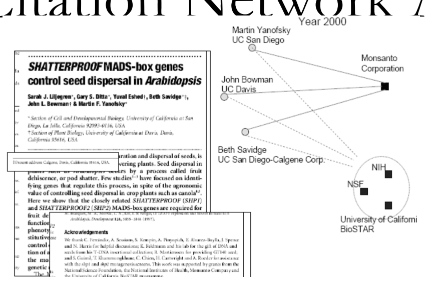
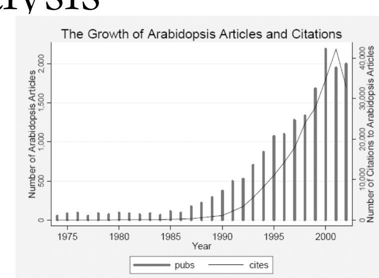
Introduction to Grid Computing

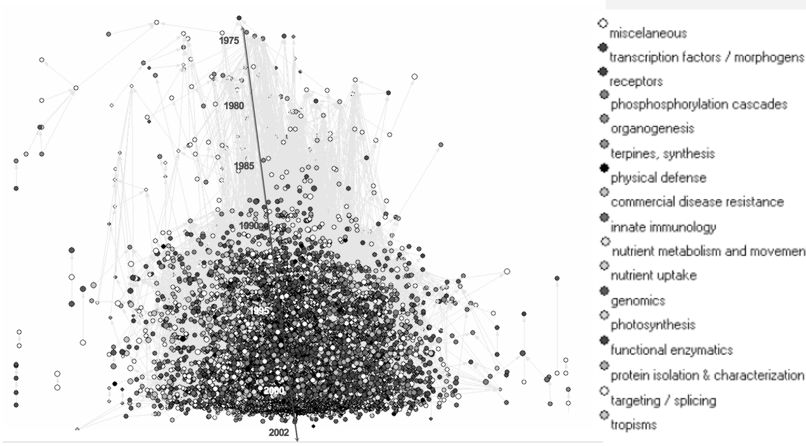
Midwest Grid Workshop Module 1

Open Science Grid

Scaling up Social Science: Citation Network Analysis





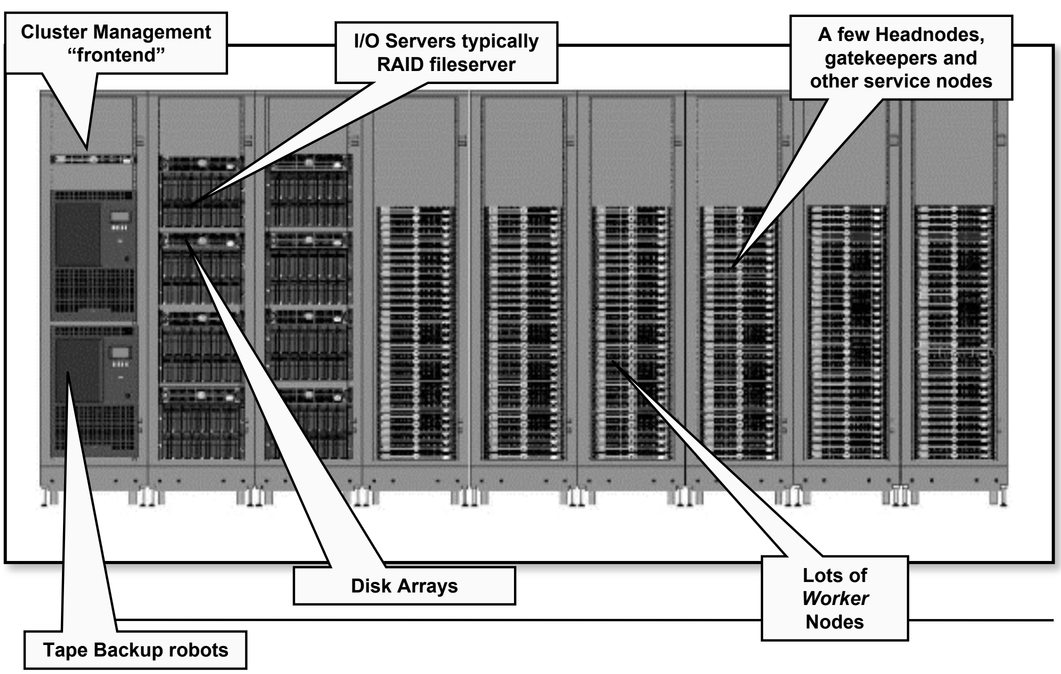


Work of James Evans, University of Chicago, Department of Sociology

Scaling up the analysis

- Database queries of 25+ million citations
- Work started on small workstations
- Queries grew to month-long duration
- With database 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*.

