

# Data Challenges on the TeraGrid

Daniel S. Katz  
[d.katz@ieee.org](mailto:d.katz@ieee.org)

TeraGrid GIG Director of Science

Senior Grid/HPC Researcher,  
Computation Institute  
University of Chicago & Argonne National Laboratory

Affiliate Faculty,  
Center for Computation & Technology, LSU

Adjunct Associate Professor,  
Electrical and Computer Engineering Department, LSU



# What is the TeraGrid?

- World's largest open scientific discovery infrastructure
- Leadership class resources at eleven partner sites combined to create an integrated, persistent computational resource
  - High-performance networks
  - High-performance computers (>1 Pflops (~100,000 cores) -> 1.75 Pflops)
    - And a Condor pool (w/ ~13,000 CPUs)
  - Visualization systems
  - Data Storage
    - Data collections (>30 PB, >100 discipline-specific databases)
    - Archival storage (10-15 PB current stored)
  - Science Gateways
  - User portal
  - User services - Help desk, training, advanced app support
- Allocated to US researchers and their collaborators through national peer-review process
  - Generally, review of computing, not science
- Extremely user-driven
  - MPI jobs, ssh or grid (GRAM) access, etc.

# TeraGrid Governance

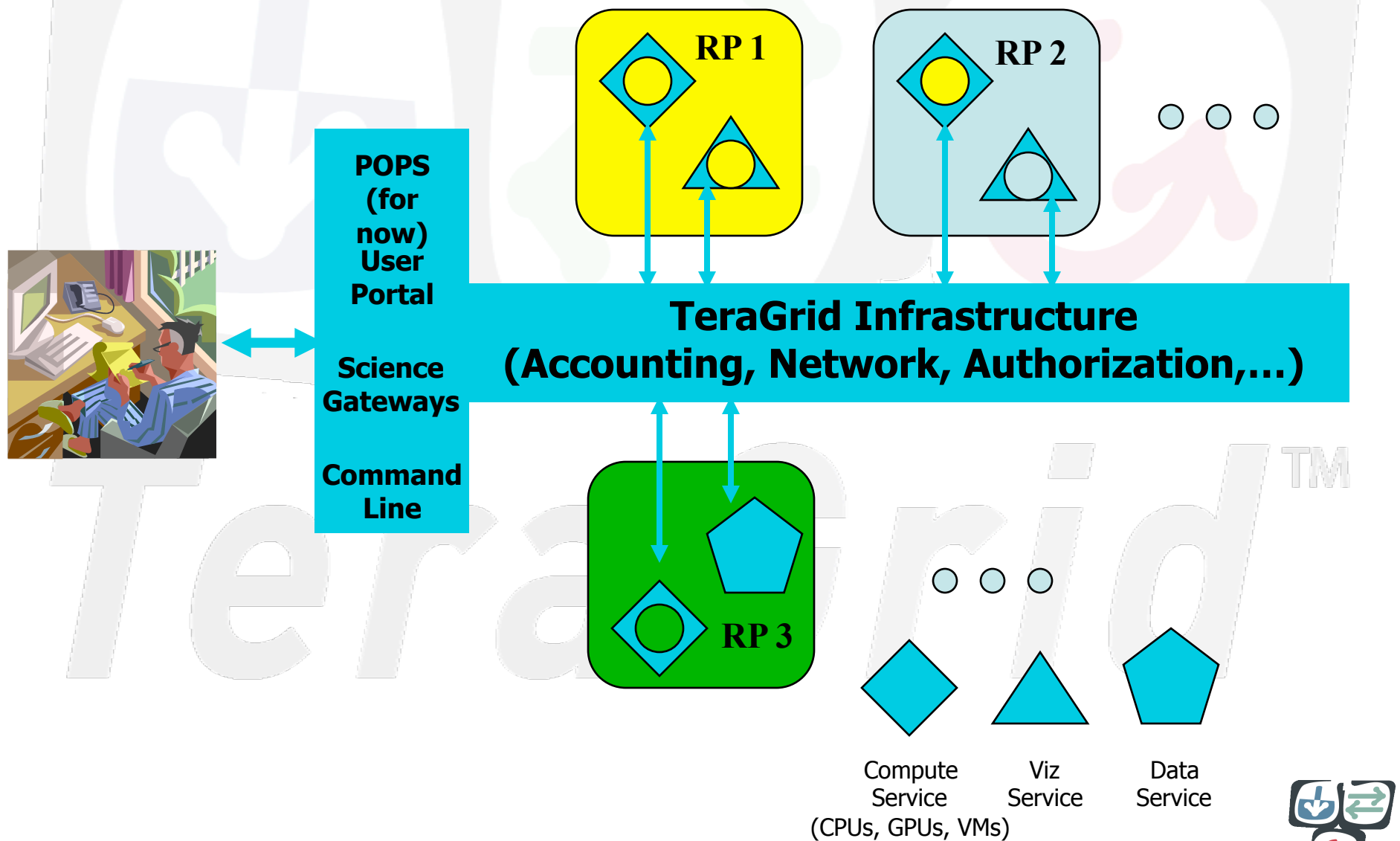
- 11 Resource Providers (RPs) funded under individual agreements with NSF
  - Mostly different: start and end dates, goals, and funding models
- 1 Coordinating Body – Grid Integration Group (GIG)
  - University of Chicago/Argonne
  - Subcontracts to all RPs and six other universities
  - ~10 Area Directors, lead coordinated work across TG
  - ~18 Working groups with members from many RPs work on day-to-day issues
  - RATs formed to handle short-term issues
- TeraGrid Forum with Chair sets policies and is responsible for the overall TeraGrid

# TeraGrid





# How One Uses TeraGrid



# TG New Large Resources

- **Ranger@TACC**

- NSF 'Track2a' HPC system
- 504 TF
- 15,744 Quad-Core AMD Opteron processors
- 123 TB memory, 1.7 PB disk

- **Kraken@NICS (UT/ORNL)**

- NSF 'Track2b' HPC system
- 170 TF Cray XT4 system
- To be upgraded to Cray XT5 at 1 PF
  - 10,000+ compute sockets
  - 100 TB memory, 2.3 PB disk

