

Grid Computing

The background of the slide features a series of vertical lines in various shades of blue and grey, creating a textured, rain-like effect. These lines are of varying heights and thicknesses, some appearing as thin, light lines and others as thicker, darker strokes. The overall effect is a modern, abstract pattern that fills the upper portion of the slide.

**THANKS TO ALL THE SOURCES FROM WHICH
THESE SLIDES ARE PREPARED.**

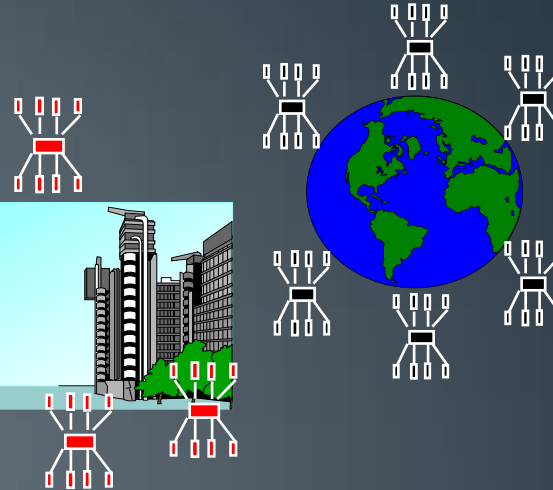
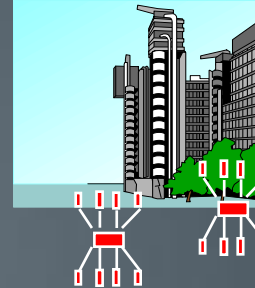
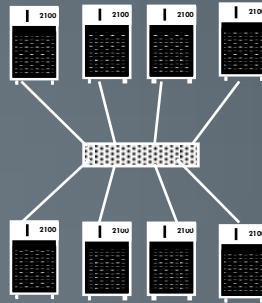
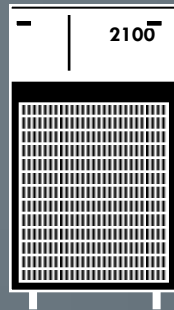
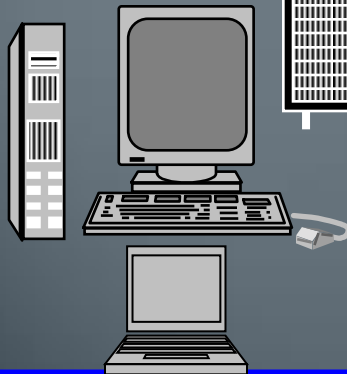
THANK YOU VERY MUCH!

Scalable Computing

PERFORMANCE
+
Q
oS

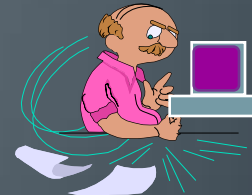
METRIC

DEVICES/SYSTEMS



Administrative Barriers

- Individual
- Group
- Department
- Campus
- State
- National
- Globe



Personal Device

SMPs or
SuperComputers

Local
Cluster

Enterprise
Cluster/Grid

Global
Grid

Grid Computing

- The term *Grid* comes from an analogy to the Electric Grid
- Grid Computing is a form of distributed computing
- Execution of large-scale resource intensive applications
- On geographically distributed systems (computing resources).

Formal Definition of Grids

- A grid is a system that:
 - Coordinates resource sharing in a de-centralized manner (i.e., different VO's).
 - Uses standard, open, general purpose protocols and interfaces.
 - Delivers non-trivial qualities of service.
 - Guaranteed bandwidth for application.
 - Guaranteed CPU cycles.
 - Guaranteed latency.

Grid Computing

- ◆ Motivation: high performance, improving resources utilization
- ◆ Aims to create illusion of a simple, yet powerful computer out of a large number of heterogeneous systems
- ◆ Tasks are submitted and distributed on nodes in the grid

Elements of Grid Computing

- Resource sharing
 - Computers, data, storage, sensors, networks, ...
 - Sharing always conditional: issues of trust, policy, negotiation, payment, ...
- Coordinated problem solving
 - Beyond client-server: distributed data analysis, computation, collaboration, ...
- Dynamic, multi-institutional *virtual organizations*
 - Community overlays on classic org structures
 - Large or small, static or dynamic