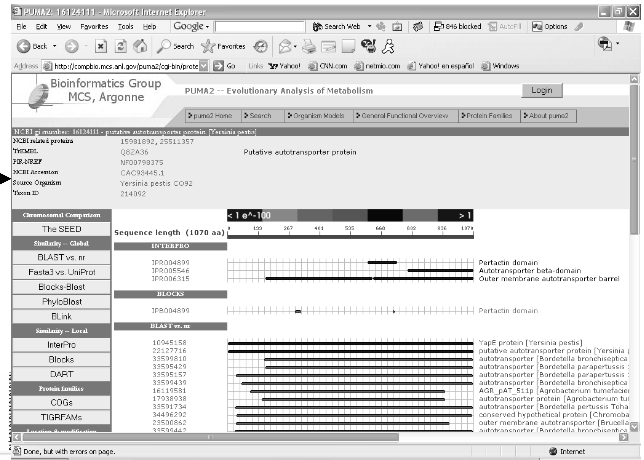
Computing clusters have commoditized supercomputing

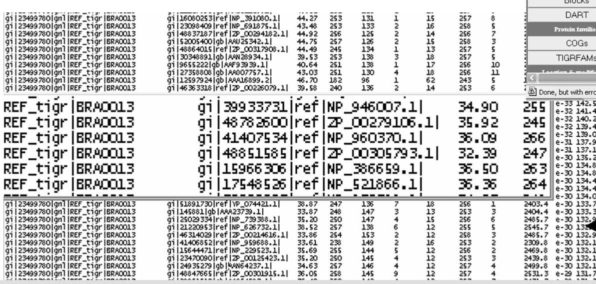


PUMA: Analysis of Metabolism

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

Initial driver: High Energy Physics

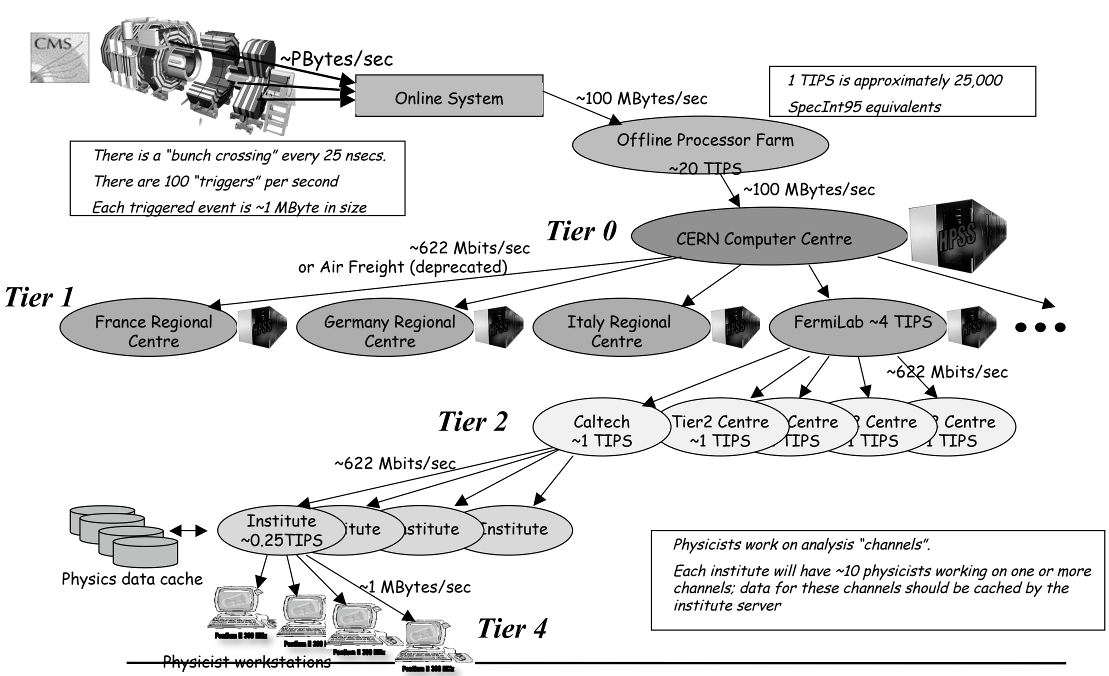
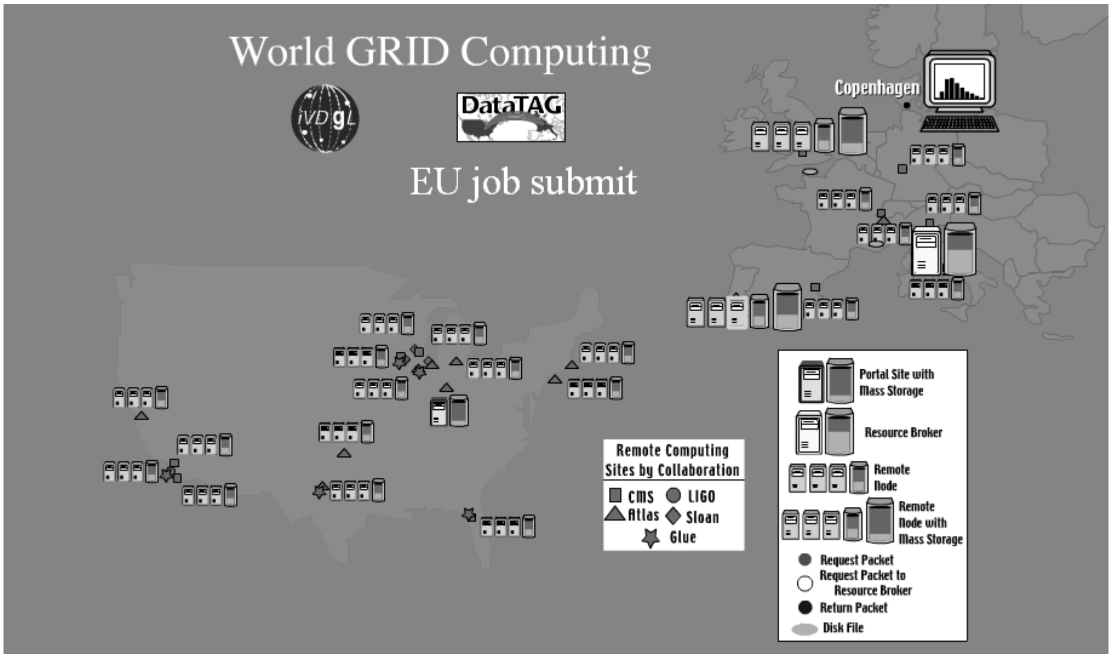


