Grid Computing

Lectured by: Dr. Pham Tran Vu

Email: ptvu@cse.hcmut.edu.vn

Grid Architecture

Outline

- Layer Architecture
- Open Grid Service Architecture

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

In short, Grid Computing is about…

- Resource sharing
  - Resource: anything that can be shared, e.g. a file, CPU cycles, storage, a program, something in your mind,…
- Interoperability
  - The ability that two different entities can work together
  - Enabled by common sets of protocols

More Definitions…

- Network Protocol
  - A formal description of message formats and a set of rules for exchange of messages
  - Rules define sequences of message exchange, and potentially resulting behavior
    - Protocol may define state-change in endpoint
- Network Enabled Services
  - An network enabled entities
  - Defines a set of capabilities can be performed over the network
  - Protocol defines interaction with service
    - All services require protocols, although not all protocols are to services

API and SDK

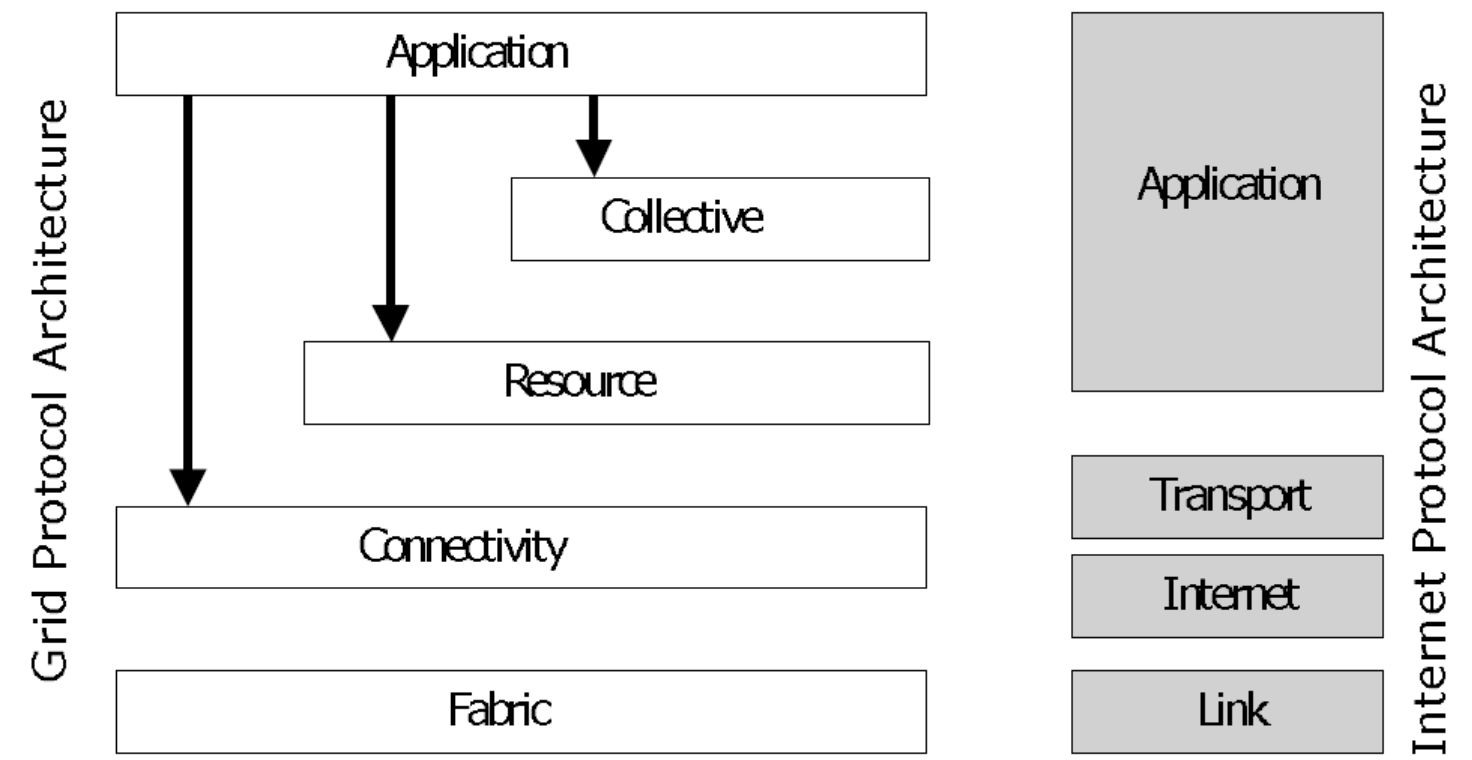
- API: Application Programming Interface
  - A set interfaces and libraries for building applications
  - To enable interoperability
- SDK: System Development Kit
  - Package that enables application development, consisting of one or more APIs, and programming tools

