

TeraGrid: National Cyberinfrastructure

Charlie Catlett

Director, TeraGrid

www.teragrid.org

Senior Fellow, Computation Institute

The University of Chicago and Argonne National Laboratory

November 2006



TeraGrid is supported by the National Science Foundation *Office of Cyberinfrastructure*

Charlie Catlett (cec@uchicago.edu)



November 2006

TeraGrid™



"NSF Cyberinfrastructure Vision for 21st Century Discovery"

4. Education and Workforce

3. Collaboratories, observatories, virtual organizations

"sophisticated" science application software

1. Distributed, scalable up to petaFLOPS HPC

includes networking, middleware, systems software

2. Data, data analysis, visualization

includes data to and from instruments

- *provide sustainable and evolving CI that is secure, efficient, reliable, accessible, usable, and interoperable*
- *provide access to world-class tools and services*

From Draft 7.1 CI Plan at www.nsf.gov/oci/

TeraGrid A National Production CI Facility



Phase I: 2001-2004 Design, Deploy, Expand

Phase II: 2005-2010 Operation & Enhancement



20+ Distinct Computing Resources
(>150TF today; ~600TF by 8/2007)
100+ Data Collections



November 2006

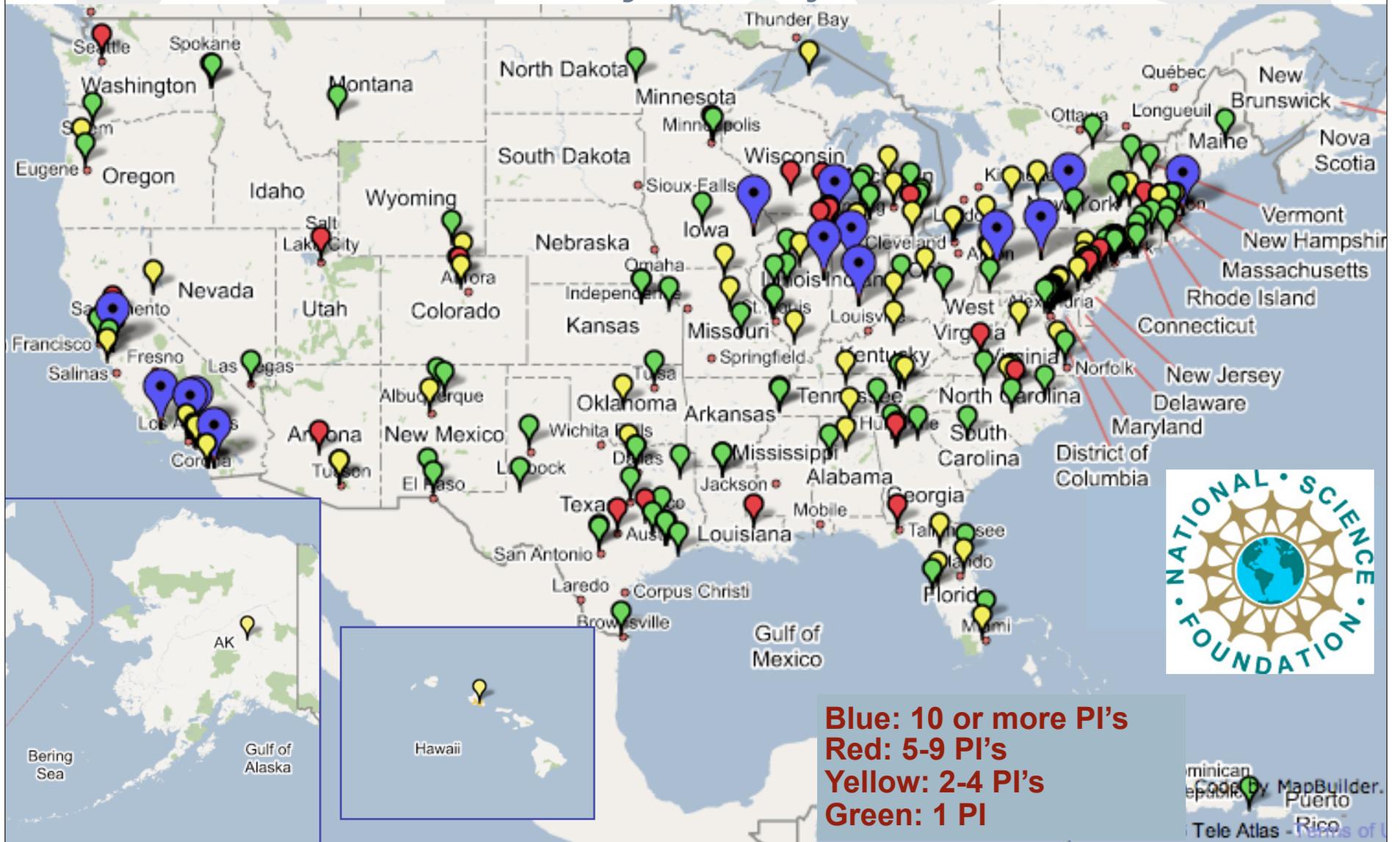
TeraGrid™

TeraGrid Mission

- TeraGrid provides integrated, persistent, and pioneering computational resources that will significantly improve our nation's ability and capacity to gain new insights into our most challenging research questions and societal problems.
 - Our vision requires an integrated approach to the scientific workflow including obtaining access, application development and execution, data analysis, collaboration and data management.
 - These capabilities must be accessible broadly to the science, engineering, and education community.*



TeraGrid Projects by Institution



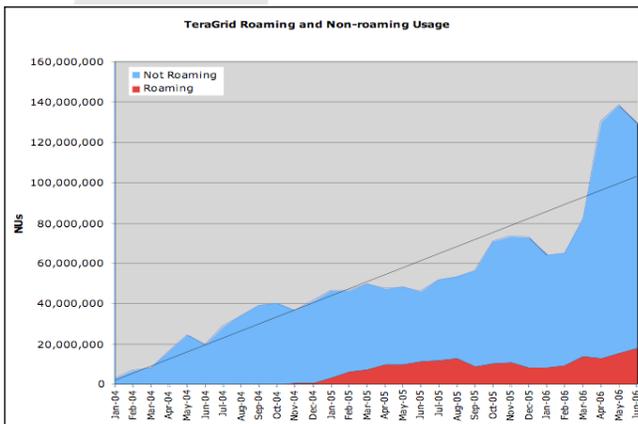
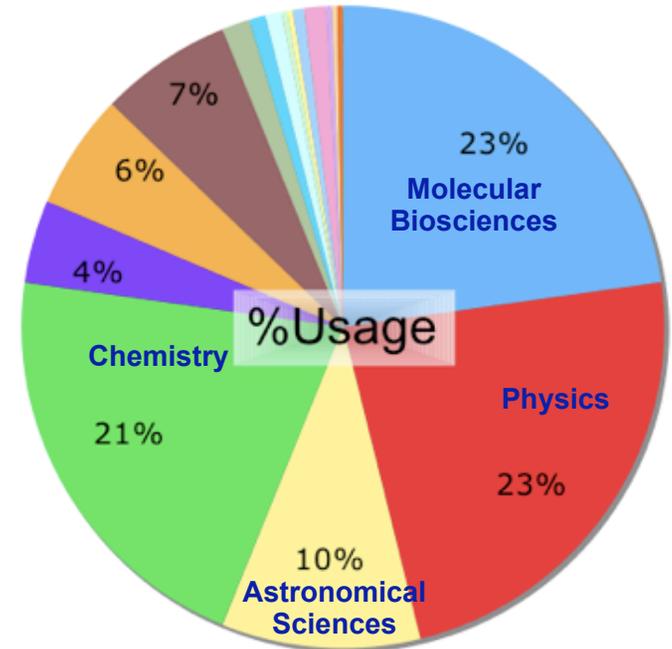
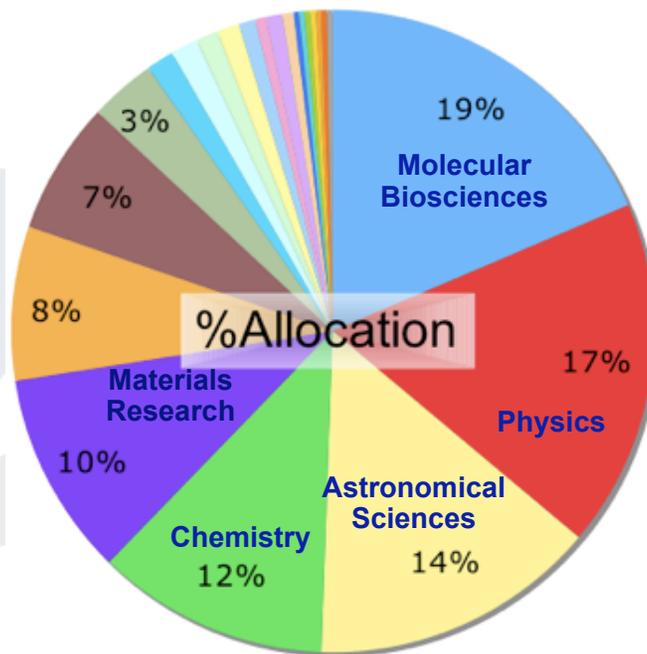
Charlie Catlett (cec@uchicago.edu)