Image courtesy Harvey Newman, Caltech

Grids Provide Global Resources To Enable e-Science



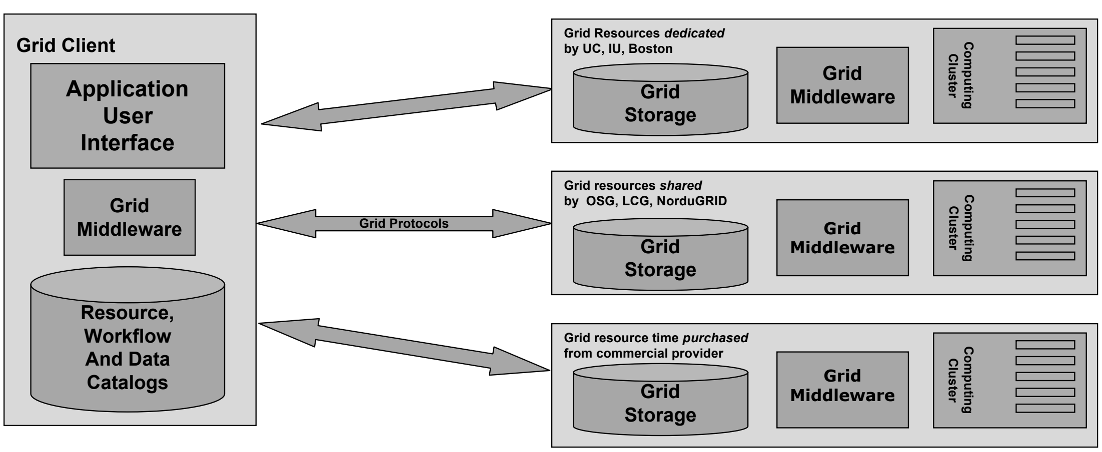
Ian Foster's Grid Checklist

- 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

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.*

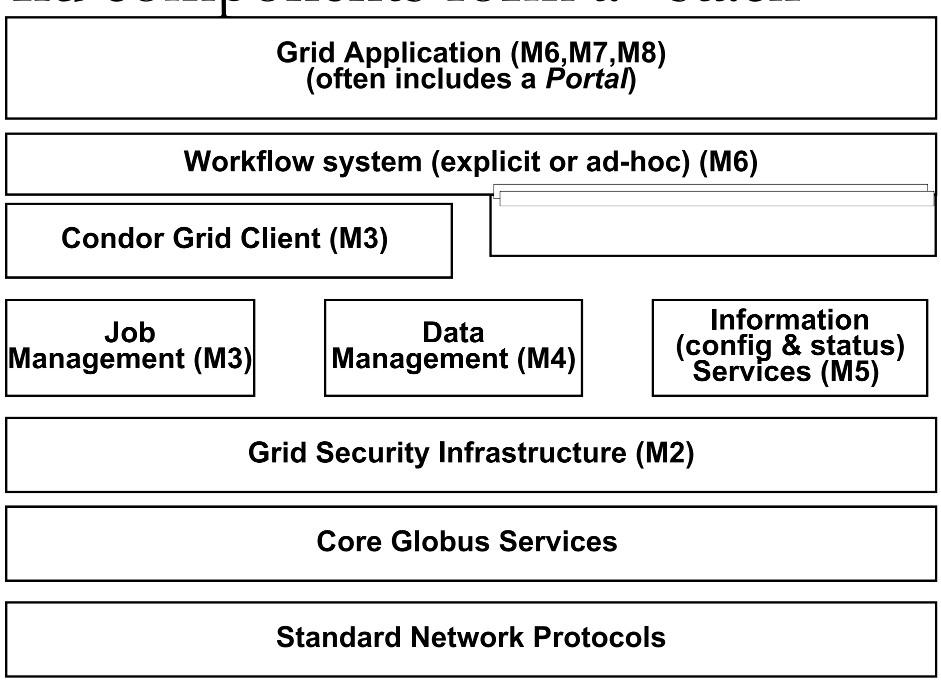


- Security to control access and protect communication (GSI)
- Directory to locate grid sites and services: (VORS, MDS)
- Uniform interface to computing sites (GRAM)
- Facility to maintain and schedule queues of work (Condor-G)
- Fast and secure data set mover (GridFTP, RFT)
- Directory to track where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

We will focus on Globus and Condor

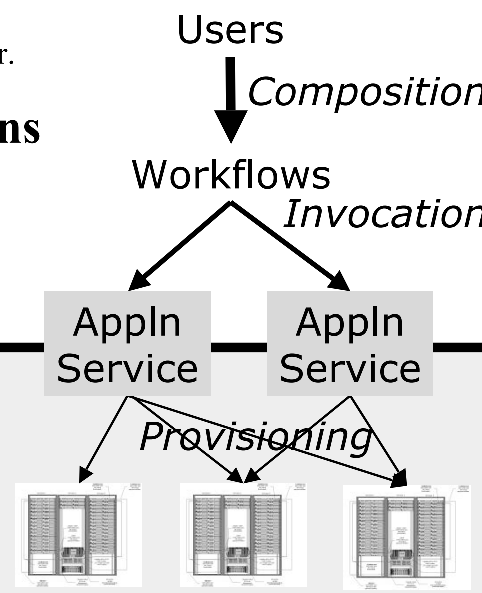
- Condor provides both client & server scheduling
  - In grids, Condor provides an agent to queue, schedule and manage work submission
- Globus Toolkit provides the lower-level middleware
  - Client tools which you can use from a command line
  - APIs (scripting languages, C, C++, Java, …) to build your 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 components form 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

