Grid Computing

Lectured by: Dr. Pham Tran Vu

Email: ptvu@cse.hcmut.edu.vn

Course Objectives

- Grid computing concepts
- Roles of Grid computing
- Grid architecture
- Grid related technologies
- Grid programming models
- Developing Grid applications

Assessment

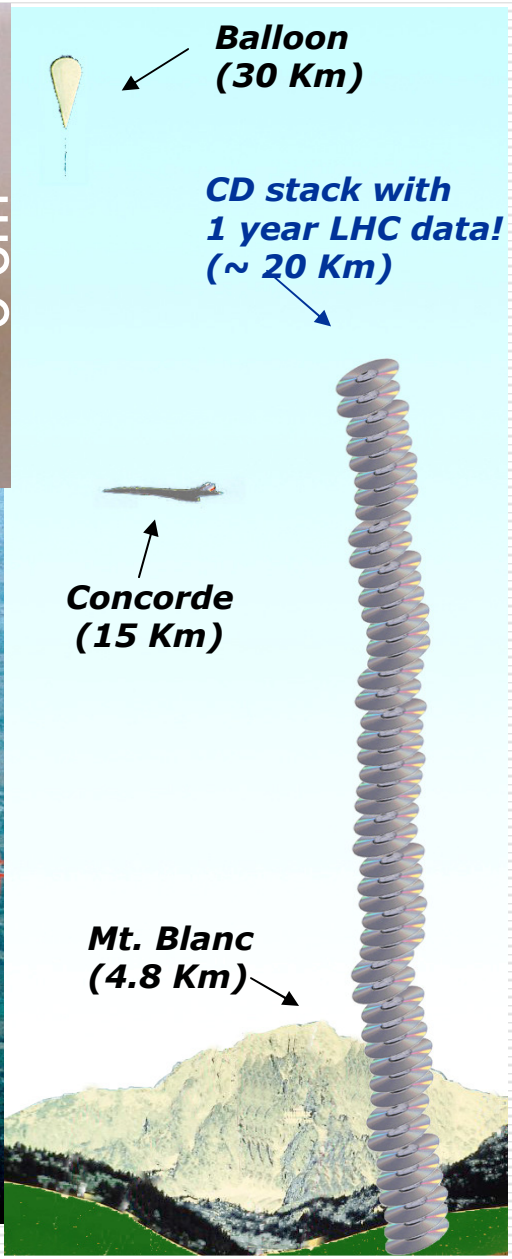
- Exercises: 30%
- Seminar: 20%
- Final exam: 50%

References

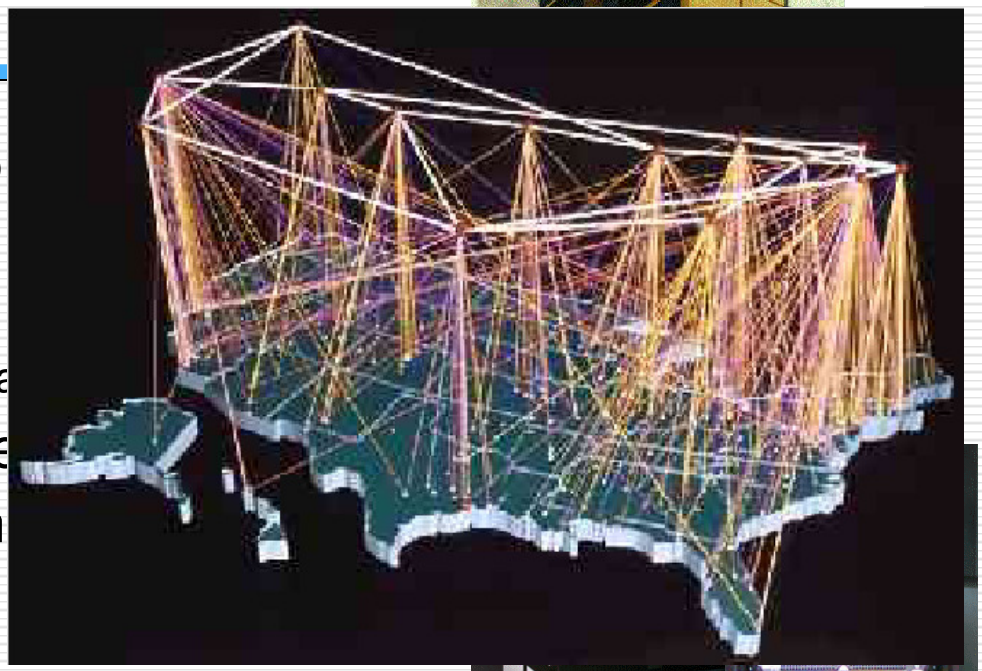
- Books
  - I. Foster and C. Kesselman, *The Grid: Blueprint for a New Computing Infrastructure*. Morgab Kaufmann Publishers, 1999.
  - Maozhen Li, Mark Baker, The Grid Core Technologies, Wiley, 2005.
  - Fran Berman, Anthony J. G. Hey and Geoffrey C. Fox, *Grid computing: Making the Global Infrastructure a Reality*. John Wiley & Sons Ltd, 2003.
  - Luis Ferrelra et al. *Grid Services Programming and Application Enablement*. IBM Redbooks, 2004.
  - Mark Endrei et al. *Patterns: Service Oriented Architecture and Web Services*. IBM Redbooks, 2004.
- Web links:
  - Globus project: http://www.globus.org/alliance/
  - Global Grid Forum: http://www.ggf.org
- Course resources: http://www.cse.hcmut.edu.vn/~ptvu/gc