- **Something@PSC**

- NSF 'Track2c' HPC system

- **FlashGordon@SDSC**

- First NSF 'Track2d' system – for data intensive computing

6 – More 'Track2d' systems to be announced



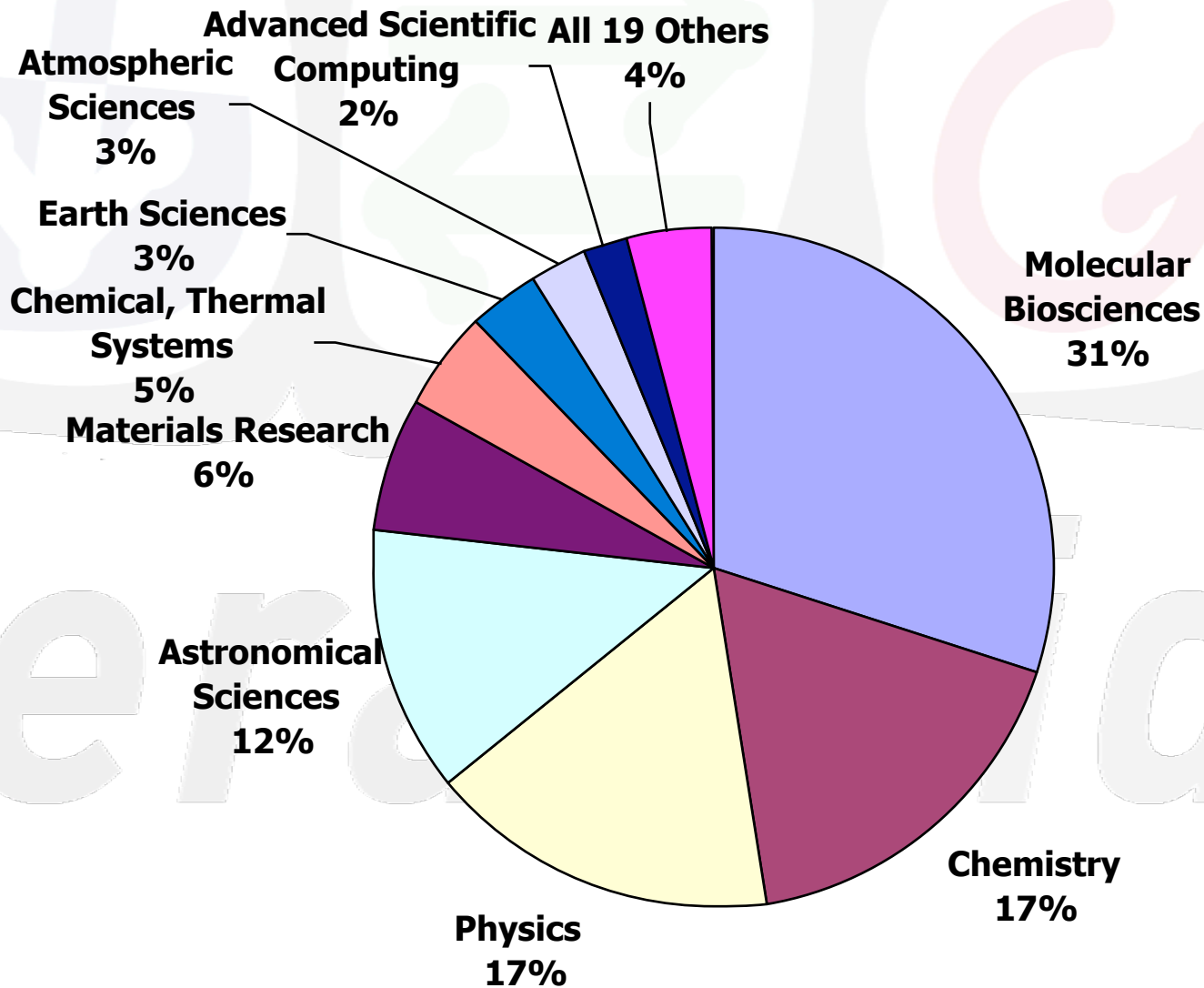
Blue Waters@NCSA  
NSF Track 1  
10 PF peak  
Coming in 2011

# How TeraGrid Is Used

Use Modality	Community Size (rough est. - number of users)
Batch Computing on Individual Resources	850
Exploratory and Application Porting	650
Workflow, Ensemble, and Parameter Sweep	250
Science Gateway Access	500
Remote Interactive Steering and Visualization	35
Tightly-Coupled Distributed Computation	10

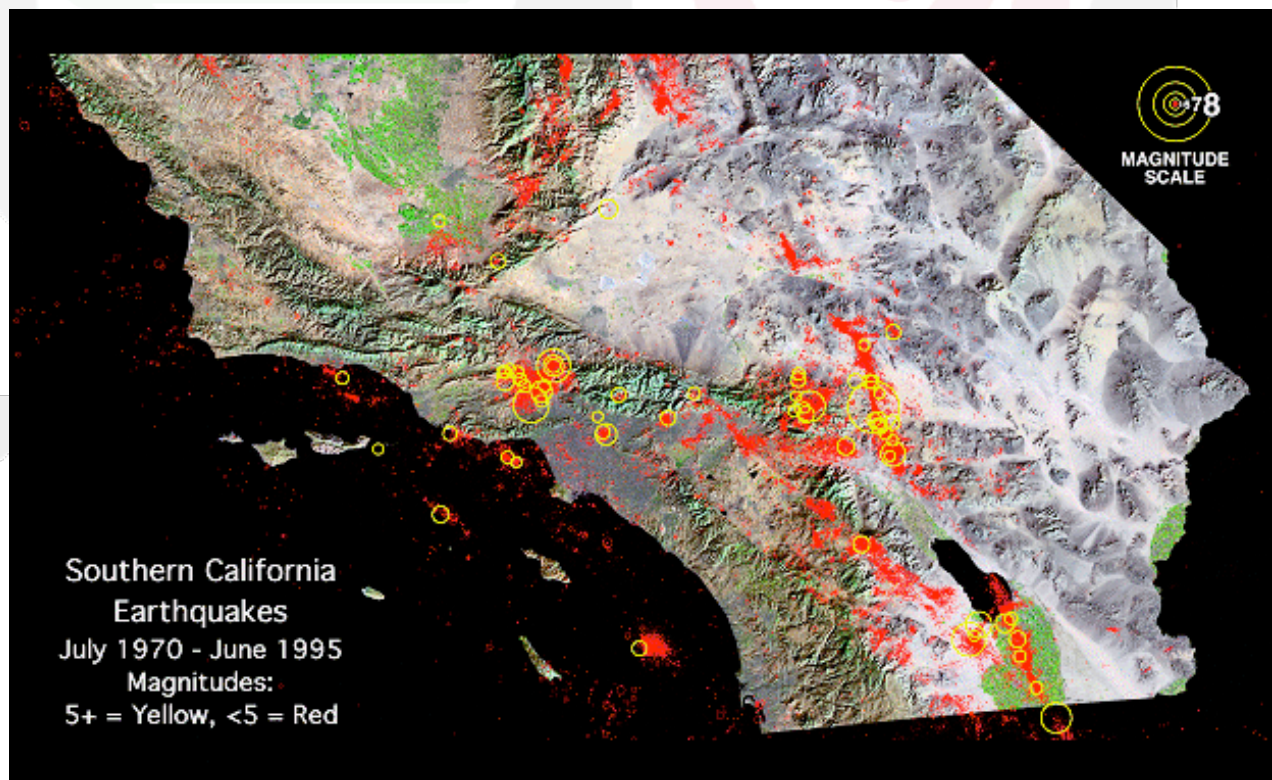


# Who Uses TeraGrid (2007)



# Two SCEC Projects

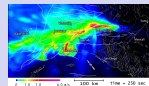
- SCEC: Southern California Earthquake Center
- PI: Tom Jordan, USC
- PetaShake:  
extend  
deterministic  
simulations of  
strong ground  
motions to 3 Hz
- CyberShake:  
compute physics-  
based probabilistic  
seismic hazard  
attenuation maps





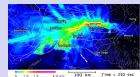
# SCEC PetaShake Simulations By Name, Maximum Frequency and Source Description

## TeraShake 1.x



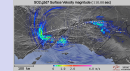
Earthquake wave propagation simulations of Mw7.7 earthquakes with max frequency of 0.5Hz run in 2004-2005 using kinematic source based on 2002 Denali earthquake rupture

