

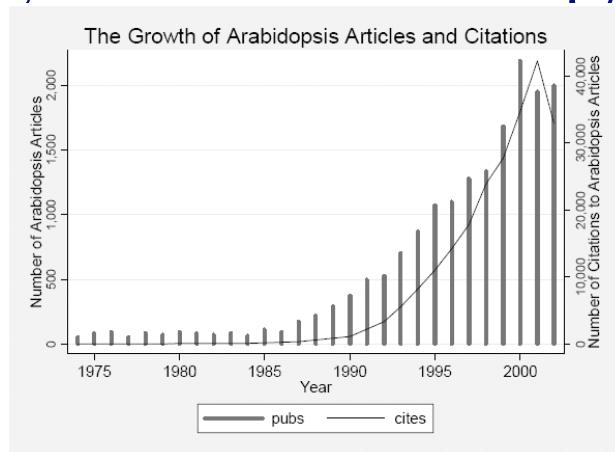
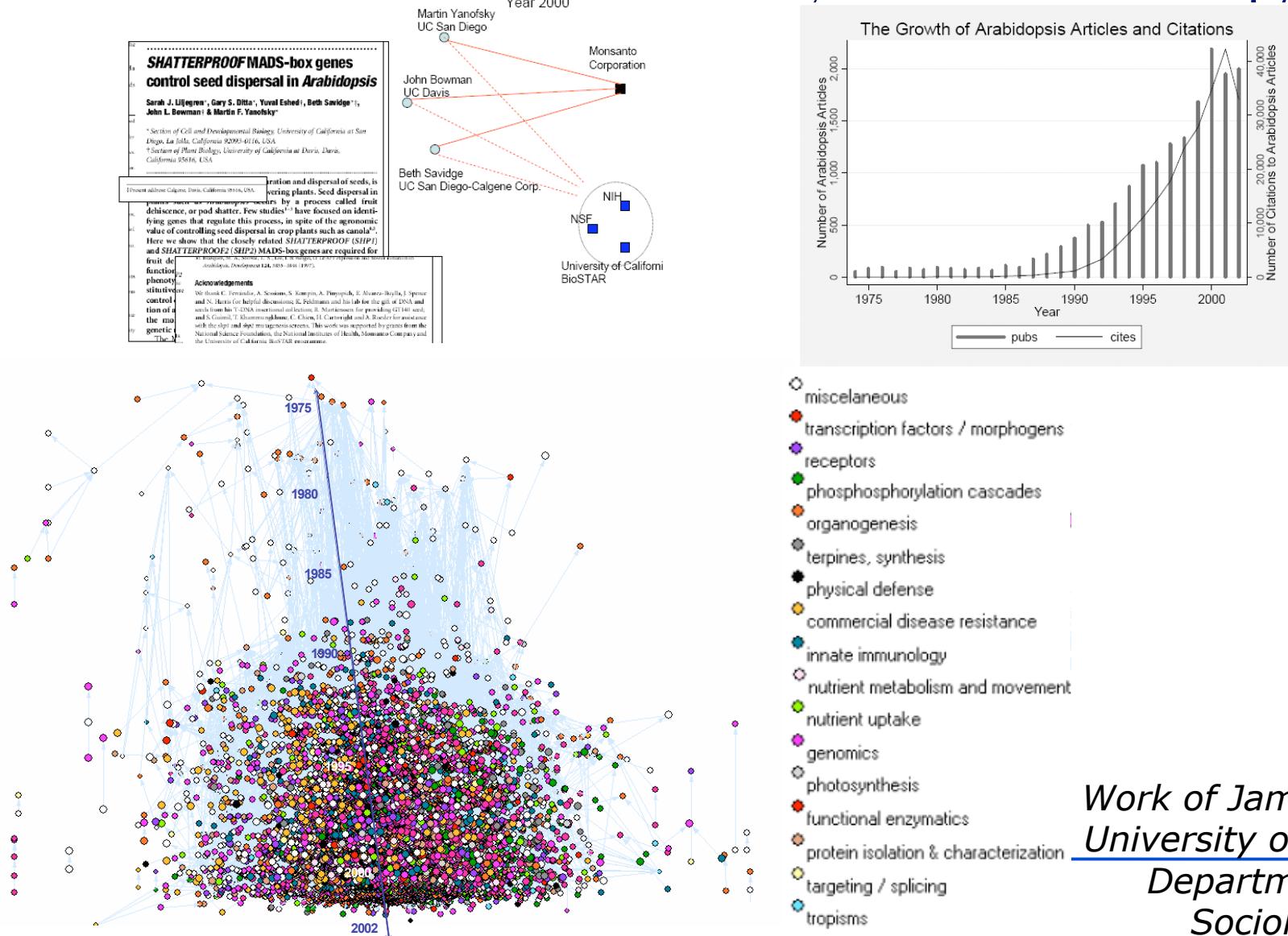
Introduction to Grid Computing

Grid School Workshop – Module 1



Open Science Grid

Scaling up Science: Citation Network Analysis in Sociology

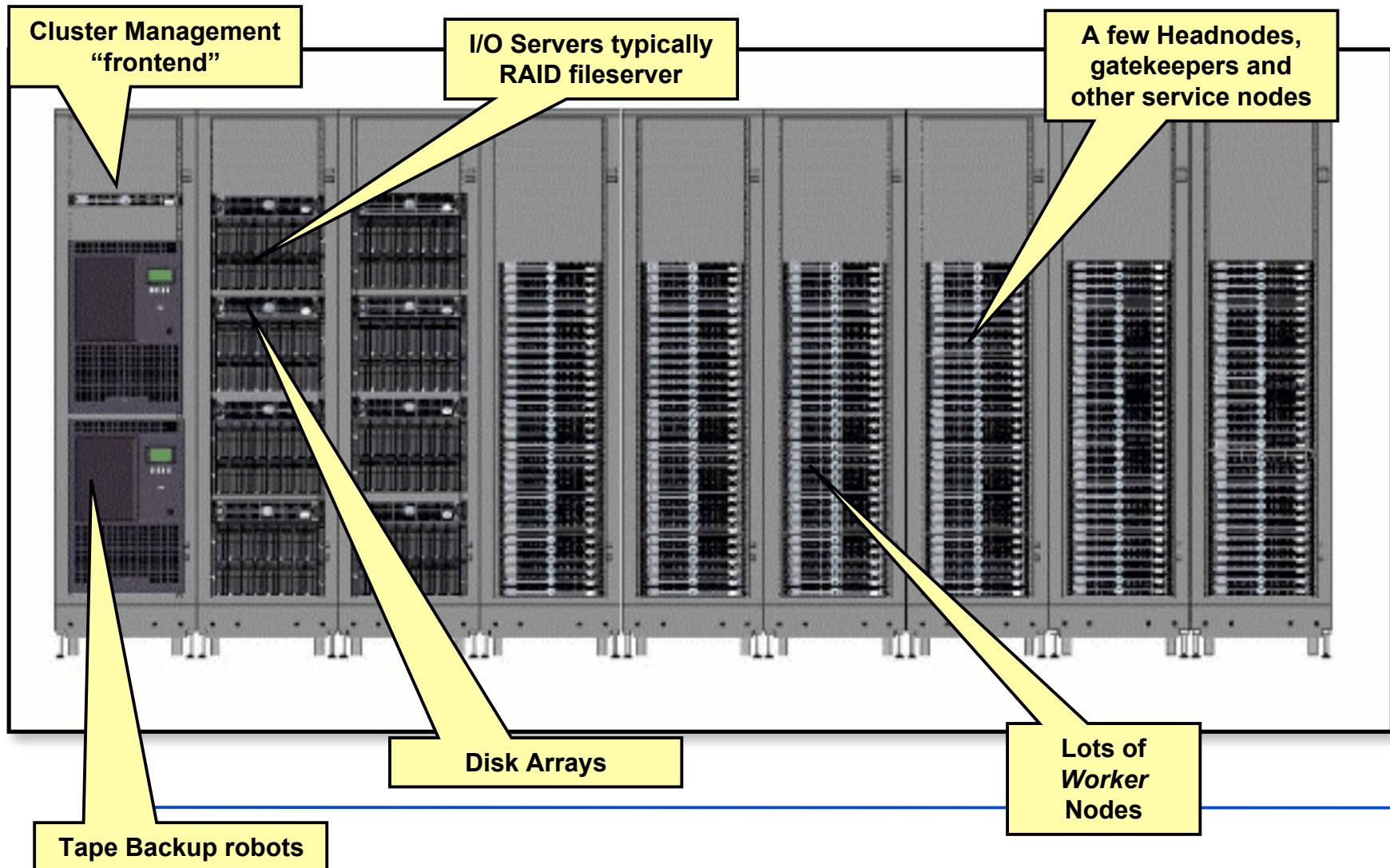


Work of James Evans,
University of Chicago,
Department of
Sociology

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*.