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

Security

Workshop Module 2

Grid security is a crucial component

- Resources are typically valuable
- Problems being solved might be sensitive
- 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

Security Services

- Forms the underlying communication medium for all the services
- Secure *Authentication* and *Authorization*
- Single Sign-on
  - User explicitly authenticates only once – then single sign-on works for all service requests
- Uniform Credentials
- Example: GSI (Grid Security Infrastructure)

Authentication means identifying that you *are* whom you claim to be

- Authentication stops imposters
- *Examples of authentication:*
  - Username and password
  - Passport
  - ID card
  - Public keys or certificates
  - Fingerprint

Authorization controls what you are allowed to *do*.

- Is this device allowed to access to this service?
- Read, write, execute permissions in Unix
- Access conrol lists (ACLs) provide more flexible control
- Special "callouts" in the grid stack in job and data management perform authorization checks.

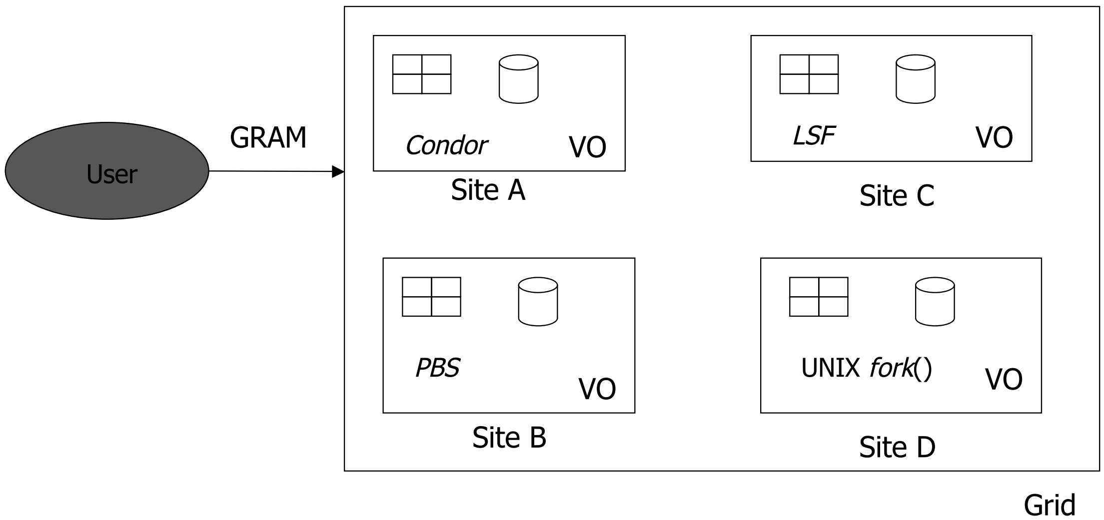
Job and resource management

Workshop Module 3

Job Management Services provide a standard interface to remote resources

- Includes CPU, Storage and Bandwidth
- Globus component is *Globus Resource Allocation Manager (GRAM)*
- The primary Condor grid *client* component is *Condor-G*
- Other needs:
  - scheduling
  - monitoring
  - job migration
  - notification

GRAM provides a uniform interface to diverse resource scheduling systems.



GRAM: What is it?

- *Globus Resource Allocation Manager*
- Given a job specification:
  - Create an environment for a job
  - Stage files to and from the environment
  - Submit a job to a local resource manager
  - Monitor a job
  - Send notifications of the job state change
  - Stream a job's stdout/err during execution

A “Local Resource Manager” is a batch system for running jobs across a computing cluster

- In GRAM
- *Examples:*
  - Condor
  - PBS
  - LSF
  - Sun Grid Engine
- Most systems allow you to access “fork”
  - Default behavior
  - It runs on the gatekeeper:
    - A bad idea in general, but okay for testing

Managing your jobs

- We need something more than just the basic functionality of the globus job submission commands
- Some desired features
  - Job tracking
  - Submission of a set of inter-dependant jobs
  - Check-pointing and Job resubmission capability
  - Matchmaking for selecting appropriate resource for executing the job
- *Options:* Condor, PBS, LSF, …

Data Management

Workshop Module 4

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

- Fast
- Flexible
- Secure
- Ubiquitous
- Grids are used for analyzing and manipulating large amounts of data
  - Metadata (data about data): *What is the data?*
  - Data location: *Where is the data?*
  - Data transport: *How to move the data?*

GridFTP is a secure, efficient and standards-based data transfer protocol

- Robust, fast and widely accepted
- Globus GridFTP server
- Globus *globus-url-copy* GridFTP client
- Other clients exist (e.g., *uberftp*)

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)
- Not all features are accessible from basic client
  - Can be embedded in higher level and application-specific frameworks

