Grids and High Performance Distributed Computing

Andrew Chien

March 31, 2004

CSE225, Spring 2004

Last Time: Vision of the Grid

- Flexible shared infrastructures that can be automatically configured and adapted to use
  - » "Utility", "Shared", "Plug in and Use", "Dependable"
  - » Efficient, flexible, low-cost use of resources
- Open infrastructures that enable federation at high levels of access and functionality
  - » Computation, Data Sharing
  - » Standards, Self-describing presentations, Security
  - » Enable composition: resources, services, semantics, all the way up!

Today's Readings

- Fox, et. al. The Grid: Past, Present, Future, GridInfrastructure2003
- Anderson, SETI@Home a Desktop Grid
  - » Entropia: Architecture of a Desktop Grid System
- The Grid2003 Production Grid:Principles and Practice

A more technology oriented view…

- The Globus Toolkit
- A taxonomical view of what's needed to enable grid applications….

Grid Architecture

Why Discuss Architecture?

- Descriptive
  - Provide a common vocabulary for use when describing Grid systems
- Guidance
  - Identify key areas in which services are required
- Prescriptive
  - Define standard "Intergrid" protocols and APIs to facilitate creation of interoperable Grid systems and portable applications

Aspects of the Systems Problem

1) Need for interoperability when different groups want to share resources
   - Diverse components, policies, mechanisms
   - E.g., standard notions of identity, means of communication, resource descriptions
2) Need for shared infrastructure services to avoid repeated development, installation
   - E.g., one port/service/protocol for remote access to computing, not one per tool/appln
   - E.g., Certificate Authorities: expensive to run

- A common need for protocols & services

Hence, a Protocol-Oriented View of Grid Architecture, that Emphasizes …

- Development of Grid protocols & services
  - Protocol-mediated access to remote resources
  - New services: e.g., resource brokering
  - "On the Grid" = speak Intergrid protocols
  - Mostly (extensions to) existing protocols
- Development of Grid APIs & SDKs
  - Interfaces to Grid protocols & services
  - Facilitate application development by supplying higher-level abstractions
- The (hugely successful) model is the Internet

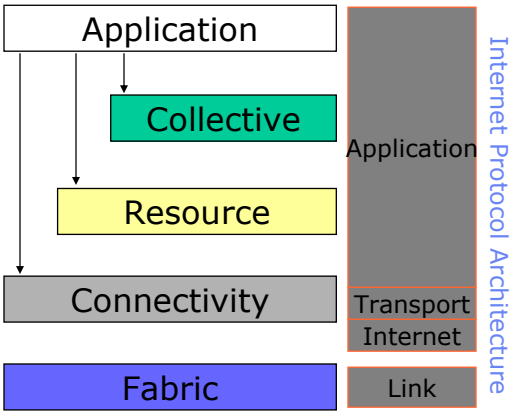
Layered Grid Architecture (By Analogy to Internet Architecture)

"Coordinating multiple resources": ubiquitous infrastructure services, app-specific distributed services

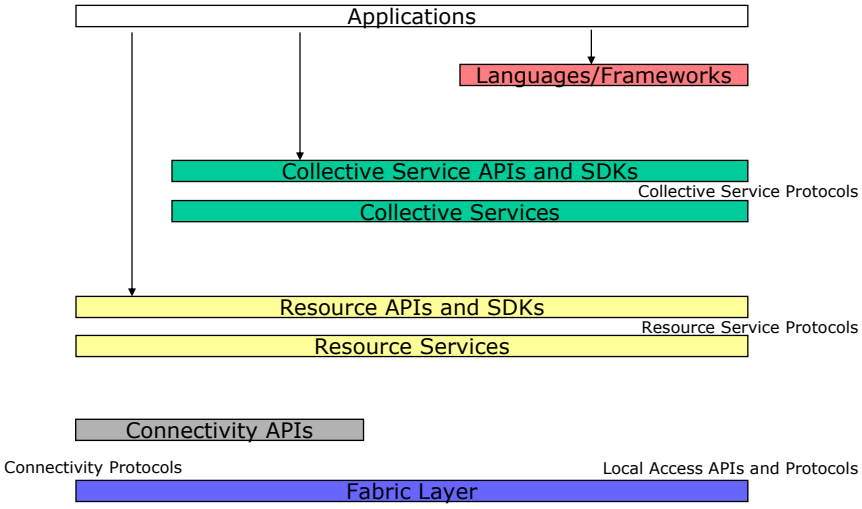
"Sharing single resources": negotiating access, controlling use