Computing clusters have commoditized supercomputing



Initial Grid driver: High Energy Physics

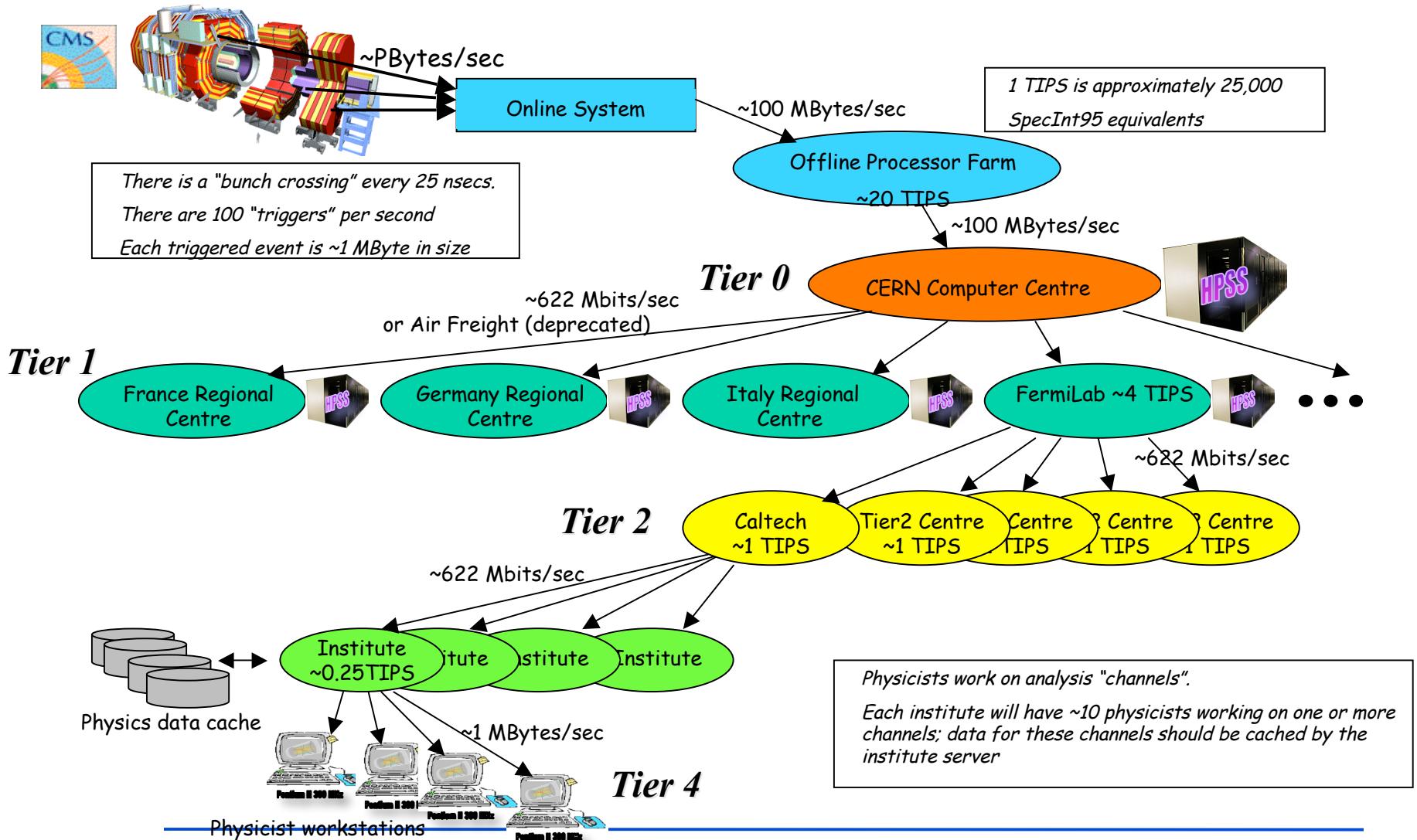


Image courtesy Harvey Newman, Caltech

Grids can process vast datasets.

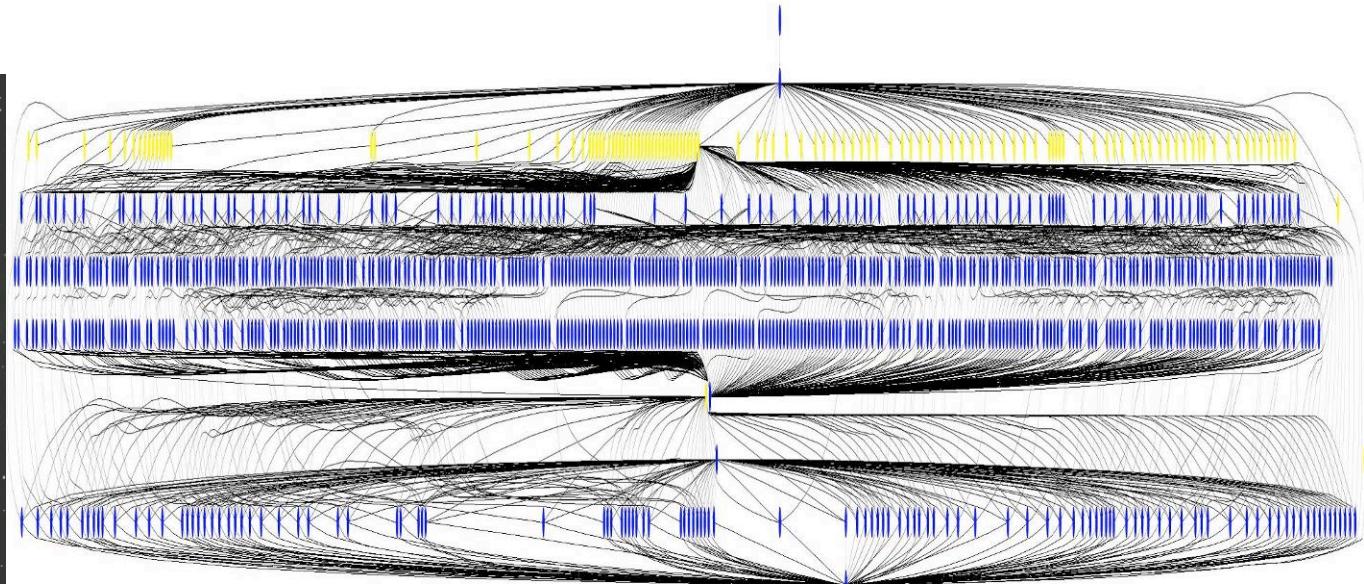
- Many HEP and Astronomy experiments consist of:
 - Large datasets as inputs (find datasets)
 - “Transformations” which work on the input datasets (process)
 - The output datasets (store and publish)
- The emphasis is on the sharing of the large datasets
- Programs when *independent* can be *parallelized*.



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

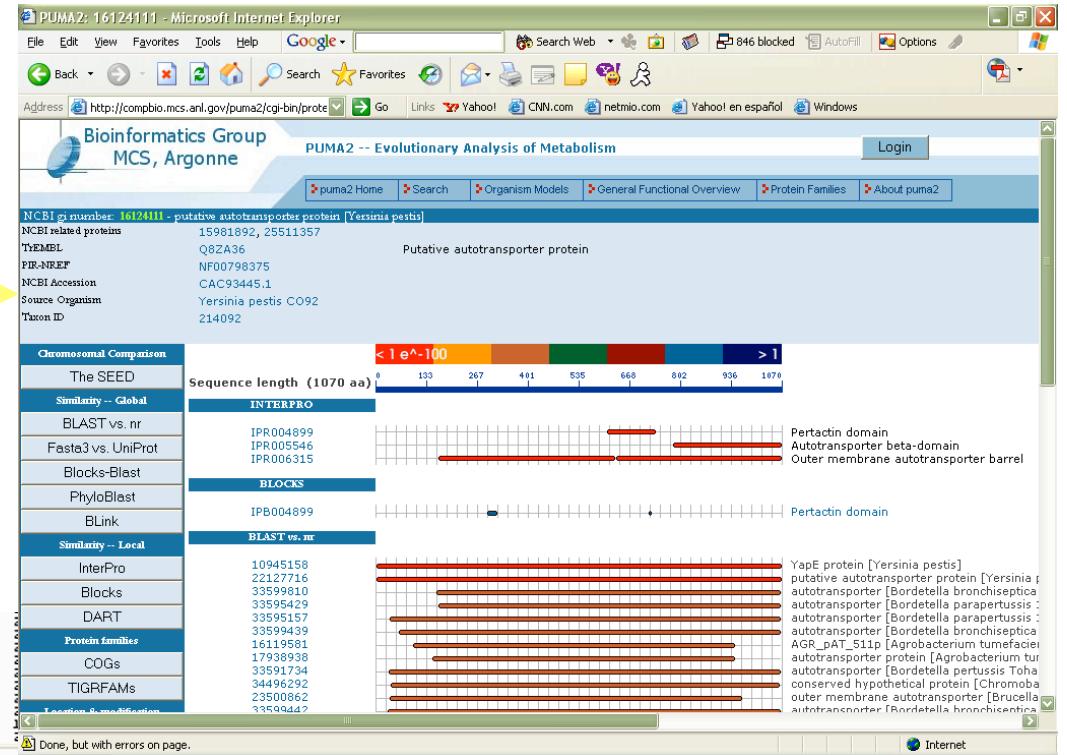
Information about proteins analyzed against ~2 million gene sequences

