TeraGrid Overview and Potential TG-OSG Interaction Areas

John Towns

Chair, TeraGrid Forum

Director, Persistent Infrastructure

National Center for Supercomputing Applications

University of Illinois

TeraGrid™

Our Vision of TeraGrid

- Three part mission:
  - support the most advanced computational science in multiple domains
  - empower new communities of users
  - provide resources and services that can be extended to a broader cyberinfrastructure
- TeraGrid is…
  - an advanced, nationally distributed, open cyberinfrastructure comprised of supercomputing, storage, and visualization systems, data collections, and science gateways, integrated by software services and high bandwidth networks, coordinated through common policies and operations, and supported by computing and technology experts, that enables and supports leading-edge scientific discovery and promotes science and technology education
  - a complex collaboration of over a dozen organizations and NSF awards working together to provide collective services that go beyond what can be provided by individual institutions

TeraGrid™

Strategic Objectives

- Objectives determined from considering numerous inputs
  - user input via various mechanisms
    - surveys, user contacts, advisory bodies, review panels, etc.
  - technical input from TG staff

- Annual planning based on identifying/updating 5 high level project strategic objectives
  - Enable science that could not be done without TeraGrid
  - Broaden the user base
  - Simplify users lives
 - Improve Operations
  - Enable connections to external resources

TeraGrid™

TeraGrid: greater than the sum of its parts…

- Single unified allocations process
- Single point of contact for problem reporting and tracking
  - especially useful for problems between systems
- Simplified access to high end resources for science and engineering
  - single sign-on
  - coordinated software environments
  - uniform access to heterogeneous resources to solve a single scientific problem
  - simplified data movement
- Expertise in building national computing and data resources
- Leveraging extensive resources, expertise, R&D, and EOT
  - leveraging other activities at participant sites
  - learning from each other improves expertise of all TG staff
- Leadership in cyberinfrastructure development, deployment and support
  - demonstrating enablement of science not possible without the TeraGrid-coordinated human and technological resources

TeraGrid™

Diversity of Resources (not exhaustive)