1000 projects, 4000 users

November 2006



TeraGrid Usage



- Molecular Biosciences
- Physics
- Astronomical Sciences
- Chemistry
- Materials Research
- Chemical, Thermal Systems
- Atmospheric Sciences
- Advanced Scientific Computing
- Earth Sciences
- Biological and Critical Systems
- Ocean Sciences
- Cross-Disciplinary Activities
- Computer and Computation Research
- Integrative Biology and Neuroscience
- Mechanical and Structural Systems
- Mathematical Sciences
- Electrical and Communication Systems, Design and Manufacturing Systems, Environmental Biology

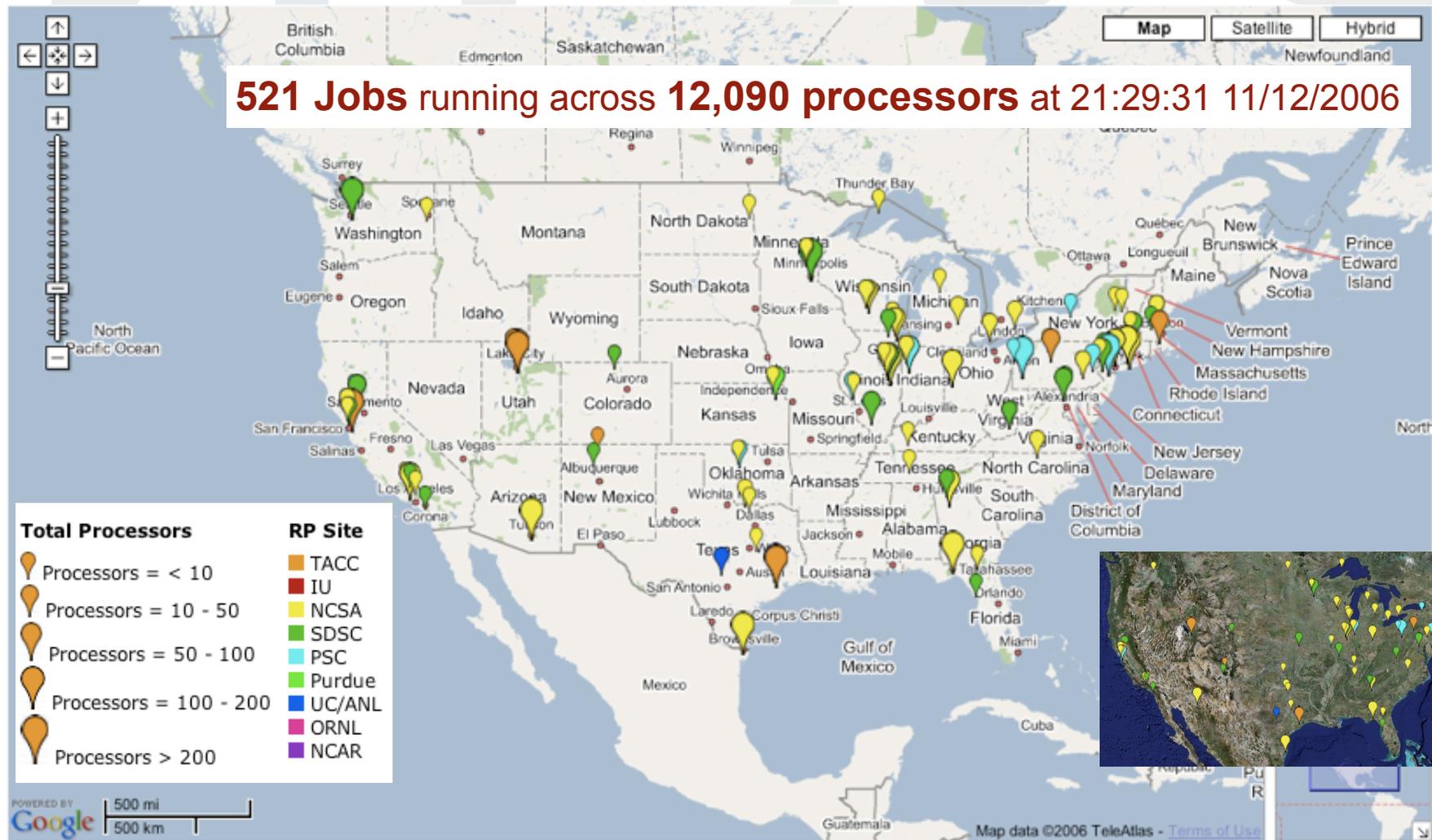
Charlie Catlett (cec@uchicago.edu)