gi 23499780 gnl REF_tigr BRA0013	gi 16080253 ref NP_391080.1	44.27	293	131	1	15	257	8
gi 23499780 gnl REF_tigr BRA0013	gi 16080253 ref NP_391085.1	43.48	253	133	2	16	258	5
gi 23499780 gnl REF_tigr BRA0013	gi 148837187 ref NP_00294182.1	44.92	256	125	2	14	256	7
gi 23499780 gnl REF_tigr BRA0013	gi 52005400 gb AU25342.1	44.75	257	126	2	15	258	3
gi 23499780 gnl REF_tigr BRA0013	gi 48864015 ref NP_00317908.1	44.49	245	134	1	13	257	5
gi 23499780 gnl REF_tigr BRA0013	gi 30398831 gb AW28934.1	38.53	253	138	3	18	257	5
gi 23499780 gnl REF_tigr BRA0013	gi 96514293 gb AW93939.1	40.64	241	138	1	17	266	10
gi 23499780 gnl REF_tigr BRA0013	gi 21258800 gb AU00772.1	42.92	251	120	4	16	256	11
gi 23499780 gnl REF_tigr BRA0013	gi 125997924 gb AU445899.2	46.70	182	96	1	62	243	2
gi 23499780 gnl REF_tigr BRA0013	gi 46383318 ref NP_00286079.1	39.58	240	136	2	14	253	6

REF_tigr BRA0013	gi 39933731 ref NP_946007.1	34.90
REF_tigr BRA0013	gi 48782600 ref NP_00279106.1	35.92
REF_tigr BRA0013	gi 41407534 ref NP_960370.1	36.09
REF_tigr BRA0013	gi 48861585 ref NP_00305793.1	32.39
REF_tigr BRA0013	gi 15966306 ref NP_386659.1	36.50
REF_tigr BRA0013	gi 17548526 ref NP_521866.1	36.36

gi 23499780 gnl REF_tigr BRA0013	gi 51891730 ref NP_074421.1	38.87	247	136	7	18	256	1
gi 23499780 gnl REF_tigr BRA0013	gi 145881 gb AA23739.1	33.87	247	147	3	13	253	3
gi 23499780 gnl REF_tigr BRA0013	gi 25029334 ref NP_739388.1	35.20	250	147	4	15	256	6
gi 23499780 gnl REF_tigr BRA0013	gi 21220953 ref NP_736732.1	38.52	257	138	6	12	255	5
gi 23499780 gnl REF_tigr BRA0013	gi 48400282 ref NP_00214674.1	33.88	244	153	2	12	256	3
gi 23499780 gnl REF_tigr BRA0013	gi 11414221 ref NP_98921.1	32.14	238	146	2	16	253	2
gi 23499780 gnl REF_tigr BRA0013	gi 15644471 ref NP_229523.1	35.69	251	144	5	12	256	2
gi 23499780 gnl REF_tigr BRA0013	gi 23420080 ref NP_00125423.1	35.20	250	145	4	12	253	3
gi 23499780 gnl REF_tigr BRA0013	gi 24935279 gb KAE64237.1	34.63	257	146	4	12	257	4
gi 23499780 gnl REF_tigr BRA0013	gi 48847865 ref NP_00301915.1	36.05	258	145	9	12	257	4

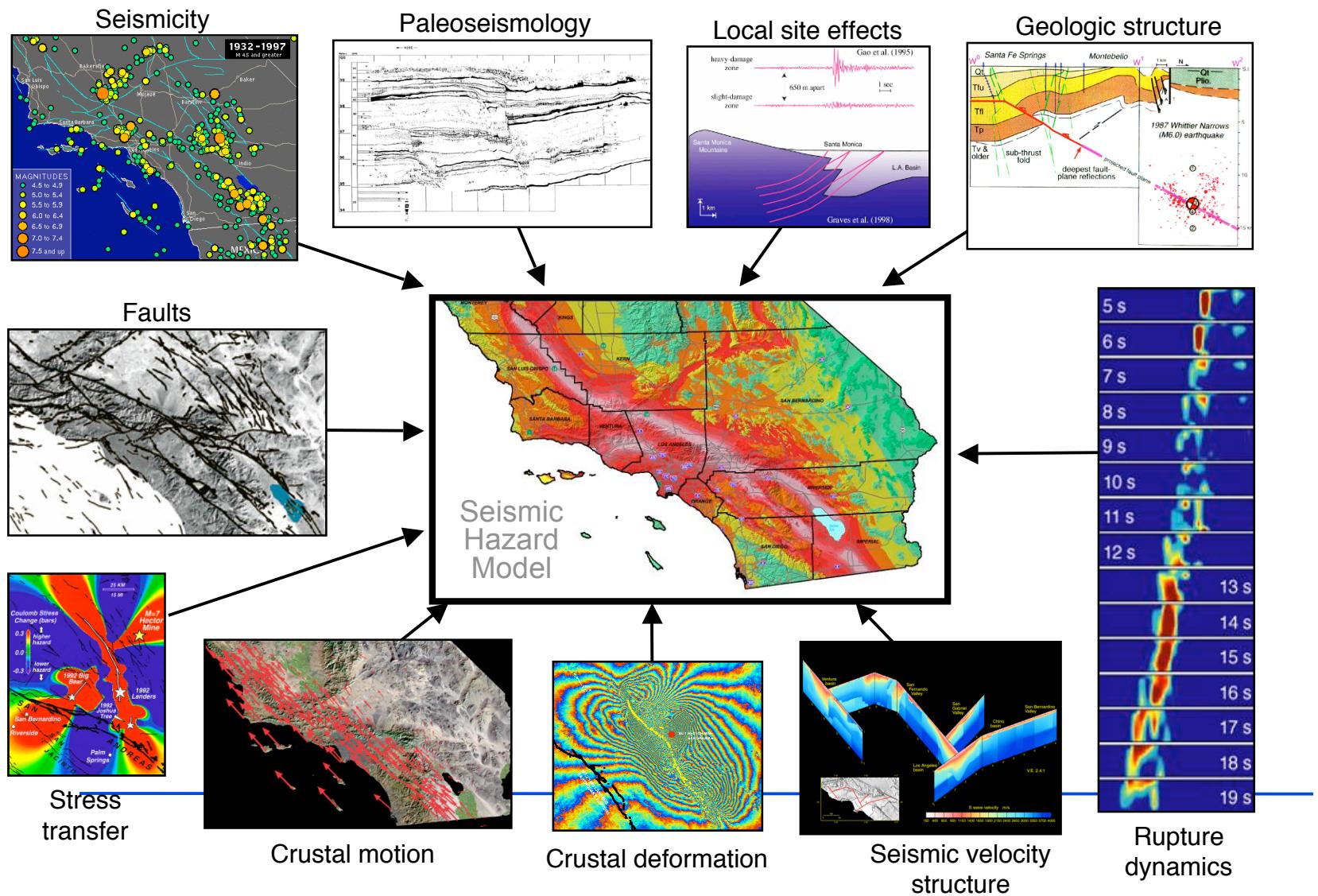
Natalia Maltsev et al.
<http://compbio.mcs.anl.gov/puma2>



