# The Evolution of the Grid

David De Roure, Mark A. Baker, Nicholas R. Jennings, Nigel R. Shadbolt

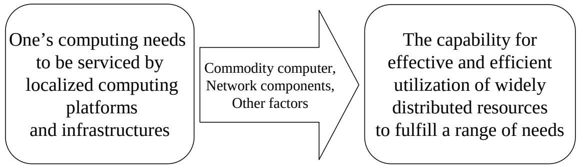
Grid Computing – Making the Global Infrastructure a Reality
Fran Berman, Geoffrey Fox, and Tony Hey
2002 John Wiley & Sons, Ltd.
Chapter 3

Based on slides by Kwangwon, Koh



# Introduction

- The situation has changed




- The issues in designing, building and deploying distributed computer systems have now been explored over many years
  - Metacomputing, Scalable computing, Global computing, Internet computing, and Grid computing



# Introduction (cont.)

- Grid problem (Foster, Kesselman, Tuecke, 2001)
  - Flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources
    - Emphasizes the importance of information aspects to essential for resource discovery and interoperability
    - XML, Resource Description Framework (RDF)



# The Evolution of the Grid: The First Generation

- The early Grid efforts started as projects to link supercomputing sites
- FAFNER and I-WAY had to overcome a number of similar hurdles, including communications, resource management, and the manipulation of remote data, to be able to work efficiently and effectively

## The First Generation: FAFNER 1995

- FAFNER – Factoring via Network-Enabled Recursion
- To factor RSA130 using a new numerical technique called the Number Field Sieve (NFS) factoring method using computational Web servers
- Contributors downloaded a sieving software daemon that used HTTP GET/POST to get values and post results
- CGI scripts supported cluster management minimizing impact on owners by regulating day/night usage
- Three factors combined to make this approach successful:
  - The NFS implementation allowed even workstations with 4Mbytes of memory to perform useful work
  - FAFNER supported anonymous registration
  - A hierarchical network of servers reduced the potential administration bottleneck



## The First Generation: I-WAY

- I-WAY was an experimental high performance network linking many high performance computers (17 sites) and advanced visualization environments
- Not to build a network but to integrate existing high bandwidth networks (ATM)
- To help standardize the I-WAY software interface and management, the key sites installed I-POP servers to act as gateways to I-WAY
  - The I-POP servers were UNIX workstations configured uniformly and possessing a standard software environment called I-Soft
  - The I-POP server mechanisms allowed uniform I-WAY authentication, resource reservation, process creation, and communication functions
- I-WAY developed
  - resource scheduler Computational Resource Broker (CRB)
  - Security using telnet modified to use Kerberos
  - Used Andrew File System (AFS)
- To support user-level tools, a low-level communications library, Nexus, was adapted to execute in the I-WAY environment



## A Summary of Early Experience

- FAFNER and I-Way attempted to produce metacomputing environments by integrating resources from opposite ends of the computing spectrum
- FAFNER was tailored to a particular factoring application that was in itself trivially parallel and was not dependent on a fast interconnect
  - Forerunner of the likes of SETI@home and Distributed.Net
- I-WAY was designed to cope with a range of diverse high performance applications that typically needed a fast interconnect and powerful resources
  - Forerunner of the likes of Globus and Legion
  - Had a few shortcomings
    - I-POP had to be specially set up and was single point of failure



## The Evolution of the Grid: The Second Generation

- Today the grid infrastructure is capable of binding together more than just a few specialized supercomputing centers
- Grid: a viable distributed infrastructure on a global scale that can support diverse applications requiring large-scale computation and data
  - Three main issues
    - Heterogeneity
    - Scalability - latency tolerance, handle authentication over organizational boundaries
    - Adaptability – tolerant of resource failure
- Middleware is used to hide the heterogeneous nature and provide users and applications with a homogeneous and seamless environment by providing a set of standardized interfaces to a variety of services on the Grid

## Requirements for the data and computation infrastructure

- Administrative Hierarchy
- Communication Services – various types needed, QoS
- Information Services – dynamic info on location/type of services
- Naming Services – uniform namespace (X500 or DNS)
- Distributed File Systems and Caching – uniform global namespace
- Security and Authorization
- System Status and Fault Tolerance - monitoring tools required
- Resource Management and Scheduling – efficient global scheduling that interact with local schedulers
- User and Administrative GUI – intuitive/heterogeneous, usually Web-based

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 9

## Second Generation Core Technologies

- Globus
  - A software infrastructure that enables applications to handle distributed heterogeneous computing resources as a single virtual machine
  - The GTK provides a bag of services which developers of specific tools or applications can use to meet their own particular needs
  - The GTK is modular and consists of the following:
    - GRAM: Globus Toolkit Resource Allocation Manager – HTTP based
    - GridFTP – security, parallelism
    - GSI: Grid Security Infrastructure- authentication
    - LDAP: Lightweight Directory Access Protocol – distributed access to structure/state information
    - GASS: Global Access to Secondary Storage
    - GEM: Globus Executable Manager – construction, caching, location of executables
    - GARA: Globus Advanced Reservation and Allocation