"Talking to things": communication (Internet protocols) & security

"Controlling things locally": Access to, & control of, resources



Protocols, Services, and APIs Occur at Each Level

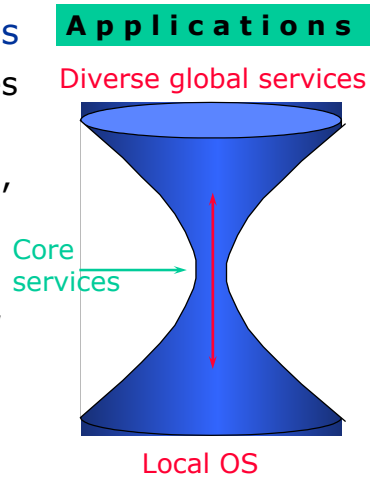


Important Points

- Built on Internet protocols & services
  - Communication, routing, name resolution, etc.
- "Layering" here is conceptual, does not imply constraints on who can call what
  - Protocols/services/APIs/SDKs will, ideally, be largely self-contained
  - Some things are fundamental: e.g., communication and security
  - But, advantageous for higher-level functions to use common lower-level functions

The Hourglass Model

- Focus on architecture issues
  - Propose set of core services as basic infrastructure
  - Use to construct high-level, domain-specific solutions
- Design principles
  - Keep participation cost low
  - Enable local control
  - Support for adaptation
  - "IP hourglass" model



Where Are We With Architecture?

- No "official" standards exist
- But:
  - Globus Toolkit™ has emerged as the de facto standard for several important Connectivity, Resource, and Collective protocols
  - GGF has an architecture working group
  - Technical specifications are being developed for architecture elements: e.g., security, data, resource management, information
  - Internet drafts submitted in security area

Introduction to the Globus Toolkit™

Globus Toolkit™

- A software toolkit addressing key technical problems in the development of Grid enabled tools, services, and applications
  - Offer a modular "bag of technologies"
  - Enable *incremental* development of grid-enabled tools and applications
  - Implement standard Grid protocols and APIs
  - Make available under liberal open source license

General Approach

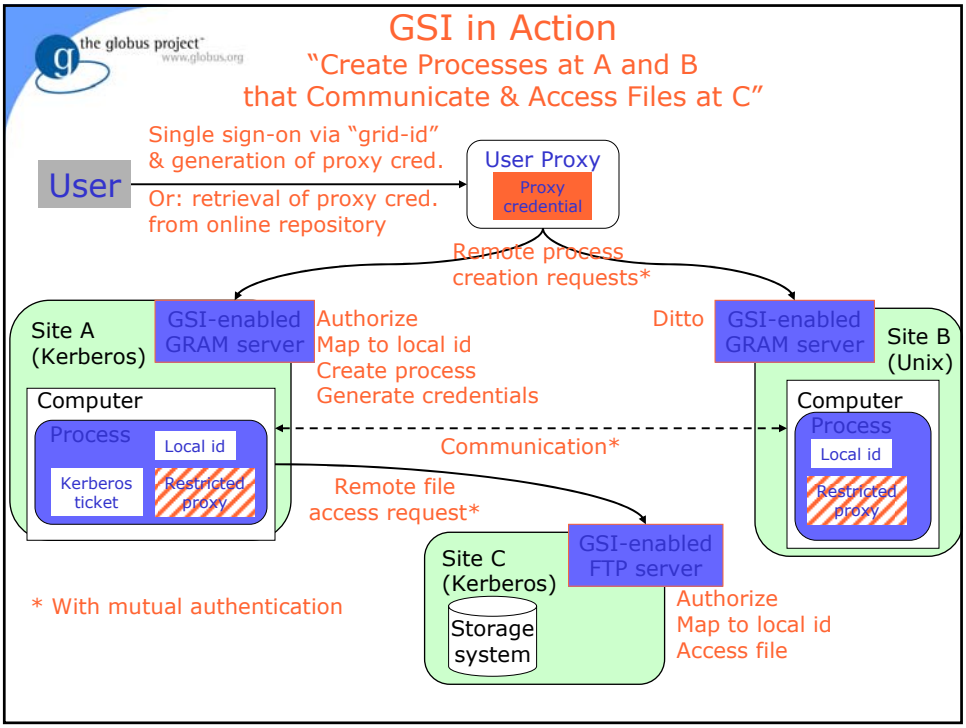
- Define Grid protocols & APIs
  - Protocol-mediated access to remote resources
  - Integrate and extend existing standards
  - "On the Grid" = speak "Intergrid" protocols
- Develop a reference implementation
  - Open source Globus Toolkit
  - Client and server SDKs, services, tools, etc.
- Grid-enable wide variety of tools
  - Globus Toolkit, FTP, SSH, Condor, SRB, MPI, …
- Learn through deployment and applications