- **Very Powerful Tightly Coupled Distributed Memory**
  - Ranger (TACC): Sun Constellation, 62,976 cores, 579 Tflop, 123 TB RAM
  - Kraken (NICS): Cray XT5, 66,048 cores, 608 Tflop, > 1 Pflop in 2009
- **Shared Memory**
  - Cobalt (NCSA): Altix, 8 Tflop, 3 TB shared memory
  - Pople (PSC): Altix, 5 Tflop, 1.5 TB shared memory
- **Clusters with Infiniband**
  - Abe (NCSA): 90 Tflops
  - Lonestar (TACC): 61 Tflops
  - QueenBee (LONI): 51 Tflops
- **Condor Pool (Loosely Coupled)**
  - Purdue- up to 22,000 cpus
- **Visualization Resources**
  - TeraDRE (Purdue): 48 node nVIDIA GPUs
  - Spur (TACC): 32 nVIDIA GPUs
- **Storage Resources**
  - GPFS-WAN (SDSC)
  - Lustre-WAN (IU)
  - Various archival resources
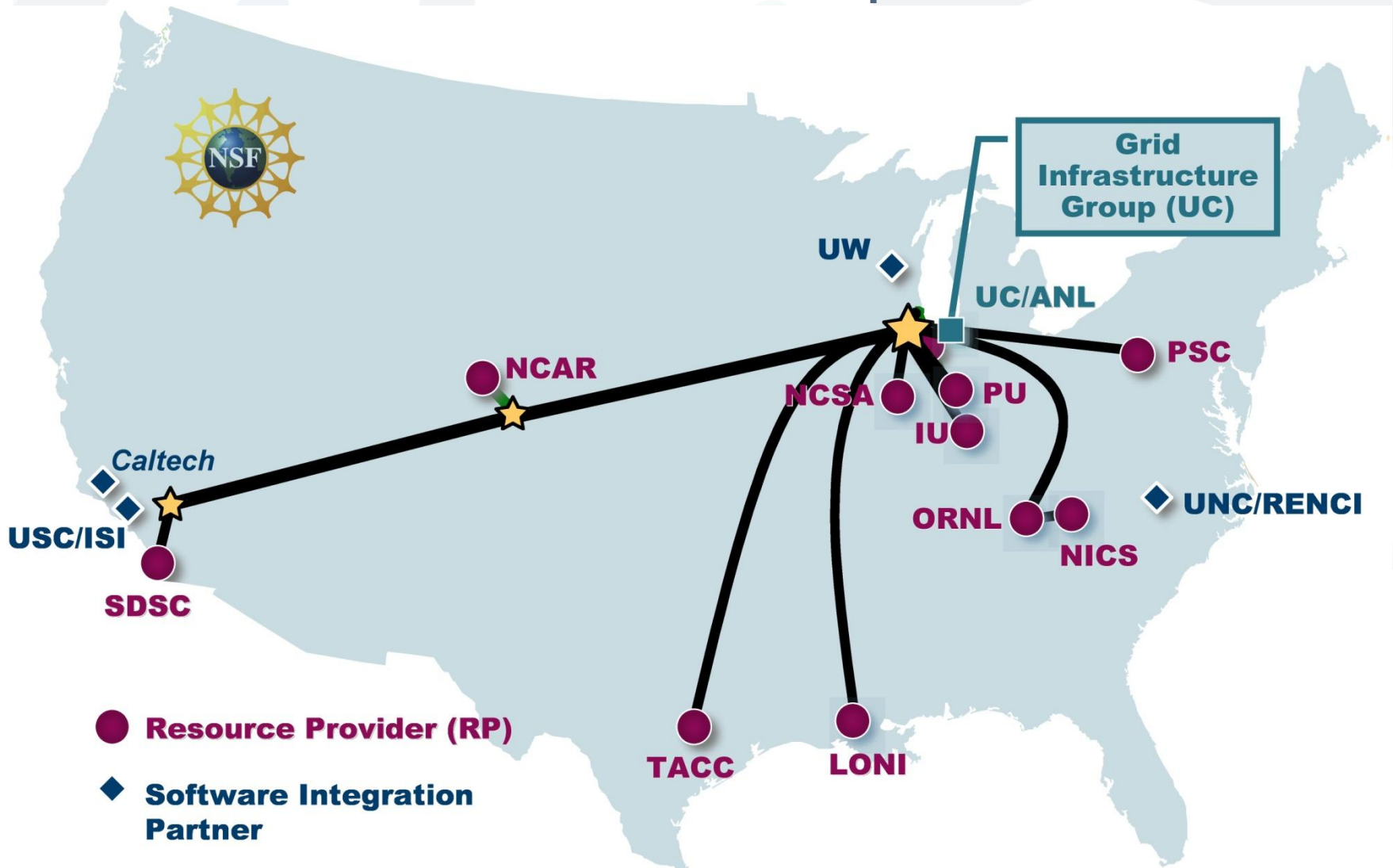
TeraGrid™
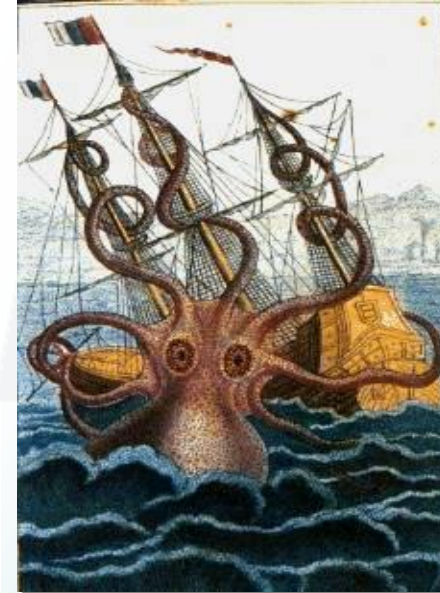
Resources to come...

- Track 2c
  - large system in 2010 ??
- Track 2d being competed
  - data-intensive HPC system
    - SDSC ??
  - experimental HPC system
    - GaTech/ORNL ??
  - pool of loosely coupled, high throughput resources
    - Purdue ??
  - experimental, high-performance grid test bed
    - Indiana ??
- eXtreme Digital (XD) High-Performance Remote Visualization and Data Analysis Services
  - service and possibly resources; 2 awards (NICS and TACC)
- Blue Waters (Track 1) @ NCSA:
  - 1 Pflop sustained on serious applications in 2011

