

---

## DISTRIBUTED SYSTEMS (COMP9243)

Slide 1

### Lecture 10b: Grid Computing

- ① Grid Computing
- 
- 

## GRID COMPUTING

Slide 2

A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities  
(Foster & Kesselman, 1998)

---

---

## MOTIVATION

Processing Power like the Electrical Grid:

- Plug an appliance into the socket
- It works "on demand"
- Don't have to worry about who is producing electricity
- Don't have to worry about how it's produced

Ideally:

Slide 3

- Plug a parallel program into the Grid. *It works*
- Don't have to worry about where it is running
- Don't have to worry about where resources are

Supercomputing for the Masses:

- Not everyone can afford their own supercomputer
  - Take advantage of scarce resources (e.g., a telescope)
  - Take advantage of the many 'idle' resources on the net
  - 'Dynamic' clusters
- 
- 

## GRID DEFINITION

A system that:

- Coordinates resources not subject to centralised control
- Uses standard, open, general-purpose protocols and interfaces
- Delivers nontrivial qualities of service

Slide 4

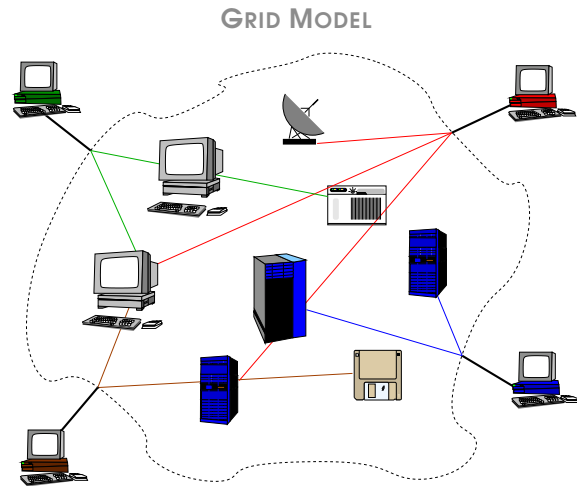
What *isn't* a Grid:

- Local clusters - centralised control
- World Wide Web - no qualities of service
- P2P file sharing - application specific protocols

*The Grid:*

- Standard protocols
  - Open Grid Services Architecture
-

Slide 5



Slide 7

**Scalability:**

- Number of resources
- Number of applications
- Geographic
- Administrative

**Dynamic:**

- Resource failure
- Resource availability fluctuates
- Addition of new resources

Slide 6

**GRID CHARACTERISTICS**

**Multiple Administrative Domains:**

- Geographically distributed
- Distributed over administrative domains
- Resource owners are autonomous

**Heterogeneity:**

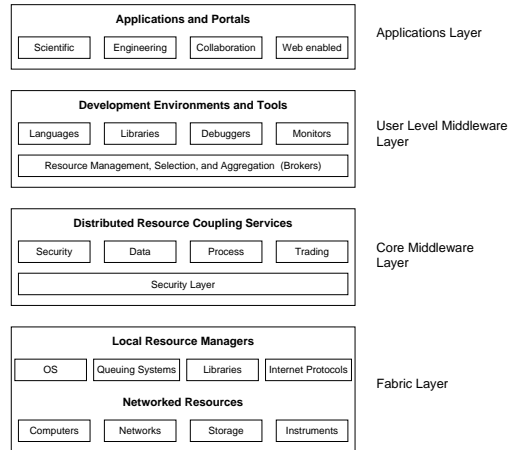
- Heterogeneous resources
- Hardware, OS, other software

Slide 8

**GRID SERVICES**

- **Computational services:** executing application jobs on computational resources
- **Data services:** providing access to and processing of distributed data
- **Application services:** providing access to remote software and libraries
- **Information services:** extraction and presentation of data
- **Knowledge services:** building new knowledge (e.g. data mining)

## GRID ARCHITECTURE



Slide 9

### Fabric:

- Computers
- Operating systems
- Storage devices
- Network
- Databases
- Sensors
- Specialised devices (e.g. telescope)

Slide 10

### Core Middleware:

- Process management
- Storage access
- Communication
- Security
- Example: Globus

Slide 11

### User-Level Middleware:

- Resource management
- Resource discovery
- Application scheduling
- Data replication
- Monitoring and diagnostics
- Application development environments and programming tools
- Example: Network Weather Service

Slide 12

Application Characteristics:

- Clearly identified problem(s) to be solved
- Knows what resources are required solve the problem
- Specifies required quality of service

Example:

→ An application that executes thousands of independent tasks, each taking as input some set of files containing events

Slide 13

- 1 Obtain necessary authentication credentials
- 2 Determine availability of computers, storage systems, and networks, and the location of required input files
- 3 Submit requests to appropriate computers, storage systems, and networks to initiate computations, move data, etc.
- 4 Monitor progress of the computations and data transfers, detect and respond to failure conditions

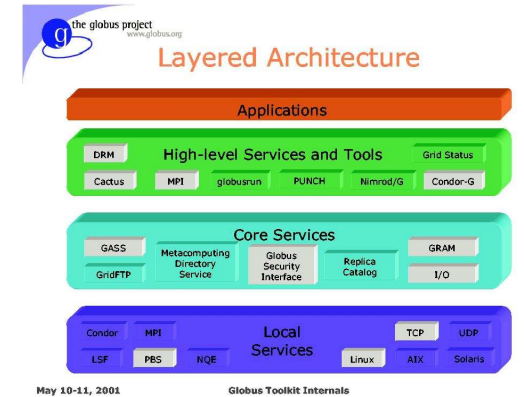
Application must be a distributed application!