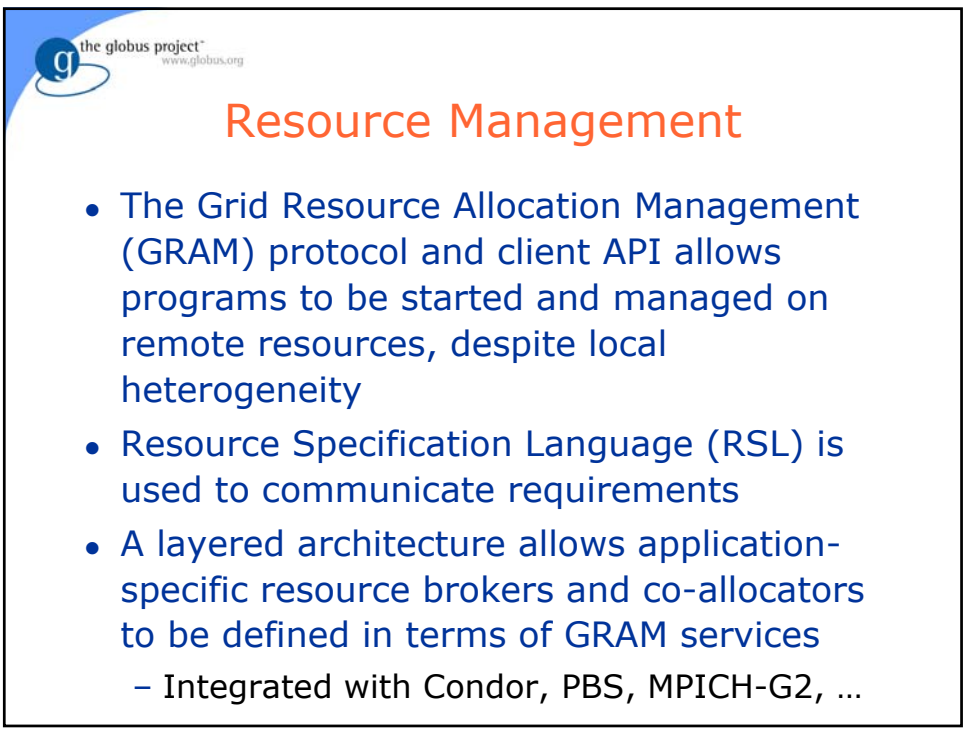
Key Protocols

- The Globus Toolkit™ centers around four key protocols
  - Connectivity layer:
    - *Security*: Grid Security Infrastructure (GSI)
  - Resource layer:
    - *Resource Management*: Grid Resource Allocation Management (GRAM)
    - *Information Services*: Grid Resource Information Protocol (GRIP)
    - *Data Transfer*: Grid File Transfer Protocol (GridFTP)
- Also key collective layer protocols
  - Info Services, Replica Management, etc.