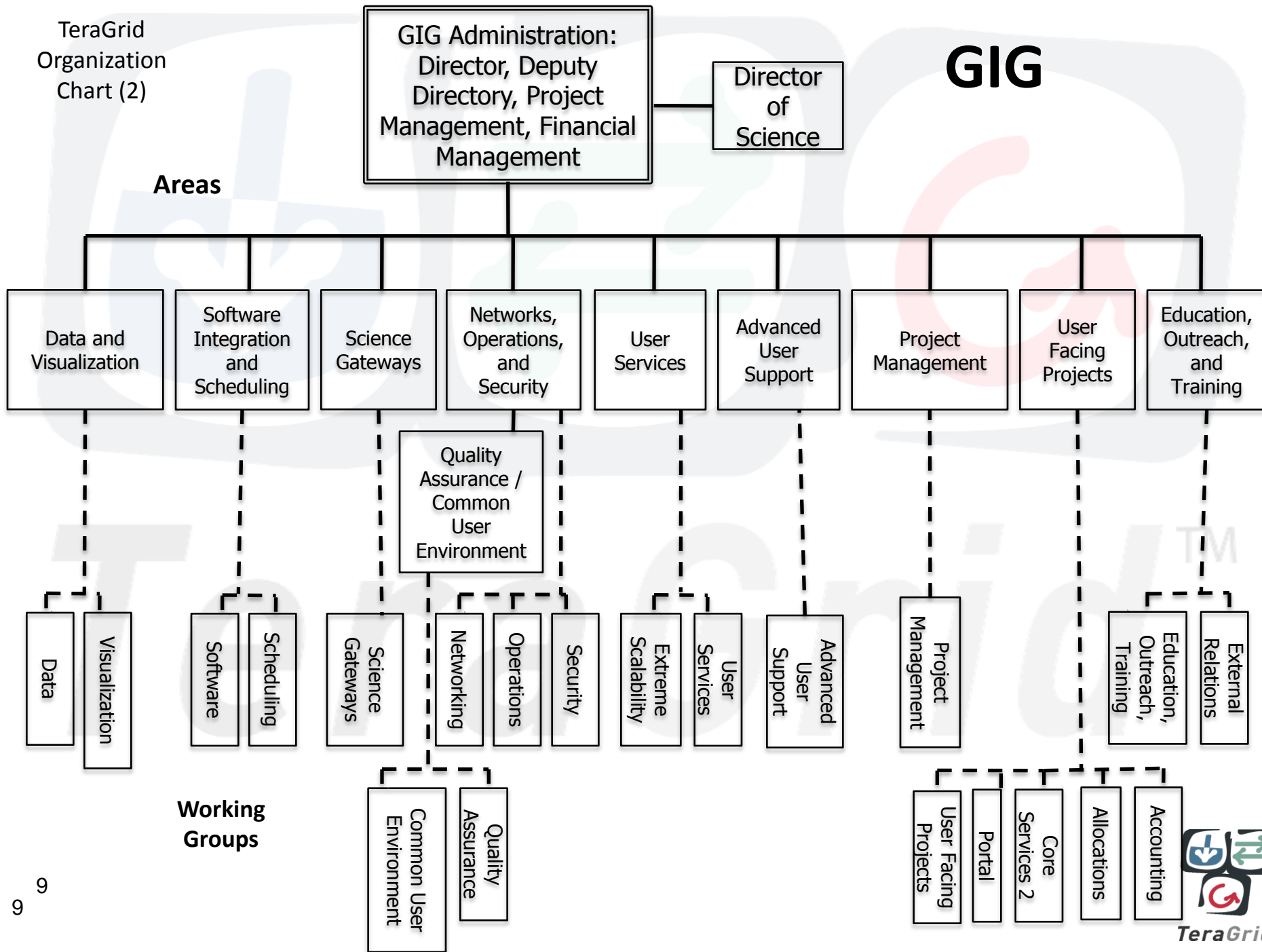TeraGrid™

How is TeraGrid Organized?