## TeraShake 2.x



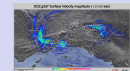
Earthquake wave propagation simulations of Mw7.7 earthquakes with max frequency of 0.5Hz run in 2006 using source descriptions generated by dynamic rupture simulations based on Landers initial stress conditions

## ShakeOut-K



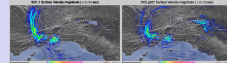
Earthquake wave propagation simulations of Mw7.8 earthquakes with max frequency of 1.0Hz run in 2007 using kinematic source based on geological observations

## ShakeOut-D



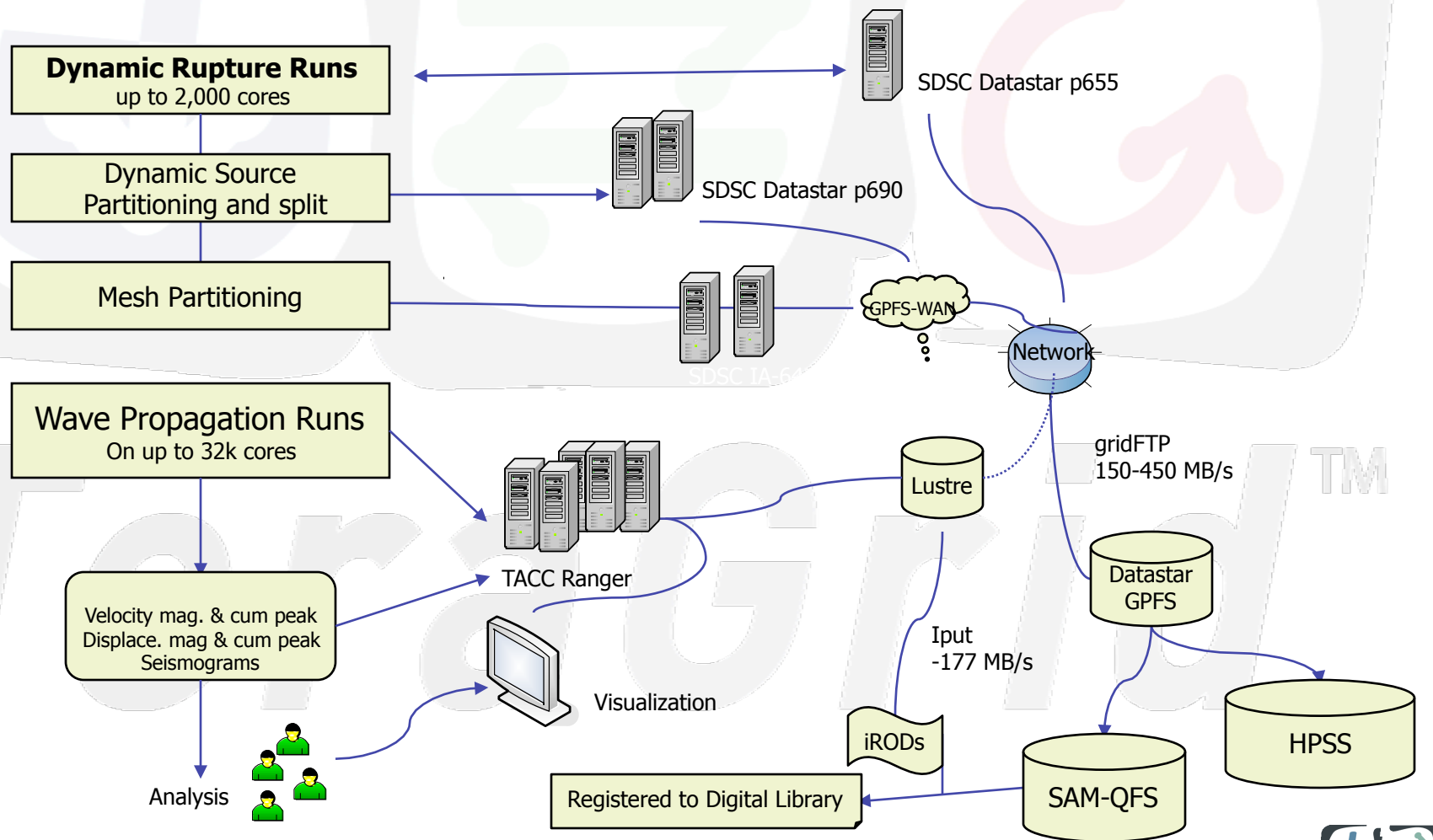
Earthquake wave propagation simulations of Mw7.8 earthquakes with max frequency of 1.0Hz run in 2008 using source descriptions generated by dynamic rupture simulations (SGSN). The SGSN Dynamic rupture simulations were constructed to produce final surface slip equivalent to the ShakOut-K

SO-K vs SO-D



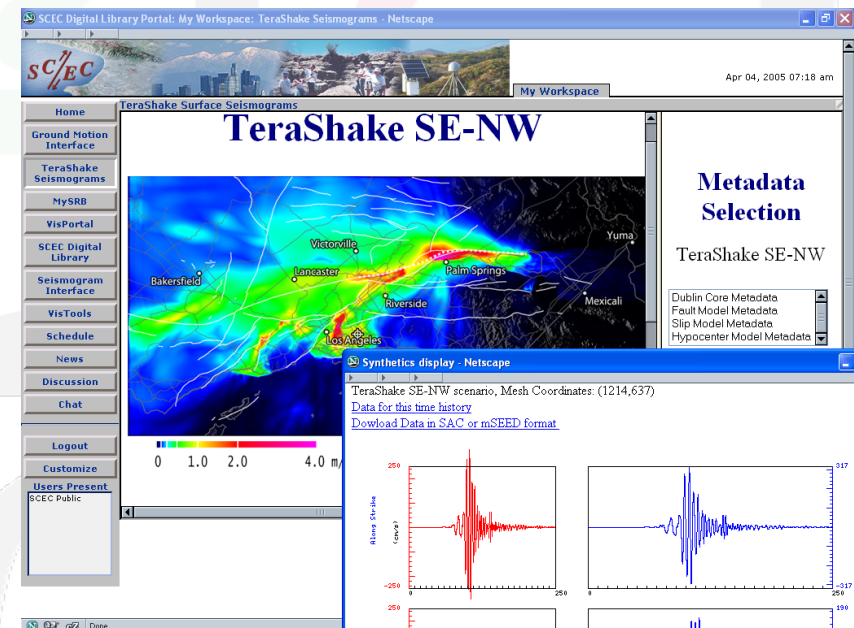
Olsen et al. SEG'08

# ShakeOut-D: Coordinated Executions on TeraGrid



# Data Transfer, Archive and Management

- Input/output data transfer between SDSC disk/HPSS to Ranger disk at the transfer rate up to 450 MB/s using GridFTP
- 90k – 120k files per simulation, 150 TBs generated on Ranger, organized as a separate sub-collection in iRODs, direct data transfer using iRODs from Ranger to SDSC SAM-QFS at 177MB/s using our data ingestion tool
- More than 200 TBs sub-collections published through digital library
- Integrated through SCEC portal into seismic-oriented interaction environments