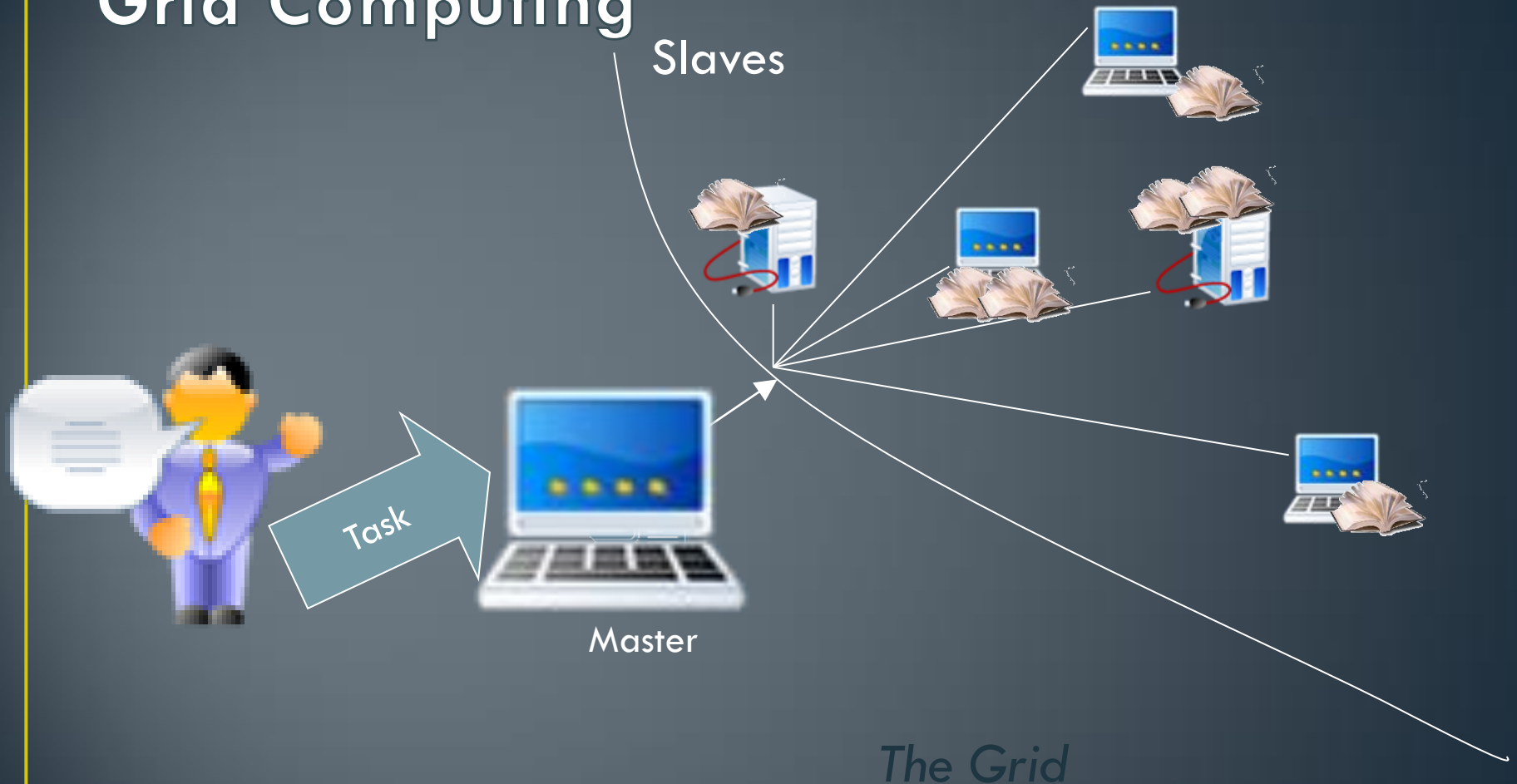
Grid Components and Services

- Communications
- Authentication and Authorization
- Naming Services and local transparency
- Distributed File System
- Resource Management
- Fault Tolerance
- Graphical User Interface

Grid working

- User submits a job requiring high computation from a low end Work Station
- A server collects requests from similar users and schedules them based on
 - Existing types of resources
 - Their capabilities
 - Current work load
- Submitted jobs are run on the selected computer nodes and results combined later

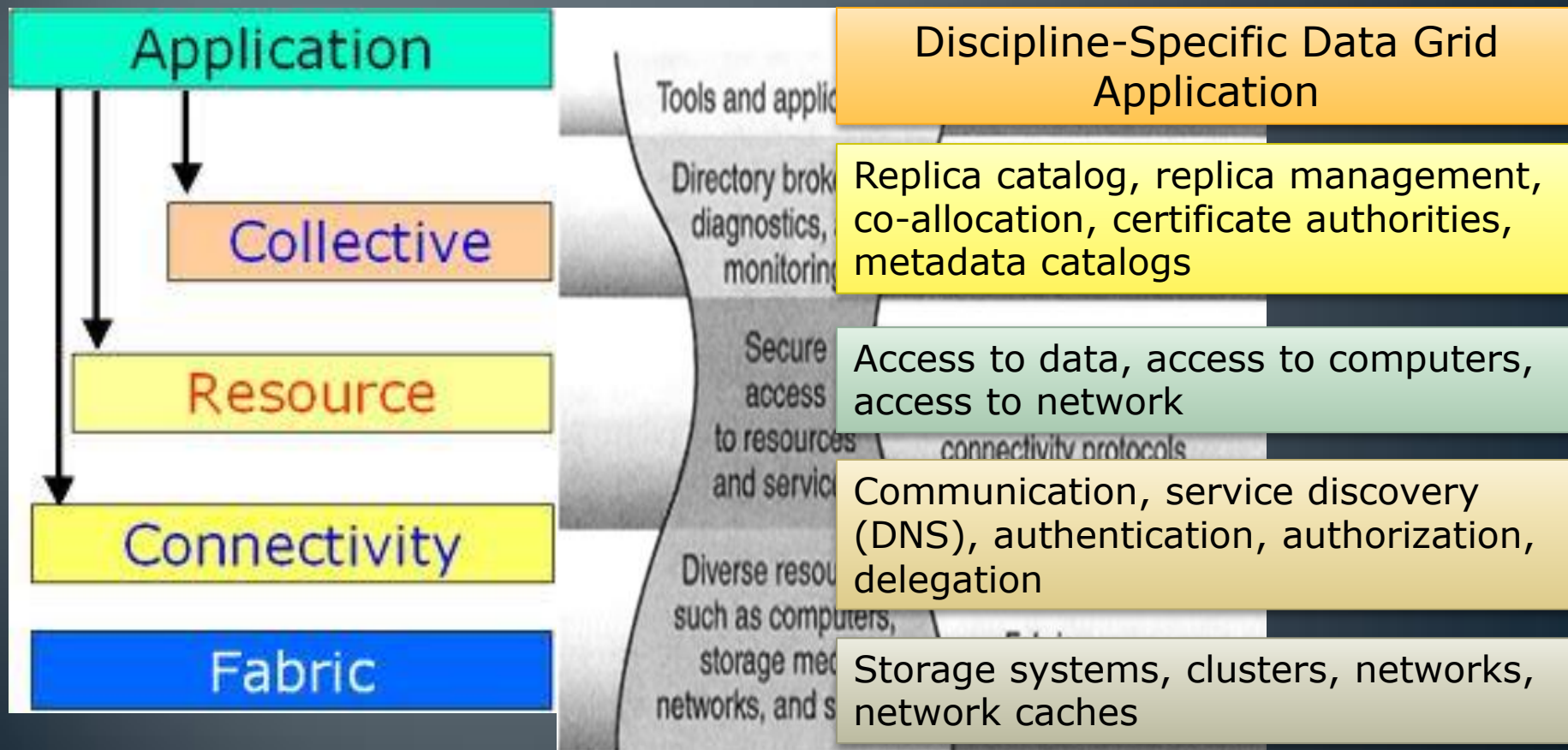
Grid Computing



The background of the slide features a series of thin, vertical, slightly wavy lines in a light blue-grey color against a light grey gradient. A solid teal horizontal band spans the width of the slide, positioned below the line pattern. The title 'Grid Architecture' is written in a white, bold, sans-serif font within this band.