Introduction to Grid Computing

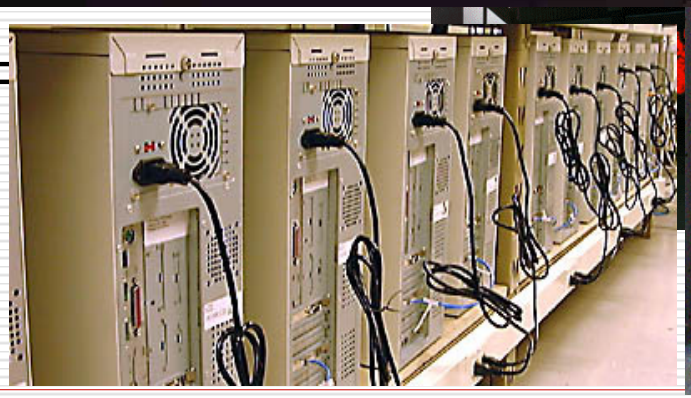
Let's Imagine…



A Bit of History …



- I.Foster and C.Kesselman co
  computing field Parallelism
  - 1970-80 : The pioneers
    - Vectorial computers (Cra
  - 1980-90 : The exuberance
    - Massively Parallel Com
      Connection Machine,…
  - 1990-2000 : The profit
    - Cluster (Beowulf, ASCII F
  - 2000 ->? The Globalization
    - GRID
    - Global Computing



Definitions…

- What is a Grid?
- Analogous to grid in electrical power grids

  (Foster & Kesselman 1999, Chetty & Buyya 2002)
- "Hardware and software infrastructures that provides dependable consistent, pervasive and inexpensive access to high-end computational capabilities"

  (Foster & Kesselman 1999)
- "Flexible, secure, coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organisations"

  (Foster et al. 2001)

… and Definitions…

- "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"

(Buyya 2002)

- The three point checklist:
  - (i) Coordinates resources that are not subject to a centralised control
  - (Ii) Uses standard, open, general purpose protocol and interfaces
  - (iii) Delivers nontrivial quality of service

(Foster 2002)

My Keywords

- “Distributed”
- “Coordinated sharing”
- “Resources”
- “Large scale”
- “Non-trivial”
- “Big problem”
- “Internet”

Grid Generations

- 1st generation - sharing high performance computing resources in distributed environment
- 2nd generation - middleware to address issues of scalability, heterogeneity and adaptability in distributed environments with the focus on large scale computational power and huge volumes of data
- 3rd generation - addresses the requirements for distributed collaboration in virtual environments; service oriented; automation; knowledge technology

De Roure, Baker, Jennings & Shadbolt (2003)

Types of Grid

- Usually one distinguishes three types of GRID
  - Knowledge GRID
    - To share knowledge
    - Idea: Virtual distributed laboratory
  - Data GRID
    - To share the data
    - Idea: Large scale data storage
  - Computing GRID
    - To share computing power
    - Idea: Alternative to supercomputers

The GRID today -1

- Knowledge GRID
  - Virtual distributed Institutes/Laboratories allowing remote people to collaborate as they would be at the same location
  - Purpose: to share knowledge, to collaborate for researches or specific activities
  - Few specific tries
    - Remote surgery
    - ….