Grid Layer Architecture

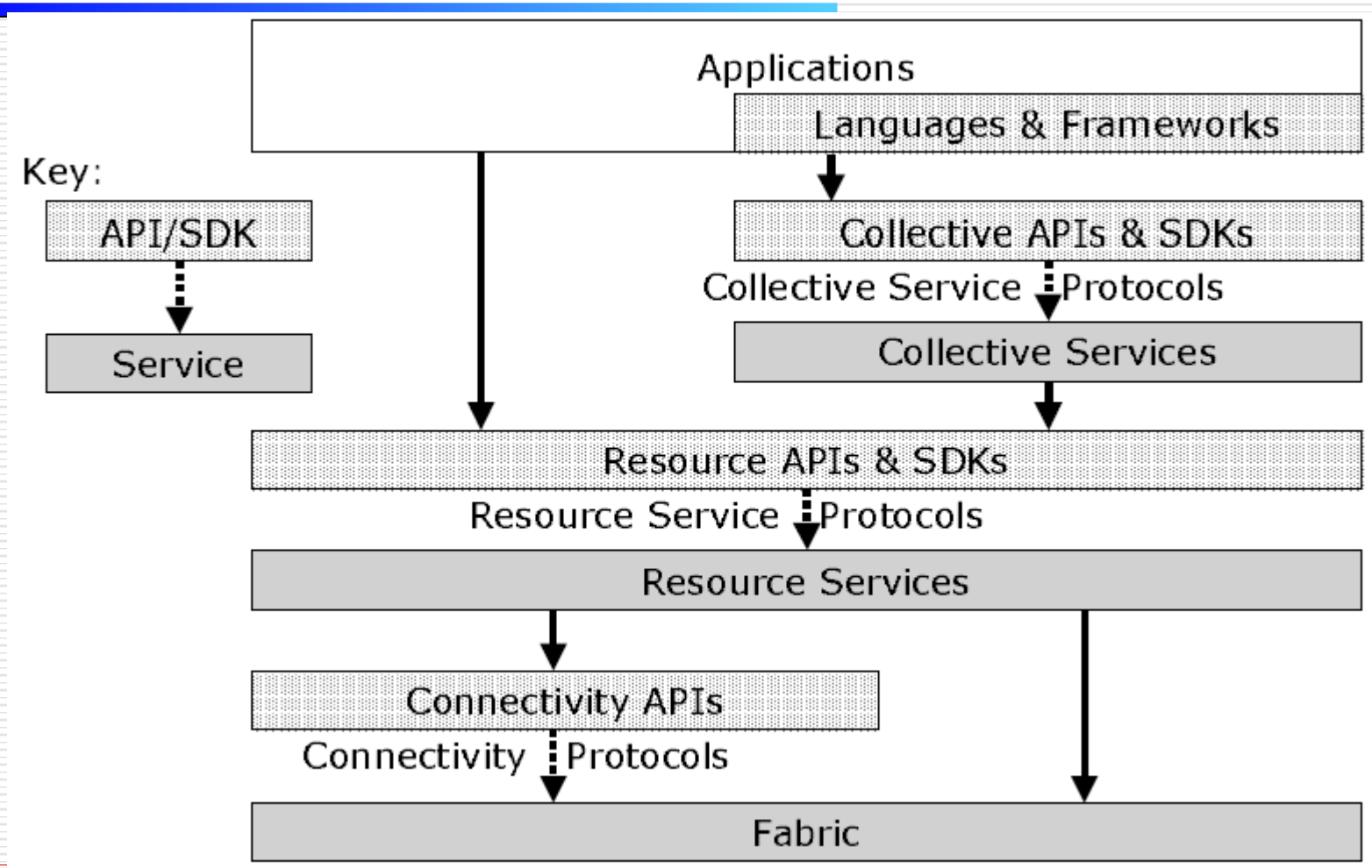
- Provides a layer view of Grid
- Based on the current Internet architecture
- Focused on protocols to enable interoperability