Grid Architecture

Basic Grid Protocol Architecture



Grid Middleware

- Grids are typically managed by grid ware - a special type of middleware that enable **sharing and manage grid components** based on user requirements and resource attributes (e.g., capacity, performance)
- Software that connects other software components or applications to provide the following functions:
 - Run **applications** on suitable available resources
 - **Brokering, Scheduling**
 - Provide uniform, high-level access to **resources**
 - **Semantic interfaces**
 - **Web Services**, Service Oriented Architectures
 - Address inter-domain **issues** of security, policy, etc.
 - Federated Identities
 - Provide application-level **status monitoring and control**

Grid Middleware

- Major functions are:
 - Optimizing use of widely dispersed resources
 - Organizing efficient access to scientific data
 - Authenticating users accessing the resources
 - Arranging interfaces to local site authorization

The background of the slide features a series of thin, vertical, slightly wavy lines in a light blue-grey color against a light grey gradient. A solid teal-colored horizontal bar spans the width of the slide, positioned in the lower half. The title 'Grid Systems' is written in a white, bold, sans-serif font with a thin black outline, located in the lower-left corner of the teal bar.

Grid Systems

Data Grid vs. Computing Grid

◆ Data Grid:

- distributed data storage
- controlled sharing and management of large amounts of distributed data.

◆ Computing Grid:

- Parallel execution
- divide pieces of a program among several computers

• Data Grid + Computing Grid

Grid Computing

Data Grid

- A **data grid** is a grid computing system that deals with data — the **controlled sharing and management of large amounts of distributed data**.
- Large datasets can be stored in repositories
- Data grid provide services to distributed data-intensive applications
 - Data replication
 - Data invalidation
 - Data backup
 - Distributed transactions
 - **Data affinity/partitioning**

Example :

**Biomedical informatics Research Network (BIRN),
the Southern California earthquake Center (SCEC).**

Computational Grid

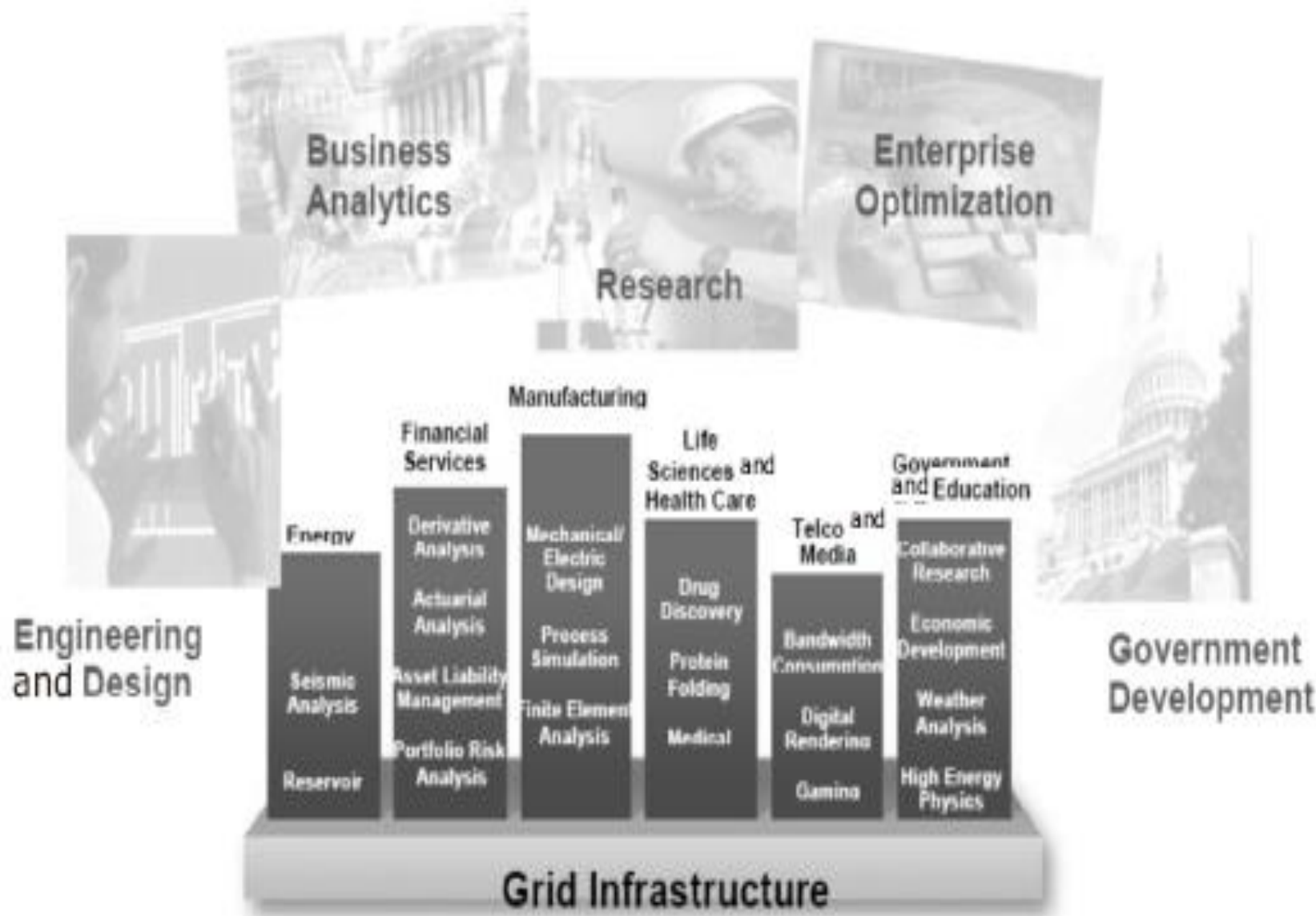
- “A computational grid is a **hardware and software infrastructure** that provides dependable, consistent, pervasive, and inexpensive access to **high-end computational capabilities**.”
- Example : Science Grid (US Department of Energy)
- **Characteristics:**
 - Made up of clusters of clusters
 - Enable CPU scavenging for better resource utilization
 - provide computational power for compute-intensive jobs
 - Provide instant access on demand