File catalogues tell you where the data is

- Replica Location Service (RLS)
- Phedex
- RefDB / PupDB

Requirements from a File Catalogue

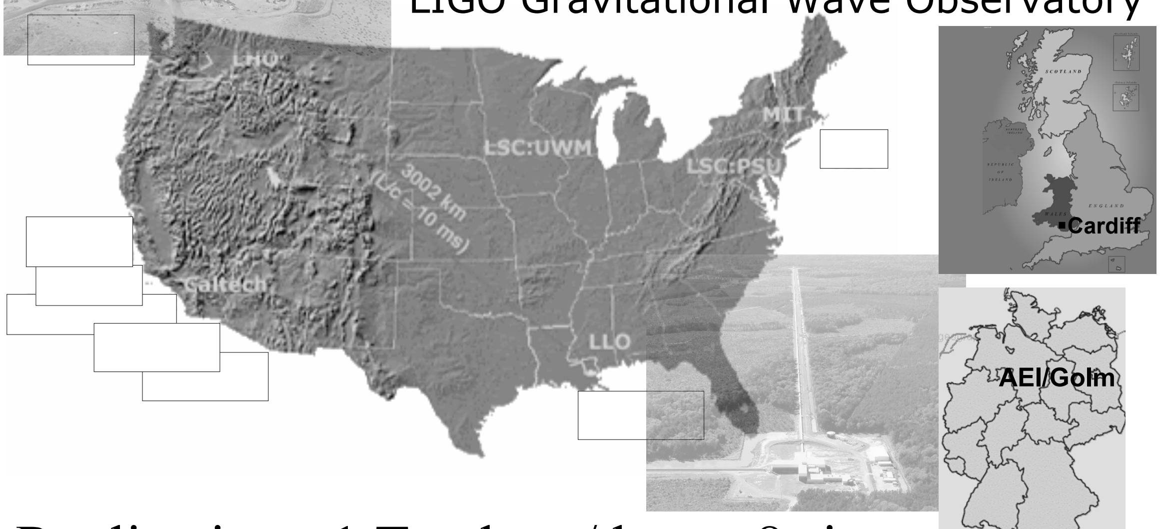
- Abstract out the logical file name (LFN) for a physical file
  - maintain the mappings between the LFNs and the PFNs *(physical file names)*
- Maintain the location information of a file

In order to avoid *"hotspots"*, replicate data files in more than one location

- Effective use of the grid resources
- Each LFN can have more than 1 PFN
- Avoids single point of failure
- Manual or automatic replication
  - Automatic replication considers the demand for a file, transfer bandwidth, etc.

The Globus-Based LIGO Data Grid

LIGO Gravitational Wave Observatory



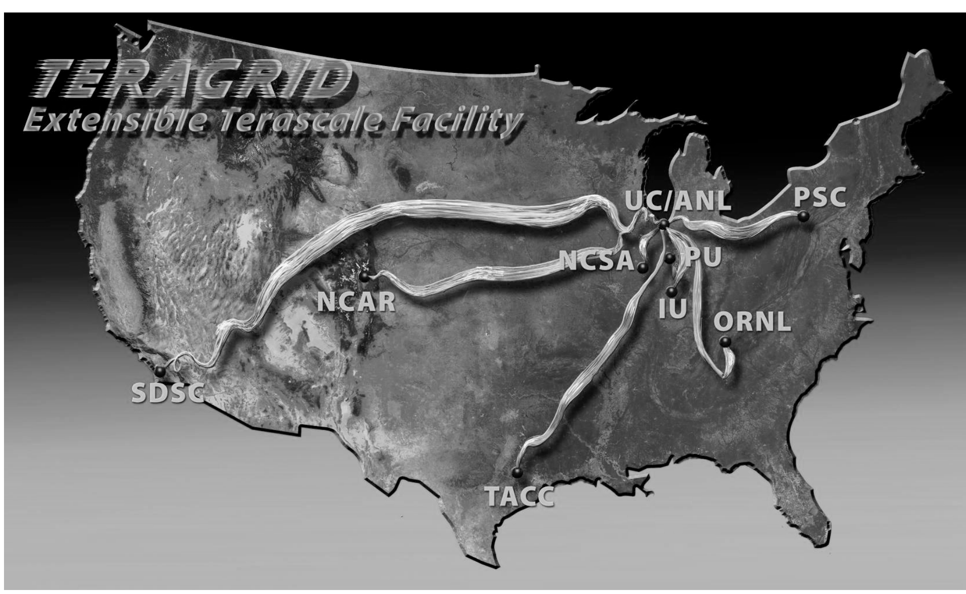
Replicating >1 Terabyte/day to 8 sites

>40 million repli[redacted]