The GRID today - 2

- Data GRID
  - Also referenced as « peer-to-peer » technology
    - Versus to the « Client/Server » technology who dominated the market of distributed systems during last decades, with this approach every participant is both client and server
  - Nobody knows exactly where the data are
  - The system figures to find the nearest requested data out
    - The requested data could be distributed on several location and same data could be located at several places (data coherence problem)
  - The pursuit of the data is transparent to the user
  - Some examples:
    - DATAGRID (CERN), The British Atmospheric Data Centre…

The GRID today - 3

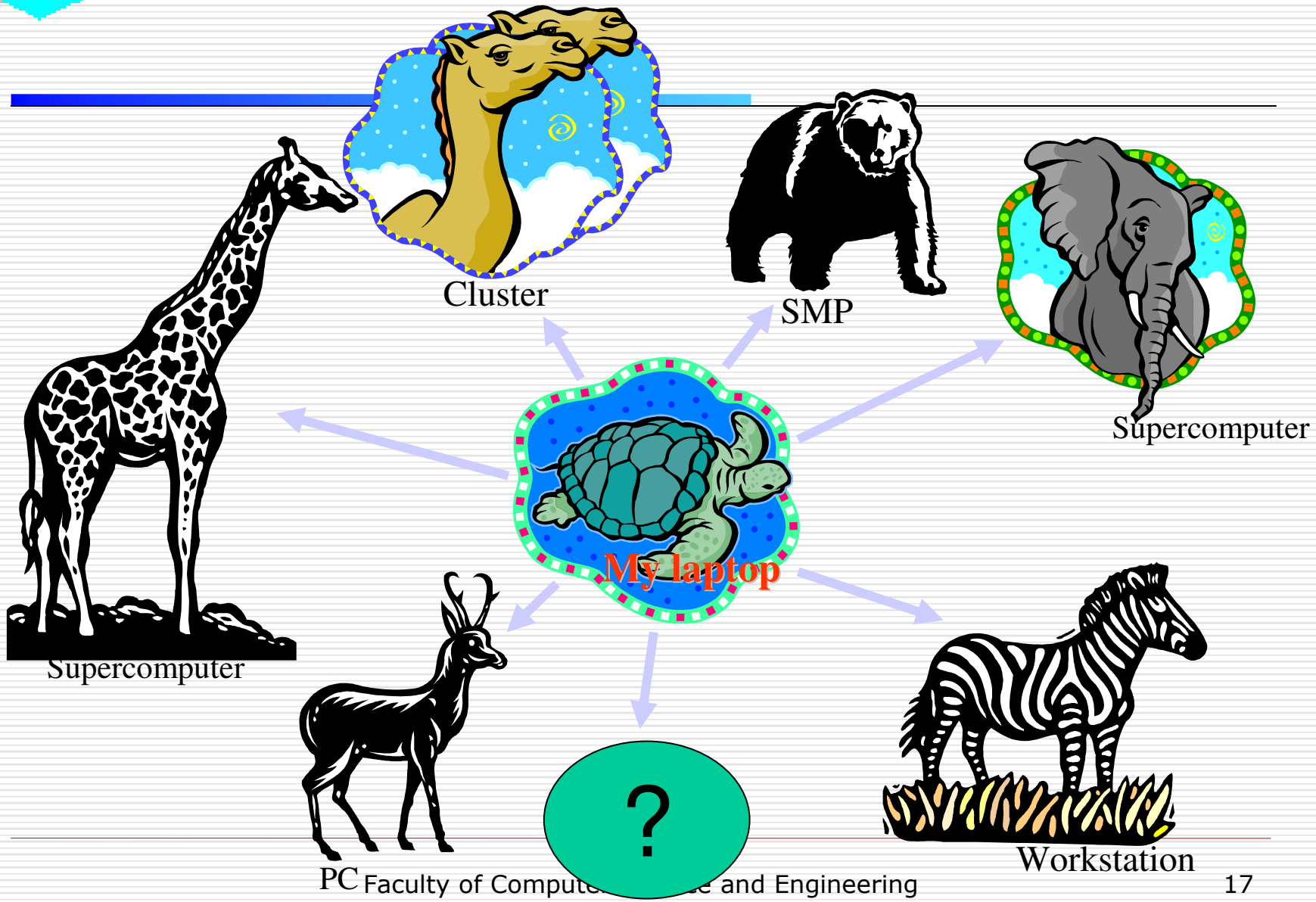
- Computational GRID
  - A supercomputer on you the table of your kitchen …
- Some facts
  - Supercomputers are very expensive and age very rapidly
  - Except for some very specific persons, we do not need a supercomputer every day…

    … but when we need one it is very frustrating not to have one !
  - The total time personal and other computer waste time not doing anything is huge
  - I am (almost) permanently connected to a lot of computers.
  - Potentially I am permanently connected to a supercomputer
    - First try to exploit this fact: Project SETI@home

“Grid” vs. cluster computing

- “Partial view” vs. “Global view” of the environment
- Large scale, geographically distributed
- Non-structured
- Heterogeneous
- Dynamic, unstable
- Multiple administrative domains