Benefits of Grid Computing

- Exploit Underutilized resources
- Resource **load Balancing**
- **Virtualize** resources across an enterprise
 - Data Grids, Compute Grids
- Enable **collaboration** for virtual organizations

Grid Computing Applications

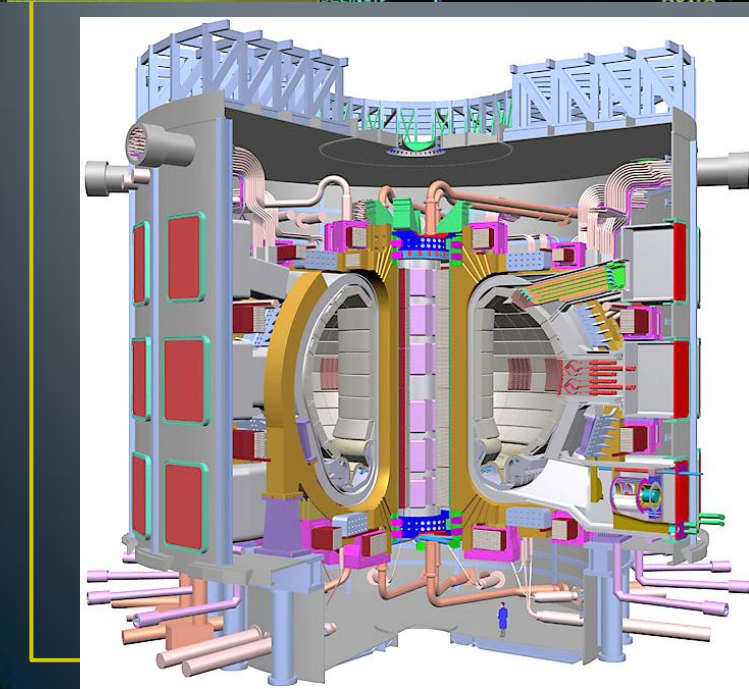
- Distributed computing
- High Throughput
- On-demand
- Data-intensive
- Collaborative



Variety of grid computing applications

The background of the slide features a series of vertical lines in various shades of blue and grey, creating a textured, rain-like effect. A solid teal horizontal band spans the width of the slide, positioned below the patterned area. The word "Examples" is written in white, bold, sans-serif font on the left side of this teal band.

Examples



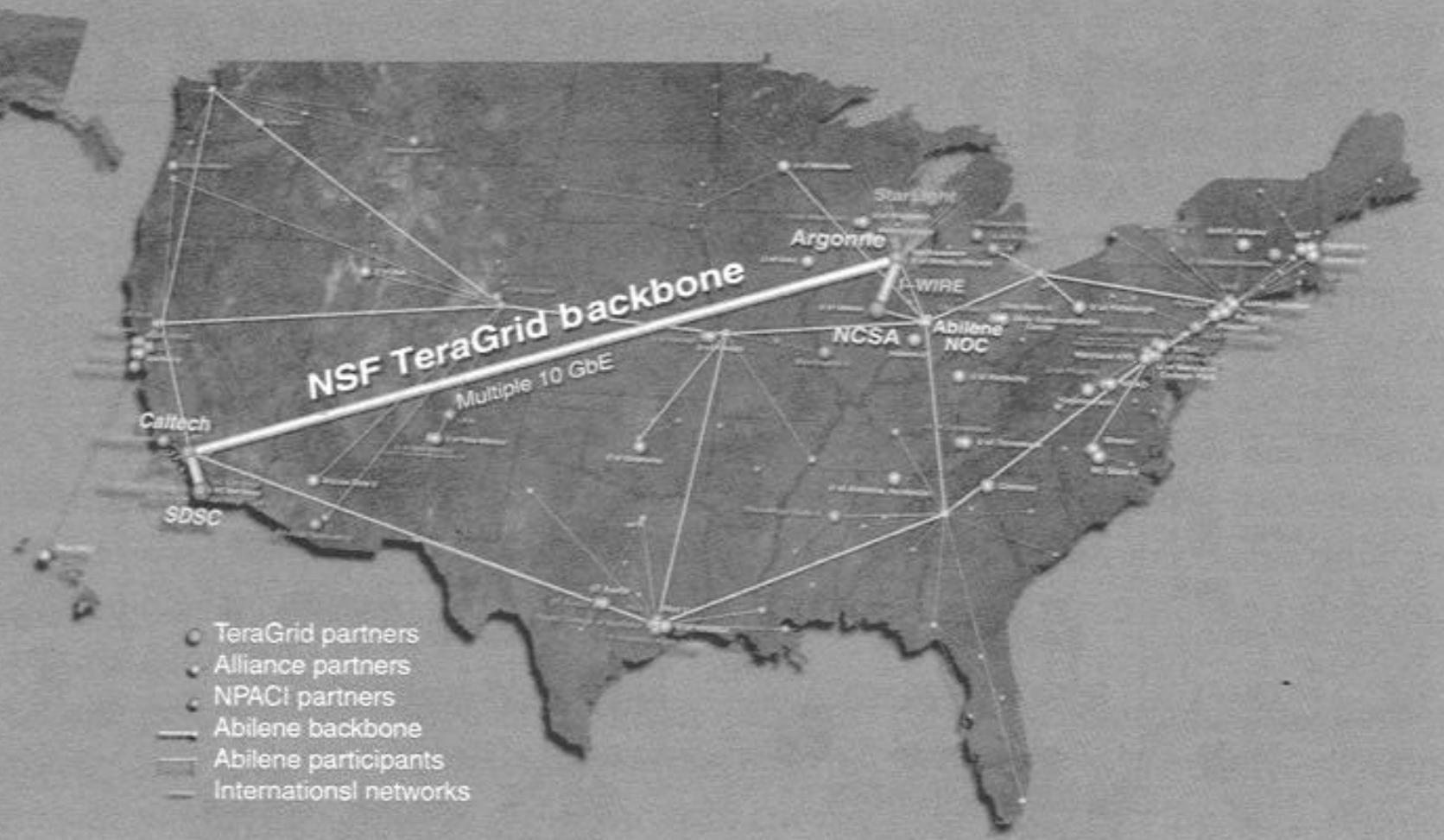
The diagram illustrates the 'Experiment Enterprise' as a global network of research groups and their resources. The background is a world map. A large black dashed line encloses the entire area. Inside, several groups are defined by dashed lines of different colors:

- Individual Research** (yellow dashed line): Located in the top left, containing icons for a storage cylinder and a computer.
- Trigger Studies Work Group** (red dashed line): Located in the bottom left, containing icons for a document, a desktop monitor, and a storage cylinder.
- Higgs Working Group** (cyan dashed line): Located in the top center, containing icons for a computer and a document.
- Shared Resources** (red dashed line): Located in the center, containing icons for documents and a storage cylinder.
- New Phenomena Work Group** (green dashed line): Located in the bottom center, containing icons for a storage cylinder, a document, and a desktop monitor.
- Supersymmetry Working Group** (blue dashed line): Located in the top right, containing icons for a computer, a document, and a storage cylinder.

A legend in the bottom right corner defines the icons used:

- Physicist: A stick figure icon.
- Computer: A vertical rectangle icon.
- Desktop: A horizontal rectangle icon with a black square in the top left corner.
- Schema: A document icon with a grid pattern.
- Storage: A cylinder icon.
- Software: A document icon with a wavy bottom edge.

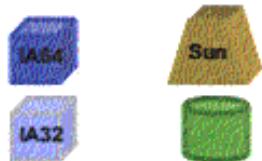
TeraGrid is an Important Project developed by the National Science Foundation (NSF).



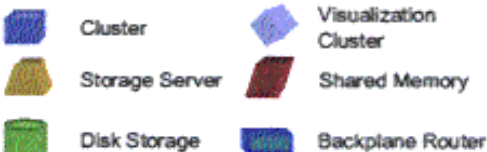
TeraGrid

Caltech: Data collection analysis

0.4 TF IA-64
IA32 Datawulf
80 TB Storage

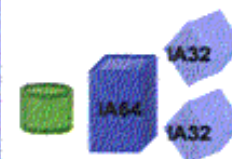


LEGEND

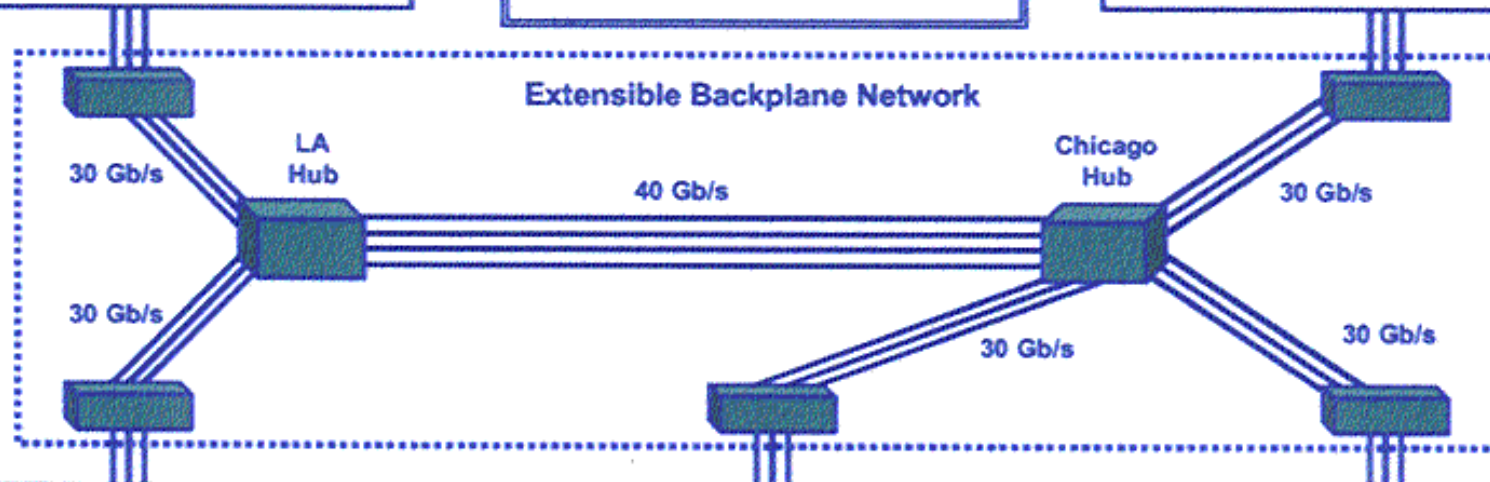


ANL: Visualization

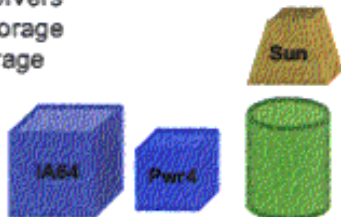
1.25 TF IA-64
96 Visualization nodes
20 TB Storage



