sharing across organizational boundaries
- Can not be handled easily by today’s Internet Technology and web infrastructure

Simple Grid

- Grid Computing differs from conventional distributed computing in its focus on Innovative Applications
- Grid aware applications also called as Multi-disciplinary Applications or Meta Applications
- These make use of coupled computational resources that are not available at a Single Site
- Distributed Supercomputing (Ultra-large virtual supercomputers constructed to solve problems too large to fit on any single computer)
Definition in 1969

Back in 1969 Len Kleinrock first suggested:

“We will probably see the spread of ‘computer utilities’, which, like present electric and telephone utilities, will service individual homes and offices across the country.”

Early definition in 1998

In 1998, Carl Kesselman and Ian Foster proposed a definition in their book “The Grid: Blueprint for a New Computing Infrastructure.”

They wrote:

“A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities”
Definition proposed in 2000
(Dr. Raj Kumar Buyya)

“Grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of geographically distributed "autonomous" resources dynamically at runtime depending on their availability, capability, performance, cost, and users' quality-of-service requirements.”

Goal

Grid enable communities (“virtual organizations”) to share geographically distributed resources as they pursue common goals -- assuming the absence of...

- central location
- central control
- omniscience

The grid virtualizes heterogeneous geographically disperse resources


TeraGrid and the Alliance

Source: TeraGrid Site
Evolution of grid computing

• Started as a form of distributed computing.

• Early distributed computing systems:
  – Later - Distributed objects systems:
    • CORBA (Common Request Broker Architecture)
    • Java RMI (Remote Method Invocation)
Evolution of grid computing

- With the use of the Internet interconnection technology, implementation now based upon Internet technologies.

- Now uses a form of web services.

- Enables using existing protocols, security mechanisms, etc.

Web Services

- Software components designed to provide specific operations (“services”) accessible using standard Internet technology.

- Machine interaction over a network is done usually through SOAP (Simple Object Access Protocol)

- These messages carry XML documents.
Grid Computing Vs. Distributed Computing

• Distributed Computing normally refers to managing or pooling the hundreds or thousands of computer systems which individually are more limited in their memory and processing power.

• On the other hand, grid computing has some extra characteristics. It is concerned to efficient utilization of a pool of heterogeneous systems with optimal workload management utilizing an enterprise’s entire computational resources (servers, networks, storage, and information) acting together to create one or more large pools of computing resources. There is no limitation of users, departments or origins in grid computing.

Grid Computing Vs. Distributed Computing

• Grid computing is focused on the ability to support computation across multiple administrative domains that sets it apart from traditional distributed computing.

• Its concept of support for multiple administrative policies and security authentication and authorization mechanisms enables it to be distributed over a local, metropolitan, or wide-area network.
Grid Computing Vs. Distributed Computing

• Grid computing is focused on the ability to support computation across multiple administrative domains that sets it apart from traditional distributed computing.

• Its concept of support for multiple administrative policies and security authentication and authorization mechanisms enables it to be distributed over a local, metropolitan, or wide-area network.

Syllabus

- **Client-Server Architecture**: Introduction to Distributed Computing, Characteristics, Basic Mechanics of Client-Server, Thin Vs Thick client, Issues, 3-tiered architectures, Distributed Objects, Background for the web services.

Syllabus

- **Grid Services**: Introduction to Grid Computing with Globus Toolkit, Overview of Grid Middleware Distributed Object Technology for Grid Computing (OGSA, WSRF) Grid Middleware, GSI. Developing Grid Services


- **Programming Enterprise Clouds using Aneka**: Introduction, Aneka Architecture, Aneka Deployment, Parallel Programming Models, Threads, Tasks, and MapReduce,