Grid Layer Architecture

From The Anatomy of the Grid (Foster, Kesselman and Tuecke, 2001)



Protocols, services and interfaces



Fabric

- Local resources to be shared
- Physical
  - computational resources, storage systems, catalogues, network resources, and sensors
- Logical
  - distributed file system, computer cluster, or distributed computer pool
- Specific to local operating environments

Connectivity

- Provide access to individual resources and services in the fabric layer
- Communication protocols
  - enable the exchange of data
- Authentication protocols
  - provide cryptographically secure mechanisms for verifying the identity of users and resources

Resource

- Sharing single resources
- Define protocols (and APIs and SDKs) for the secure negotiation, initiation, monitoring, control, accounting, and payment of sharing operations on individual resources
- Information protocols
  - to obtain information about the structure and state of a resource, for example, its configuration, current load, and usage policy
- Management protocols
  - negotiate access to a shared resource

Collective

- Coordinating multiple resources
- Directory services
  - Resource discovery
- Co-allocation, scheduling, and brokering services
- Monitoring and diagnostics services
- …

Application

- User applications that operate within the Grid environment
- May call sophisticated frameworks and libraries
  - Eg. CORBA, Workflow systems

Examples

	Multidisciplinary Simulation	Ray Tracing
Collective (application-specific)	Solver coupler, distributed data archiver	Checkpointing, job management, failover, staging
Collective (generic)	Resource discovery, resource brokering, system monitoring, community authorization, certificate revocation	
Resource	Access to computation; access to data; access to information about system structure, state, performance.	
Connectivity	Communication (IP), service discovery (DNS), authentication, authorization, delegation	
Fabric	Storage systems, computers, networks, code repositories, catalogs	

Open Grid Service Architecture (OGSA)

- A standard for building Grid infrastructures and applications
- Based on Service Oriented Architecture and Web Service standards
- Resources are represented as services