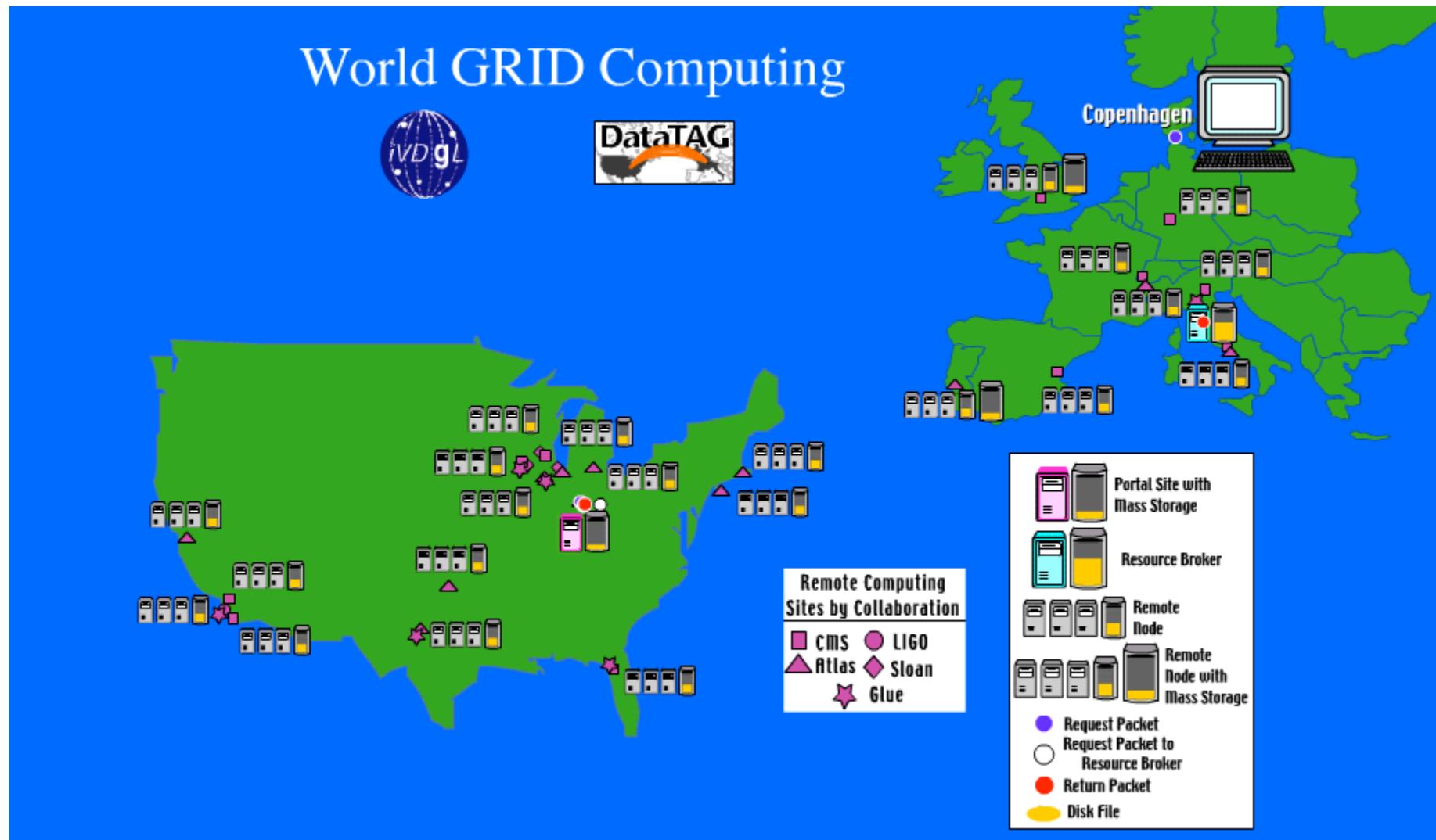
Analysis on Grid

Involves millions of
BLAST, BLOCKS, and
other processes

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



Grids Provide Global Resources To Enable e-Science



Ian Foster's Grid Checklist

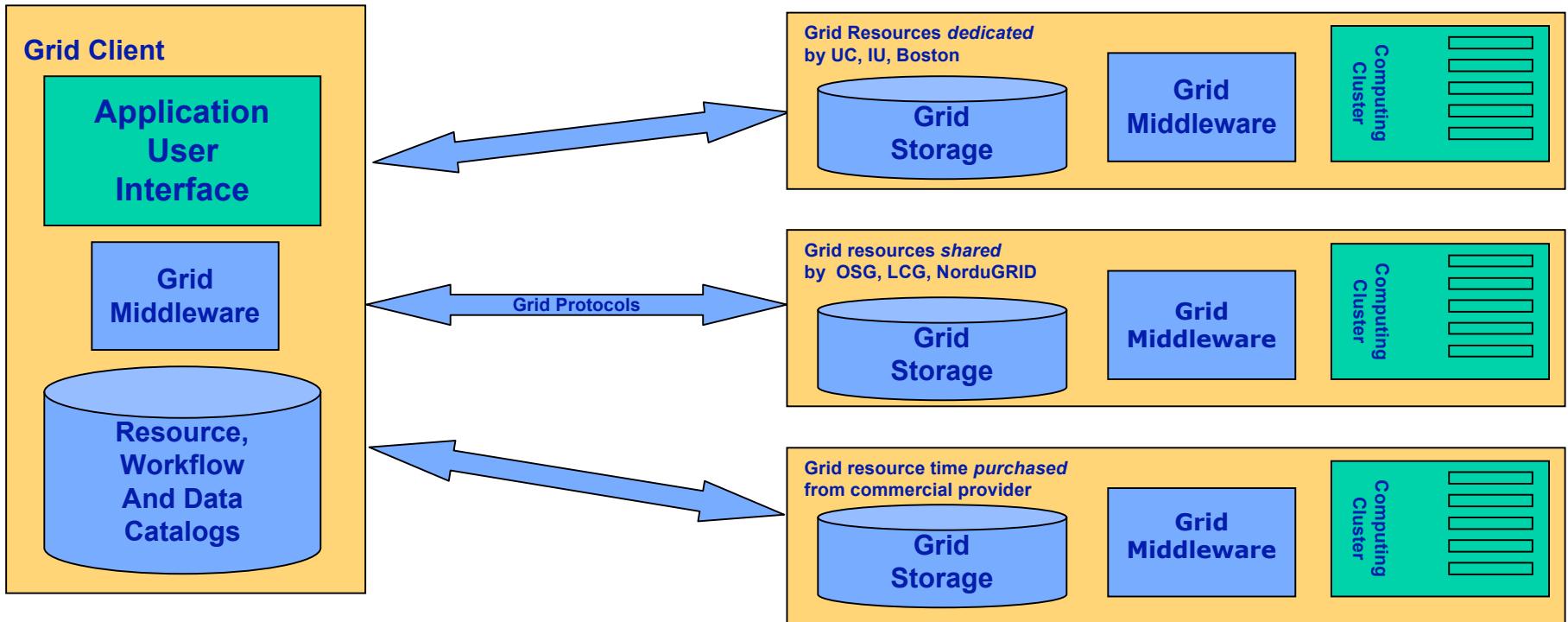
- A Grid is a system that:
 - Coordinates resources that are not subject to centralized control
 - Uses standard, open, general-purpose protocols and interfaces
 - Delivers non-trivial qualities of service