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 10

## Second Generation Core Technologies (cont.)

- Legion - UVa
  - An object-based 'metasystem'
  - A system to enable heterogeneous, geographically distributed, high performance machines to interact seamlessly
  - Provide users with a single integrated infrastructure regardless of scale, physical location, language and underlying operating system
  - Difference from Globus
    - Encapsulate all of its components as objects
  - Core object types : classes/metaclasses, host objects, vault objects, implementation objects/caches, binding agents, context objects/spaces

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 11

## Distributed Object Systems

- Common Object Resource Broker Architecture (CORBA)
  - Open distributed object-computing infrastructure
  - Automates network programming tasks such as object registration, location activation; request de-multiplexing; framing/error handling; parameter marshalling; operation dispatching
  - Negatives: may be difficulty in crossing organizational boundaries; real-time/multimedia not in original design
- Java/JVM
  - Single implementation framework; mask heterogeneity
  - Can have Java wrappers around non-Java code
  - Negatives: computational speed; concurrency

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 12

## Distributed Object Systems

- Jini and RMI
  - To provide a software infrastructure that can form a distributed computing environment that offers network plug and play
  - In Jini, applications will normally be written in Java and communicate using the Java Remote Method Invocation (RMI) mechanism
  - Concepts
    - Lookup: search for a service and download the code needed to access it;
    - Discovery: spontaneously find a community and join;
    - Leasing: time-bounded access to a service;
    - Remote Events: service A notifies service B of A's state change.
    - Transactions: used to ensure that a system's distributed state stays consistent
  - Does not provide file system, scheduling or logins



## Grid Resource Brokers and Schedulers

- Batch and Scheduling Systems
  - Condor
    - A software package for executing batch jobs on a variety of UNIX platforms, in particular those that would otherwise be idle
    - Features: resource location, job allocation, check pointing, migration
  - PBS (Portable Batch System)
    - A batch queuing and workload management system
    - Allows sites to determine own scheduling policies
  - SGE (Sun Grid Engine)
    - Jobs (with requirements profile) wait in a holding area and queues located on servers provide the services for jobs
  - LSF (Load Sharing Facility)
    - A commercial system from Platform Computing Corp.
    - LSF comprises distributed load sharing/ batch queuing and monitoring of resources/workloads on a network of heterogeneous computers



## Grid Resource Brokers and Schedulers (cont.)

- Storage Resource Broker (SRB)
  - To provide "uniform access to distributed storage" across a range of storage devices
  - Key feature
    - SRB supports metadata associated with a distributed file system, such as location, size and creation date information
  - SRB is attractive for Grid applications in that it deals with large volumes of data



## Grid Resource Brokers and Schedulers (cont.)

- Nimrod/G Resource Broker and GRACE
  - Nimrod-G : a Grid broker that performs resource management and scheduling of parameter sweep and task-farming applications
  - Components
    - A task farming engine: allows user defined schedulers, customized applications or problem solving environments
    - A scheduler: support resource discovery, selection, scheduling, and the execution of user jobs on remote resources
    - A dispatcher: uses Globus for deploying Nimrod G agents on remote resources in order to manage the execution of assigned jobs
    - Resource agents
  - Supports user-defined deadline and budget constraints for scheduling optimization and manages the supply and demand of resources in the Grid using a set of resource trading services called GRACE (GRid Architecture for Computational Economy)
    - Scheduling algorithms: Cost, time, cost-time optimization, conservative

## Grid Portals

- A Grid portal provides access to grid resources specific to a particular domain of interest
- The NPACI HotPage hotpage.npaci.edu
  - A user portal that has been designed to be a single point-of-access to computer-based resources, to simplify access to resources that are distributed across member organizations and allows them to be viewed either as an integrated Grid system or as individual systems
  - Provides information, resource access/management
- The SDSC Grid Port Toolkit gridport.npaci.edu
  - A reusable portal toolkit that uses HotPage infrastructure
  - Web portal to allow the execution of secure portal services and an application API to enable development of customized science portals

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 17

## Integrated Systems

- Cactus
  - An open source problem-solving environment
  - Modular structure
  - Provide a front-end to many core backend services
    - Globus, HDF5 file library, PETSc, visualization
- EU DataGrid eu-datagrid.Web.cern.ch
  - A computational and data-intensive Grid of resources for the analysis of data coming from scientific exploration
    - LHC (Large Hadron Collider) , which will operate at CERN from about 2005 to 2015
    - Main challenge of the LHC is providing the means to share many petabytes of distributed data over the current network infrastructure