- **TG is set up like a large cooperative research group**
  - evolved from many years of collaborative arrangements between the centers
  - still evolving!
- **Federation of 12 awards**
  - Resource Providers (RPs)
    - provide the computing, storage, and visualization resources
  - Grid Infrastructure Group (GIG)
    - central planning, reporting, coordination, facilitation, and management group
- **Strategically lead by the TeraGrid Forum**
  - made up of the PI's from each RP and the GIG
  - led by the TG Forum Chair, who is responsible for coordinating the group (elected position)
    - John Towns – TG Forum Chair
  - responsible for the strategic decision making that affects the collaboration
- **Day-to-Day Functioning via Working Groups (WGs):**
  - each WG under a GIG Area Director (AD), includes RP representatives and/or users, and focuses on a targeted area of TeraGrid

TeraGrid™

TeraGrid Participants

TeraGrid Organization Chart (2)

GIG

Areas

Working Groups

GIG Administration: Director, Deputy Directory, Project Management, Financial Management

Director of Science

- Data and Visualization
  - Data
  - Visualization
- Software Integration and Scheduling
  - Software
  - Scheduling
- Science Gateways
  - Science Gateways
- Networks, Operations, and Security
  - Quality Assurance / Common User Environment
    - Common User Environment
    - Quality Assurance
  - Networking
  - Operations
  - Security
- User Services
  - Extreme Scalability
  - User Services
- Advanced User Support
  - Advanced User Support
- Project Management
  - Project Management
    - User Facing Projects
    - Portal
    - Core Services 2
    - Allocations
    - Accounting
- User Facing Projects
- Education, Outreach, and Training
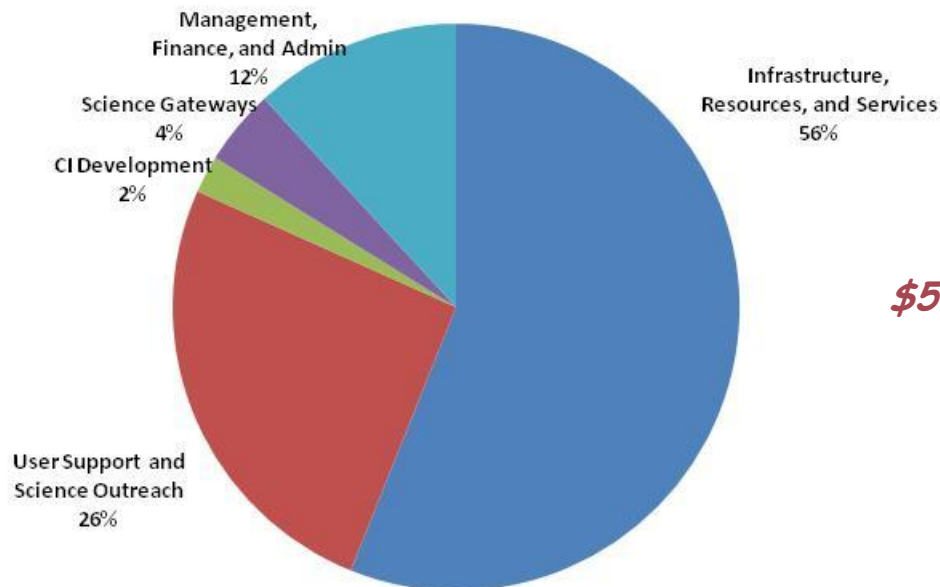 - Education, Outreach, Training
  - External Relations

9

9

9

CY08 Total TeraGrid Expenditures Distribution

- Distribution of total TeraGrid expenditures most closely resembles the RP expenditures simply because the RP expenditures are about four times those of the GIG