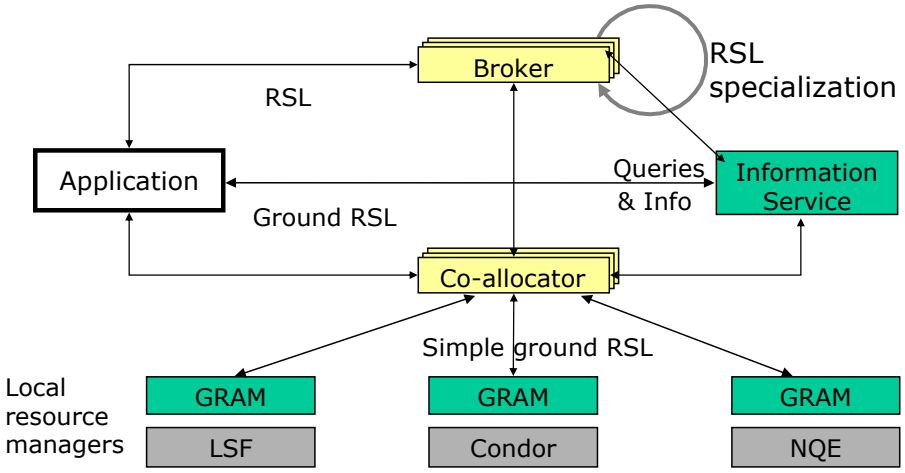
Grid Security Infrastructure (GSI)

- Globus Toolkit implements GSI protocols and APIs, to address Grid security needs
- GSI protocols extends standard public key protocols
  - Standards: X.509 & SSL/TLS
  - Extensions: X.509 Proxy Certificates & Delegation
- GSI extends standard GSS-API


the globus project
www.globus.org
GSI in Action
"Create Processes at A and B
that Communicate & Access Files at C"
User
Single sign-on via "grid-id"
& generation of proxy cred.
Or: retrieval of proxy cred.
from online repository
User Proxy
Proxy credential
Remote process
creation requests*
Site A
(Kerberos)
GSI-enabled
GRAM server
Authorize
Map to local id
Create process
Generate credentials
Ditto
GSI-enabled
GRAM server
Site B
(Unix)
Computer
Process
Local id
Kerberos ticket
Restricted proxy
Communication*
Computer
Process
Local id
Restricted proxy
Remote file
access request*
Site C
(Kerberos)
GSI-enabled
FTP server
Storage system
Authorize
Map to local id
Access file
* With mutual authentication


the globus project
www.globus.org
Resource Management
• The Grid Resource Allocation Management (GRAM) protocol and client API allows programs to be started and managed on remote resources, despite local heterogeneity
• Resource Specification Language (RSL) is used to communicate requirements
• A layered architecture allows application-specific resource brokers and co-allocators to be defined in terms of GRAM services
– Integrated with Condor, PBS, MPICH-G2, …