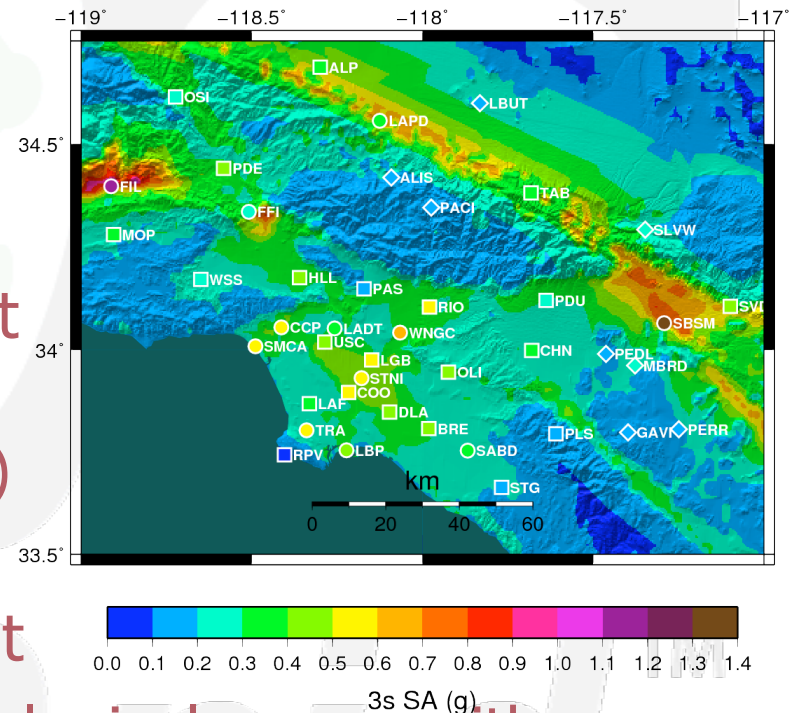
# Challenges of enabling large scale earthquake simulations

- Short-term disk storage for Tera3D project alone will be more than 600 TB this fiscal year
- Consider management of 100s TB data: bit-error-rates on the order of  $10^{-14}$ , 100 TB collection will inherently contain multiple errors, plus other source of corruption such as media failure etc.
- Automate administrative tasks huge challenge such as replication, distribution, access controls, metadata extraction. Data virtualization and grid technology to be integrated. With iRODS, for example, can write rules to track administrative functions such as integrity monitoring
  - provide logical name space so the data can be moved without the access name changing
  - provide metadata to support discovery of files and track provenance
  - provide rules to automate administrative tasks (authenticity, integrity, distribution, replication checks)
  - provide micro-services for parsing data sets (HDF5 routines)



# SCEC CyberShake

- Using the large scale simulation data, estimate probabilistic seismic hazard (PSHA) curves for sites in southern California (probability that ground motion will exceed some threshold over a given time period)
- Used by hospitals, power plants etc. as part of their risk assessment
- Plan to replace existing phenomenological curves with more accurate results using new CyberShake code. (better directivity, basin amplification)
- Completed 40 locations  $\leq 2008$ , targeting 200 in 2009, and 2000 in 2010



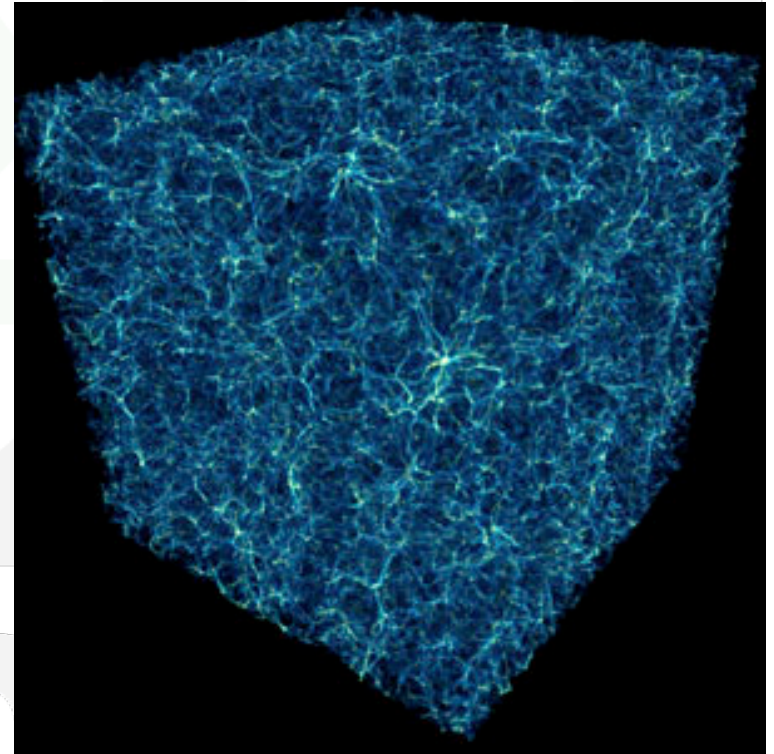


# SCEC CyberShake – PSHA Computing

- A. Generate rupture variations – 48 hours of sequential processing at USC
  - Output: 1 TB of data describing 600,000 potential earthquakes, each with a unique hypocenter and slip distribution
- For each location:
  - B. Generate Strain Green Tensors (SGTs) – two 18 hour 400-core jobs
    - Output: 25 GB of data
  - C. Generate Hazard Curve – 840,000 sequential jobs, takes  $O(10)$  hours on 1000 cores
    - Output: Hazard Curve (small amount of data)
  - Distribute the work for the locations across USC, TACC, NCSA
- 1 TB of data is zipped, GridFTP'ed, and unzipped at each site
- Takes about 3 days
- Generating all curves takes a few weeks
- Managing the sequential jobs for each hazard curve requires effective grid workflow tools for job submission, data management and error recovery, using Pegasus (ISI) and DAGman (U of Wisconsin)

# ENZO

- ENZO simulated cosmological structure formation
- Big current production simulation:
  - 4096x4096x4096 non-adaptive mesh, 16 fields per mesh point
  - 64 billion dark matter particles
  - About 4000 MPI processes, 1-8 OpenMP threads per process
  - Reads 5 TB input data
  - Writes 8 TB data files
    - A restart reads latest 8 TB data file
  - All I/O uses HDF5, each MPI process reading/writing their own data
  - Over a few months for the simulation, >100 data files will be written, and about >20 will be read for restarts
  - 24 hour batch runs
    - 5-10 data files output per run
    - Needs ~100 TB free disk space at start of run