## GLOBAL TOOLKIT

Overview:

- De facto Grid middleware
- Virtual machine view of heterogeneous infrastructure
- Provides basic services and capabilities for a Grid
- Well defined service APIs and protocols
- Security done right
- Layered architecture

Slide 14



Slide 15

Core Globus Services:

- Globus Security Interface (GSI)
- Metacomputing Directory Service (MDS)
- Globus Resource Access Manager (GRAM)

Other Globus Services:

- GridFTP: secure file transfer
- Globus Access to Secondary Storage (GASS): cache files locally (where they will be accessed)
- Replica Catalog: LDAP registry for managing replicated files

Slide 16

Globus Enabled Tools:

- MPICH-G: Globus enabled MPI
- Condor-G: Condor for Globus

---

## GLOBUS SECURITY INTERFACE (GSI)

### Overview:

- Certificate based (X.509 certificates)
- All users require a certificate (public key signed by certificate authority)
- Send certificate to resource's authenticator
- Authenticator checks certificate authenticity
- Authenticator returns encrypted challenge

Slide 17

### Trust Assumptions:

- Administrator who installed Globus
- Certificate authority

### Other functions of GSI:

- Delegation: create traceable authentication chain
  - Integrity: ensure integrity of communication
  - Confidentiality: ensure confidentiality of communication
- 

## METACOMPUTING DIRECTORY SERVICE (MDS)

### Overview:

- Provide information about the Grid
- Uniform format
- Standard protocols
- LDAP based

Slide 18

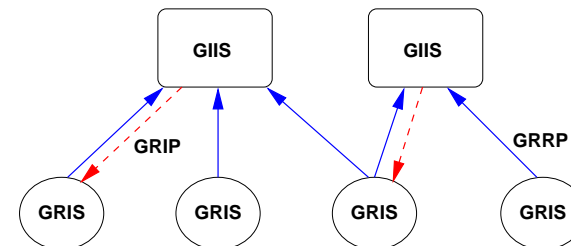
### Requirements:

- Universal naming scheme
  - Query support
  - Update support
  - Access control (for information in MDS)
- 

### Architecture:

- Hierarchy of servers
  - Grid Resource Information Server (GRIS)
    - Gathers and maintains raw data
    - Data provided by information providers
  - Grid Information Index Server (GIIS)
    - Indexes raw data
    - Different GIISes provide indexes that optimise for different views of data
  - Grid Resource Registration Protocol (GRRP)
    - Protocol to register information about data
    - GRIS registers information about its data with GIIS
    - Soft-state - registered information must be renewed
  - Grid Resource Interrogation Protocol (GRIP)
    - Protocol to retrieve data
    - GIIS request data from GRIS
- 

Slide 19

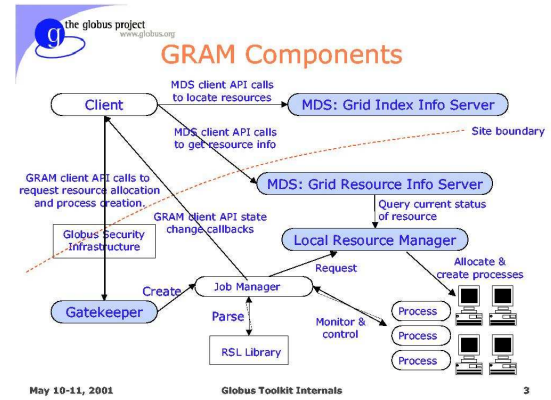
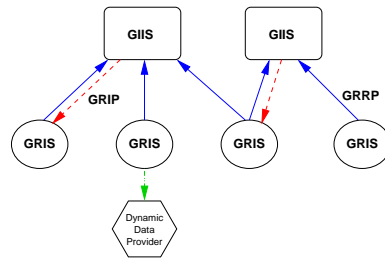


Slide 20

Dynamically Generated Data:

- Information provider protocol
- Data generator registers itself with GRIS
- GRIS contacts information provider when data required

Slide 21



Slide 23

**GLOBUS RESOURCE ACCESS MANAGER (GRAM)**

Overview:

- Enables the launching of jobs
- Resource Specification Language (RSL) to specify attributes of resources
- Job manager
- Local resource manager

Slide 22

**GLOBUS TOOLKIT 4**

Newest release of Globus Toolkit

- April 2005
- Added a bunch of Web Services (WS) stuff
  - WS (e.g., SOAP, WSDL) based protocols
  - WS standard based interfaces (in particular WSRF)
  - WS versions of Globus services
- Redesign of GRAM - more scalable
- Redesign of GridFTP
  - replication
- Extended security infrastructure
- Java, Python, C core language support
- Improved documentation

Slide 24