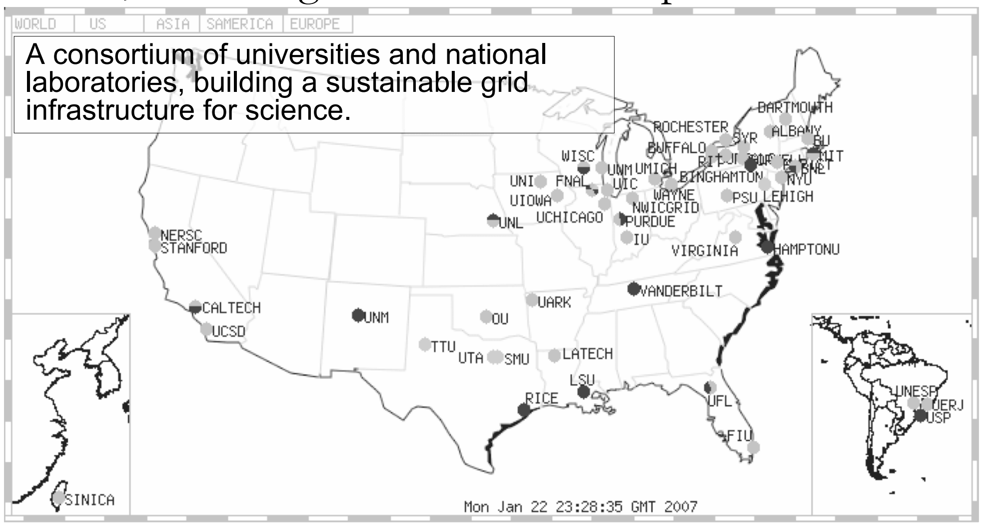
National Grid Cyberinfrastructure

Workshop Module 5

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



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

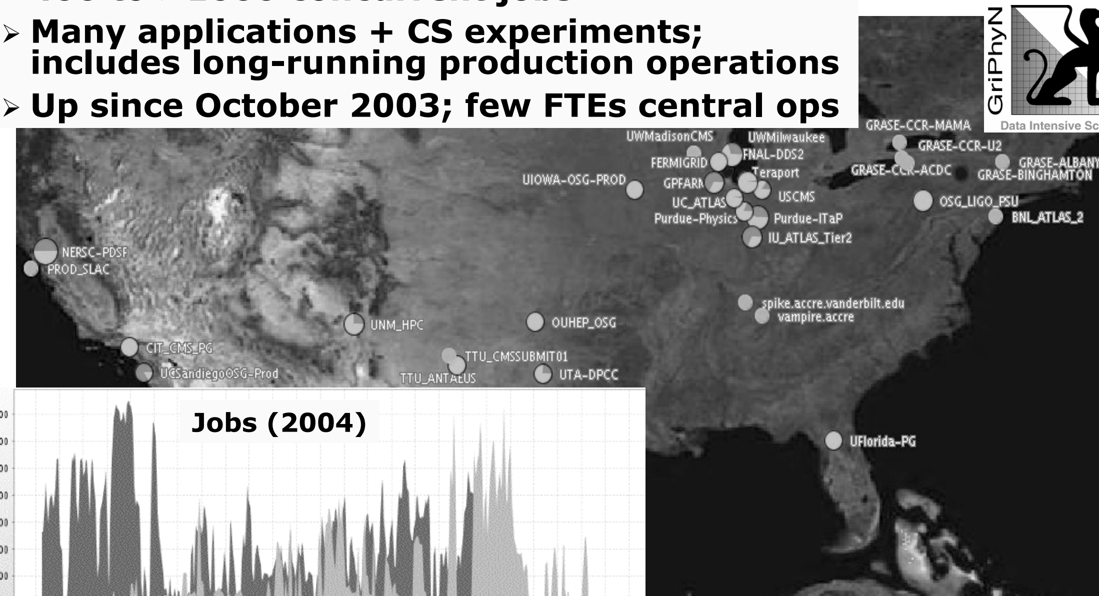
- **50 sites (15,000 CPUs) & growing**
- **400 to >1000 concurrent jobs**
- **Many applications + CS experiments; includes long-running production operations**
- **Up since October 2003; few FTEs central ops**

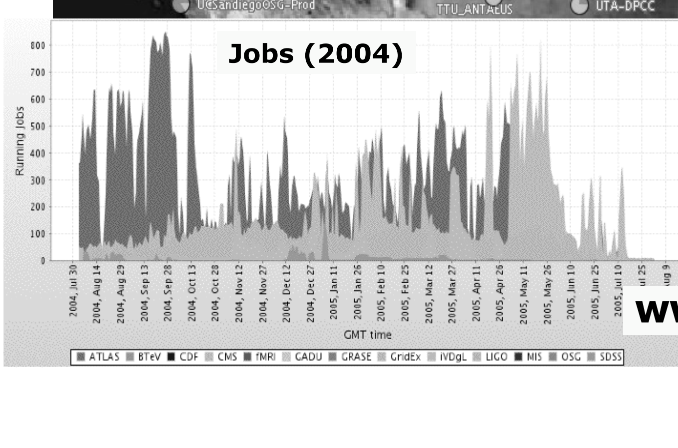
PPDG

iVDgL

GriPhyN

Data Intensive Science





www.opensciencegrid.org

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

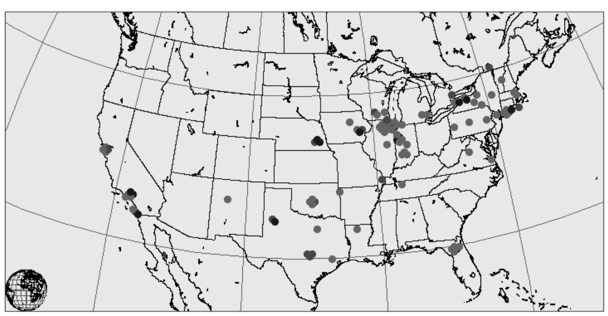
- Check the availability of different grid sites
- Discover different grid services
- Check the status of “jobs”
- Make better scheduling decisions with information maintained on the “health” of sites