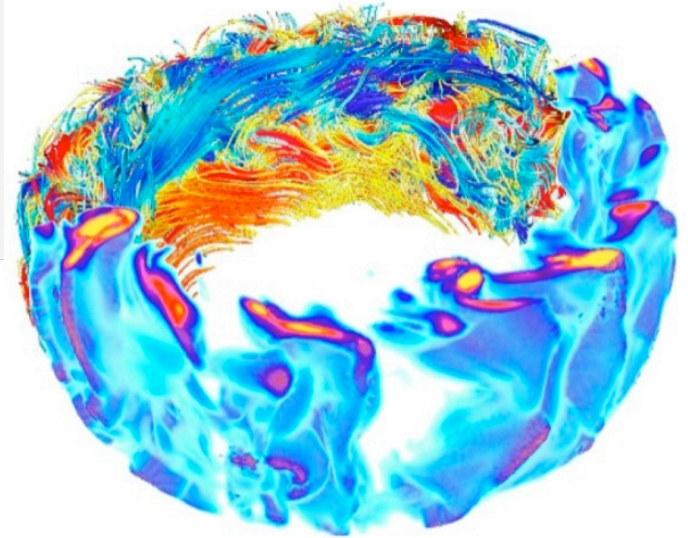
# ENZO Calculation Stages

1. Generate initial conditions for density, matter velocity field, dark matter particle position and velocity (using parallel HDF5)
  - Using NICS Kraken w/ 4K MPI processes, though TACC Ranger is reasonable alternative
  - About 5 TB of initial data are created in 10 0.5-TB files
2. Decompose the initial conditions for the number of MPI tasks to be used for the actual simulation (using sequential HDF5)
  - The decomposition of the mesh into the "tiles" needed for MPI tasks requires strided reads in a large data cube
    - This is very costly on NICS Kraken but can be done more efficiently on TACC Ranger
    - If done on Ranger, then 2 TB of data (4 512-GB files) have to be transmitted from NICS to TACC and after running the MPI decomposition task (with 4K MPI tasks) there are 8K files (another 2 TB) which must be returned to NICS
  - The dark matter particle sort onto the "tiles" is most efficient on NICS Kraken because it has a superior interconnect
    - The sort is usually run in 8 slices using 4K MPI tasks
3. Evolve time (using sequential HDF5)
  - Dump data files during run
  - Archive data files (8 TB every couple of hours -> 1100 MB/sec but NICS HPSS only reaches 300 MB/sec)
4. Derive data products
  - Capture 5-6 fields from each data file (~256 GB each)
  - Send to ANL or SDSC for data analysis or viz
  - Archive output of data analysis or viz (back at NICS)


This overall run will produce at least 1 PB of data, with at least 100 TB needed to be archived, and requires 100 TB of free disk space

# Visualizing the Interior of the Sun

- Benjamin Brown (U. Colorado & JILA) using a visualization tool (VAPOR) developed at NCAR in collaboration with the UC Davis and Ohio State to help the AMNH Hayden Planetarium produce a movie about stars
- Both computation and visualization using same disk resources at TACC
  - Computation output: 5 TB
    - Didn't have to be moved
  - Visualization output was moved from TACC to AMNH
- The movie, which will reach an estimated one million people each year, is slated to be released in 2009
- The sequences will include simulated “flybys” through the interior of the Sun, revealing the dynamos and convection that churn below the surface



# User Portal: portal.teragrid.org

**TeraGrid™  
User Portal**

[Login](#)  
Welcome, Guest User

[Home](#) [Resources](#) [Documentation](#) [Training](#) [Consulting](#) [Allocations](#)  
[About](#) [Team](#) [Changes and Plans](#) [Feedback](#) [Citation Info](#)

## Welcome to the TeraGrid User Portal

### About

The TeraGrid User Portal is a Web interface for making TeraGrid account management easier, for getting information about TeraGrid resources, and for accessing many of the existing TeraGrid services in a single place.

While users may utilize many features of the User Portal without logging in, authenticating provides access to a full set of services available on the TeraGrid. All new users will receive a "New User Form" via U.S. postal mail containing a User Portal username and password along with their other TeraGrid system account usernames and passwords.

### Login


User Name

Password


☐ Remember my login

[Forget your password?](#)

[Need help logging in?](#)

 [Report Security Incident](#)

### Feature Spotlight



**TeraGrid '08**  
June 9-13, 2008 Las Vegas





# Access to resources

- Terminal: ssh, gsissh
- Portal: TeraGrid user portal, Gateways
  - Once logged in to portal, click on “Login”
- Also, SSO from command-line

Home My TeraGrid Resources Documentation Training Consulting Allocations

Accounts and Usage Profile Registered DNs Change Portal Password GSI SSH [Beta] Add/Remove User Co

### System Accounts

Resource	Username	Connect
<b>IU</b>		
login.bigred.iu.teragrid.org	tg-danie	<a href="#">Login</a>
<b>LONI</b>		
queenbee.loni-lsu.teragrid.org	dsk	<a href="#">Login</a>
<b>NCAR</b>		
tg-login.frost.ncar.teragrid.org	dsk	<a href="#">Login</a>
<b>NCSA</b>		
login-abe.ncsa.teragrid.org	dsk	<a href="#">Login</a>
login-co.ncsa.teragrid.org	dsk	<a href="#">Login</a>
tg-login.ncsa.teragrid.org	dsk	<a href="#">Login</a>
login-w.ncsa.teragrid.org	dsk	<a href="#">Login</a>
<b>ORNL</b>		
tg-login.ornl.teragrid.org	dsk	<a href="#">Login</a>
<b>PSC</b>		
tg-login.rachel.psc.teragrid.org	dsk	<a href="#">Login</a>
tg-login.bigben.psc.teragrid.org	no account	
<b>Purdue</b>		
tg-login.purdue.teragrid.org	ux455239	<a href="#">Login</a>
<b>SDSC</b>		
bglogin.sdsc.edu	ux455239	<a href="#">Login</a>
dslogin.sdsc.edu	ux455239	<a href="#">Login</a>

### Allocations

Start Date (YYYY-MM-DD)	End Date (YYYY-MM-DD)	Resource	R
<b>Project Title:</b> The NSF National Virtual Observatory <b>Charge No.:</b> TG-MCA04N009 <b>Grant No.:</b> MCA04N009 <b>Project PI?</b> No			
2004-03-22	2005-03-31	teragrid	
<b>Project Title:</b> TeraGrid: Montage Portal Development <b>Charge No.:</b> TG-AST060006T <b>Grant No.:</b> AST060006T <b>Project PI?</b> No			
2005-11-04	2006-11-30	teragrid_roaming	
<b>Project Title:</b> TG RP LSU <b>Charge No.:</b> TG-STA080000N <b>Grant No.:</b> STA080000N <b>Project PI?</b> No			
2007-10-10	2008-05-31	teragrid_roaming	
<b>Project Title:</b> NON-ROAMING machines only -- TG <b>Charge No.:</b> TG-STA080001N <b>Grant No.:</b> STA080001N <b>Project PI?</b> No			
2008-01-01	2008-12-31	abe-queenbee.teragrid	

