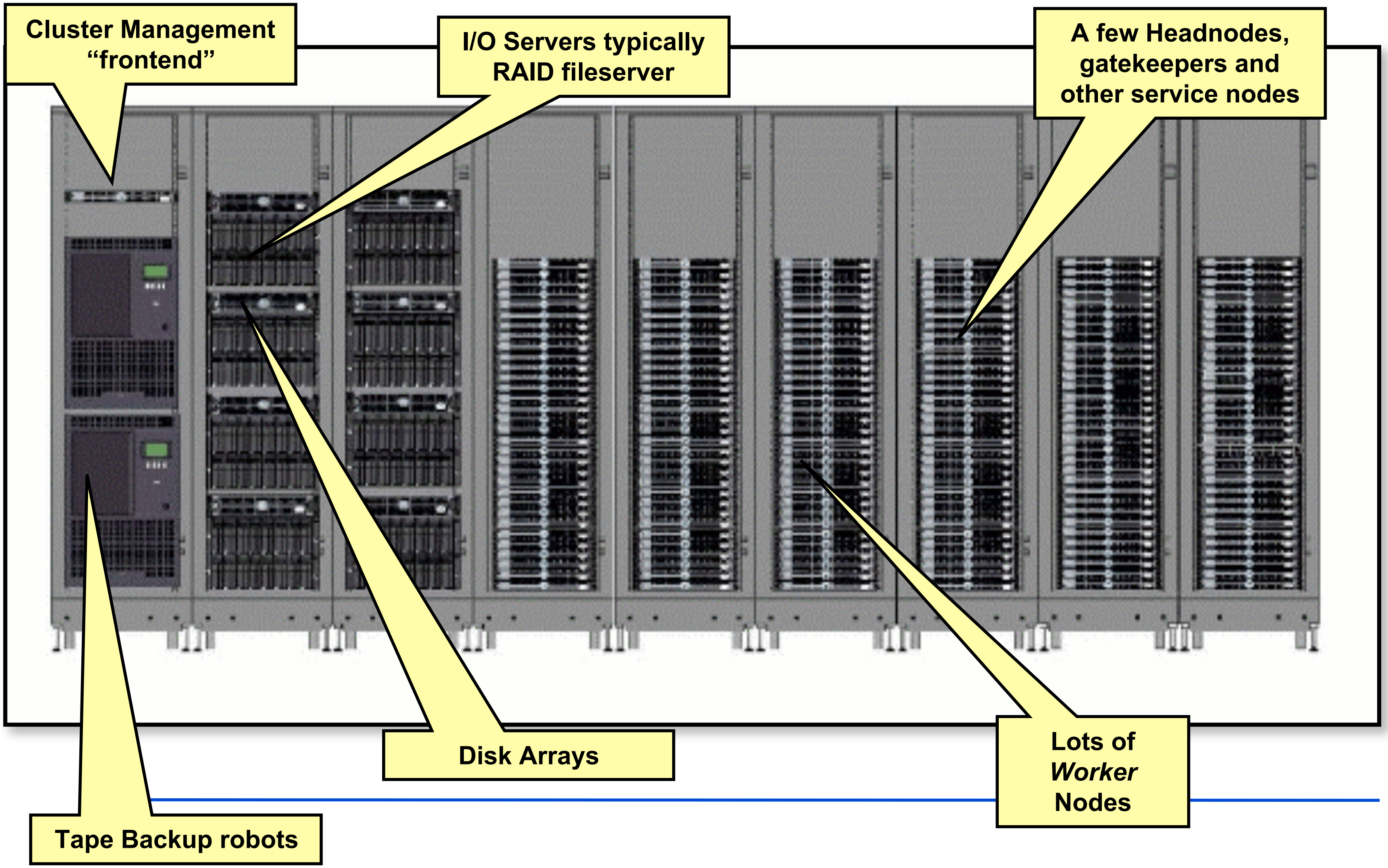
Introduction to Grid Computing

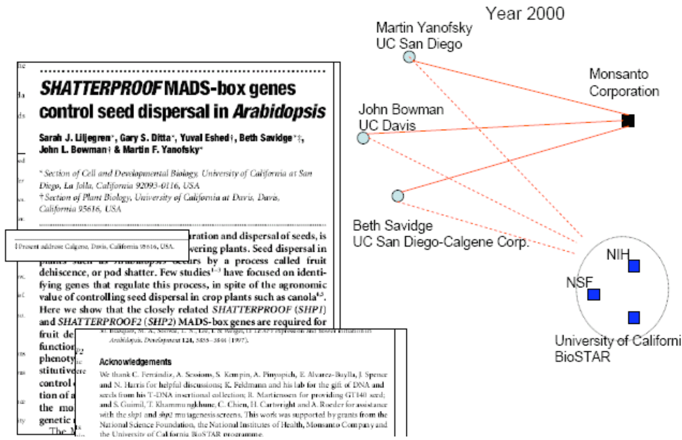
Grid School Workshop – Module 1

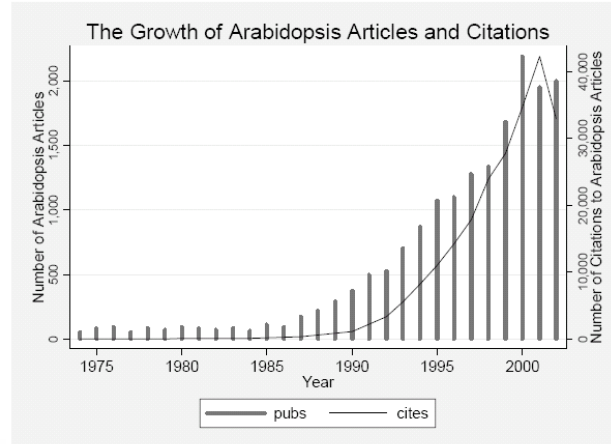
Open Science Grid

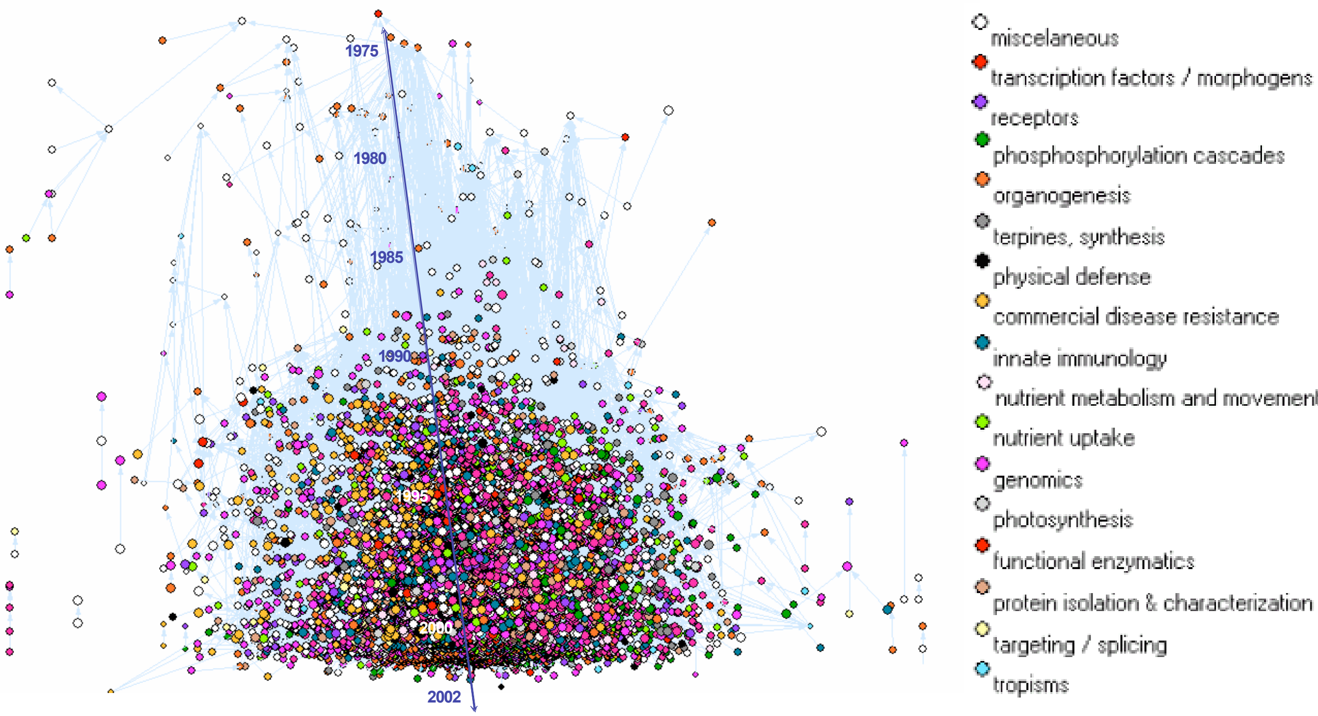
Computing clusters have commoditized supercomputing



Scaling up Science:
Citation Network Analysis in Sociology







Work of James Evans, University of Chicago, Department of Sociology

Scaling up the analysis

- Query and analysis of 25+ million citations
- Work started on desktop workstations
- Queries grew to month-long duration
- With data distributed across U of Chicago TeraPort **cluster**:
  - 50 (faster) CPUs gave 100 X speedup
  - Many more methods and hypotheses can be tested!
- Higher *throughput* and *capacity* enables *deeper analysis* and *broader community access*.