- This figure makes it most apparent that the bulk of TeraGrid funds go directly to providing facilities to scientific researchers and making sure that they have the support needed so that they can make productive use of them
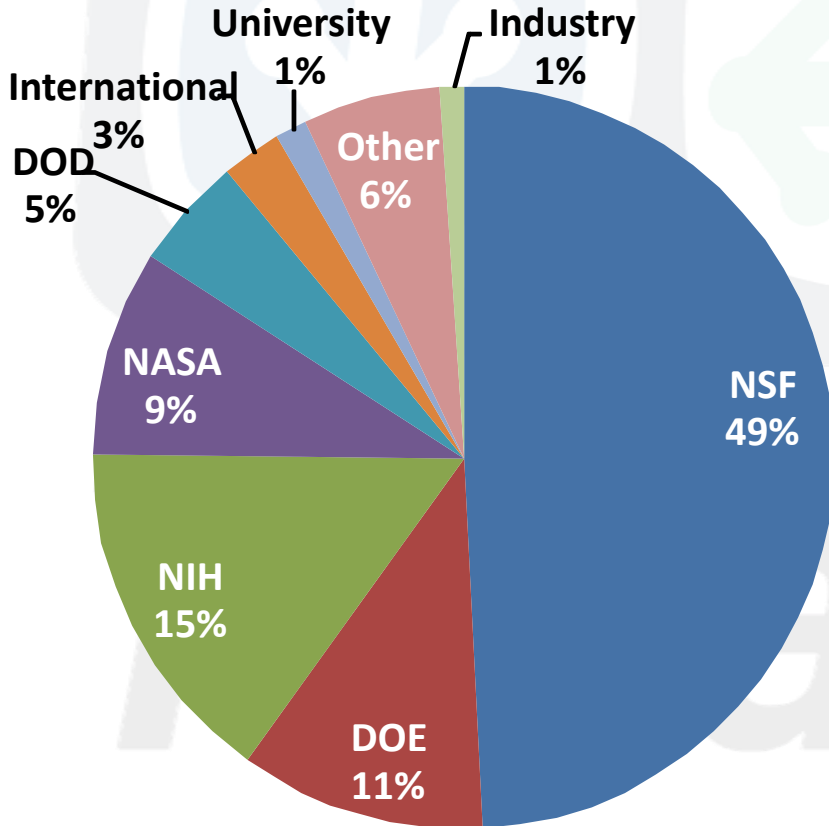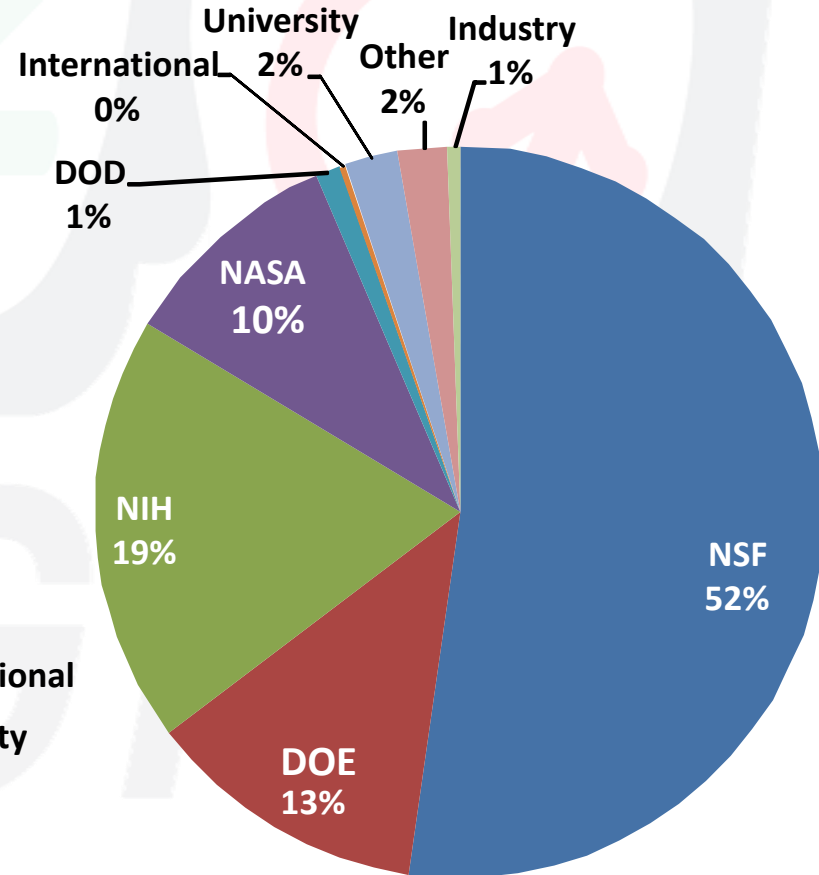
Management, Finance, and Admin 12%

Science Gateways 4%

CI Development 2%

Infrastructure, Resources, and Services 56%

$50.2M Total (CY2008)

User Support and Science Outreach 26%

Impacting Many Agencies

Supported Research Funding by Agency

Resource Usage by Agency

Legend:
- NSF
- DOE
- NIH
- NASA
- DOD
- International
- University
- Other
- Industry

$91.5M Direct Support of Funded Research

10B NUs Delivered

NOTE: OSG Resources deliver ~4B NUs/year

11

TeraGrid HPC Usage, 2008

In 2008,
- Aggregate HPC power *increased by 3.5x*
- NUs requested and awarded *quadrupled*
- NUs delivered *increased by 2.5x*

3.8B NUs in Q4 2008

Kraken, Aug. 2008

Ranger, Feb. 2008

3.9B NUs in 2007

Specific
Roaming

NUs (millions)