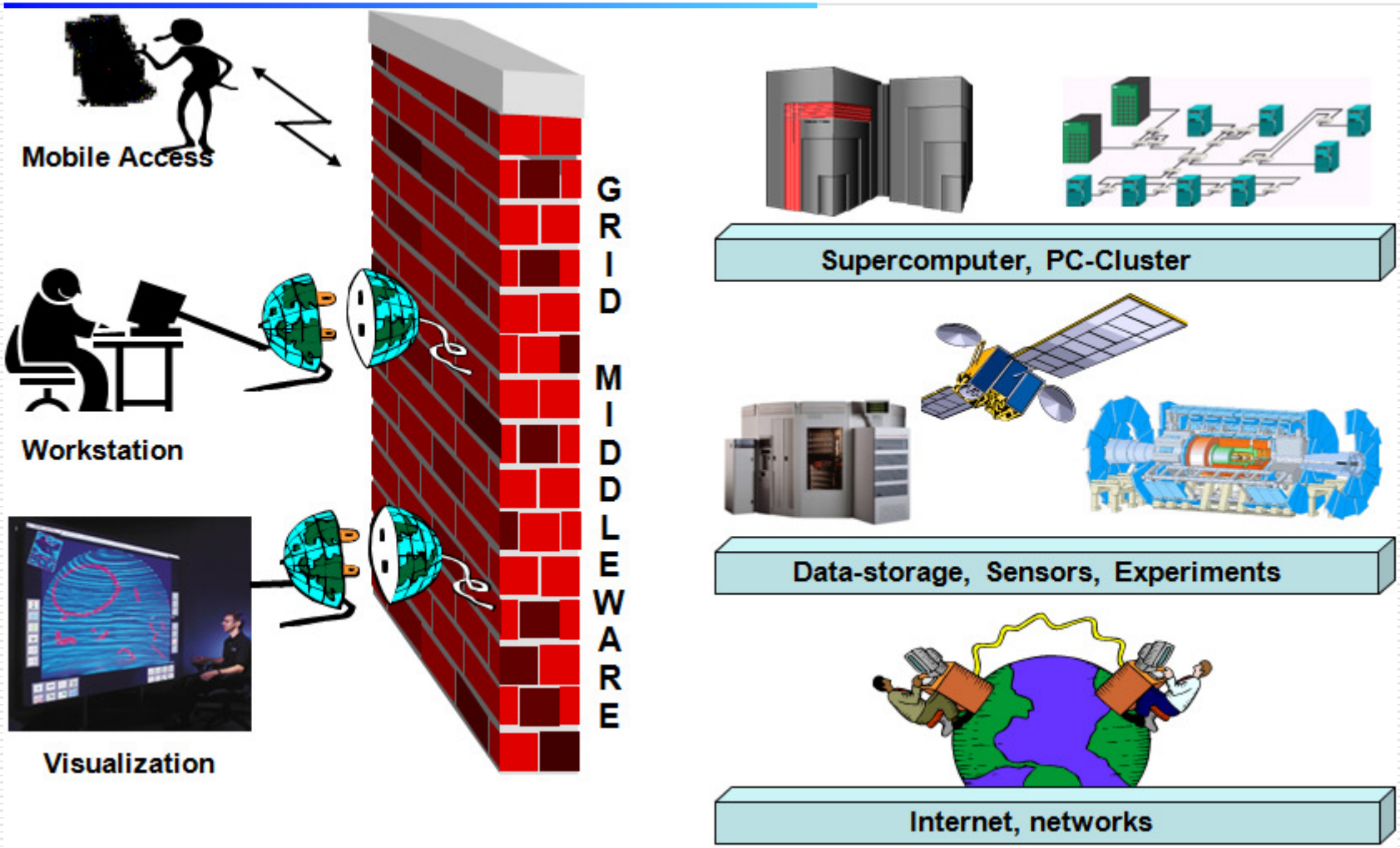
Computing Grid: our goal



The Dream

- To make computers user friendly: To switch
  - from a program centric view…
    - « Execute this program on this (these) machine(s)»
  - to a service centric view…
    - « Provide me this service with this QoS level »

The Grid We are Building…



Grid Characteristics

- Large-scale
  - Need for dynamic selection
  - Partial view of the environment
- Heterogeneity
  - Hardware, OS, network, software environments (languages, libraries, tools...)
- Complex
  - unpredictable structure
- Dynamic
  - unpredictable behaviour
- Multiple administrative domain
  - no centralized control

Challenges (1)

- Efficient exploitation of Grid performance
  - New programming models
  - Service-centric vision
- Resource connectivity
  - Middleware services
  - Applications
- Interoperability vs. performance
  - Good interoperability -> slow performance
  - Fast -> not so interoperable

Challenges (2)

- Grid scalability
  - Resource management mechanisms/models
    - Large number of heterogeneous resources
- Grid infrastructure and Grid application evaluations
  - How to benchmark Grid systems and applications
- Trust and security
  - How to trust the environment
  - How to protect sensitive data