1000 projects, 4000 users

November 2006



Real-Time Usage Mashup



The TeraGrid job map displays the current running jobs across TeraGrid. Each pin location denotes the location of the job owner, the color of the pin denotes the RP site of the job(s), the size of the pin denotes the total number of processors for the jobs. By clicking on the pin you can see the user's job information - RP site, total number of jobs running, total number of processors - in addition to the user's location, department, and institution.

TeraGrid User Portal

Resource	Username
IU	
tg-login-ia32.iu.teragrid.org	tg-catle
tg-login1.iu.teragrid.org	tg-catle
NCSA	
login-co.ncsa.teragrid.org	catlett
login-cu.ncsa.teragrid.org	catlett
tg-login.ncsa.teragrid.org	catlett
login-w.ncsa.teragrid.org	catlett
ORNL	
tg-login.ornl.teragrid.org	catlett
PSC	
tg-login1.lemieux.psc.teragrid.org	ccatlett
tg-login.rachel.psc.teragrid.org	ccatlett
tg-login.bigben.psc.teragrid.org	no account
Purdue	
tg-login.purdue.teragrid.org	catlett
SDSC	
bglogin.sdsc.edu	catlett
dslogin.sdsc.edu	catlett
tg-login.sdsc.teragrid.org	catlett
TACC	
tg-login.tacc.teragrid.org	catlett
maverick.tacc.utexas.edu	catlett
UC/ANL	
tg-login.uc.teragrid.org	catlett

Managing Credentials

Current State of all Resources

Start Date	End Date	Resource	Project Allocation (SU) Remaining / Awarded	My Usage (SU)	Alloc. Type
Project Title: TG Staff Project: Project Management Charge No.: TG-STA040014N Grant No.: STA040014N Project PI? No					
2004-10-19	2013-12-31	teragrid_roaming	99999 / 99999	0.0	new
2003-12-19	2013-12-31	teragrid	99999 / 99999	0.0	new
Project Title: TG RP UC/ANL Charge No.: TG-STA060016N Grant No.: STA060016N Project PI? No					
2006-05-19	2007-05-31	teragrid_roaming	4993 / 5000	0.0	new