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



What is a grid made of? *Middleware.*

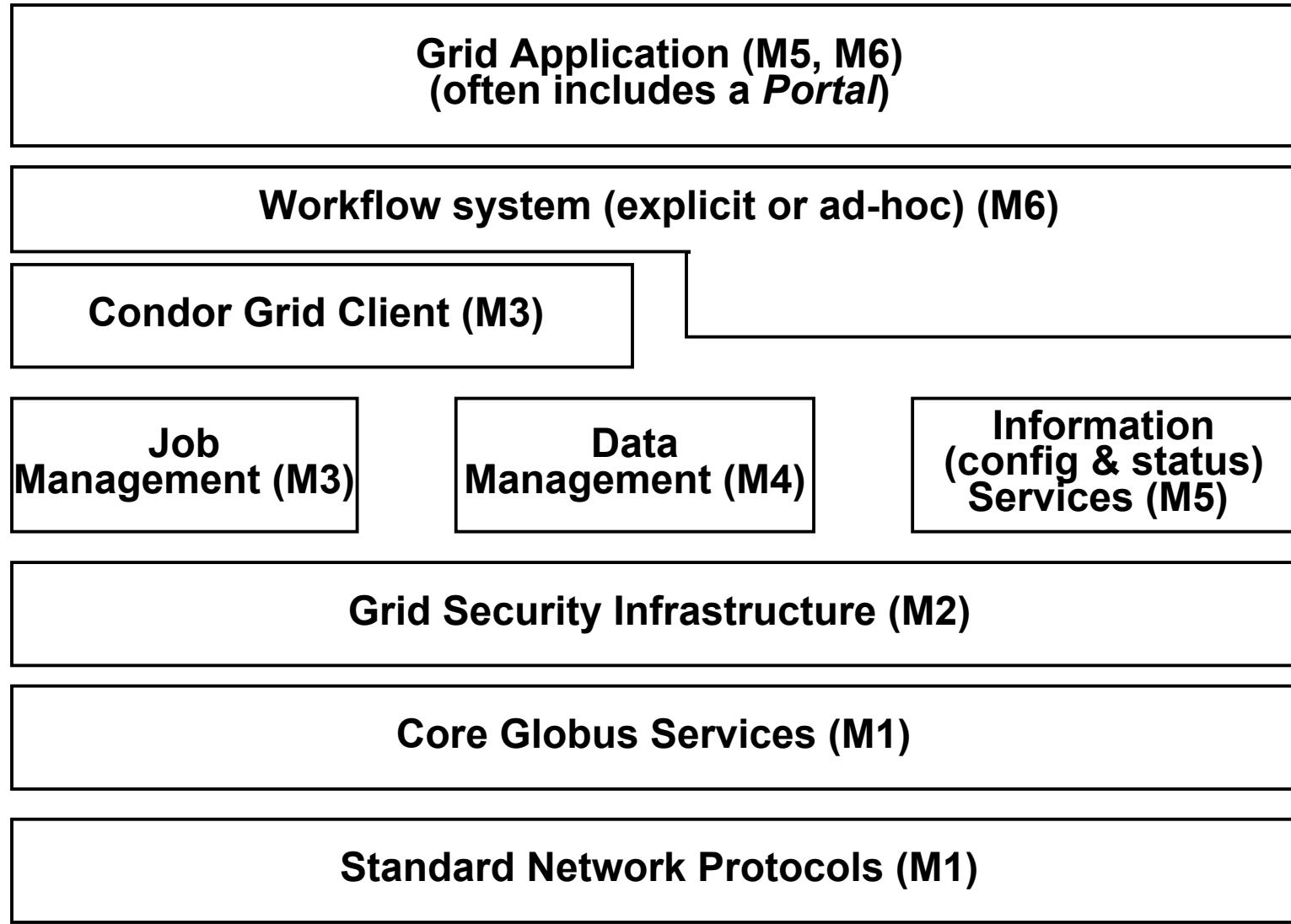


- Security to control access and protect communication (GSI)
- Directory to locate grid sites and services: (VORS, MDS)
- Uniform interface to computing sites (GRAM)
- Facility to maintain and schedule queues of work (Condor-G)
- Fast and secure data set mover (GridFTP, RFT)
- Directory to track where datasets live (RLS)

We will focus on Globus and Condor

- Condor provides both client & server scheduling
 - In grids, Condor provides an agent to queue, schedule and manage work submission
- Globus Toolkit provides the lower-level 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)

Grid software is like a “stack”



Grid architecture is evolving to a Service-Oriented approach.

...but this is beyond our workshop's scope.

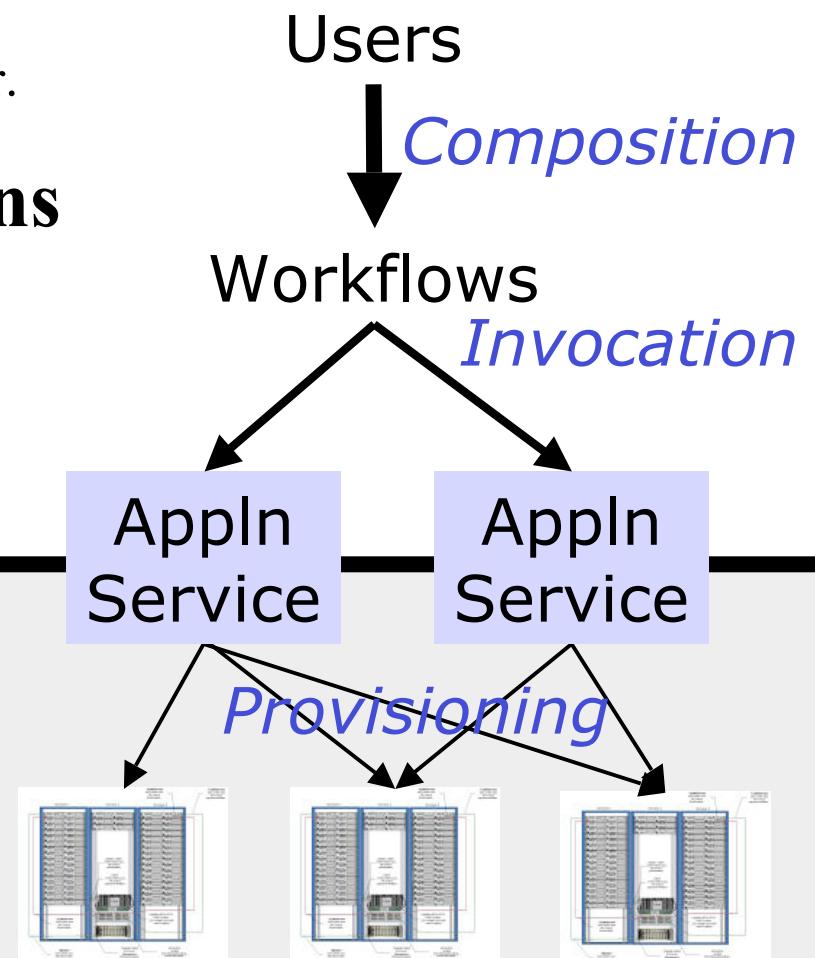
See "Service-Oriented Science" by Ian Foster.