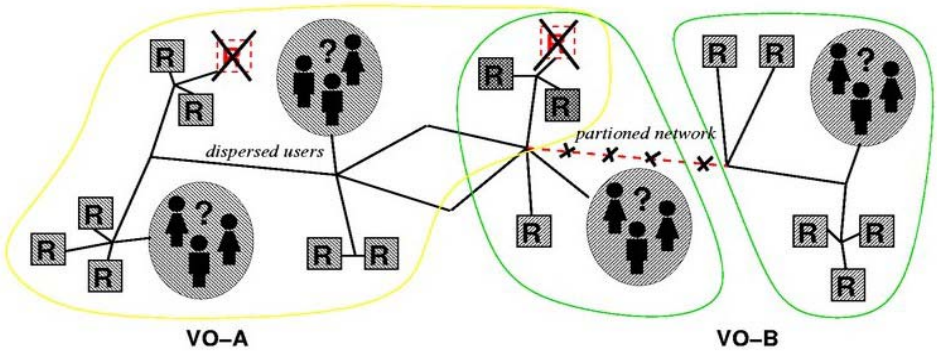
Resource Management Architecture



Data Access & Transfer

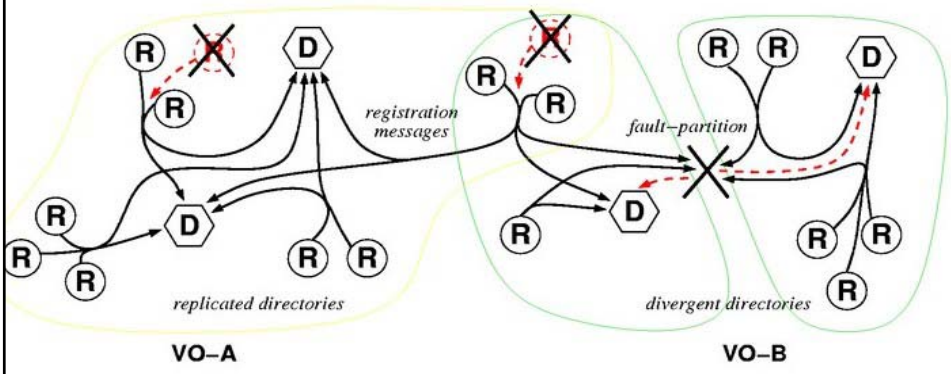
- GridFTP: extended version of popular FTP protocol for Grid data access and transfer
- Secure, efficient, reliable, flexible, extensible, parallel, concurrent, e.g.:
  - Third-party data transfers, partial file transfers
  - Parallelism, striping (e.g., on PVFS)
  - Reliable, recoverable data transfers
- Reference implementations
  - Existing clients and servers: wuftpd, ncftp
  - Flexible, extensible libraries in Globus Toolkit

The Grid Information Problem



- Large numbers of distributed "sensors" with different properties
- Need for different "views" of this information, depending on community membership, security constraints, intended purpose, sensor type

The Globus Toolkit Solution: MDS-2



Registration & enquiry protocols, information models, query languages

- Provides standard interfaces to sensors
- Supports different "directory" structures supporting various discovery/access strategies

Summary

- Globus architecture
  - » Four layers (fabric, connectivity, resource, collective)
- Key services
  - » Connectivity: communication and security
  - » Resource: Resource Allocation, Data movement, GRIS, others
  - » Collective: Index servers, Resource Brokers, Replica Catalogs, Co-reservation and co-allocation
- Information Services
  - » GRIS – local info provider
  - » GIIS / MDS – aggregating infromation server
  - » All LDAP based; Multi-level filtering and resource selection
  - » Most access is thru other resource managers and brokers

Discussion

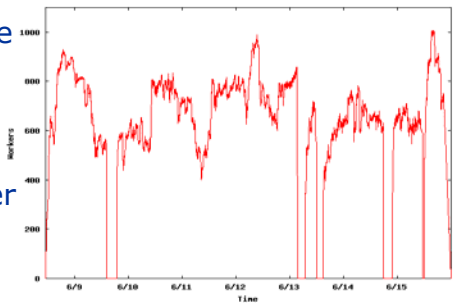
- What is the idea behind Grids?
- Does Globus address these issues?
  - » Which ones?
  - » All of them?
- What functionalities are actually provided?
- The drive to standards and research
  - » Analogy to the internet
  - » Many questions are unanswered

Real Grid Examples

the globus project™
www.globus.org

Mathematicians Solve NUG30

- Looking for the solution to the NUG30 quadratic assignment problem
- An informal collaboration of mathematicians and computer scientists
- Condor-G delivered 3.46E8 CPU seconds in 7 days (peak 1009 processors) in U.S. and Italy (8 sites)



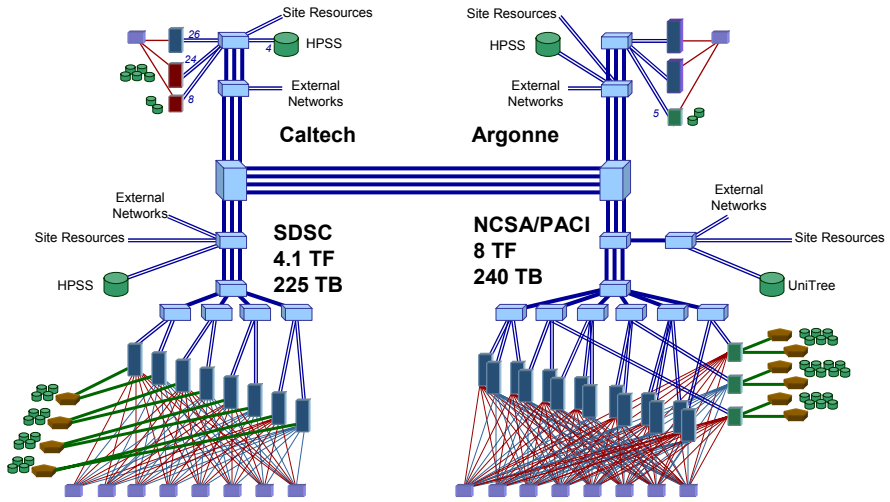
14,5,28,24,1,3,16,15,
10,9,21,2,4,29,25,22,
13,26,17,30,6,20,19,
8,18,7,27,12,11,23