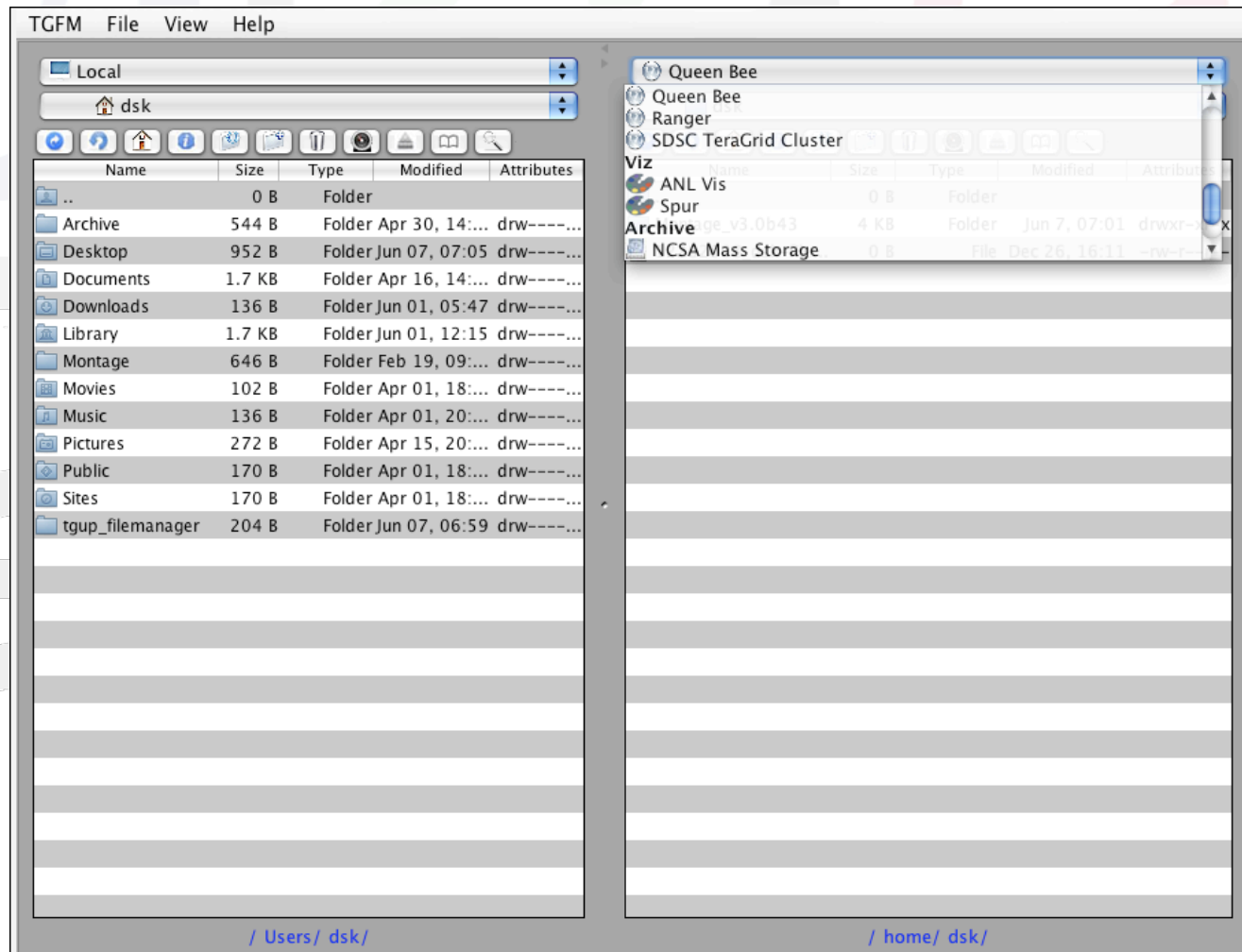
[\[User Responsibility Form\]](#)  
[Need to cite TeraGrid?](#)  
[Report Security Incident](#)

# Data Storage Resources

- Global File System
  - GPFS-WAN
    - 700 TB disk storage at SDSC, accessible from machines at NCAR, NCSA, SDSC, ANL
    - Licensing issues prevent further use
  - Data Capacitor (Lustre-WAN)
    - Mounted on growing number of TG systems
    - 535 TB storage at IU, including databases
    - Ongoing work to improve performance and authentication infrastructure
    - Another Lustre-WAN implementation being built by PSC
  - pNFS is a possible path for global file systems, but is far away from being viable for production
- Data Collections
  - Storage at SDSC (files, databases) for collections used by communities
- Tape Storage
  - Available at IU, NCAR, NCSA, SDSC
- Access is generally through GridFTP (through portal or command-line)

# TGUP Data Mover

- Drag and drop java applet in user portal
  - Uses GridFTP, 3<sup>rd</sup>-party transfers, RESTful services, etc.



# Data Kits

- Coordinated TeraGrid Software Stack (CTSS) is made of kits
  - All but one are optional for RPs
  - Kits define a set of functionality and provide an implementation
- 3 data kits now exist and are in production
  - Data movement kit – on 20 TG resources
  - Data management kit (SRB) – 4 TG resources
  - Wide area file systems kits - GPFS-WAN (5), LUSTRE-WAN (2)
- Currently reworking data kits to include:
  - new client-level kits to better express functionality and accessibility
  - new server-level kits to report more accurate information on server configurations
  - broadened use cases
  - requirements for more complex functionality (managing, not just moving, data – using iRODS)
  - improved information services to support science gateways and automated resource selection

# TeraGrid Data Architecture

- Two kinds of data movement in the TeraGrid
  - Users moving data to or from a location outside the TeraGrid
    - Tend to be smaller numbers of files and less overall data to move
    - Primarily encounter problems with usability due to availability or ease-of-use
  - Users moving data between TeraGrid resources
    - Datasets tend to be larger
    - Users are more concerned with performance, high-reliability and ease of use
  - (Frequently, users will need to do both within the span of a given workflow)
- General trend: as need for data movement has increased, both the complexity of the deployments and the frustrations of users have increased





# Data Architecture: Goals

- Meet user desires: reliability, ease of use, and in some cases high performance
- Make the technology details/implementation transparent to the user
- Enable/simplify user-initiated data movement, particularly on large systems where it has proven to create problems with contention for disk resources



# Data Movement Requirements

- R1: Users need reliable, easy to use file transfer tools for user moving data from outside the TeraGrid to resources inside the TeraGrid.
  - SSH/SCP with the High-performance networking patches (HPN-SCP)
- R2: Users need reliable, high performance, easy to use file transfer tools for moving data from one TeraGrid resource to another.
  - SCP-based transfers to gridFTP nodes – RSSH
- R3: Tools for providing transparent data movement are needed on large systems with low storage to flops ratio.
  - TGUP Data mover



# Wide Area File System Requirement

- Network architecture on the petascale systems is proving to be a challenge – only a few router nodes are connected to wide area networks directly and the rest of the compute nodes are routed through these
- Wide area file systems often need direct connect access
- It has become clear that no single solution will provide a production global wide area network file system.
  - R4: The “look and feel” or the appearance of a global wide area file system with high availability and high reliability (LUSTRE-WAN, pNFS).



# Data Analysis/Visualization Requirements

- Until recently, visualization and in many cases, data analysis have been considered a post-processing task requiring some sort of data movement
- With the introduction of petascale systems, we are seeing data set sizes that prohibit data movement or make it necessary to minimize the movement
- It is anticipated that scheduled data movement is one way in which to guarantee that the data is present at the time it is needed
  - R5: Ability to schedule data availability for post-processing tasks. (DMOVER)
- Visualization and data analysis tools have not been designed to be data aware and have made assumptions that the data can be read into memory and that the applications and tools don't need to be concerned with exotic file access mechanisms.
  - R6: Availability of data mining/data analysis tools that are more data aware. (Currently working with VisIt developers to modify open source software. Leveraging work done on parallel Mesa)



## Archival Data Requirement

- Many TeraGrid sites provide effectively unlimited archival storage to compute-allocated users
- The volume of data flowing into and out of particular archives is already increasing drastically, in some cases exponentially, beyond the ability of the disk caches and tape drives currently allocated
  - R7: The TeraGrid must provide better organized, more capable, and more logically unified access to archival storage for the user community





# Data Archive Issues

- Specific problem: SDSC and NCSA have archived data, but after TG ends, they will not have support to keep this as part of TG XD
- In the future, sites will come and go from TeraGrid
- Questions:
  - What is the TG's responsibility regarding archival data?
  - What is a site's responsibility?
  - What if that site leaves TeraGrid?
  - Who pays for keeping/moving the data?
    - The site? The user? NSF?