■ Service-oriented applications

- Wrap applications as services
- Compose applications into workflows

■ Service-oriented Grid infrastructure

- Provision physical resources to support application workloads



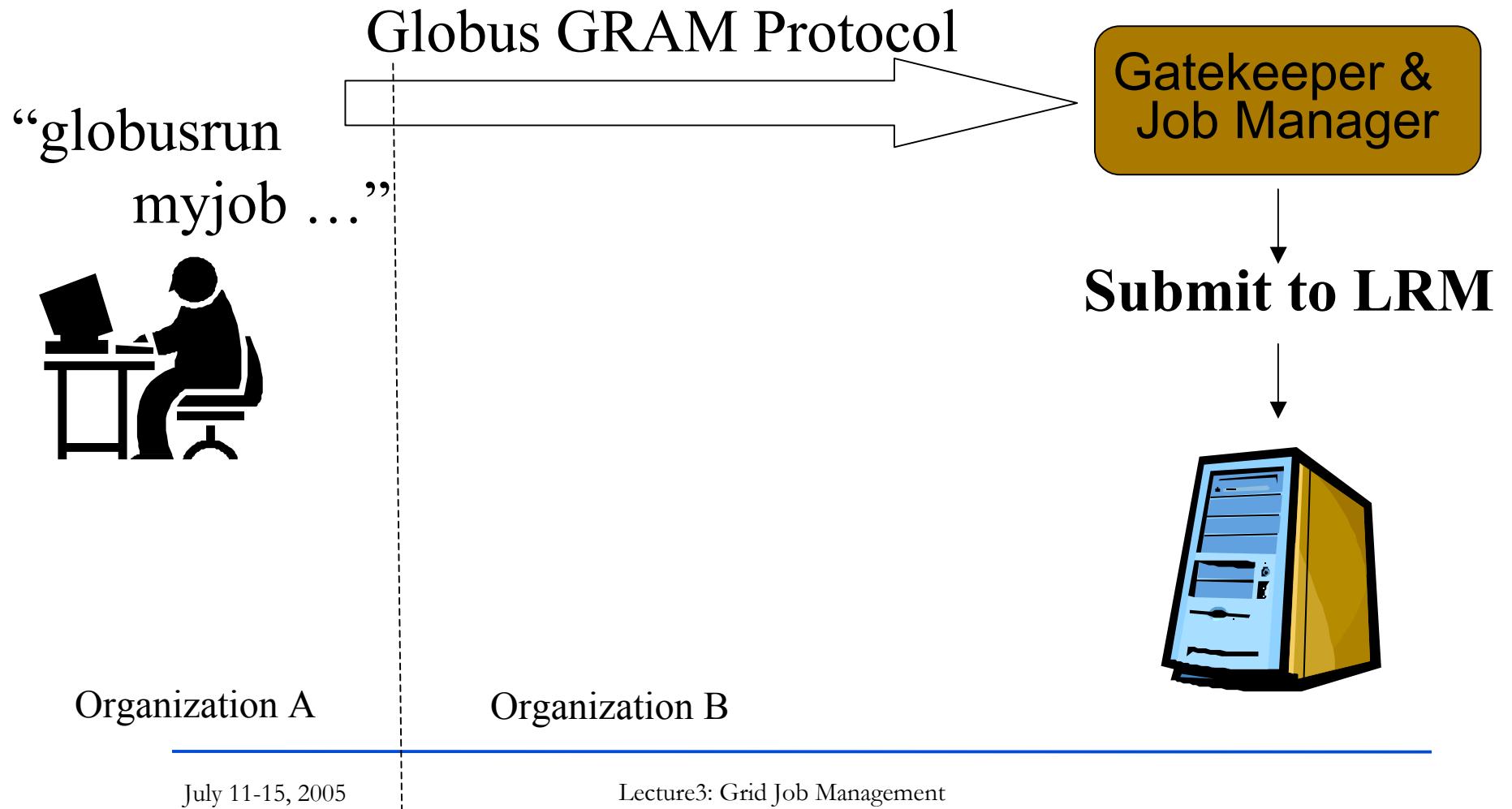
Job and resource management

Workshop Module 2

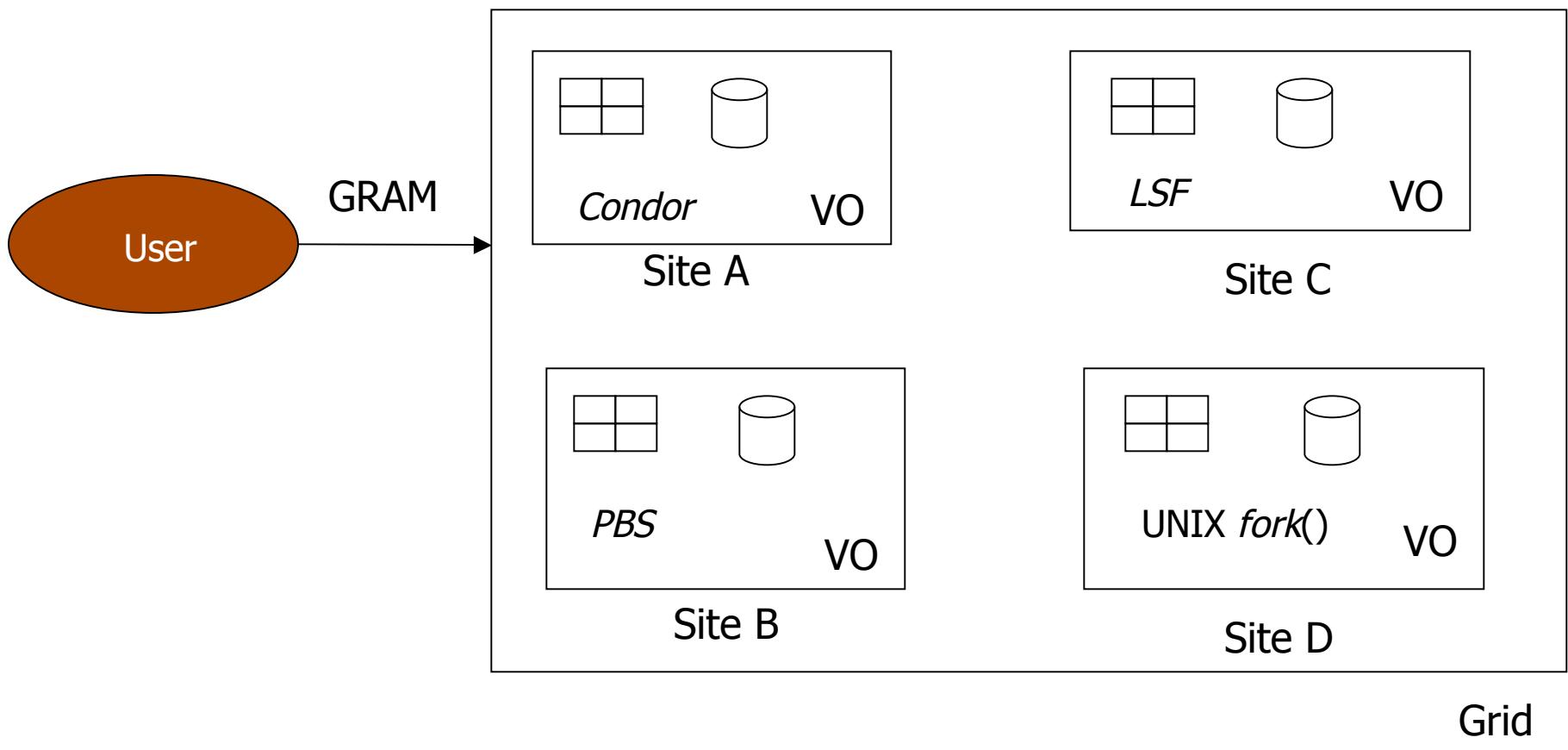
Local Resource Manager: a batch scheduler for running jobs on a computing cluster

- 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”

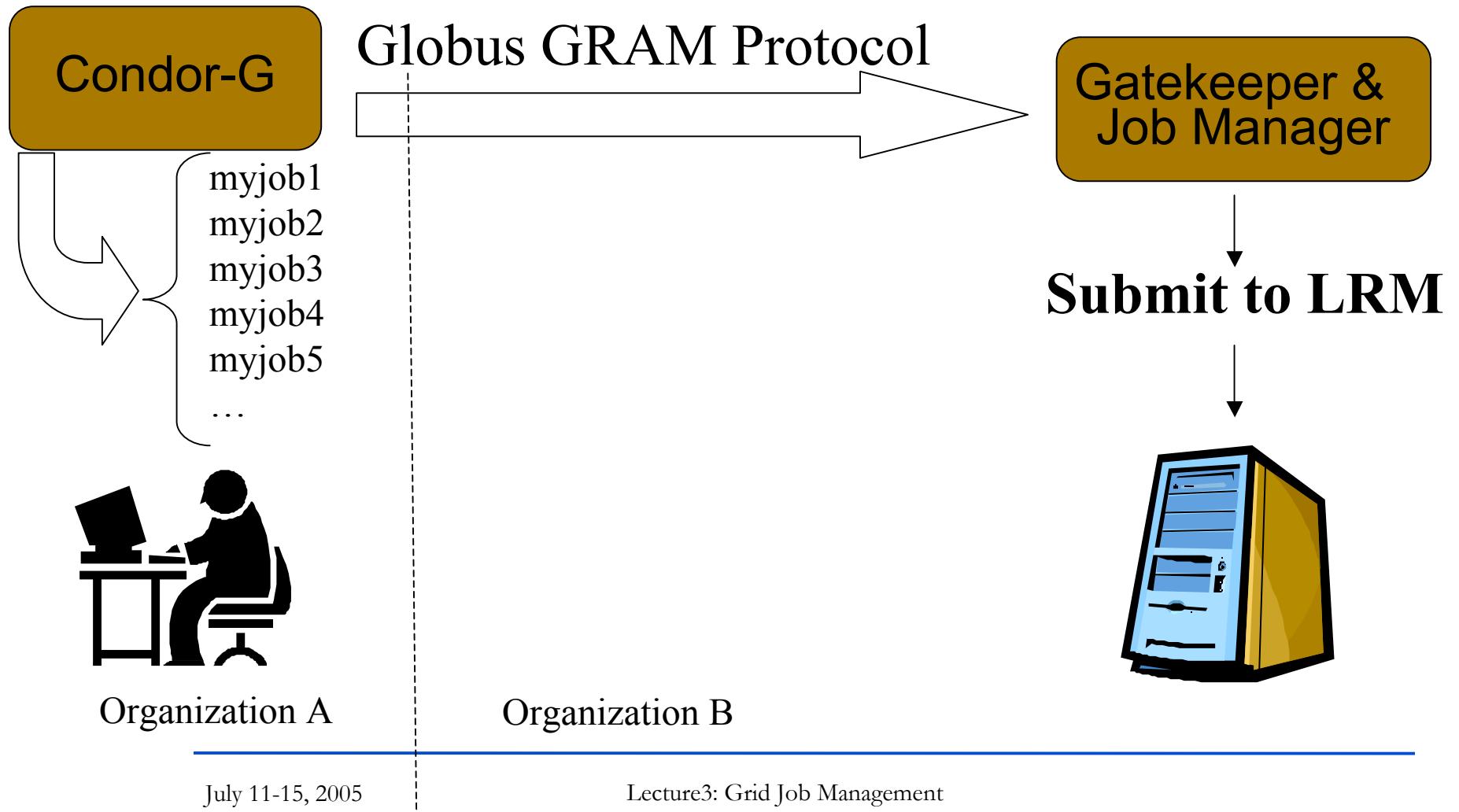
GRAM – Globus Resource Allocation Manager



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
 - Emerging services integrate replication and transport
- 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