MetaNEOS: Argonne, Iowa, Northwestern, Wisconsin

Home Computers Evaluate AIDS Drugs

- Community =
  - 10,000s of home computer users
  - Philanthropic computing vendor (Entropia)
  - Research group (Scripps)
- Common goal= advance AIDS resear

The 13.6 TF TeraGrid: Computing at 40 Gb/s



TeraGrid/DTF: NCSA, SDSC, Caltech, Argonne www.teragrid.org

Data Grids for High Energy Physics

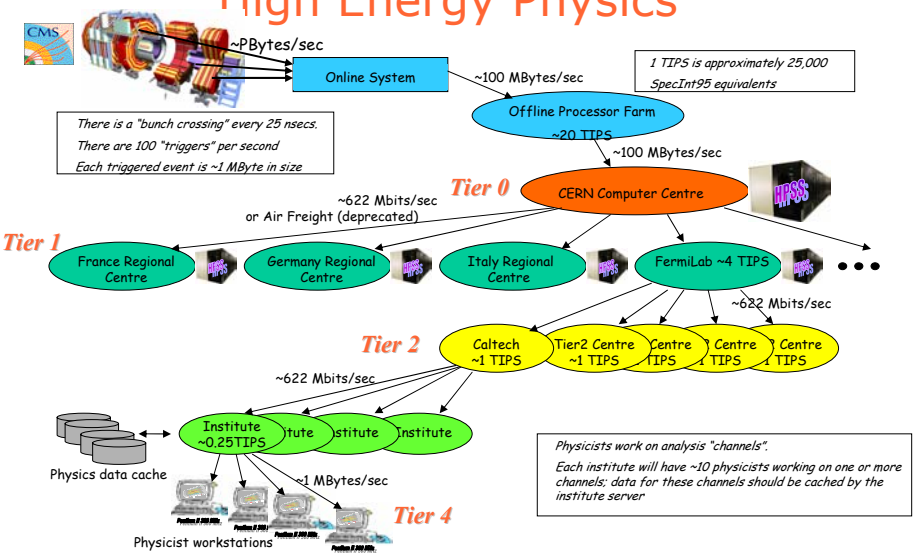
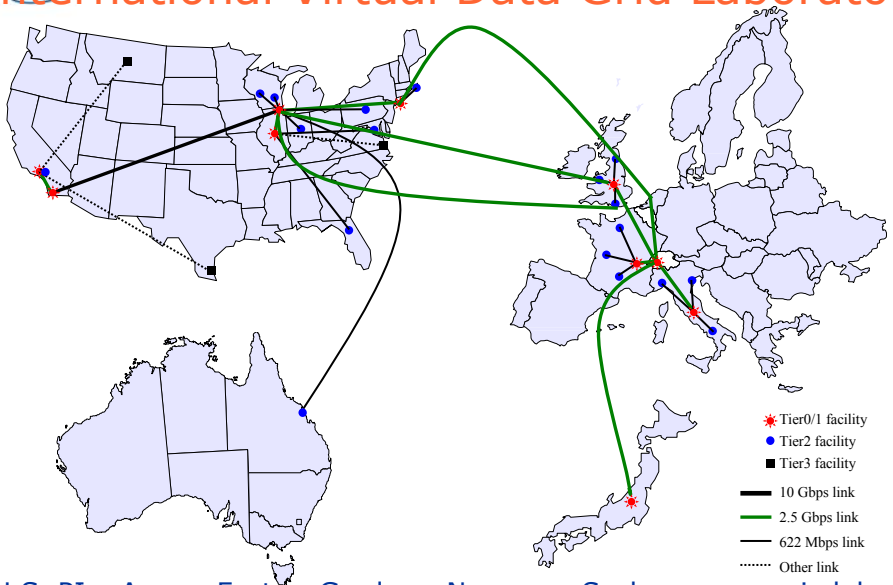


Image courtesy Harvey Newman, Caltech

iVDGL: International Virtual Data Grid Laboratory



U.S. PIs: Avery, Foster, Gardner, Newman, Szalay

www.ivdgl.org

Integrated Large Scale Grid Facility

- Multiple Virtual Organizations
- Large-scale resources
- Large-scale resource sharing (many applications)
- Metrics for Function
- Operating today

What Applications?