Initial Grid driver: High Energy Physics

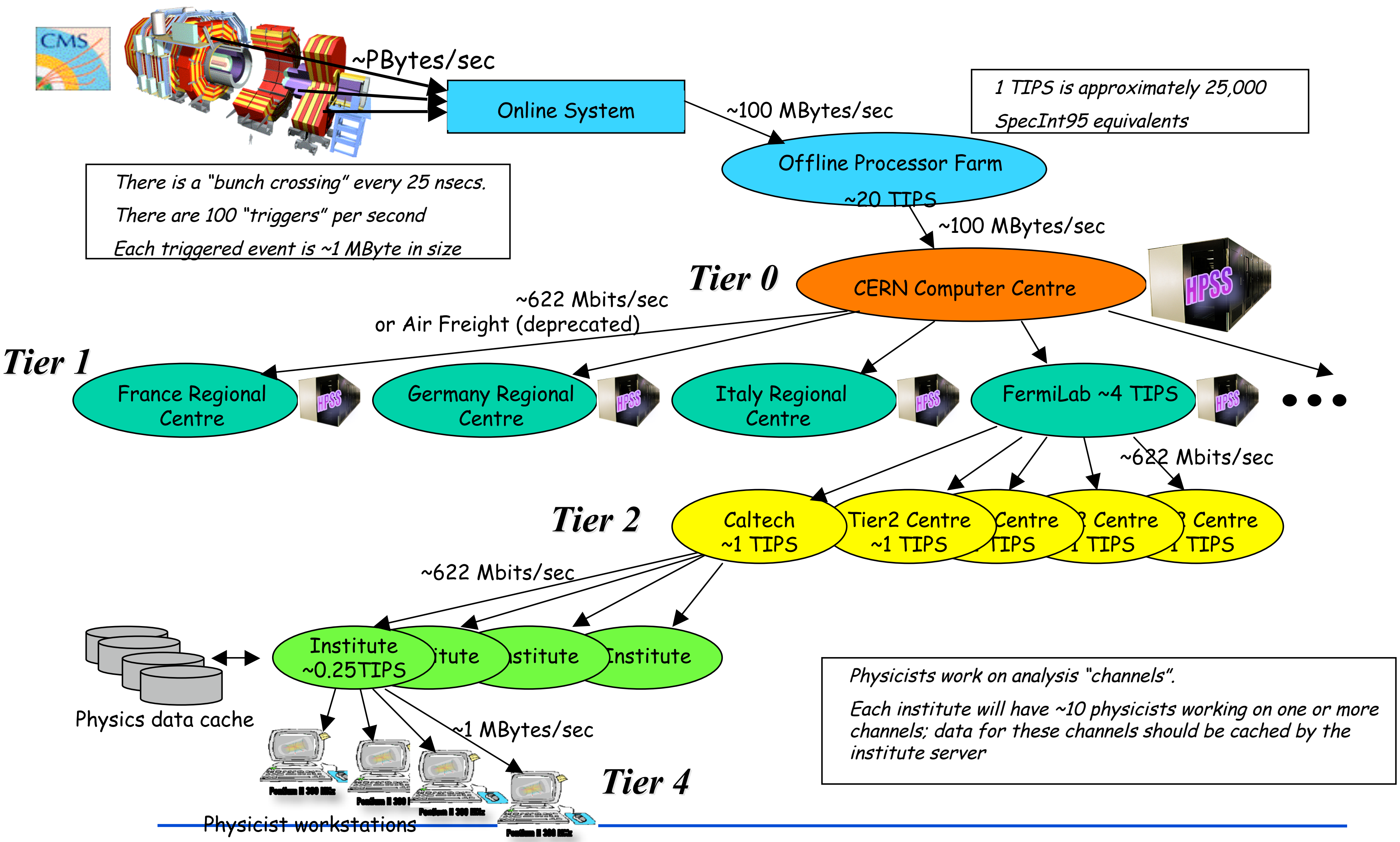
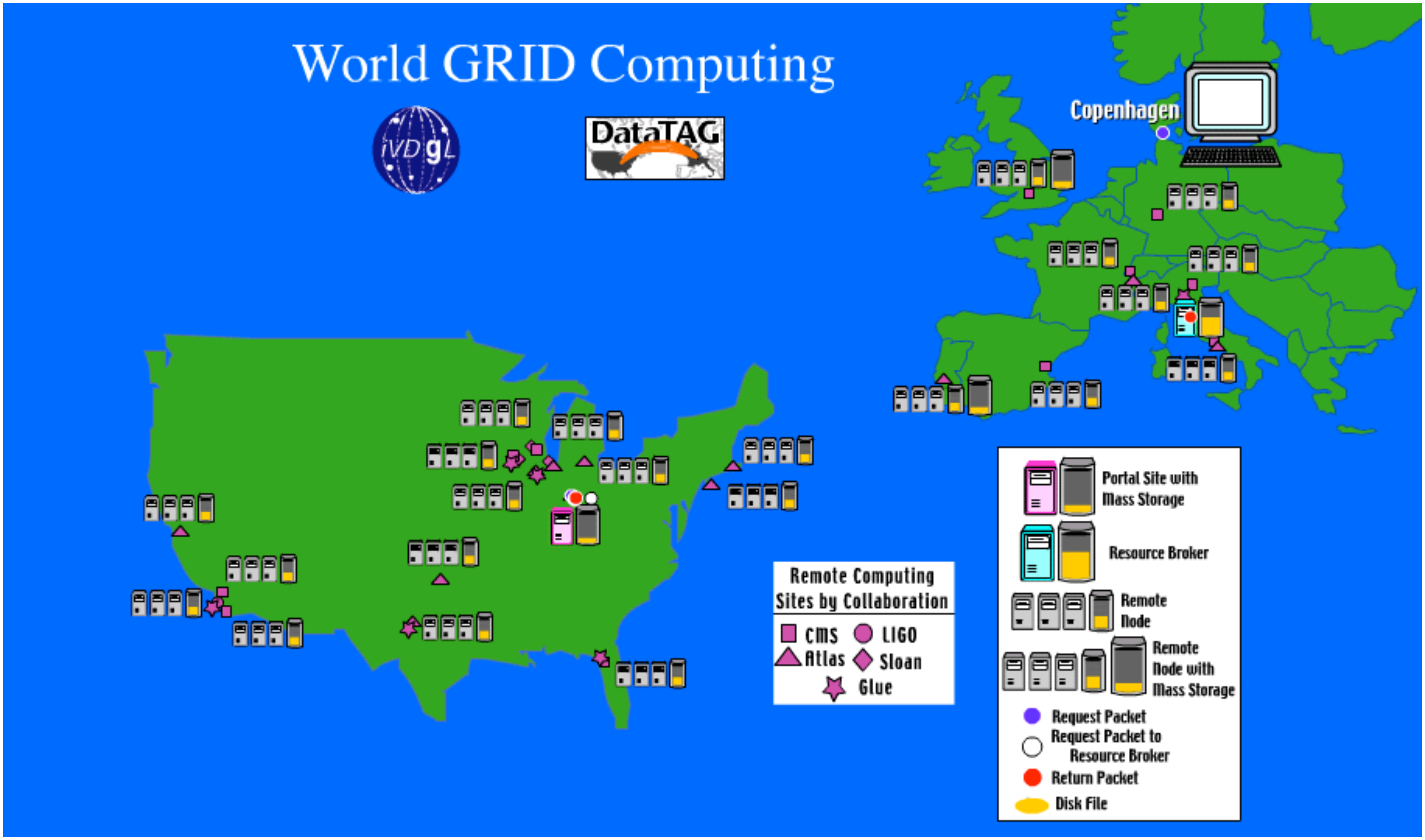