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 18

## Integrated Systems (cont.)

- UNICORE - German Ministry of Education& Research
  - Design goal
    - A uniform and easy to use GUI
    - An open architecture based on the concept of an abstract job
    - A consistent security architecture
    - Minimal interference with local administrative procedures
    - Exploitation of existing and emerging technologies through standard Java and Web technologies
  - Protocol
    - Between the components is defined in terms of Java objects
    - Low level layer called UPL(UNICORE Protocol Layer) handles authentication, SSL communication and transfer of data as in lined byte streams
    - High level layer contains the classes to define UNICORE jobs, tasks and resource requests

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 19

## Integrated Systems (cont.)

- WebFlow
  - A computational extension of the Web model that can act as a framework for wide-area distributed computing
  - Design goal
    - Build a seamless framework for publishing and reusing computational modules on the Web
    - End users can compose distributed applications using WebFlow modules via Web browser
  - The front-end uses applets for authoring, visualization, and control of the environment
  - The backend of WebFlow was implemented using the Globus Toolkit

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 20

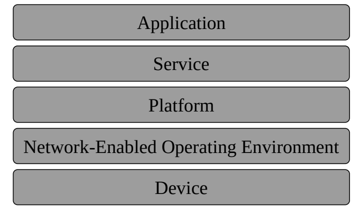
## Peer-to-Peer Computing

- Decentralisation
- To overcome a performance bottleneck and a single point of failure of client-server model
- P2P Computing
  – Napster, Gnutella, Freenet, JXTA
- Internet Computing
  – SETI@home, Parabon, Entropia
- Machines share data and resources via Internet or private networks
- Obstacles
  – Require more computing power to handle communication & security
  – Security
  – Heterogeneous
  – Find resources and services



## Peer-to-Peer Computing (cont.)

- JXTA
  – An open source development community for P2P infrastructure and applications




  – A specification rather than software

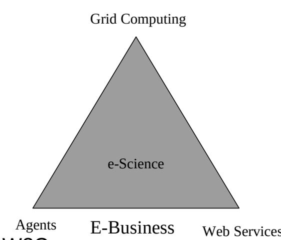


## The Evolution of the Grid: The Third Generation

- Emphasize ‘autonomic’
  – Needs detailed knowledge of its components and status
  – Must configure and reconfigure itself dynamically
  – Seeks to optimize its behaviour to achieve its goal
  – Is able to recover from malfunction
  – Protect itself against attack
  – Be aware of its environment
  – Implement open standards
  – Make optimized use of resources



## Service-oriented architectures




- Web Services - W3C
  – SOAP - Simple Object Access Protocol
    - Envelope for XML data, RPC
  – Web Services Description Language (WSDL)
    - describes services in XML
  – Universal Description Discovery and Integration (UDDI)
    - Distributed registeries

## Service-oriented architectures (cont.)

- The Open Grid Services Architecture (OGSA) Framework (Globus/IBM)
  - Supports the creation, maintenance, and application of ensembles of services maintained by VOs (virtual organization)
  - The standard interfaces
    - Discovery
    - Dynamic service creation
    - Lifetime management
    - Notification
    - Manageability
    - Simple hosting environment
  - Globus
    - GRAM - resource allocation & management
    - MDS 2- metadirectory services
    - GSI - Grid security infrastructure (single sign on)

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 25

## Service-oriented architectures (cont.)

- Agents
  - Characteristics
    - Autonomy
    - Social ability – interact with other agents
    - Reactivity – perceive/respond to environment
    - Pro-activeness - goal directed behavior
  - FIPA (Foundation for Intelligent Physical Agents)
    - Agents communicate by exchanging messages which represent speech acts, and which are encoded in an agent-communication-language
    - Services provide support agents
    - Services may be implemented either as agents or as software

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 26

## Information aspects: relationship with the WWW

- The Web as a Grid Information Infrastructure
  - Does information distribution architecture meet grid requirements?
    - Version control - none in www
    - QoS – link fragility
    - Provenance - no certified publication date
    - Digital Rights Management - no copy protection
    - Curation – www delivery oriented, no content management

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 27

## Information aspects: relationship with the WWW (cont.)

- Expressing Content and Meta-content
  - XML
    - Information exchange is facilitated by the XML
    - Designed to mark up documents
    - The tags are defined for each application using a DTD (Document Type Definition) or an XML Schema
  - RDF
    - A standard way of expressing metadata
    - RDF Schema permit definition of a vocabulary
  - XML and RDF enable the standard expression of content and metacontent

DiSCoV KENT STATE September 2004 Paul A. Farrell Grid Computing 28

## Live Information Systems

- Access Grid
  - A collection of resources that support human collaboration across the Grid, including large-scale distributed meetings and training
  - Current Access grid infrastructure is based on IP multicast
- The combination of Semantic Web technologies with live information flows is highly relevant to grid computing



## Summary and Discussion

- First generation systems involved proprietary solutions for sharing high performance computing resources
- Second generation systems introduced middleware to cope with scale and heterogeneity, with a focus on large scale computational power and large volumes of data
- Third generation systems are adopting a service-oriented approach, adopt a more holistic view of the e-Science infrastructure, are metadata-enabled and may exhibit autonomic features



## The Semantic Grid