OSG Resource Selection Service: VORS

Virtual Organization Resource Selector

All	OSG	TeraGrid	EGEE	OSG-ITB

Open Science Grid



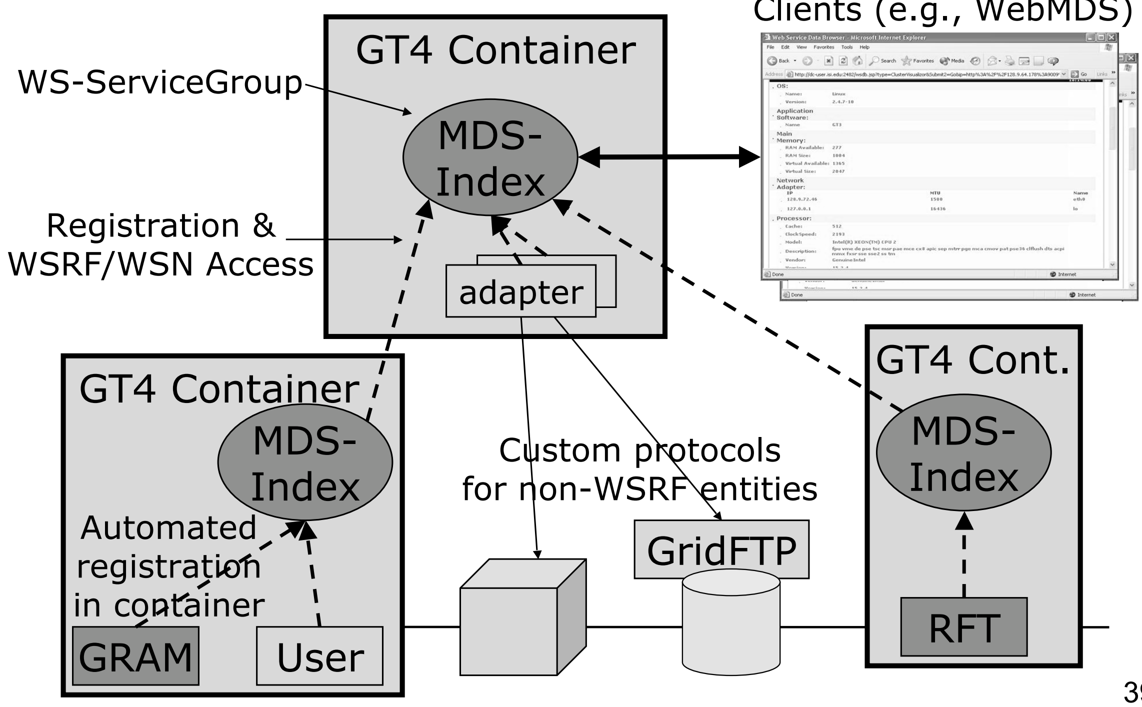
Virtual Organization Selection

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

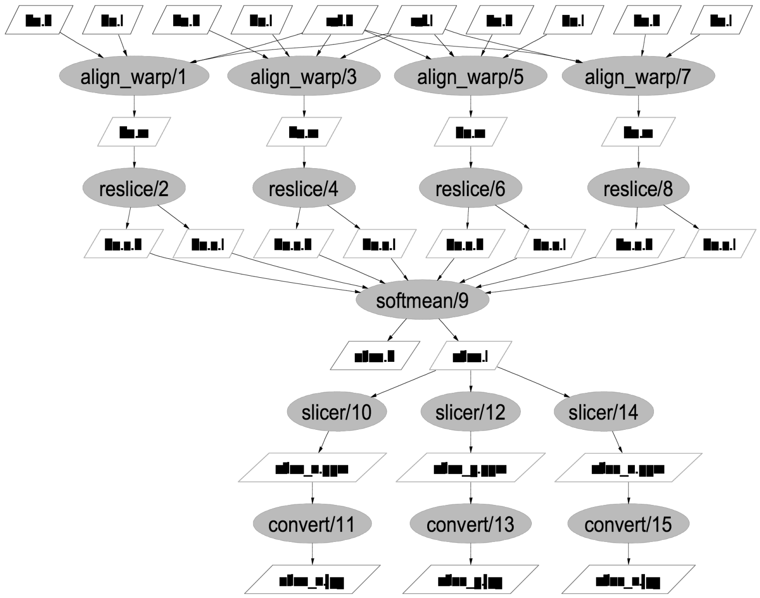
Monitoring and Discovery Service - MDS

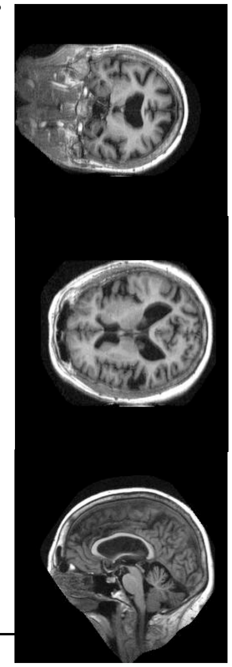


Grid Workflow

Workshop Module 6

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



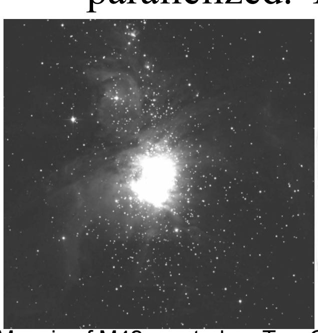


Workflow courtesy James Dobson, Dartmouth Brain Imaging Center

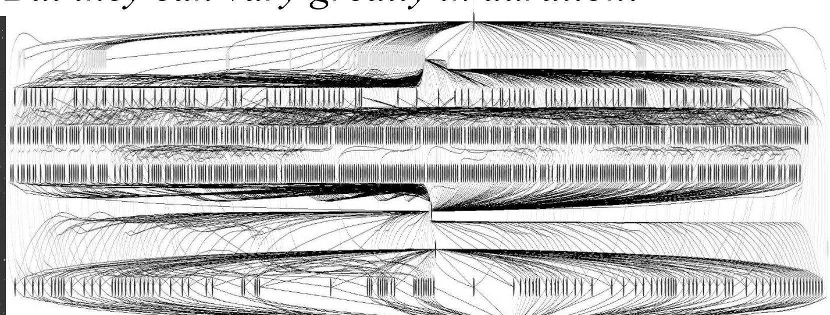
41

Workflows 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
- Transformations are usually independent and can be parallelized. *But they can vary greatly in duration.*