Image courtesy Harvey Newman, Caltech

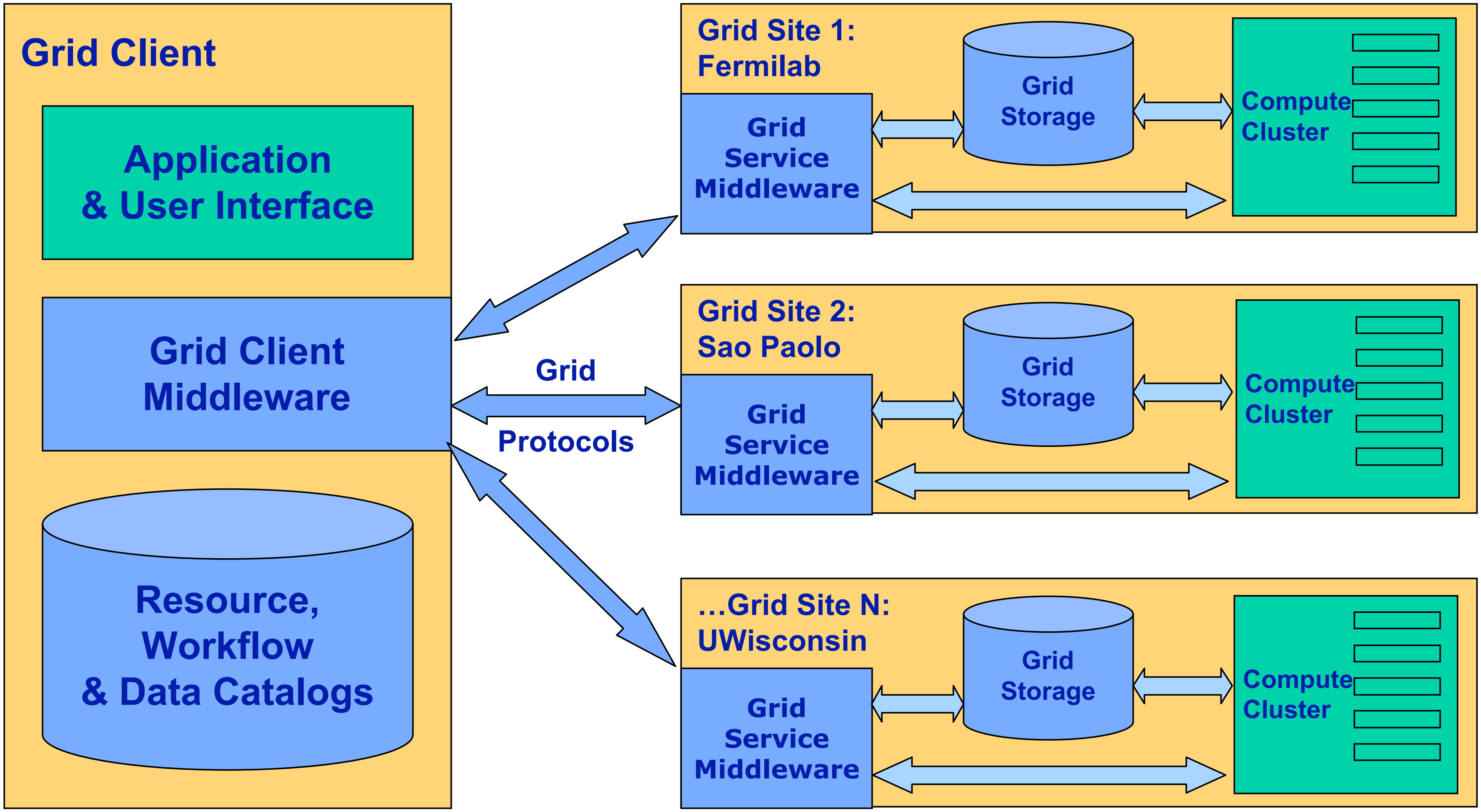
Grids Provide Global Resources To Enable e-Science



Ian Foster's Grid Checklist

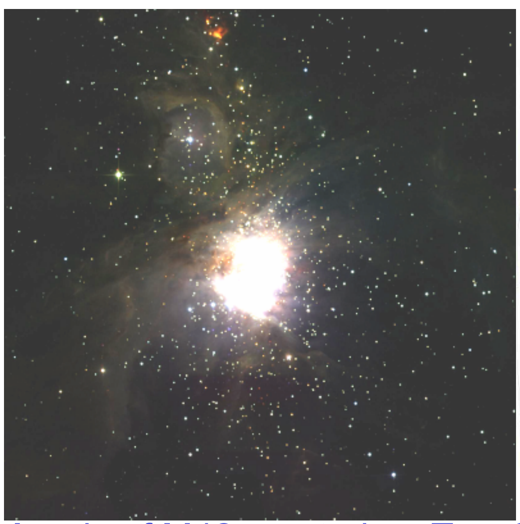
- A Grid is a system that:
  - Coordinates resources that are not subject to centralized control
  - Uses standard, open, general-purpose protocols and interfaces
  - Delivers non-trivial qualities of service

Grids consist of distributed clusters

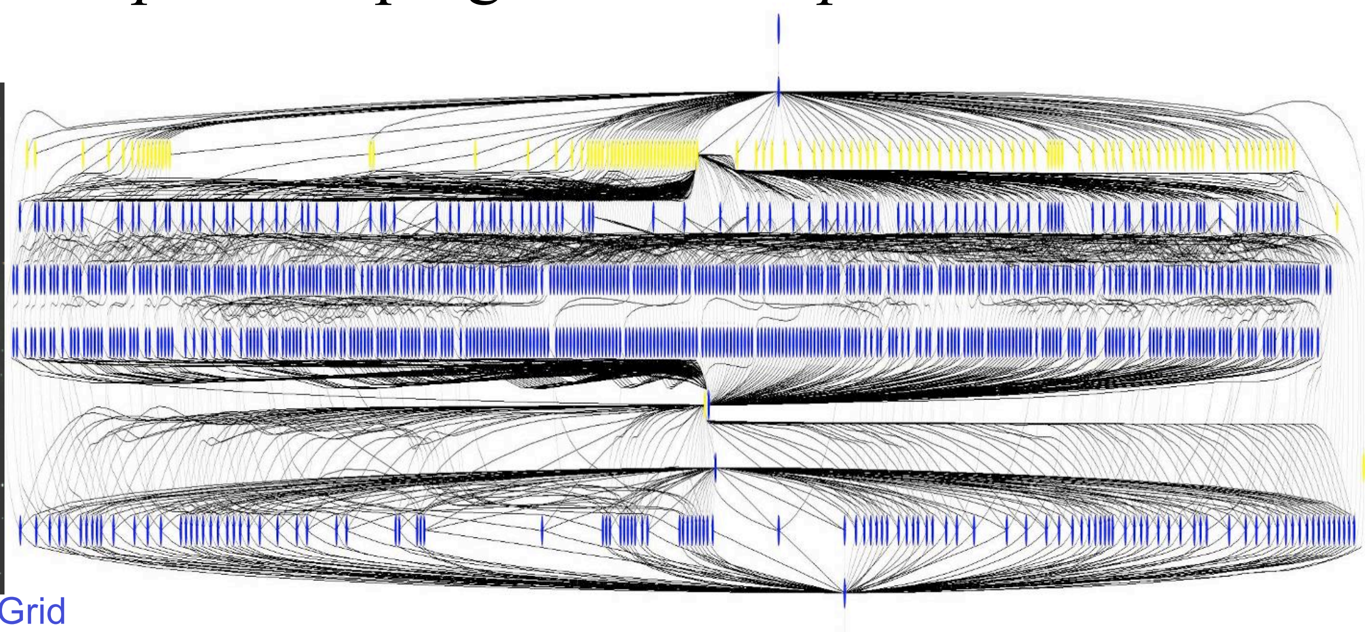


Grids can process vast datasets.

- Many HEP and Astronomy experiments consist of:
  - Large datasets as inputs (find datasets)
  - "Transformations" which work on the input datasets (process)
  - The output datasets (store and publish)
- The emphasis is on the sharing of these large datasets
- *Workflows* of *independent* program can be *parallelized*.