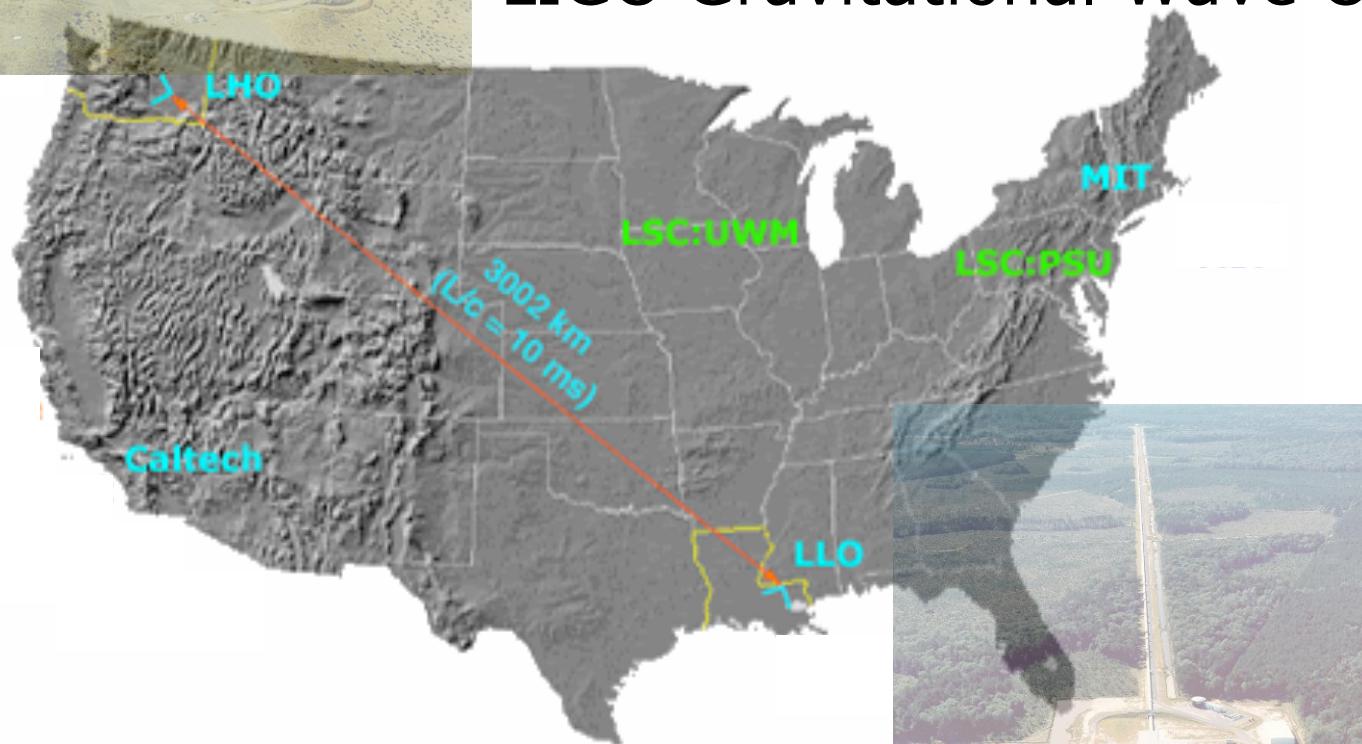
- Requirements

- 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

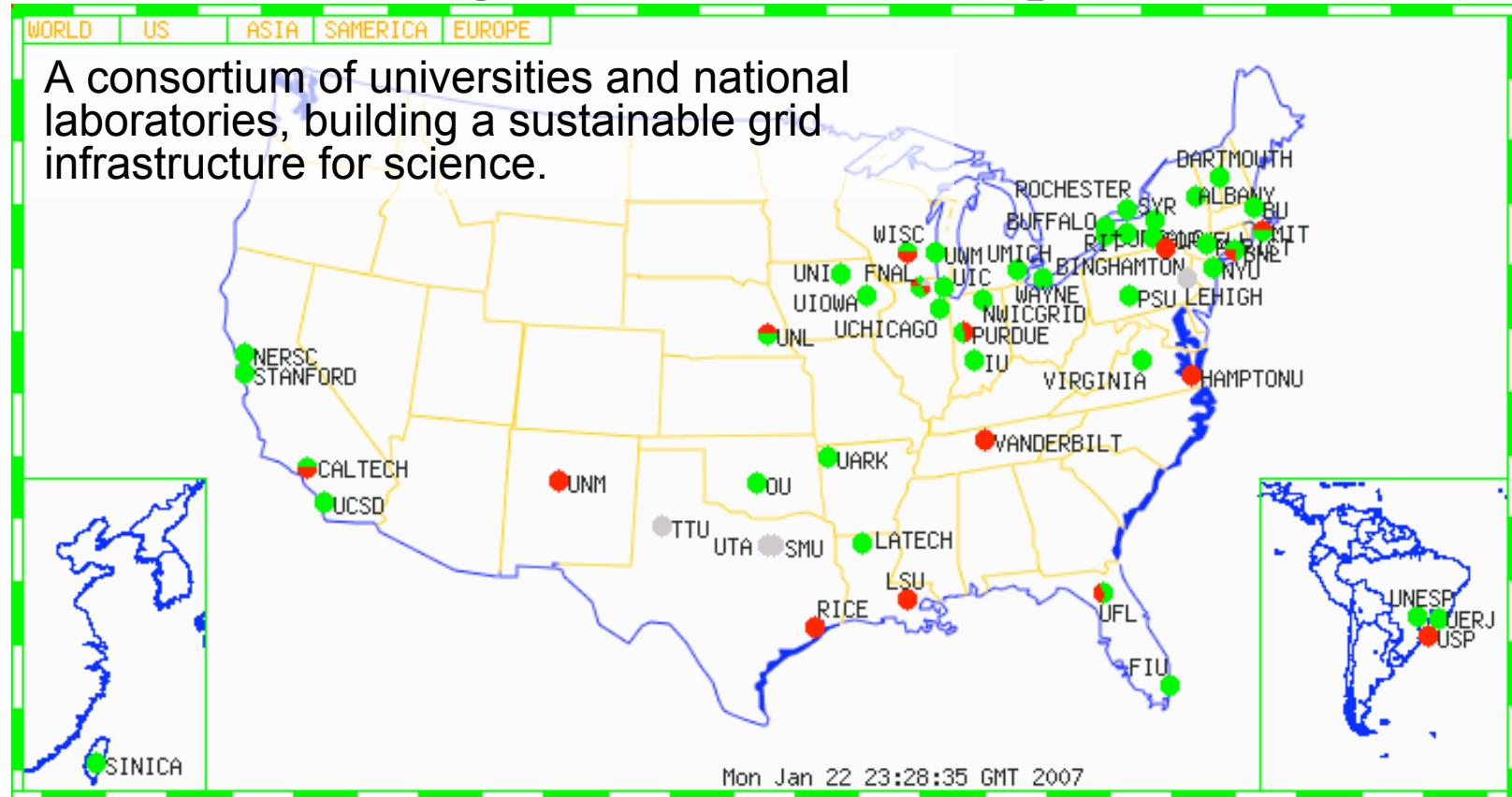
Grid Security Infrastructure - GSI

- 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

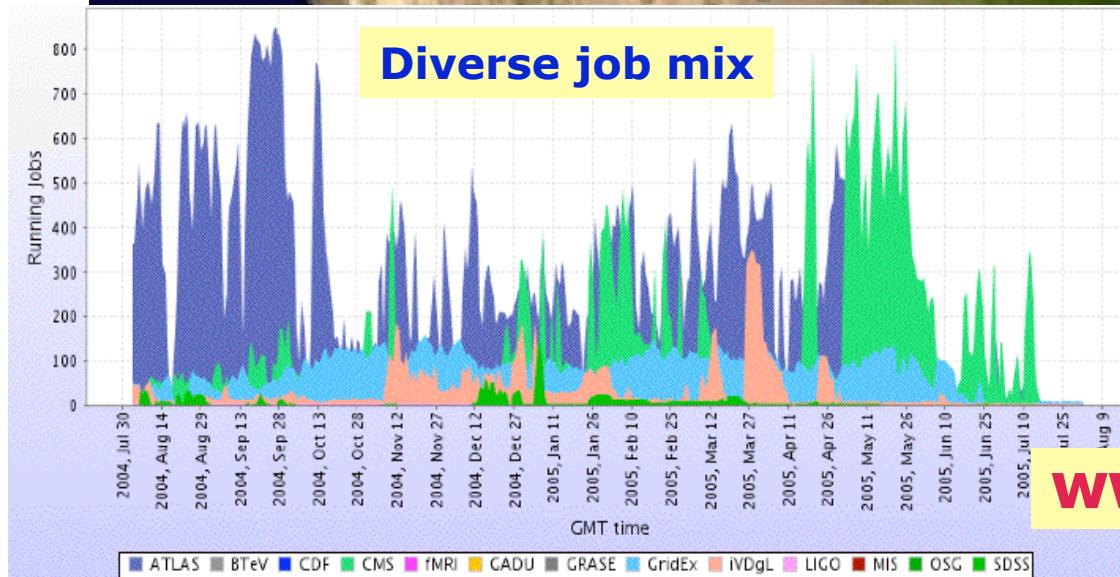
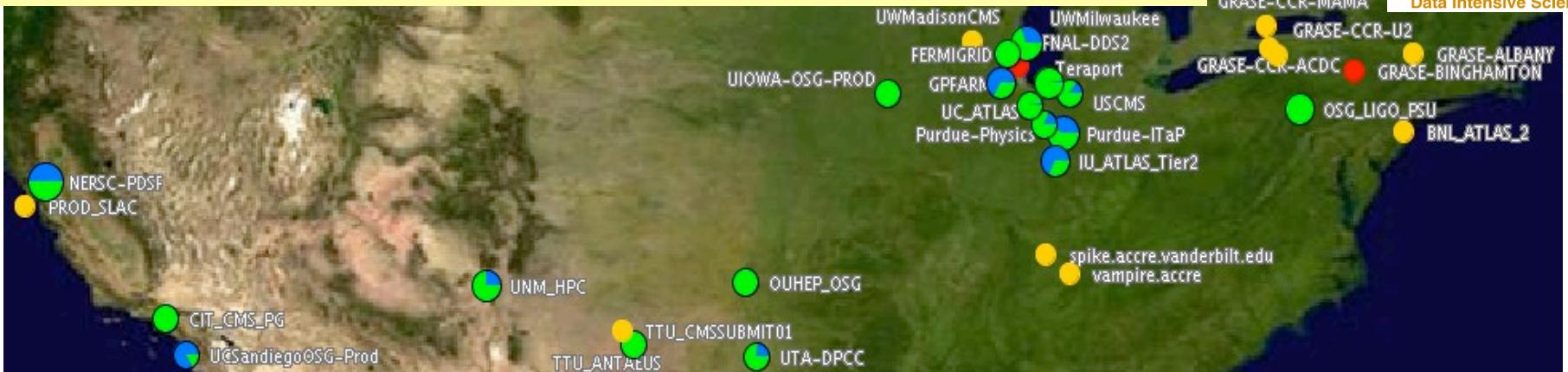
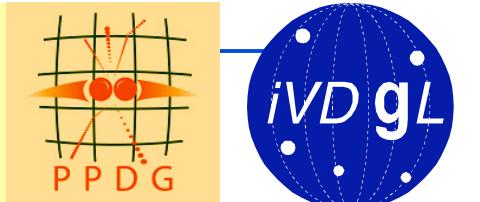