## How Much Archival Data is There?

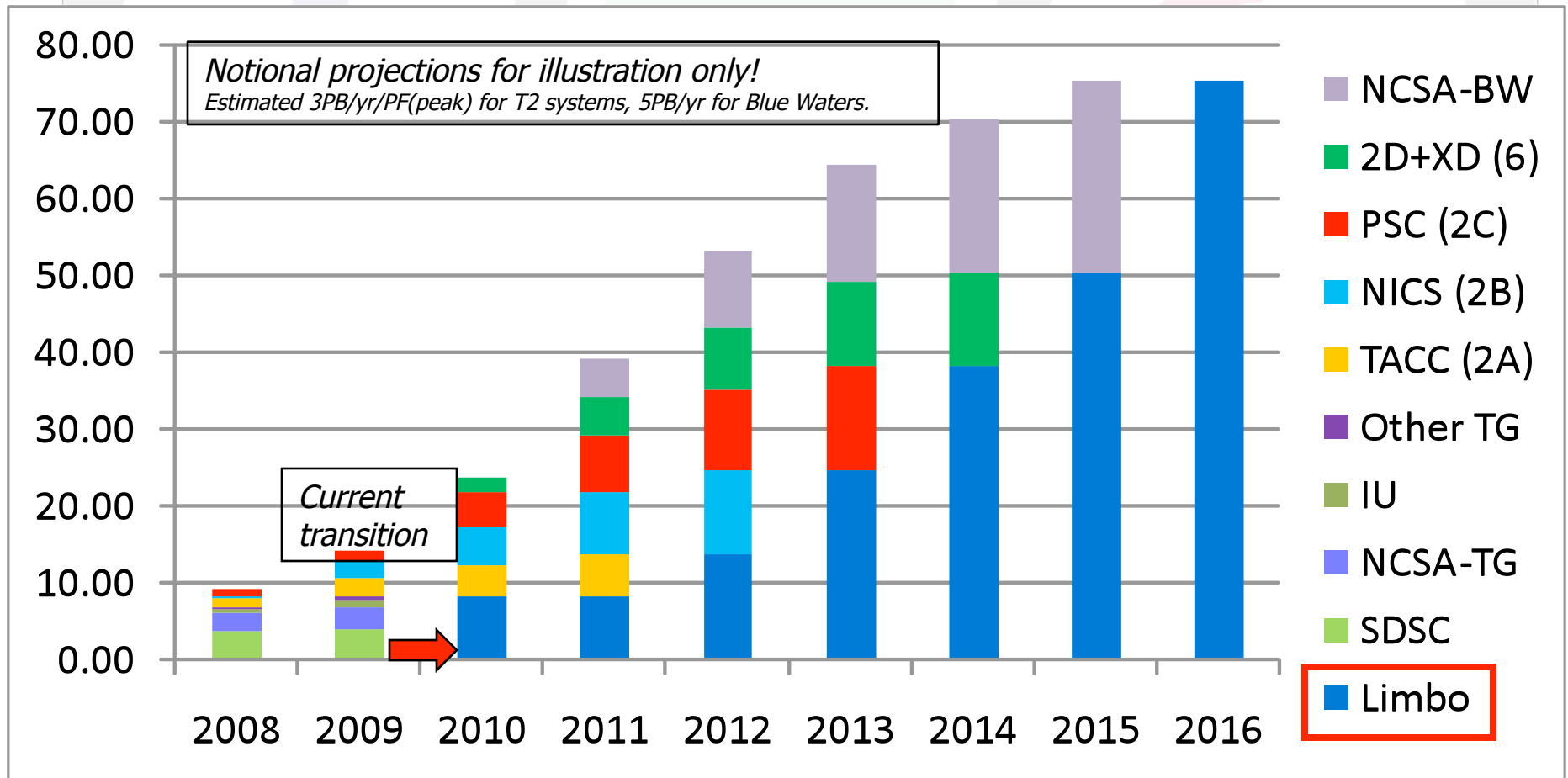
- TG performed a data census of all sites
- Critical sites are SDSC and NCSA
- SDSC and NCSA dominate the total XD unfunded archival storage
- SDSC has 3.7PB of data that needs continued support (3.1 PB in HPSS, 0.6 PB in SAM)
- NCSA has about 2.5PB of unique data
- There is continuing growth



## What To Do About It?

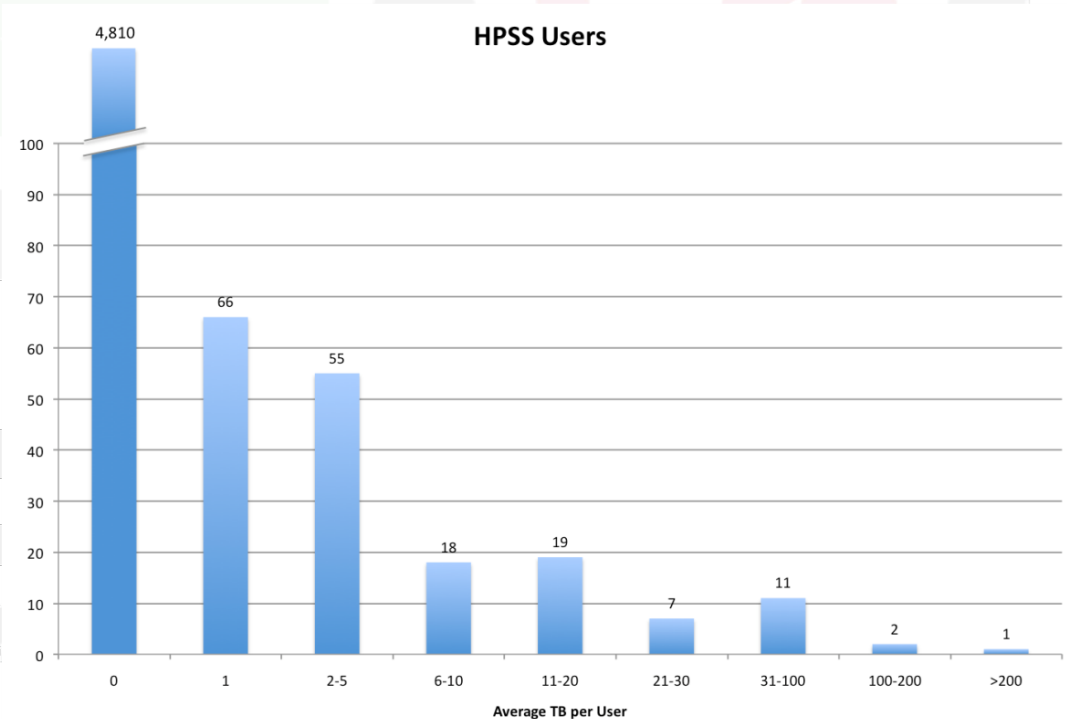
- Data is the essential life force of some efforts
- These transitions are ongoing and will continue
- TeraGrid needs to be seen as a responsible data guardian, independently of RP changes
- Rather than a one-off solution, want to increase TeraGrid's data offerings to handle this requirement
- Data replication between sites and a single persistent archive are possible solutions
- Replication capability is a necessity in any solution

# Looking Ahead ... This Will Be a Recurring and Growing Issue



# The 1/99 Experience at SDSC

- 1% of the users use 99% of the storage
- Other site experiences are 10/90 or 20/80
- Can use this to help with replication (contacting users, complete tapes, etc.)