Mosaic of M42 created on TeraGrid



= Data Transfer

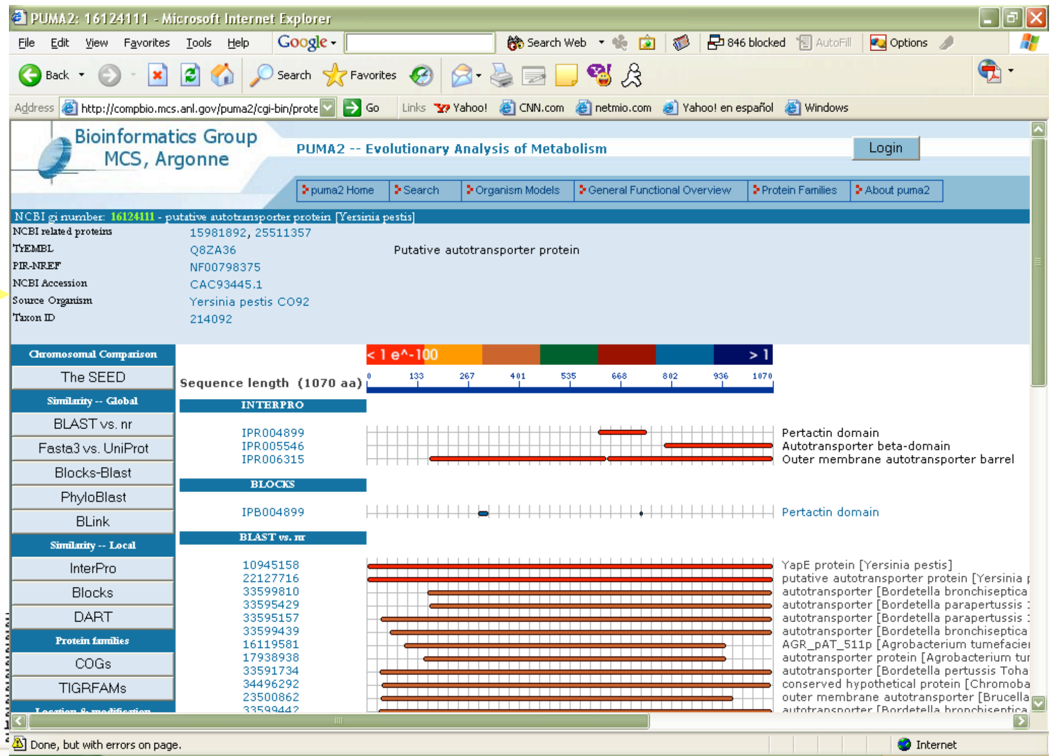
= Compute Job

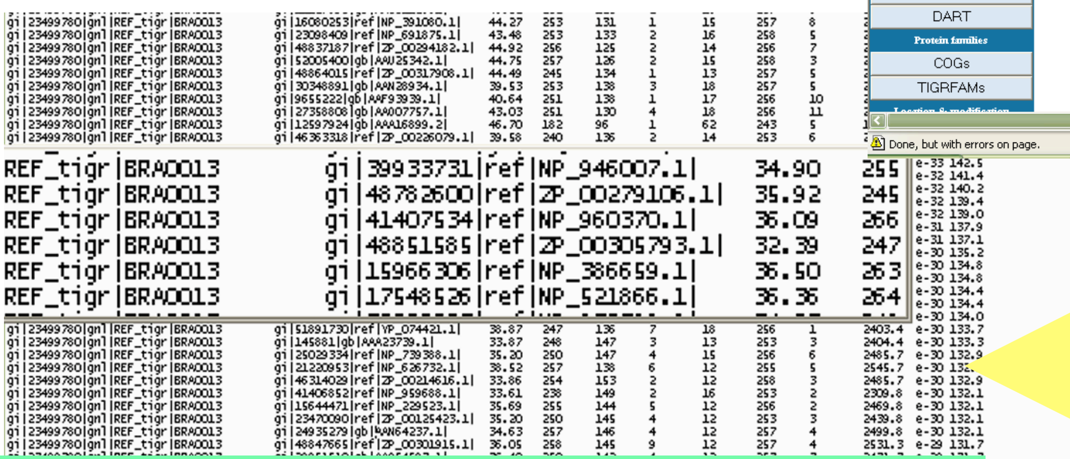
Montage Workflow: ~1200 jobs, 7 levels
NVO, NASA, ISI/Pegasus - Deelman et al.

PUMA: Analysis of Metabolism

PUMA Knowledge Base

Information about proteins analyzed against ~2 million gene sequences



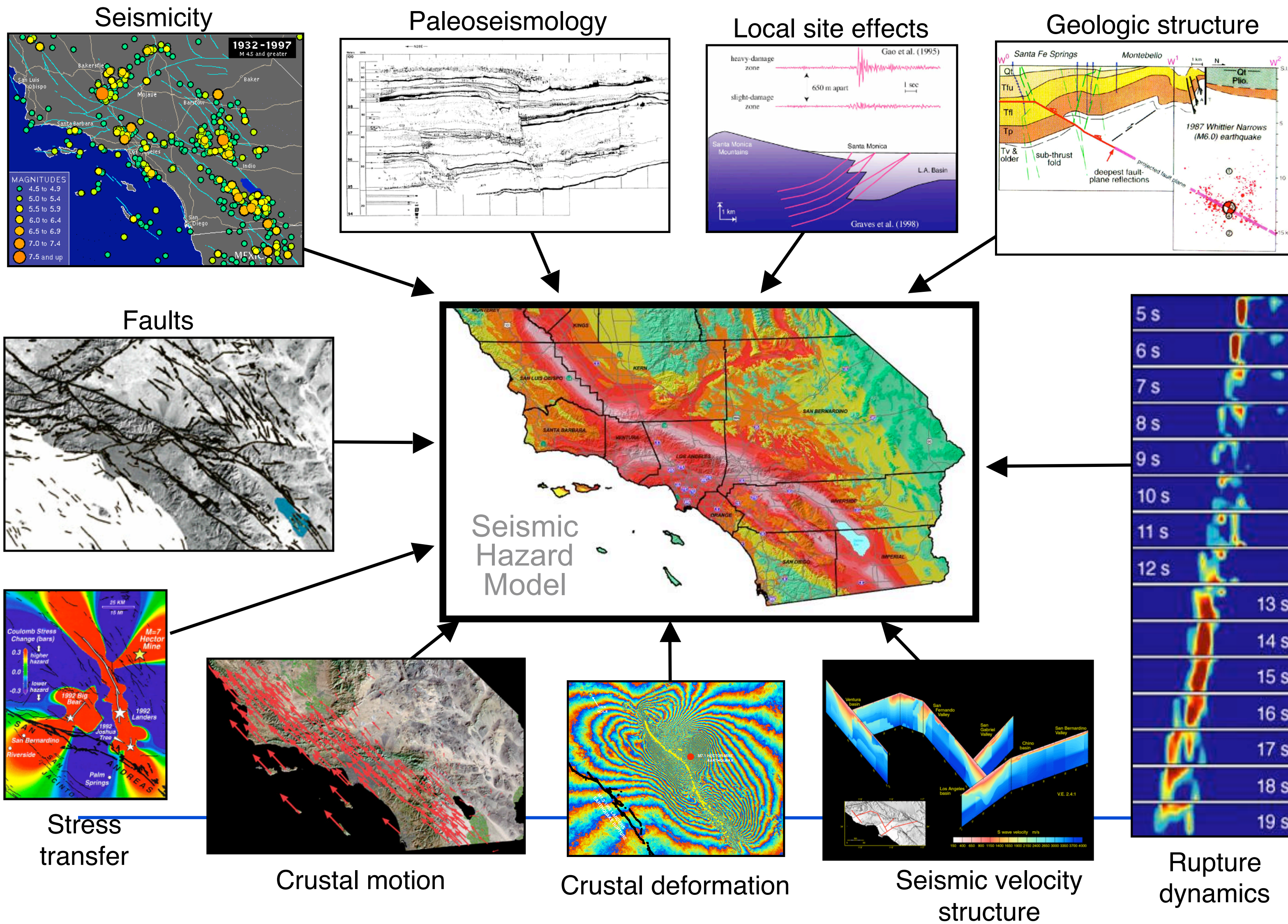


Analysis on Grid

Involves millions of BLAST, BLOCKS, and other processes

Natalia Maltsev et al.
http://compbio.mcs.anl.gov/puma2

Mining Seismic data for hazard analysis (Southern Calif. Earthquake Center).



Virtual Organizations

- Groups of organizations that use the Grid to share resources for specific purposes
- Support a single community
- Deploy compatible technology and agree on working policies
  - Security policies - difficult
- Deploy different network accessible services:
  - Grid Information
  - Grid Resource Brokering
  - Grid Monitoring
  - Grid Accounting

The Grid Middleware Stack *(and course modules)*

Grid Application (M5)
(often includes a *Portal*)

Workflow system (explicit or *ad-hoc*) (M6)

Job Management (M2)	**Data Management (M3)**	**Grid Information Services (M5)**

Grid Security Infrastructure (M4)

Core Globus Services (M1)

Standard Network Protocols and *Web Services* (M1)

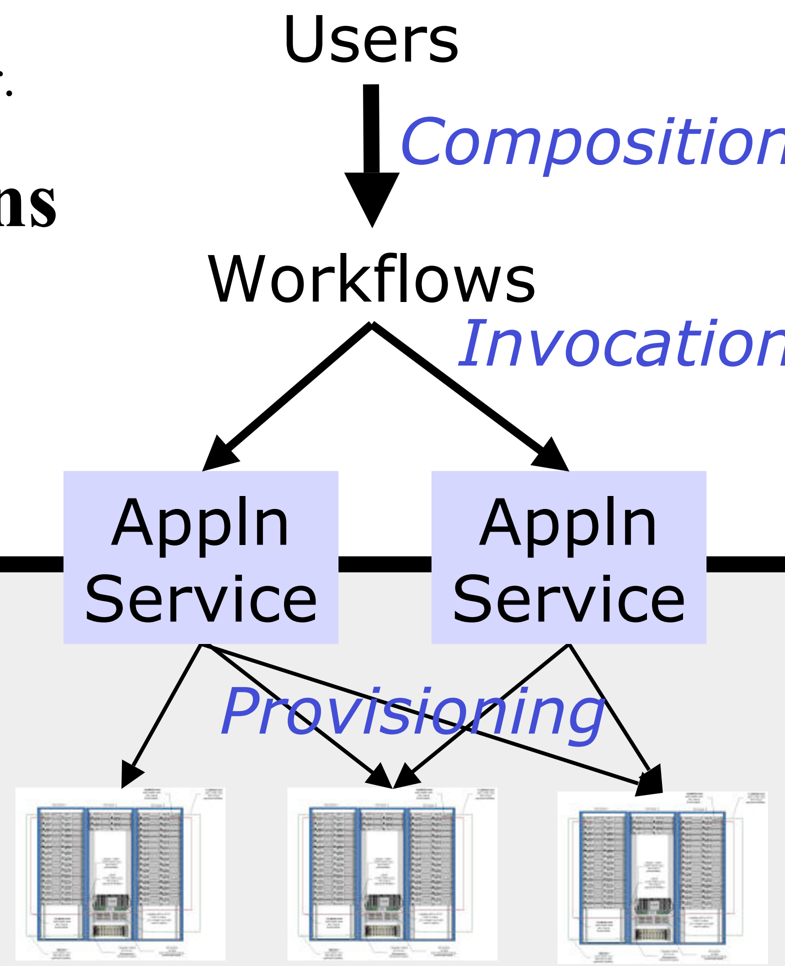
Globus and Condor play key roles

- Globus Toolkit provides the base middleware
  - Client tools which you can use from a command line
  - APIs (scripting languages, C, C++, Java, …) to build your own tools, or use direct from applications
  - Web service interfaces
  - Higher level tools built from these basic components, e.g. Reliable File Transfer (RFT)
- Condor provides both client & server scheduling
  - In grids, Condor provides an agent to queue, schedule and manage work submission

Grid architecture is evolving to a Service-Oriented approach.

...but this is beyond our workshop's scope.
See "Service-Oriented Science" by Ian Foster.