- OSG incorporates advanced networking and focuses on general services, operations, end-to-end performance
- Composed of a large number (>50 and growing) of shared computing facilities, or “sites”

<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



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

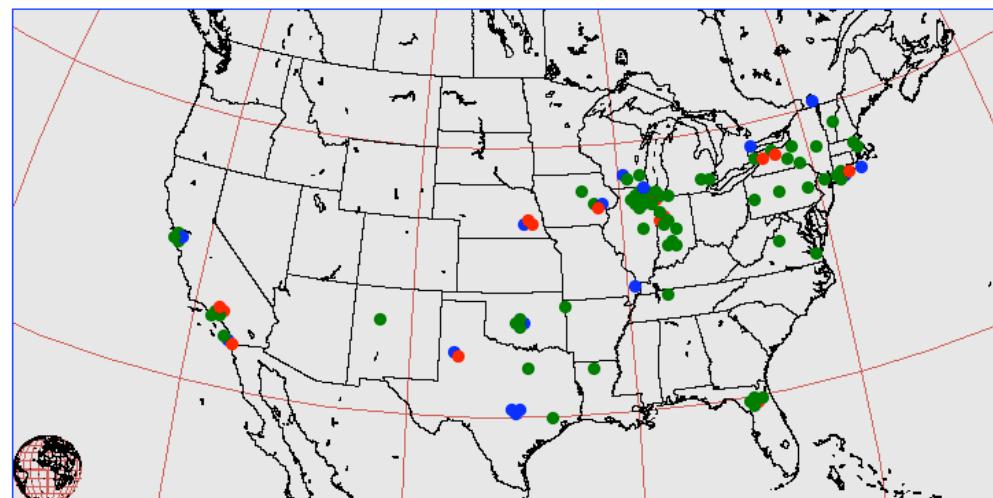


OSG Resource Selection Service: VORS

AllOSGTeraGridEGEEOSG-ITB



Open Science Grid



Virtual Organization Selection

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

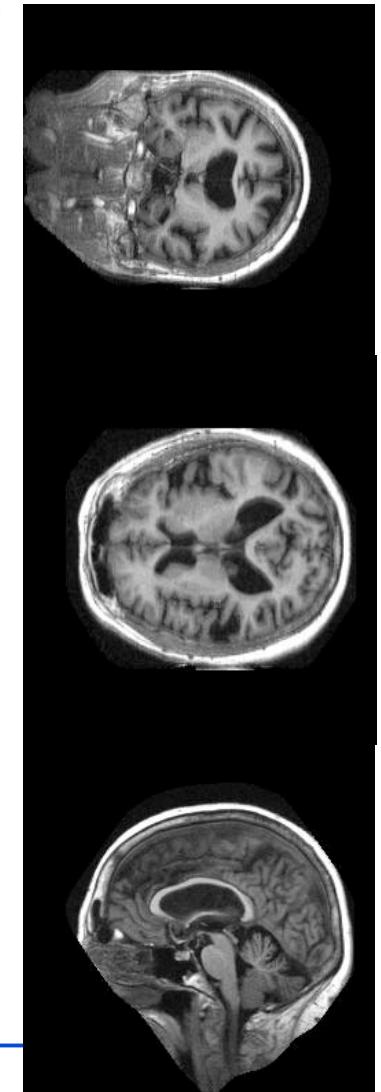
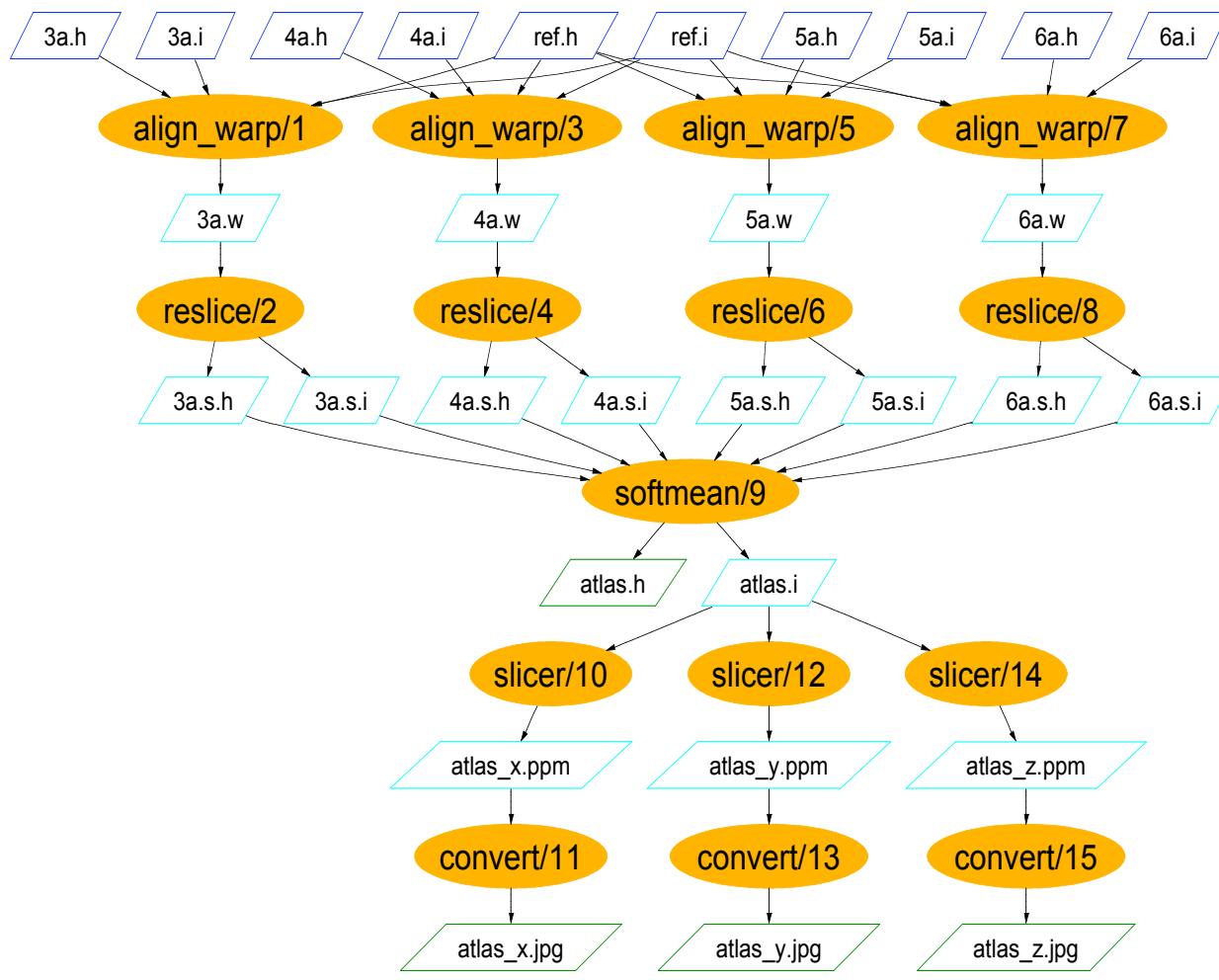
Resources

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