- MOP: USCMS Simulation: Monte-Carlo Simulation of Particle Accelarator
- GCE: USAtlas Simulation: Another accelerator simulation
- MaxBCG: Sloan Digital Sky Survey – search for interesting elements
- LIGO: Gravitational Waves Search (wide-area, wide frequency) – seti-like
- DIAL: Atlas Analysis
- BTeV Simulation: Tevatron Simulation, Hadron Colliders
- SnB Biomolecular Analysis: Analysis of X-ray Diffraction Data to Solve Large Protein Structures (search)

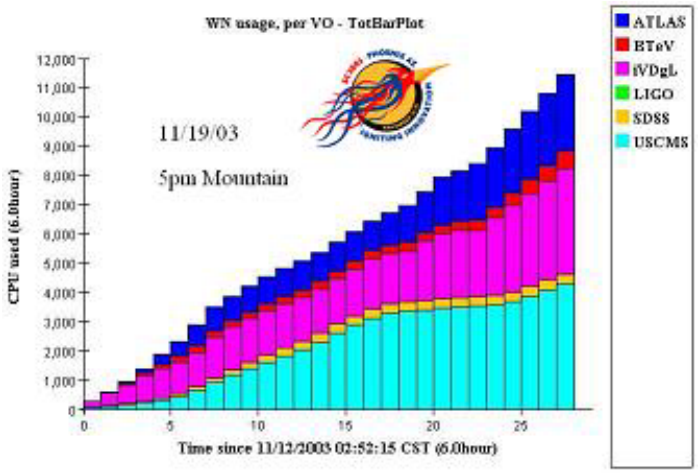
How Challenging are these Applications?

- All extremely large computations
- Some with very large quantities of data
- Many have workflows (complex sequences of actions)
- Most are completely uncoupled…

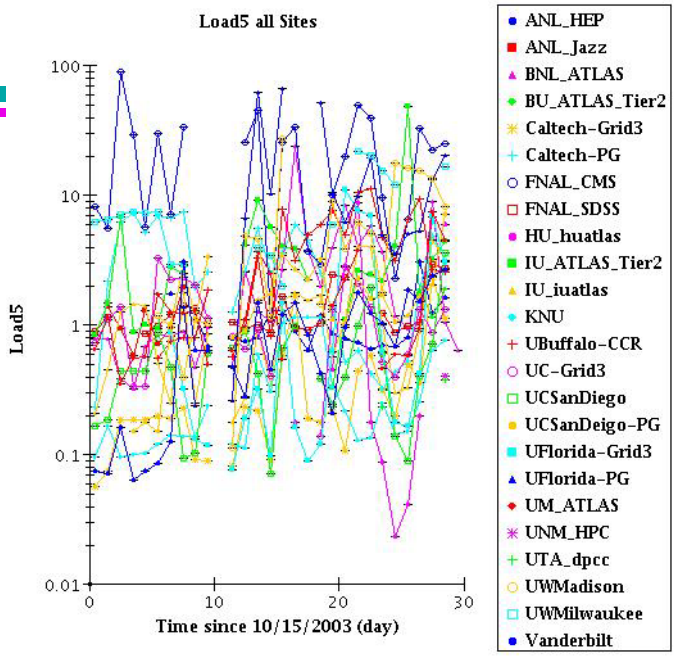
What Resources?

- 25 sites across the USA and Korea
- Many different organizations
- ~2000 total CPU's
- Hundreds of jobs running at a time

Delivered Computation



Dynamic Load



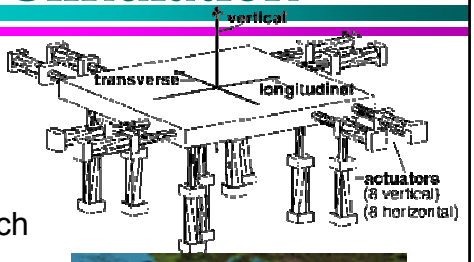
What Achieved?