- Service-oriented **applications**
  - Wrap applications as services
  - Compose applications into workflows
- Service-oriented **Grid infrastructure**
  - Provision physical resources to support application workloads



"The Many Faces of IT as Service", Foster, Tuecke, 2005

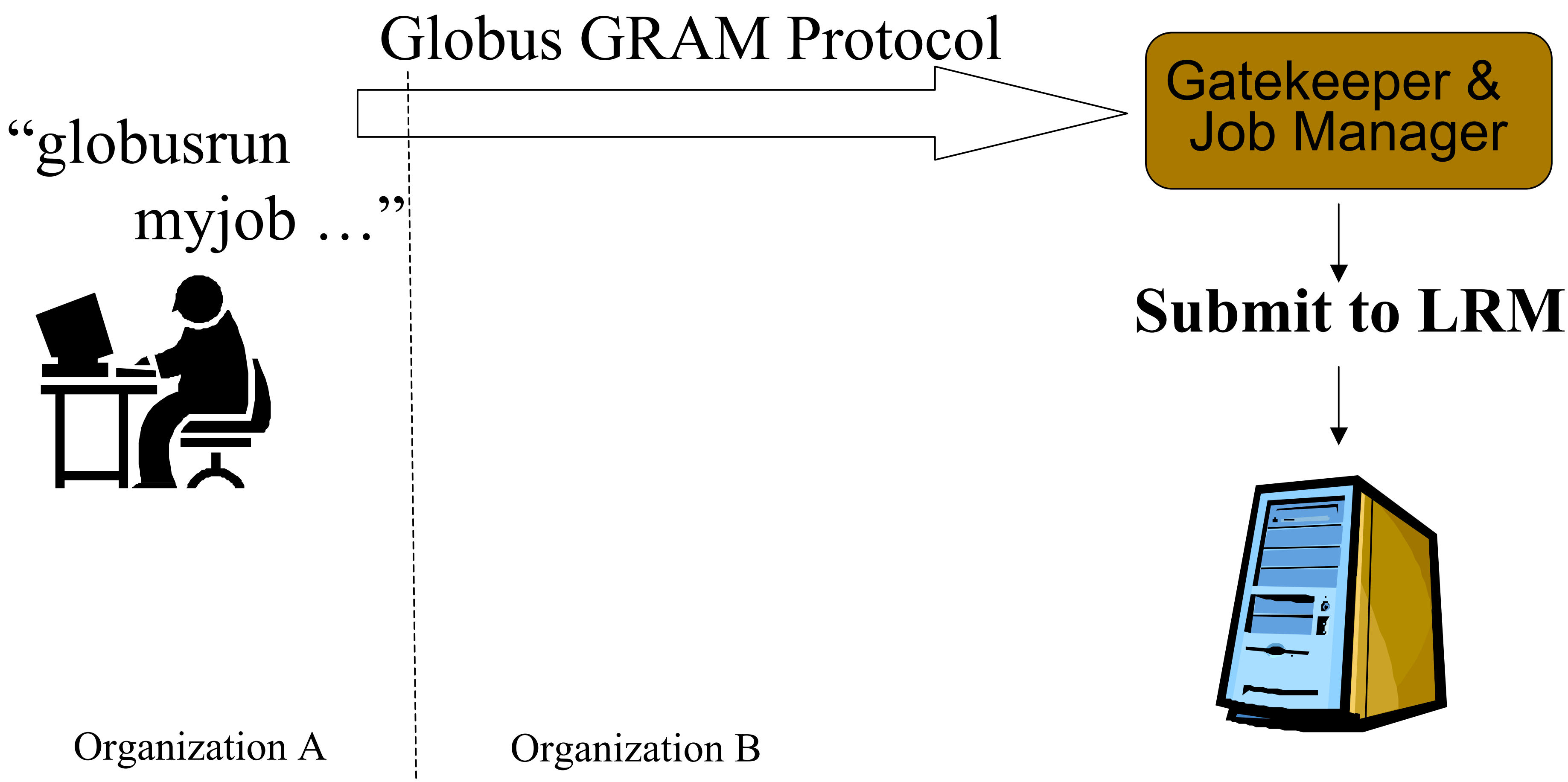
Job and resource management

Workshop Module 2

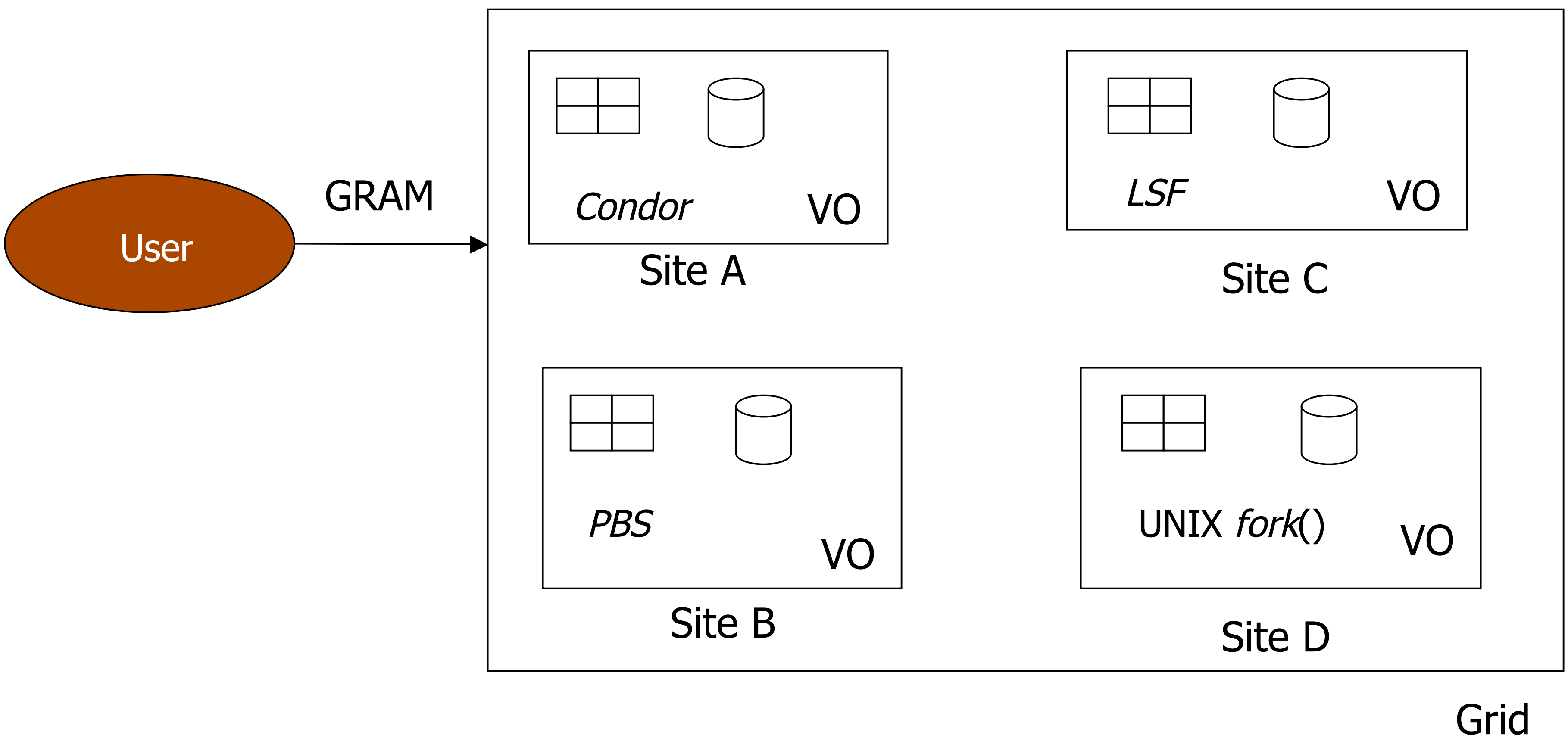
Local Resource Manager: a batch scheduler for running jobs on a computing cluster

- Popular LRMs include:
  - PBS – Portable Batch System
  - LSF – Load Sharing Facility
  - SGE – Sun Grid Engine
  - Condor – Originally for cycle scavenging, Condor has evolved into a comprehensive system for managing computing
- LRMs execute on the cluster's *head node*
- Simplest LRM allows you to "fork" jobs quickly
  - Runs on the head node (*gatekeeper)* for fast utility functions
  - No queuing (but this is emerging to "throttle" heavy loads)
- In GRAM, each LRM is handled with a "job manager"