Extensible Backplane Network

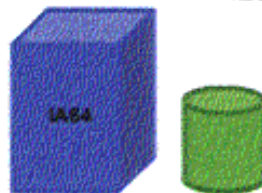


4 TF IA-64
DB2, Oracle Servers
500 TB Disk Storage
6 PB Tape Storage
1.1 TF Power4



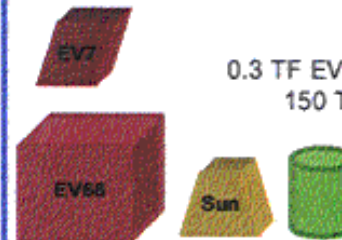
SDSC: Data-Intensive

10 TF IA-64
128 large memory nodes
230 TB Storage



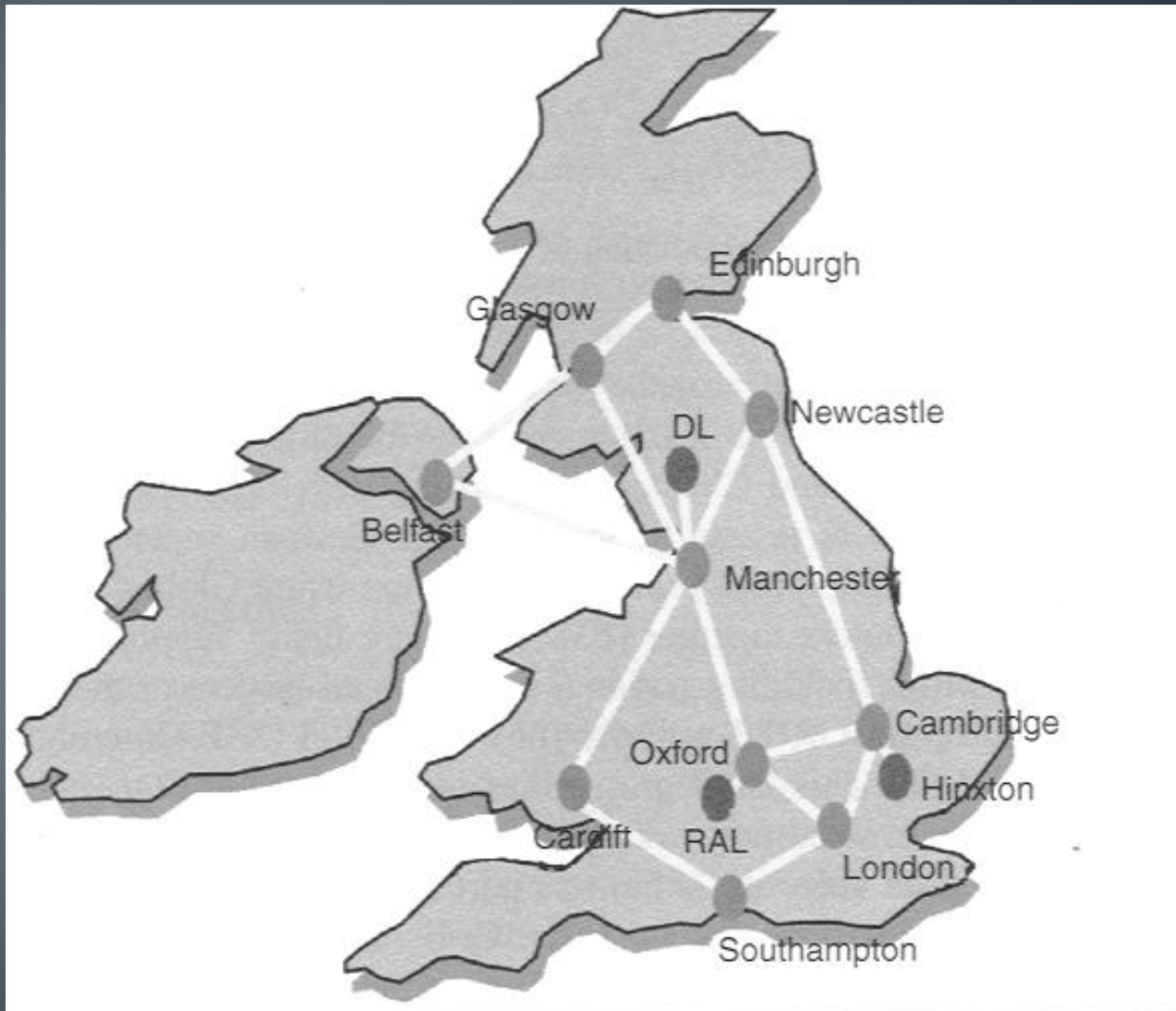
NCSA: Compute-Intensive

6 TF EV68
71 TB Storage
0.3 TF EV7 shared-memory
150 TB Storage Server



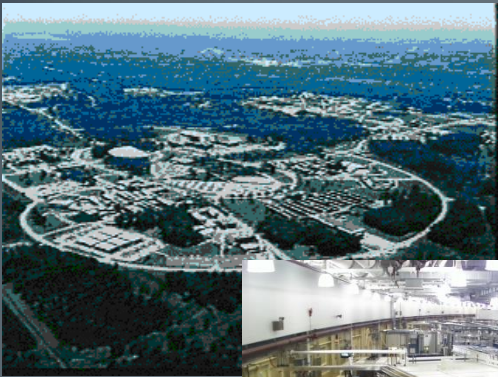
PSC: Compute-Intensive

UK e-Science Grid

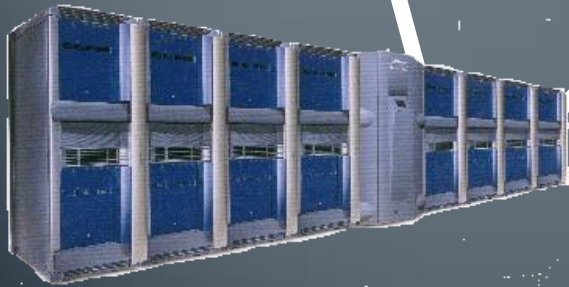


Online Access to Scientific Instruments

Advanced Photon Source

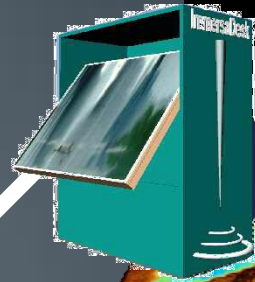


real-time
collection