1,400
1,300
1,200
1,100
1,000
900
800
700
600
500
400
300
200
100

J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D

2004   2005   2006   2007   2008
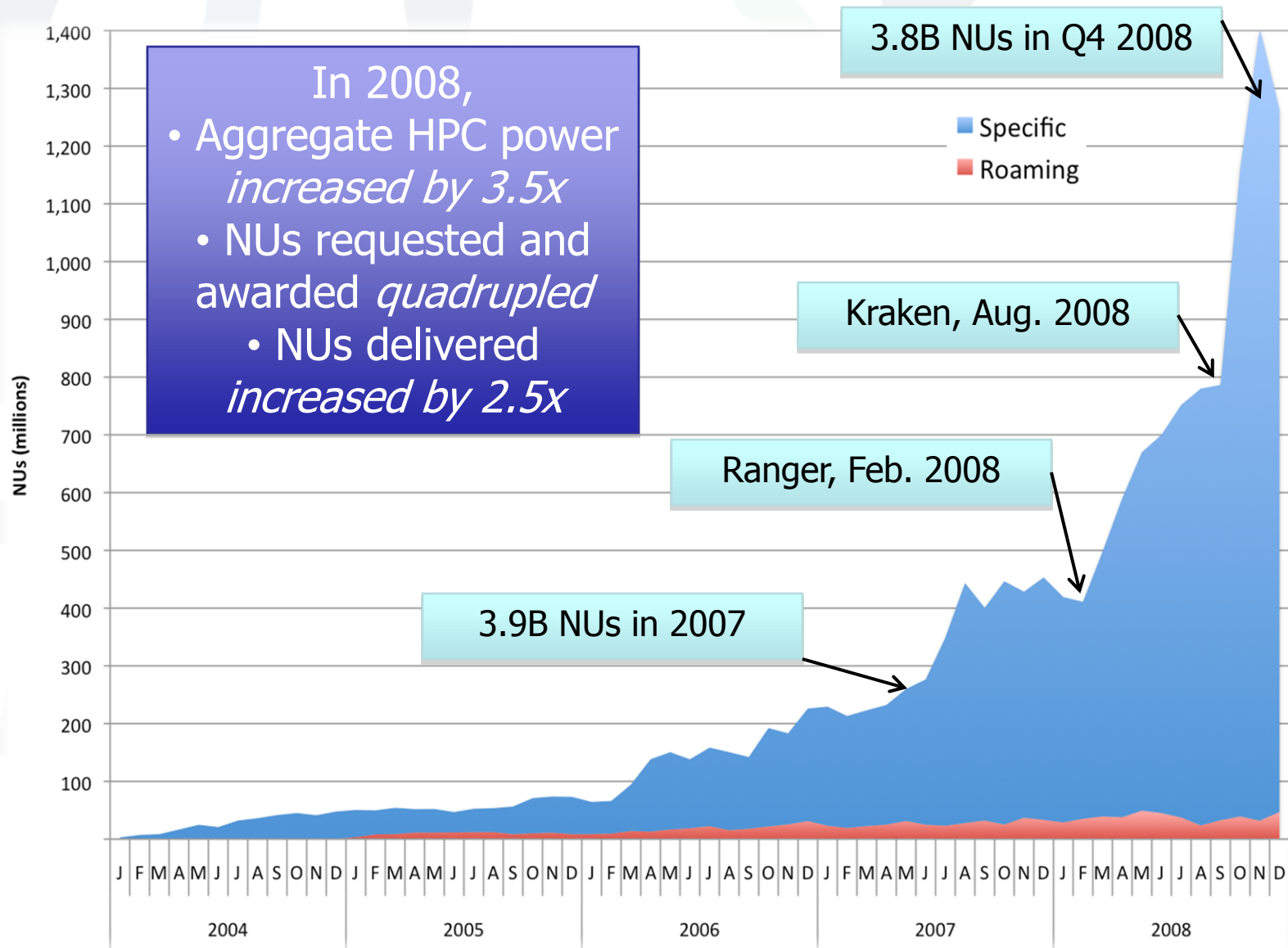
HPC User Community Continues to Grow

Legend:
- TeraGrid Users
- Current Accounts
- Active Users
- Users Charging Jobs
- New Accounts
- Gateway Users
- Target