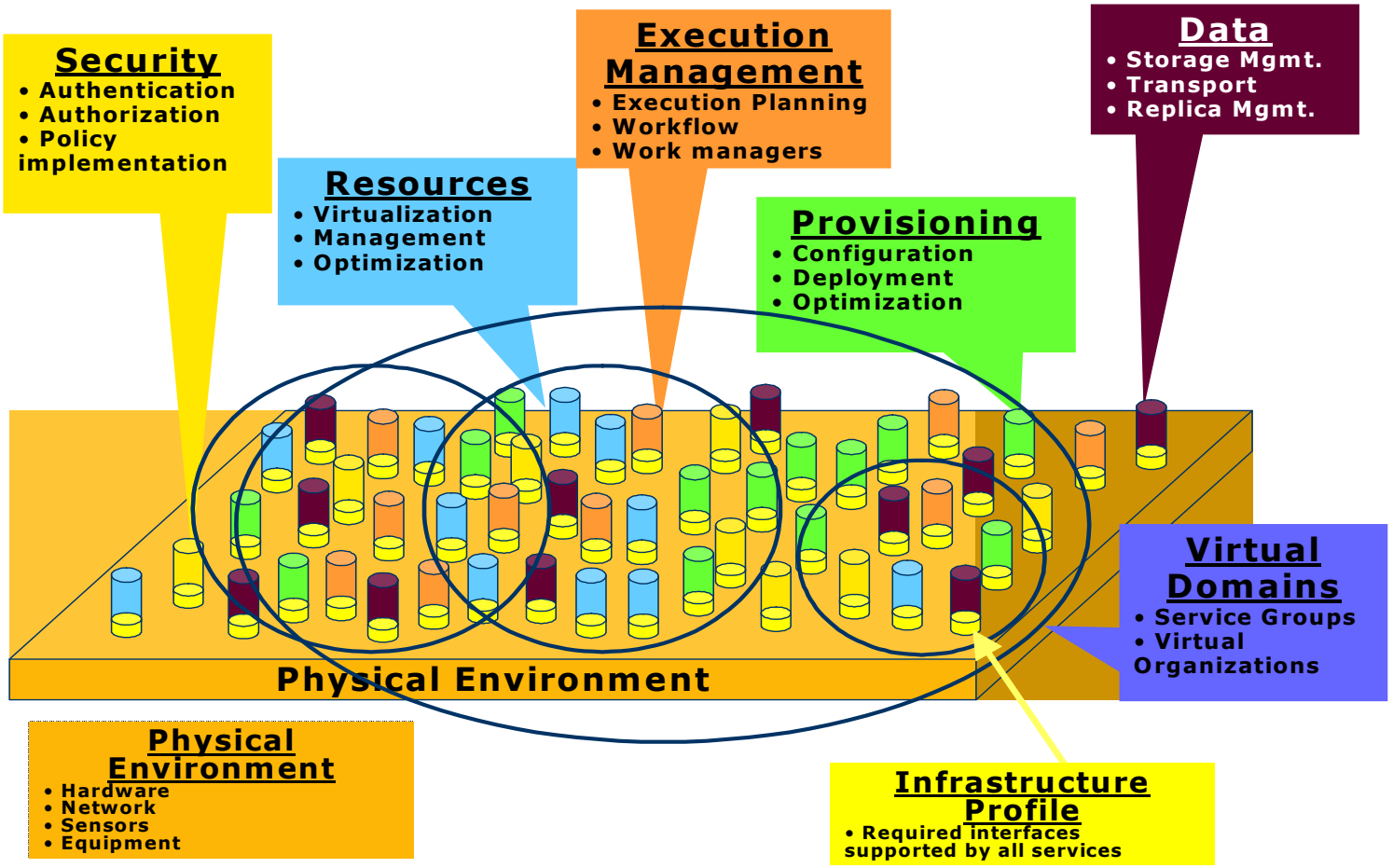
OGSA Capabilities (1)

- *Execution Management Services*
  - address problems with executing a unit of work, including their placement, provisioning and lifetime management.
- *Data Services*
  - used to move data, manage replicated copies, run queries, update and transform data to new format.
- *Resource Management Services*
  - deal with the management of resources themselves (e.g. rebooting a host), the resources on Grid (e.g. resource reservation and monitoring) and OGSA infrastructure.

OGSA Capabilities (2)

- *Security Services*
  - facilitate the enforcement of security related policy within Grid environments.
- *Self-Management Services*
  - help reduce the cost and complexity of owning and operating IT infrastructure.
- *Information Services*
  - access and manipulate information about applications, resources and services in Grid environments.

Service-centric View of the Grid



Seminar Topics

- Seminar 1: Potential Applications of Grid Computing in Vietnam
- Seminar 2: Grid Computing Middleware
- Seminar 3: Grid Resource Management
- Seminar 4: Grid Computing Security
- Seminar 5: Applications of SOA and Web Services in Grid Computing
- Seminar 6: Semantic Grid: From Concepts to Implementation
- Seminar 7: Grid Computing & e-Science
- Seminar 8: Grid Economics