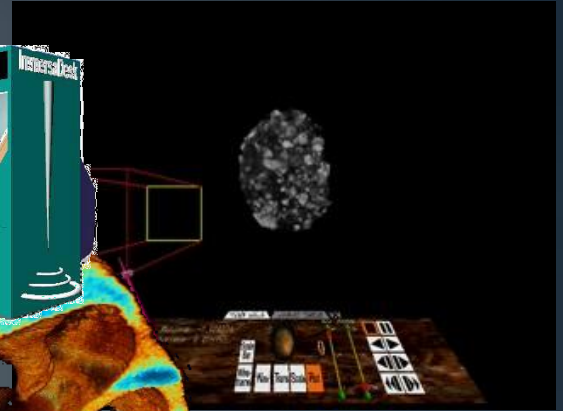
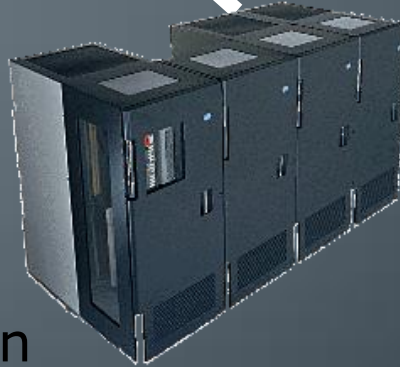


tomographic reconstruction

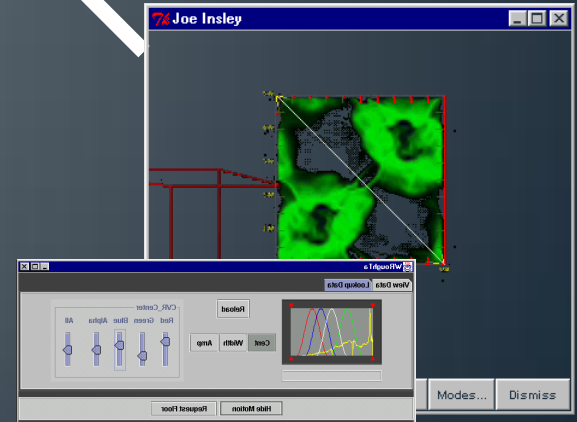
wide-area
dissemination



archival
storage



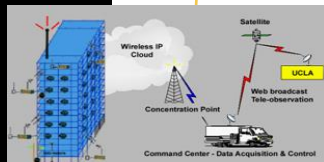
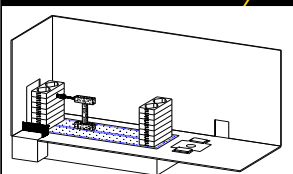
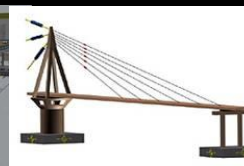
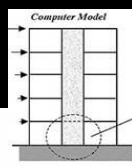
desktop & VR clients
with shared controls



DOE X-ray grand challenge: ANL, USC/ISI, NIST, U.Chicago

NSF Network for Earthquake Engineering Simulation (NEES)

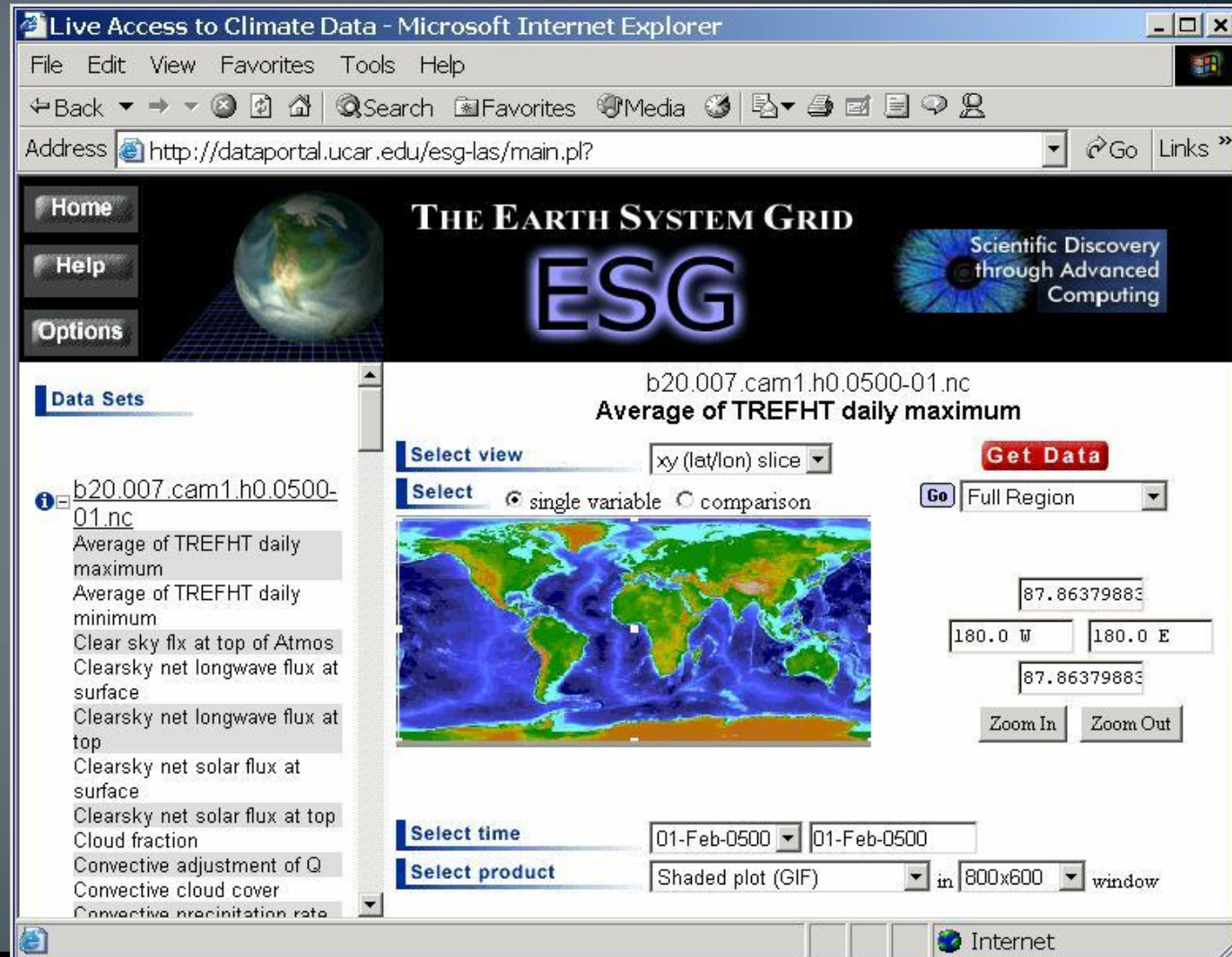
Transform our ability to carry out research vital to reducing vulnerability to catastrophic earthquakes