13

Source: TeraGrid Central Database
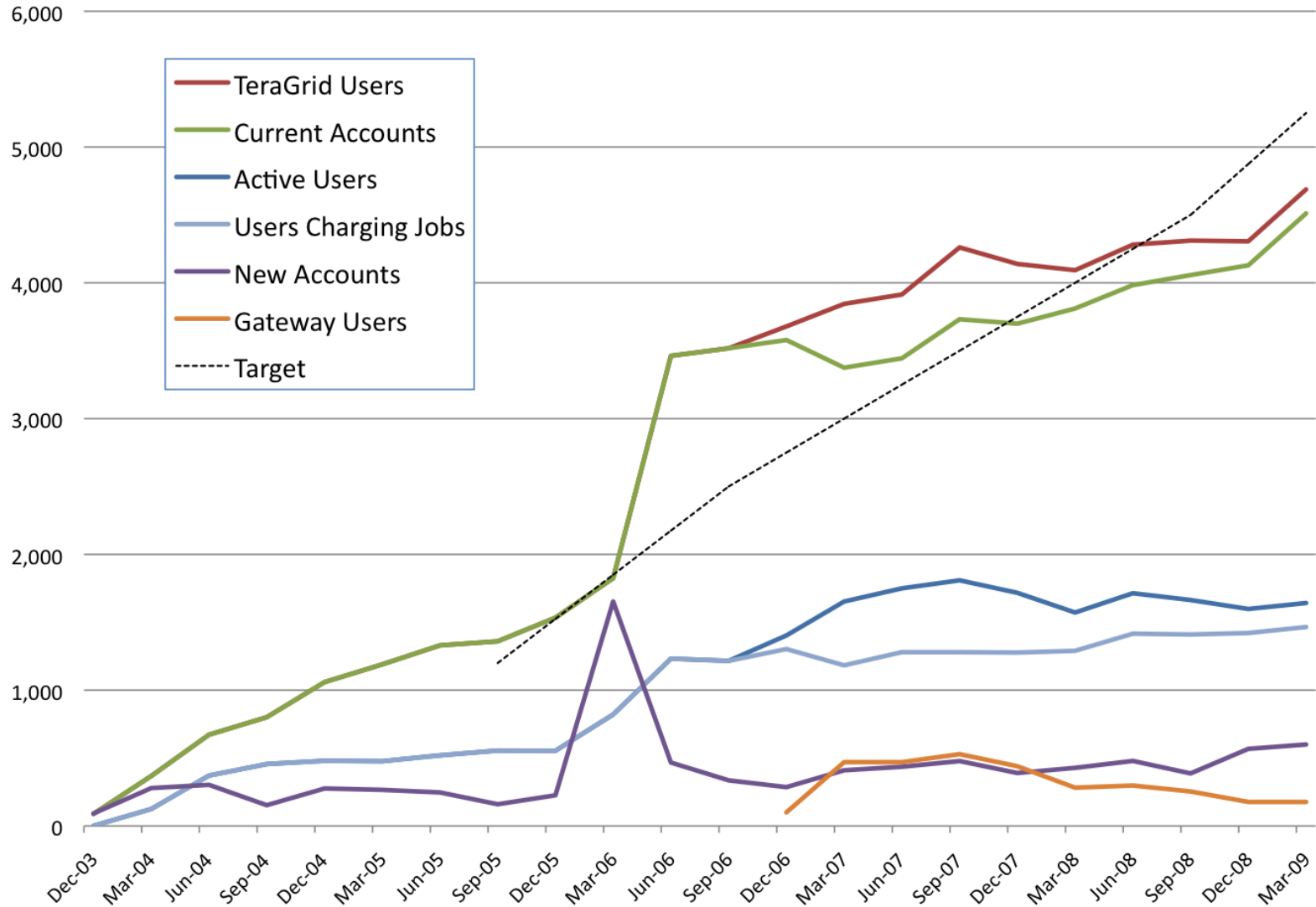
Usage by Institution

14,533,271 SUs used at 178 institutions from the past 7 days up to Aug-10-2009

Today, there are approximately 35 gateways using the TeraGrid

Gateway CPU use up five fold in 2008

- 0.5M hours used on community accounts in 2007
- 2.5M hours used on community accounts in 2008
- Gateways using CPU most heavily
  - SCEC tera3d
    - over 1M hours for hazard map calculations
  - GridChem
    - computational chemistry
  - Robetta
    - protein structure prediction using David Baker's award winning Rosetta code
- Up and coming groups with large awards
  - SIDGrid, 1M hours
- Number of end gateway users is down though
  - manual process, some under-counting
    - GridShib will solve this in 2009
  - end of ITR funding for many gateways
  - reduced class use by LEAD