Tracking usage for my allocations



Logout
Welcome, Charles E Catlett

Home My TeraGrid Resources Documentation Consulting Allocations

Systems Monitor Science Gateways Data Collections

TeraGrid Systems Monitor										
High Performance Computing Systems										
Name	Institution	System	CPUs	Peak TFlops	Memory TBytes	Disk TBytes	Load	R	Jobs* Q	O
Lonestar	TACC	Dell PowerEdge Linux Cluster	5200	55.00	10.40	94.90		186	64	0
Big Red	IU	IBM e1350	2048	20.40	8.20	266.00		1	126	1892
Tungsten	NCSA	Dell Xeon IA-32 Linux Cluster	2560	16.38	3.75	109.00		76	1149	34
DataStar p655	SDSC	IBM Power4+ p655	2176	14.30	5.75	115.00		34	48	33
TeraGrid Cluster	NCSA	IBM Itanium2 Cluster	1744	10.23	4.47	60.00		57	430	70
Bigben	PSC	Cray XT3	2090	10.00	2.02	48.80		1	32	59
Lear	Purdue	Dell EM64T Linux Cluster	1024	6.60	2.00	28.00		218	385	4
Cobalt	NCSA	SGI Altix	1024	6.55	3.00	100.00		33	262	12
Lemieux	PSC	HP Alpha Cluster	3000	6.00	2.93	78.13		46	2	38
Blue Gene	SDSC	IBM Blue Gene	2048	5.70	0.50	19.50		3	5	7
TeraGrid Cluster	SDSC	IBM Itanium2 Cluster	524	3.10	1.02	48.80		6	246	68
Copper	NCSA	IBM Power4 p690	384	2.00	1.44	30.00		88	108	2
DataStar p690	SDSC	IBM Power4+ p690	192	1.30	0.88	115.00		18	3	12
TeraGrid Cluster	UC/ANL	IBM Itanium2 Cluster	124	0.61	0.24	4.00		7	10	0
NSTG	ORNL	IBM IA-32 Cluster	56	0.34	0.07	2.14		0	0	0
Rachel	PSC	HP Alpha SMP	128	0.31	0.50	6.00		20	75	1
Total:			24322	158.82	47.17	1125.27		794	2945	2232
Advanced Visualization Systems										
Name	Institution	System	CPUs	Peak TFlops	Memory TBytes	Disk TBytes	Graphics HW			
TeraGrid Cluster	UC/ANL	Intel Xeon Cluster	192	0.61	0.38	4.00	nVIDIA GeForce 6600GT AGP graphics cards			
Maverick	TACC	Sun E25K	128	0.27	0.50	0.56	16 nVIDIA QuadroFX 3000G graphics cards			
Total:			320	0.88	0.88	4.56				

TeraGrid User Portal -
Eric Roberts, Texas Advanced
Computing Center
(ericrobe@tacc.utexas.edu)



Continuing to Improve User Tools



TeraGrid™
User Portal

[Login](#)

Welcome, Guest
User

[Home](#) [Resources](#) [Documentation](#) [Consulting](#) [Allocations](#)

[System Monitor](#) [Science Gateways](#) [Data Collections](#) **[Queue Prediction](#)**

Deadline Prediction

Deadline prediction helps you answer the following question: With a 95% confidence, if I submit my job on 'X' # of nodes and a run time of 'Y' minute(s) or hour(s) to machine 'Z' what probability will it start within my specified deadline of 'H' minute(s) or hour(s) ?

System & Queue Information

- TACC Lonestar ▼
- IU Tiger ▼
- UC/ANL TeraGrid Cluster ▼
- NCSA TeraGrid Cluster ▼
- SDSC Datastar ▼
- SDSC TeraGrid Cluster ▼

Job Information

Number of Nodes:

Runtime: ▼

Deadline: ▼

Wait Time Prediction

The predicted wait time query can help you answer the following question: With a 'X' % quantile, if I submit my job on 'A' # of nodes, a run time of 'B' minute(s) or hour(s) on machine 'C', the answer you get back will tell you with a 95% confidence that X % of the jobs submitted to the queue with your specified node will take less than that time to exit the queue.

System & Queue Information

- TACC Lonestar ▼
- IU Tiger ▼
- UC/ANL TeraGrid Cluster ▼
- NCSA TeraGrid Cluster ▼
- SDSC Datastar ▼
- SDSC TeraGrid Cluster ▼

Job Information

Number of Nodes:

Runtime: ▼

Quantile: %

Alpha-test



Deadline prediction - Network Weather Service "Batch Queue Prediction" (BQP) - Rich Wolski (rich@cs.ucsb.edu)
Wait time prediction - Warren Smith (wsmith@tacc.utexas.edu)

November 2006

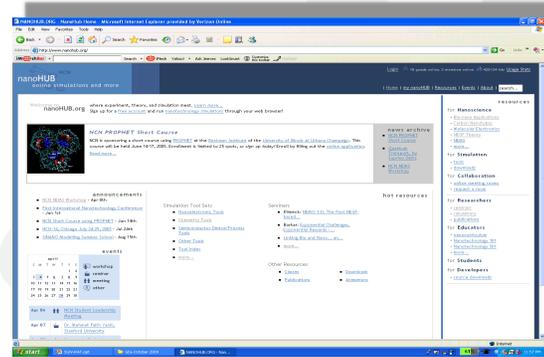
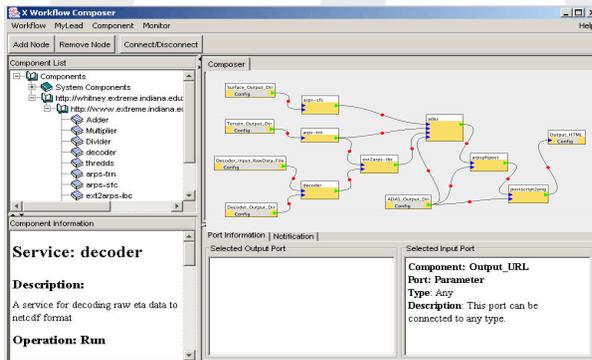
TeraGrid™