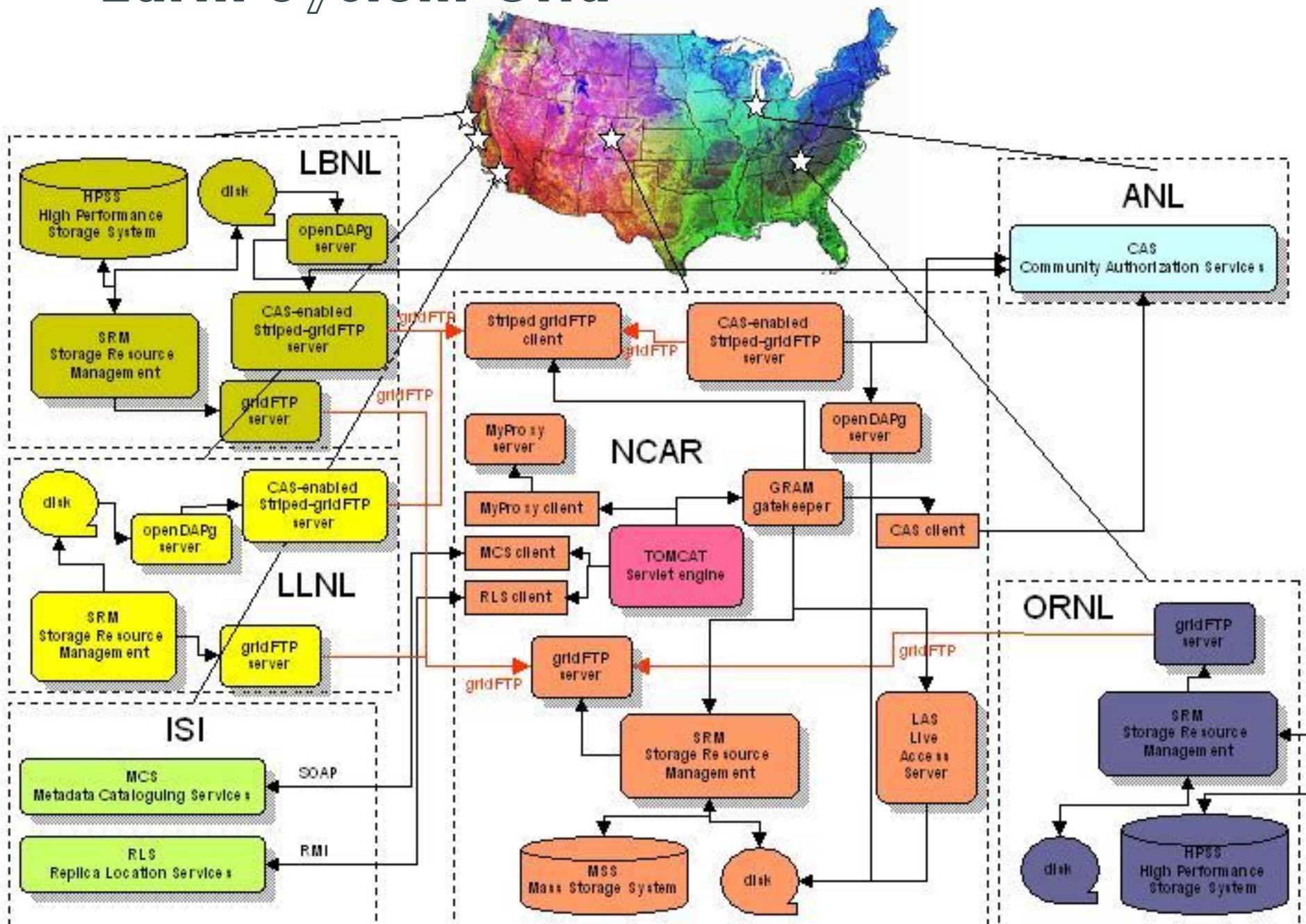
I. Foster

DOE Earth System Grid

Goal: address technical obstacles to the sharing & analysis of high-volume data from advanced earth system models



Earth System Grid



Simulation tools

- GridSim – job scheduling
- SimGrid – single client multiserver scheduling
- Bricks – scheduling
- GangSim- Ganglia VO
- OptoSim – Data Grid Simulations
- G3S – Grid Security services Simulator – security services

Middleware

- Globus –chicago Univ
- Condor – Wisconsin Univ – High throughput computing
- Legion – Virginia Univ – virtual workspaces-
collaborative computing
- IBP – Internet back pane – Tennessee Univ – logistical
networking
- NetSolve – solving scientific problems in heterogeneous
env – high throughput & data intensive

Standards Bodies

The primary standards-setting body is¹:

- Global Grid Forum (GGF)
 - Started in 1998
 - Meets three times a year, GGF1, GGF2, GGF3 ...
 - More than 40 organizations involved and growing ...

Others:

- W3C consortium (Worlds Wide Web Consortium)
 - Working on standardization of web-related technologies such as XML
 - See <http://www.w3.org>
- OASIS (Organization for the Advancement of Structured Information Standards)
- IETF, DMTF

¹ “The Grid Core Technologies” by M. Li and M. Baker, 2005, page 4.

Standards in the Web Services World

- XML introduced (ratified) in 1998
- SOAP ratified in 2000
- Web services developed
- Subsequently, standards have been are continuing to be developed:
 - WSDL
 - WS-* where * refers to names of one of many standards

Standards in the grid computing world

- Open Grid Services Architecture (OGSA)
- First announced at GGF4 in Feb 2002
- OGSA does not give details of implementation.