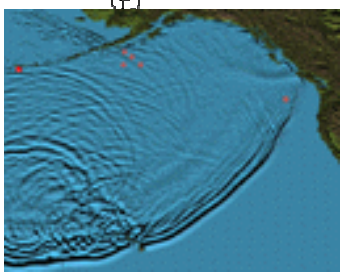
- Demonstrated that large-scale Grids can work
- Demonstrated a multiple-Virtual Organization system (one for each application)
- Captured reasonable statistics to show usage, progress, and data for large computations
- A good first step

- => but not any non-trivial quality of service…

Network for Earthquake Engineering Simulation

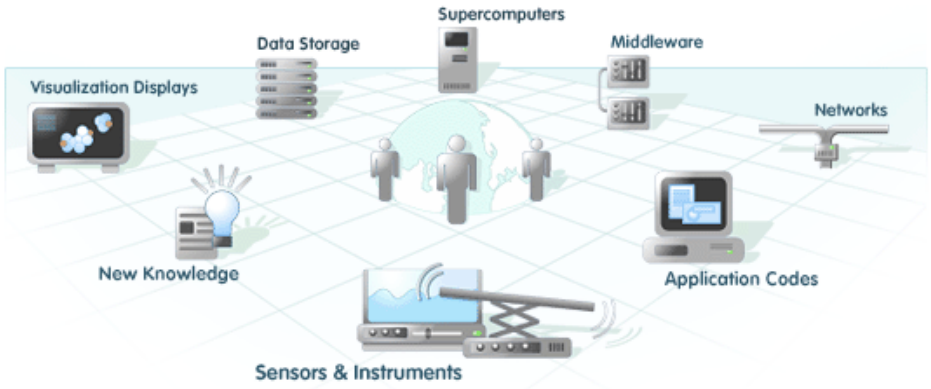
- NEESgrid: US national infrastructure to couple earthquake engineers with experimental facilities, databases, computers, & each other
- On-demand access to experiments, data streams, computing, archives, collaboration





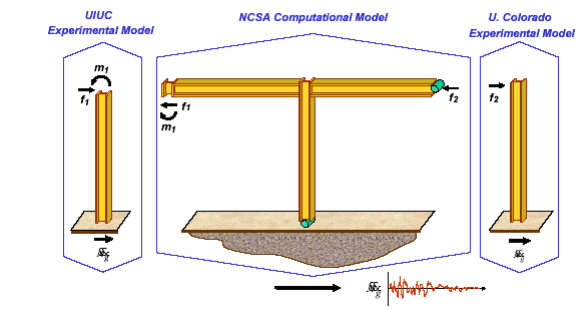
 NEESgrid: Argonne, Michigan, NCSA, UIUC, USC

NEESGrid (Earthquake Engineering)



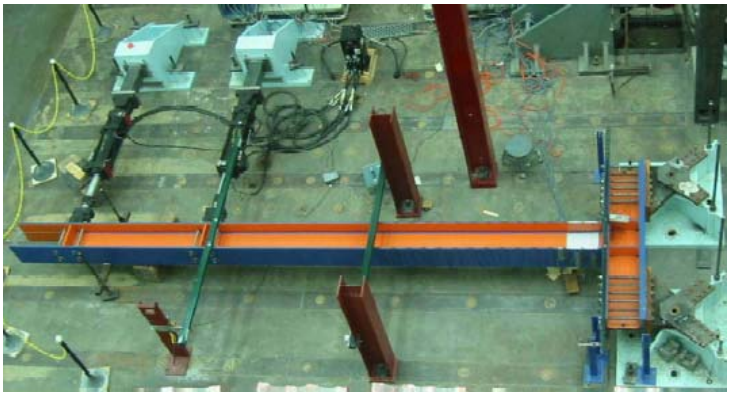
- Distributed Interactive Hybrid Simulation

MOST Experiment



- Integrated Computational and Physical Modelling
- Real-time Coupling
- Sensor and Actuators

These things are Real, and you can't Retry, Abort, Fail



Summary

- Real Grid Systems are Varied
- iVDGL: Large quantities of resources, massive throughput
- NEESGrid: Distributed Hybrid Simulation
- Seti@Home, Entropia: Massive Throughput

- … and many others…
- Focus is a little different from the business view…