Broader Engagement

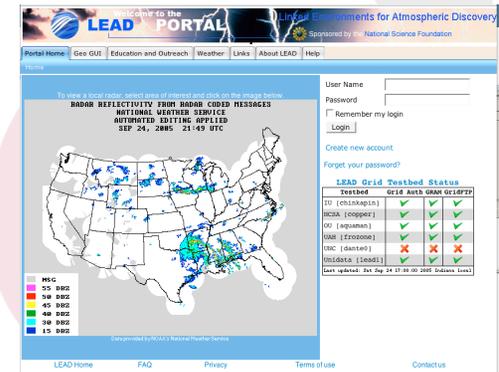
- *How can TeraGrid engage the broader science, engineering, and education community in leveraging and creating discovery?*
- TeraGrid Strategies for Engagement
 - Science Gateways
 - Empower community-designed, led, supported infrastructure
 - Education and Training Initiatives
 - Create and support a community of educators and faculty
 - Campus Partnerships
 - New partnership programs (TeraGrid, OSG, Internet2)



TeraGrid Science Gateways Initiative: Community Interface to Grids



Web Services



Grid-X

TeraGrid

Grid-Y

- Common Web Portal or application interfaces (database access, computation, workflow, etc), exploit standards (primarily web services)
- “Back-End” use of grid services such as computation, information management, visualization, etc.
- Standard approaches so that science gateways may readily access resources in any cooperating Grid without technical modification.



Science Gateway Partners

Science Gateways

Below is a complete list of current science gateways, to see a detailed project description please click on the name science gateway.

Title	Field of Science	Portal Homepage
Biology and Biomedicine Science Gateway	Molecular Biosciences	Visit Portal
Computational Chemistry Grid	Chemistry	Visit Portal
Computational Science and Engineering Online	Chemistry	Visit Portal
GEON(GEOsciences Network)	Earth Sciences	Visit Portal
GIScience Gateway	Geography and Regional Science	Visit Portal
Grid Analysis Environment	Physics	N/A
Linked Environments for Atmospheric Discovery	Atmospheric Sciences	Visit Portal
National Biomedical Computation Resource	Integrative Biology and Neuroscience	Visit Portal
National Virtual Observatory	Astronomical Sciences	Visit Portal
Network for Computational Nanotechnology and nanoHUB	Emerging Technologies Initiation	Visit Portal
Network for Earthquake Engineering Simulation	Earthquake Hazard Mitigation	Visit Portal
Neutron Science Instrument Gateway	Physics	Visit Portal
Open Life Sciences Gateway	Molecular Biosciences	Visit Portal
Open Science Grid	Advanced Scientific Computing	N/A
SCEC Earthworks Project	Earthquake Hazard Mitigation	Visit Portal
Special PRiority and Urgent Computing Environment	Advanced Scientific Computing	Visit Portal
TeraGrid Visualization Gateway	Visualization, Graphics, and Image Processing	Visit Portal
The Earth System Grid	Global Atmospheric Research	Visit Portal
The Telescience Project	Neuroscience Biology	Visit Portal
Virtual Laboratory for Earth and Planetary Materials	Materials Research	Visit Portal

For more information on the science gateways effort please visit the [Science Gateways program page](#).

Science Gateways access via the TeraGrid User Portal (portal.teragrid.org).
 Additional gateways are currently working with TeraGrid to join this list of active gateways.
 For more information please contact **Nancy Wilkins-Diehr** (wilkinsn@sdsc.edu)

Charlie Catlett (cec@uchicago.edu)



November 2006

TeraGrid™

Visit us on the Show floor!

Grid Infrastructure Group
Argonne National Laboratory/ University of Chicago
Indiana University

National Center for Atmospheric Research

National Center for Supercomputing Applications

Oak Ridge National Laboratory

Pittsburgh Supercomputing Center

Purdue University

San Diego Supercomputer Center

Texas Advanced Computing Center