Mosaic of M42 created on TeraGrid



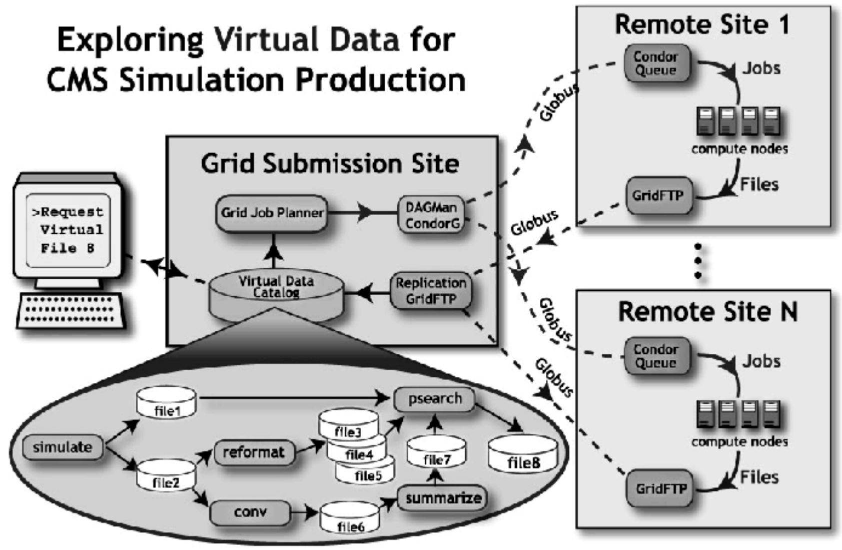
= Data Transfer

= Compute Job

Montage Workflow: ~1200 jobs, 7 levels

42

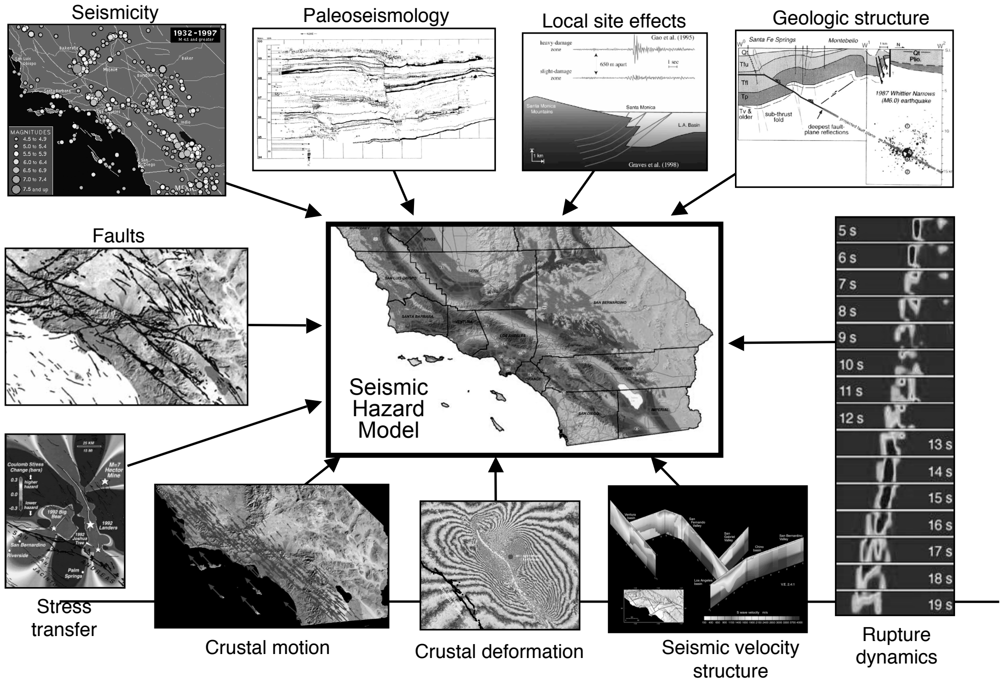
Virtual data model enables workflow to abstract grid details.



An important application pattern: Grid Data Mining

Workshop Module 7

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



Data mining in the grid

- **Decompose** across network
- Clients **integrate** dynamically
  - Select & compose services
  - Select "best of breed" providers
  - Publish result as a new service
- Decouple **resource** & **service** providers

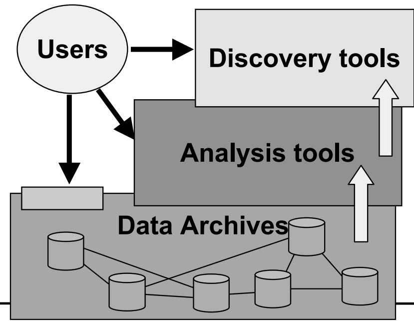


Fig: S. G. Djorgovski

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 Takes a Village …

- 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