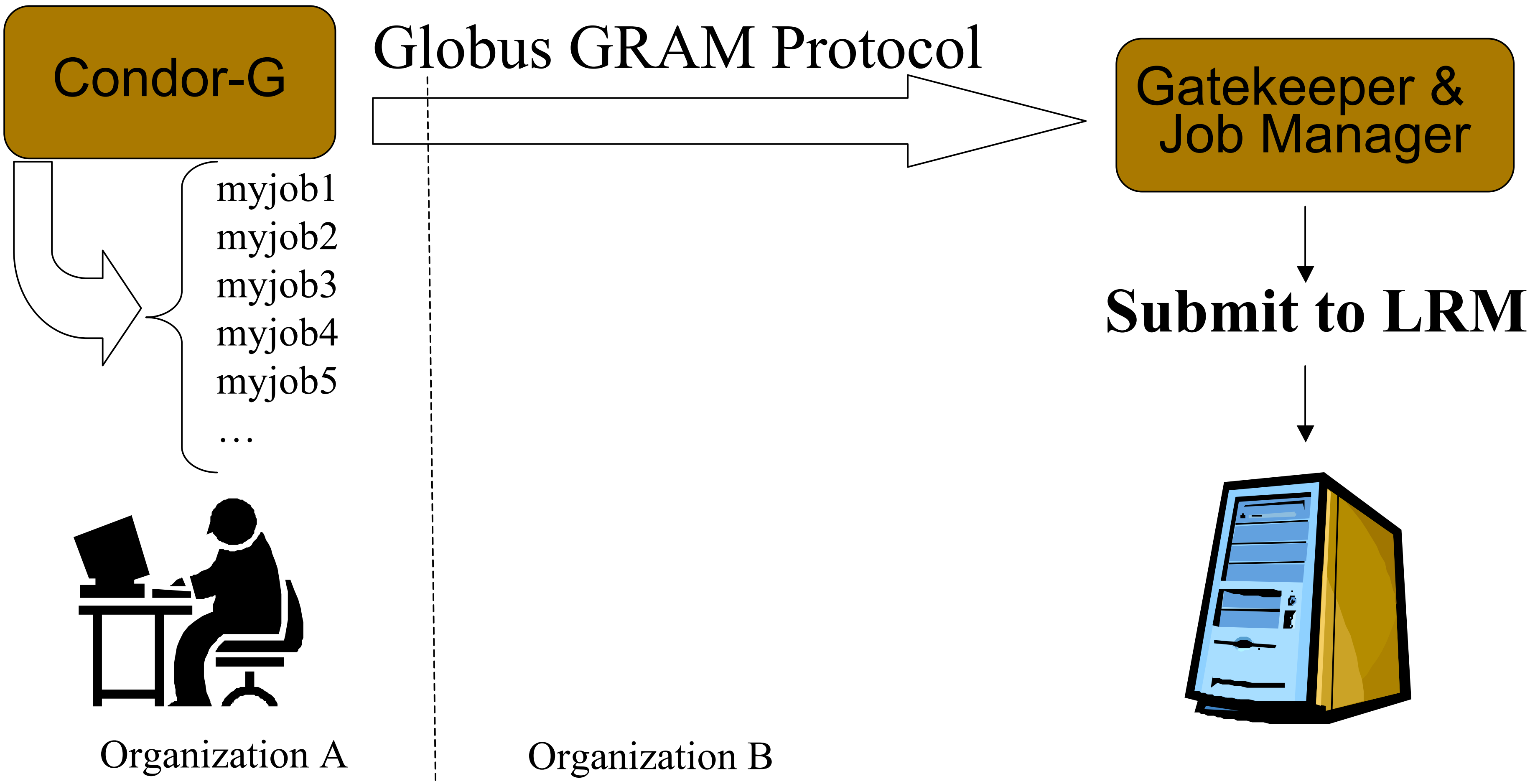
GRAM – Globus Resource Allocation Manager



GRAM provides a uniform interface to diverse cluster schedulers.



Condor-G: Grid Job Submission Manager



Data Management

Workshop Module 3

Data management services provide the mechanisms to find, move and share data

- GridFTP
  - Fast, Flexible, Secure, Ubiquitous data transport
  - Often embedded in higher level services
- RFT
  - Reliable file transfer service using GridFTP
- Replica Location Service
  - Tracks multiple copies of data for speed and reliability
- Storage Resource Manager
  - Manages storage space allocation, aggregation, and transfer
- Metadata management services are evolving

GridFTP is secure, reliable and fast

- Security through GSI
  - Authentication and authorization
  - Can also provide encryption
- Reliability by restarting failed transfers
- Fast
  - Can set TCP buffers for optimal performance
  - Parallel transfers
  - Striping (multiple endpoints)
- Client Tools

  globus-url-copy, uberftp, custom clients

Grids replicate data files for faster access

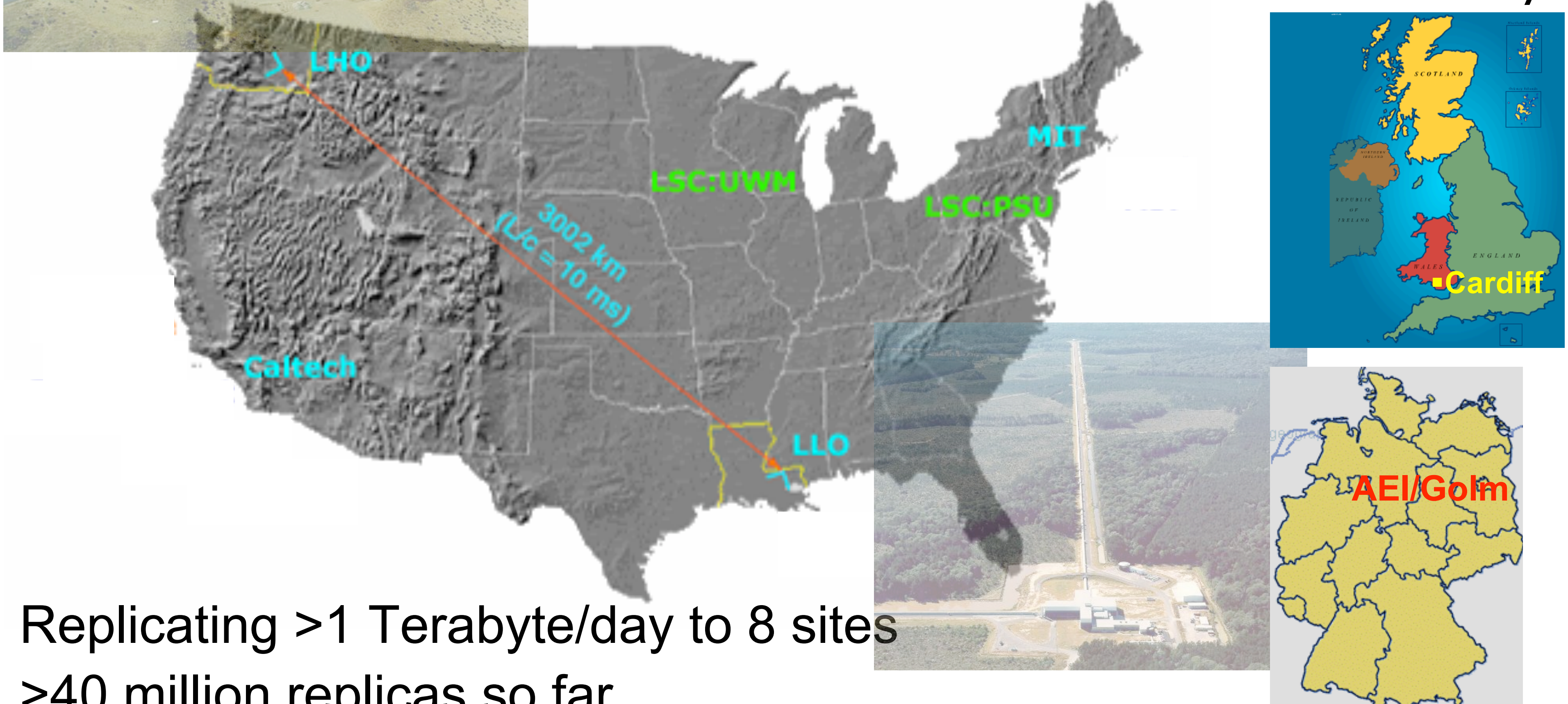
- Effective use of the grid resources – more parallelism
- Each *logical* file can have multiple *physical* copies
- Avoids single points of failure
- Manual or automatic replication
  - Automatic replication considers the demand for a file, transfer bandwidth, etc.

File catalogues tell you where the data is

- File Catalog Services
  - Replica Location Service (RLS)
  - Phedex
  - RefDB / PupDB
- Requirements
  - Abstract the logical file name (LFN) for a physical file
  - maintain the mappings between the LFNs and the PFNs *(physical file names)*
  - Maintain the location information of a file

The Globus-Based LIGO Data Grid

LIGO Gravitational Wave Observatory



Replicating >1 Terabyte/day to 8 sites

>40 million replicas so far

MTBF = 1 month

Grid Security

Workshop Module 4

Grid security is a crucial component

- Problems being solved might be sensitive
- Resources are typically valuable
- Resources are located in distinct administrative domains
  - Each resource has own policies, procedures, security mechanisms, etc.
- Implementation must be broadly available & applicable
  - Standard, well-tested, well-understood protocols; integrated with wide variety of tools