# What are the high-end archival users doing?

- (From 12/18/08 interview with Mike Norman, SDSC)
- Needs data from the last 3-5 years
  - But this is generally the largest data
  - Experimental data must be stored long-term, dual copies
- Hero runs can generate 100+ TB, 1 year to generate, 1-2 years to analyze, need easy access during period
  - Can lots of disk replace an archive with that usage model? Yes, if there's enough of it for 1-3 years!
- “Deep” user, runs cannot be done elsewhere, data cannot be stored elsewhere
- We need to meet these needs

## How much will replication cost?

- Tape costs are approximately \$100 per TB
- Assume 10 PB to be replicated
- Tape libraries, drives, etc., approximately equal tape cartridge costs
- 10 PB equates to about \$2M in hardware
- Some people costs

# How long would a network transfer take?

- Assume 10 PB to be moved via network
- 10 Gb/s = 10 PB/100 days (will achieve less)
- Most archives cannot move data at 1 GB/s
- Overhead in SRB is high
- 6 months would be a good achievement
- 12 months should be the top end
- Need to start soon!

## How should it be done?

- Multiple replication approaches are both possible and necessary
- SRB and iRODS support replication via middleware.
- GPFS is working with HPSS for archival federation.
- Slash2 from PSC would federate multiple Linux-mounted sites, including Lustre sites



# Science Gateways

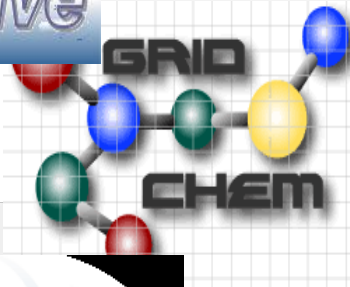
- A natural extension of Internet & Web 2.0
- Idea resonates with Scientists
  - Researchers can imagine scientific capabilities provided through familiar interface
    - Mostly web portal or web or client-server program
- **Designed by communities;** provide interfaces understood by those communities
  - Also provide access to greater capabilities (back end)
  - Without user understand details of capabilities
  - Scientists know they can undertake more complex analyses and that's all they want to focus on
  - TeraGrid provides tools to help developer
- Seamless access doesn't come for free
  - Hinges on very capable developer





# Current Science Gateways

- Biology and Biomedicine Science Gateway
- Open Life Sciences Gateway
- The Telescience Project
- Grid Analysis Environment (GAE)
- Neutron Science Instrument Gateway
- TeraGrid Visualization Gateway, ANL
- BIRN



- Open Science Grid (OSG)
- Special PRiority and Urgent Computing Environment (SPRUCE)
- National Virtual Observatory (NVO)
- Linked Environments for Atmospheric Discovery (LEAD)
- Computational Chemistry Grid (GridChem)
- Computational Science and Engineering Online (CSE-Online)
- GEON(GEOsciences Network)
- Network for Earthquake Engineering Simulation (NEES)
- SCEC Earthworks Project
- Network for Computational Nanotechnology and nanoHUB
- GIScience Gateway (GISolve)
- Gridblast Bioinformatics Gateway
- Earth Systems Grid
- Astrophysical Data Repository (Cornell)



# Social Informatics Data Grid

Collaborative access to large, complex datasets

- SIDGrid is unique among social science data archive projects
  - Focused on streaming data which change over time
    - Voice, video, images (e.g. fMRI), text, numerical (e.g. heartrate, eye movement)
  - Provides the ability to investigate/annotate multiple datasets, collected at different time scales, simultaneously
    - Large datasets result
  - Sophisticated analysis tools
  - Avenue for collaboration



<http://www.ci.uchicago.edu/research/files/sidgrid.mov>

# SIDGrid: Why a gateway?

- Social scientists have traditionally worked in isolated labs without the capability to share data or insights with others.
  - Data that is expensive to collect can now be shared with others
  - Geographically distant researchers can collaborate
  - Complex analysis tools and workflows available for all
  - Researchers have access to high performance computational resources
- TeraGrid used for computationally-intensive tasks such as media transcoding algorithms for pitch analysis of audio tracks and fMRI image analysis



Source: Dr. Steven Boker, Notre Dame

## Interesting TG Capability being developed

- **OGCE's (Gateway) Resource Prediction Service**

- Presented to TG by Gopi Kandaswamy at RENCi
- Service ranks available resources based total time to run an application (input data transfer time, batch queue wait time and the wall time for the application).
- NWS used for data transfer calculations
  - But doesn't calculate transfer time back to the gateway and NWS not installed everywhere
  - Also would be interesting to calculate success rates and use the most reliable sites
- BQP used for batch queue wait
- Performance model for the app used for wall time
- Used by LEAD

# Intensive TeraGrid use by LEAD

## Motivates closer look at reliability under load

- A successful gateway program may result in bursty loads on the TeraGrid
  - Both compute and data transfer
  - We must be able to handle this successfully
- 6-month intensive debugging effort
  - Weekly telecons with sys admins, grid software developers, gateway developers
    - Having the right people who could solve problems on the spot was key
- Clear wiki documentation of problems
  - Rather than emails and telephone calls that were difficult to track
- Testbed for debugging
  - Inca tests that simulated gateway behavior
  - Tests run more frequently on the testbed, results reviewed weekly
- Code improvements made to both LEAD gateway and Globus
- Work transitioned to ops-wg for routine monitoring
- Track 2 awards for testbeds and interactive access can help





# TeraGrid: Both Operations and Research

- Operations

- Facilities/services on which users rely
- Infrastructure on which other providers build

AND

- R&D

- Learning how to do distributed, collaborative science on a global, federated infrastructure
- Learning how to run multi-institution shared infrastructure
- Understanding and trying to solve large-scale data challenges

# TeraGrid: Data Lessons/Issues

- Data is becoming more of a problem than computing
- Make it possible to not move data
  - Analyze/visualize in place
- If data must be moved, make it as transparent as possible
  - Wide Area File system across TeraGrid would be ideal; not quite there yet
  - GridFTP works, but not as easy as users want
  - TGUP File Mover works, may need UI improvements
  - Data movement as part of application workflow not automated in practice (unlike within workflow applications)
  - Scheduled data movement?
- Don't know what to do with archival data over long term
  - "Not my job", "can't afford it", ...
- Science gateways can hide data complexity as well as they hide compute complexity
- Non TeraGrid-specific issues:
  - Data size & multiplicity growing (more, larger files)
  - Data complexity growing (more formats, more complex structures)
  - Data heterogeneity growing (more kinds of data and few universal standards for metadata)



# TeraGrid 10

Pittsburgh, PA August 2-5, 2010

[www.teragrid.org/tg10](http://www.teragrid.org/tg10)