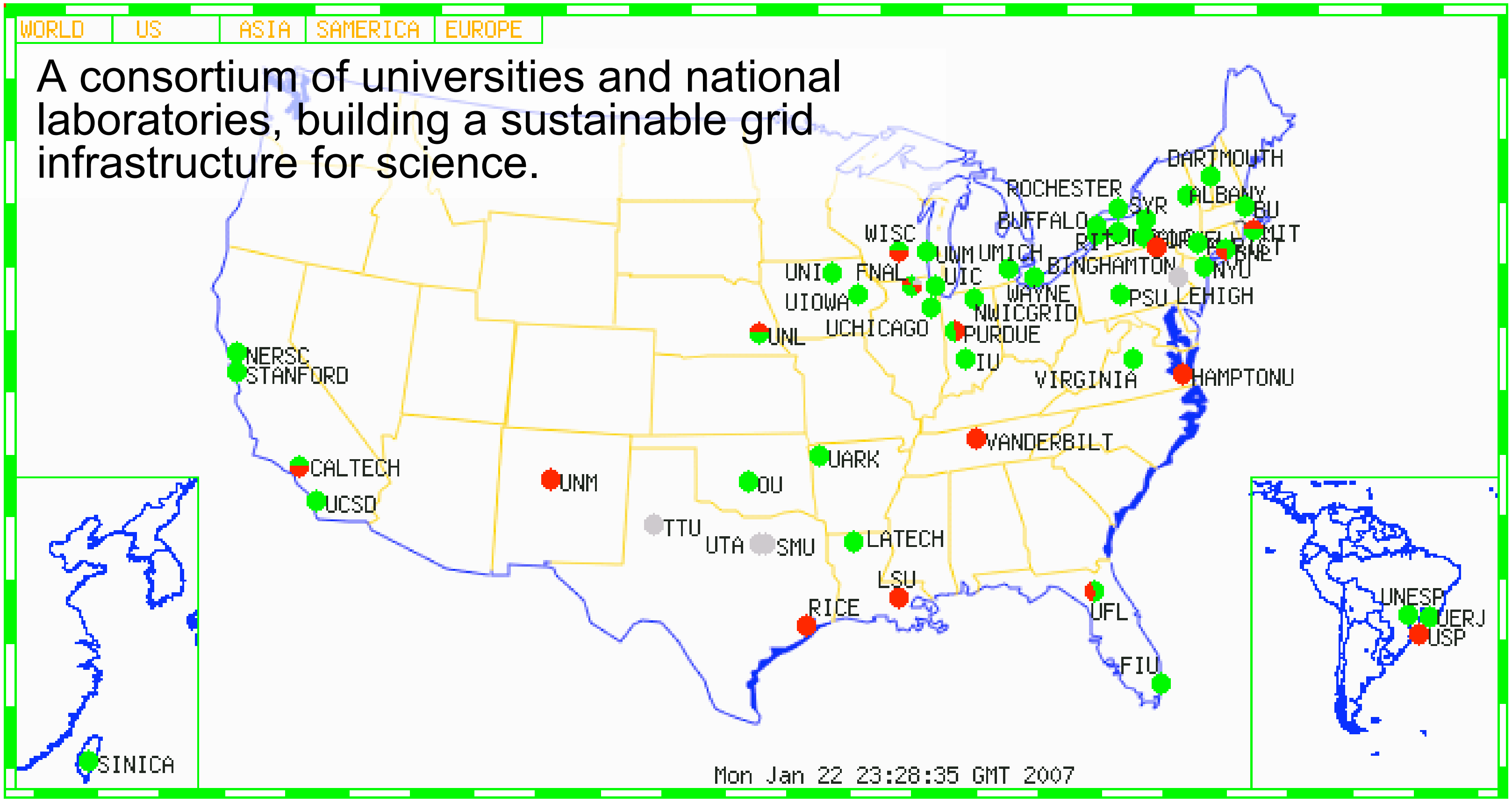
Grid Security Infrastructure - GSI

- Provides secure communications for all the higher-level grid services
- Secure *Authentication* and *Authorization*
  - Authentication ensures you *are* whom you claim to be
    - *ID card, fingerprint, passport, username/password*
  - Authorization controls what you are permitted to *do*
    - *Run a job, read or write a file*
- GSI provides Uniform Credentials
- Single Sign-on
  - User authenticates once – then can perform many tasks

National Grid Cyberinfrastructure

Workshop Module 5

Open Science Grid (OSG) provides shared computing resources, benefiting a broad set of disciplines

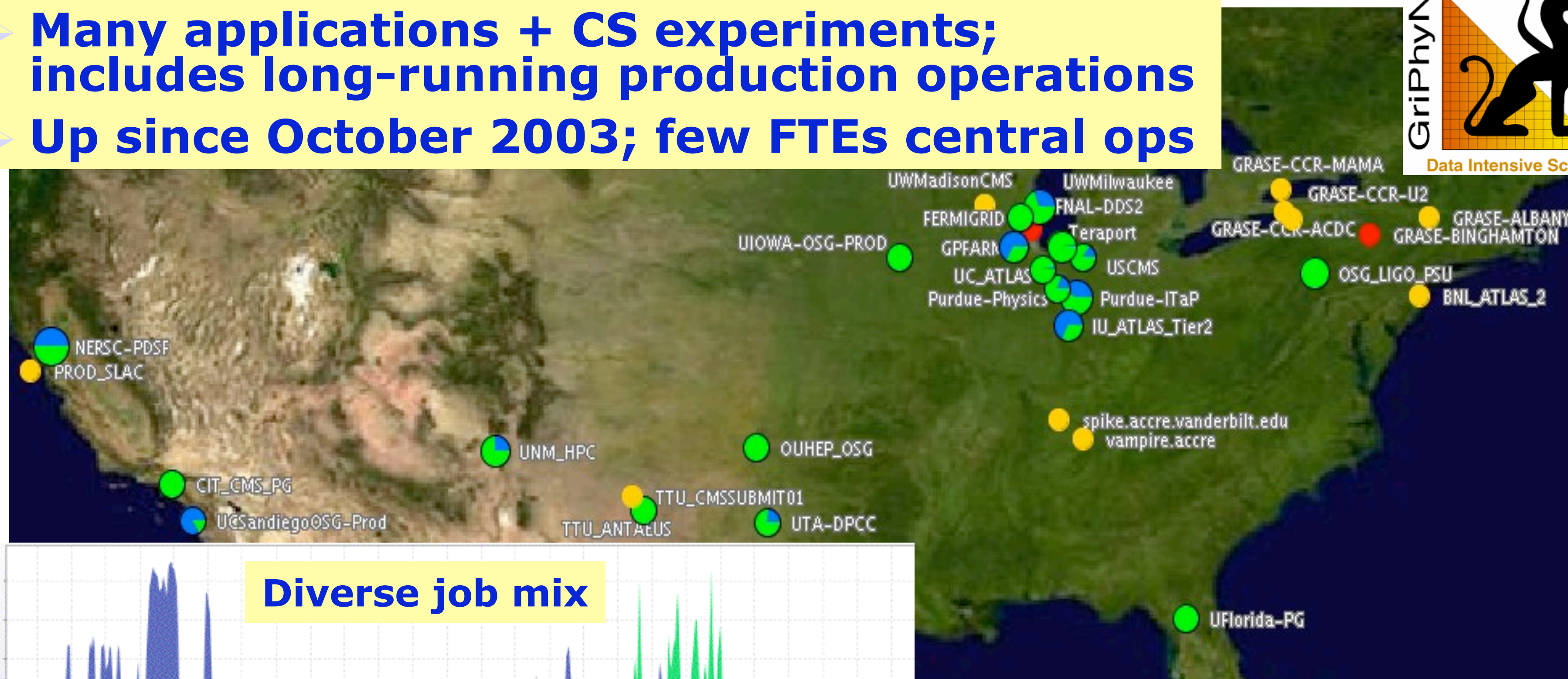


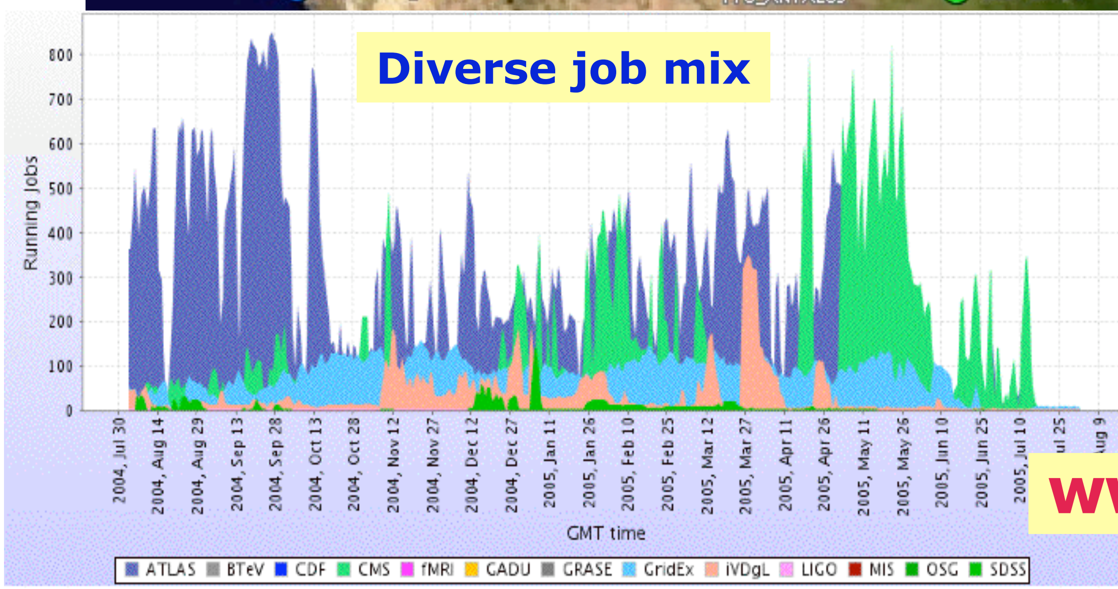
- OSG incorporates advanced networking and focuses on general services, operations, end-to-end performance
- Composed of a large number (>50 and growing) of shared computing facilities, or “sites”

http://www.opensciencegrid.org/

Open Science Grid

- 50 sites (15,000 CPUs) & growing
- 400 to >1000 concurrent jobs
- Many applications + CS experiments; includes long-running production operations
- Up since October 2003; few FTEs central ops





Diverse job mix

www.opensciencegrid.org

TeraGrid provides vast resources via a number of huge computing facilities.

To efficiently use a Grid, you must locate and monitor its resources.

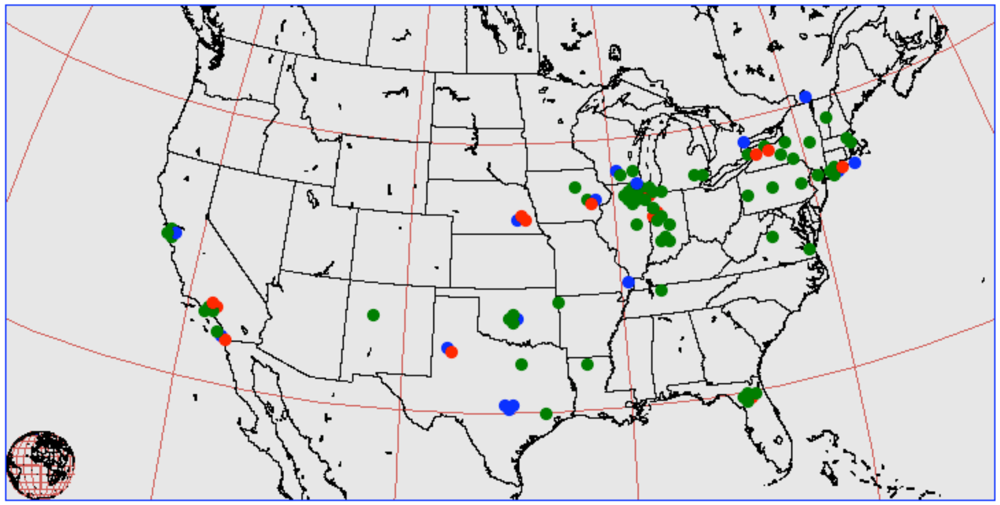
- Check the availability of different grid sites
- Discover different grid services
- Check the status of "jobs"
- Make better scheduling decisions with information maintained on the "health" of sites

OSG Resource Selection Service: VORS

Virtual Organization Resource Selector

Open Science Grid

All	OSG	TeraGrid	EGEE	OSG-ITB



Virtual Organization Selection

All	CDF	CMS	CompBioGrid	DES	DOSAR	DZero	Engage	Fermilab	fMRI	GADU
	geant4	GLOW	GPN	GRASE	GridChem	GridEx	GROW	i2u2	iVDGL	LIGO
mariachi	MIS	nanoHUB	NWICG	Ops	OSG	OSGEDU	SDSS	STAR	USATLAS	

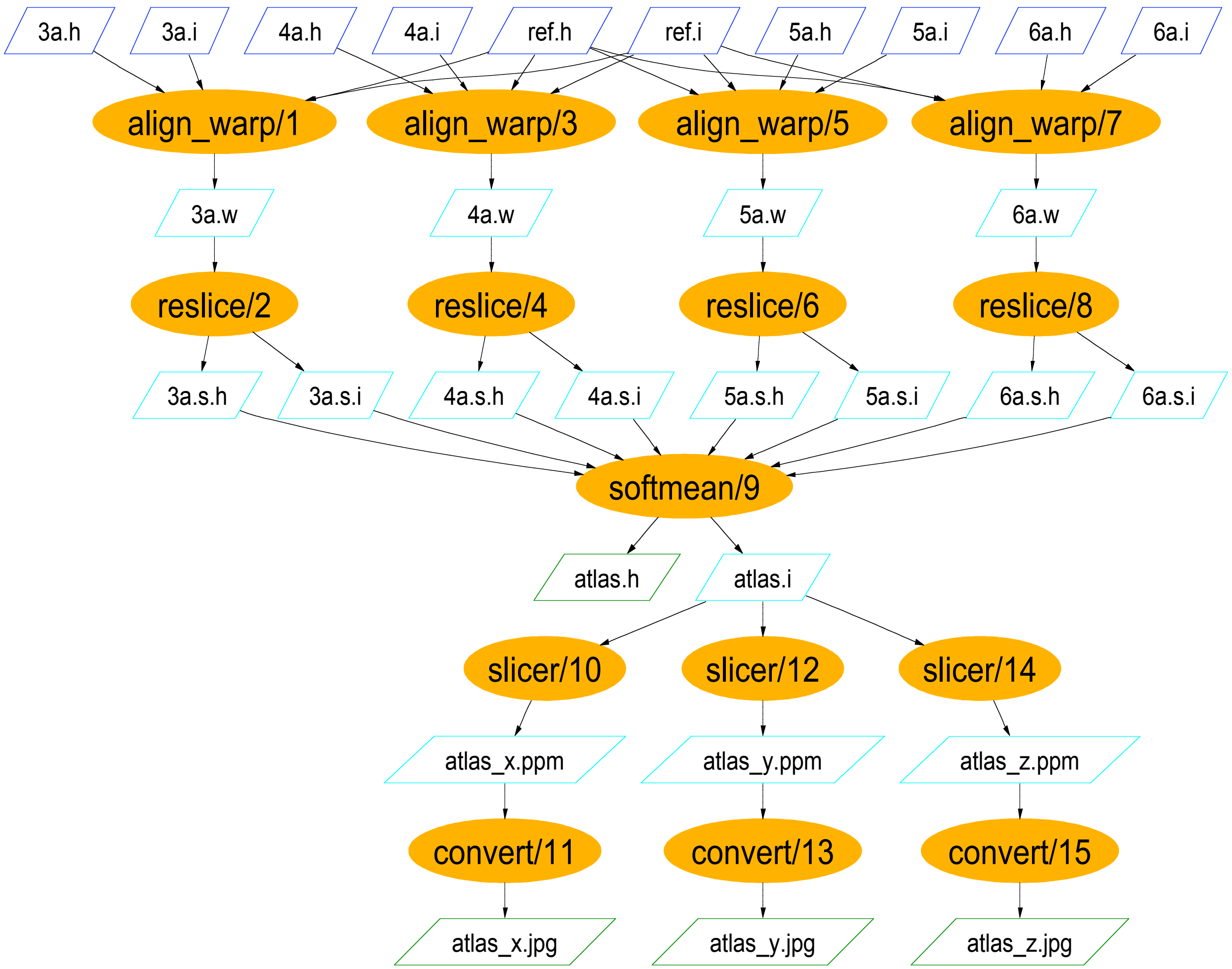
Resources

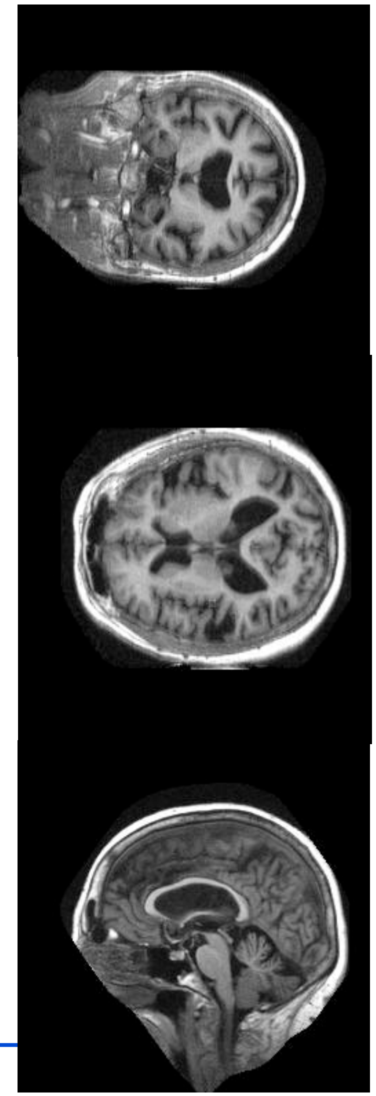
Name	Gatekeeper	Type	Grid	Status	Last Test Date
BNL_ATLAS_1	gridgk01.racf.bnl.gov:2119	compute	OSG	PASS	2006-12-08 14:57:13
BNL_ATLAS_2	gridgk02.racf.bnl.gov:2119	compute	OSG	PASS	2006-12-08 14:58:43
BU_ATLAS_Tier2	atlas.bu.edu:2119	compute	OSG	PASS	2006-12-08 15:00:44

Grid Workflow

Workshop Module 6

A typical workflow pattern in image analysis runs many filtering apps.





Workflow courtesy James Dobson, Dartmouth Brain Imaging Center

Conclusion: Why Grids?

- New approaches to inquiry based on
  - Deep analysis of huge quantities of data
  - Interdisciplinary collaboration
  - Large-scale simulation and analysis
  - Smart instrumentation
  - ***Dynamically assemble the resources to tackle a new scale of problem***
- Enabled by access to resources & services without regard for location & other barriers

Grids:
Because Science needs community …

- Teams organized around common goals
  - People, resource, software, data, instruments…
- With diverse membership & capabilities
  - Expertise in multiple areas required
- And geographic and political distribution
  - No location/organization possesses all required skills and resources
- Must adapt as a function of the situation
  - Adjust membership, reallocate responsibilities, renegotiate resources

Based on:

Grid Intro and Fundamentals Review

Dr. Gabrielle Allen

Center for Computation & Technology

Department of Computer Science

Louisiana State University

gallen@cct.lsu.edu