TeraGrid™

The Starting Point for TeraGrid Users

- Enhanced user access
  - access RP resources and special multi-site services

- Information sharing & RP integration
  - current, up-to-date information about TG environment

- Accounting and monitoring
  - manage and monitor progress via common tools

- User information and presentation
  - first line of support for users
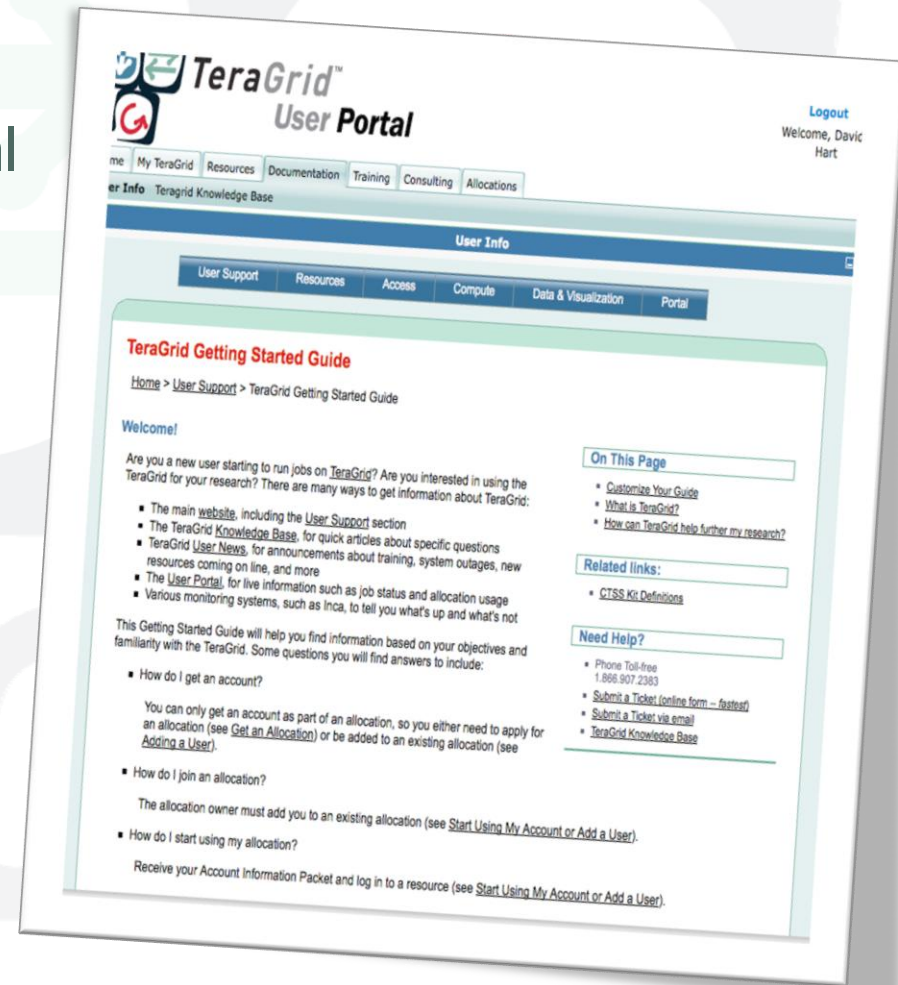
Enhanced User Access

- TeraGrid User Portal is primary common access point for TeraGrid users.
  - **https://portal.teragrid.org/**
- TeraGrid provides an environment that spans RP sites to simplify access to resources.
- Includes single-sign-on (SSO) capabilities to all TG resources.
  - Login to TGUP, one-click GSI-SSH to TG systems

User Portal Plans

- Customization & personalization
  - personalized TGUP home page
  - domain views
  - user discussion forums
- Migrate to integrated backend for Portal, Web, and Wiki
  - migrate existing and add new capabilities
  - more monitoring, measurement for users
- IIS-based Software Catalog
  - combining CTSS/3rd-party software information

Jobs from the past 7 days up to Jan-23-2009

< 100,000   < 500,000   < 1,000,000   +1,000,000

TeraGrid™
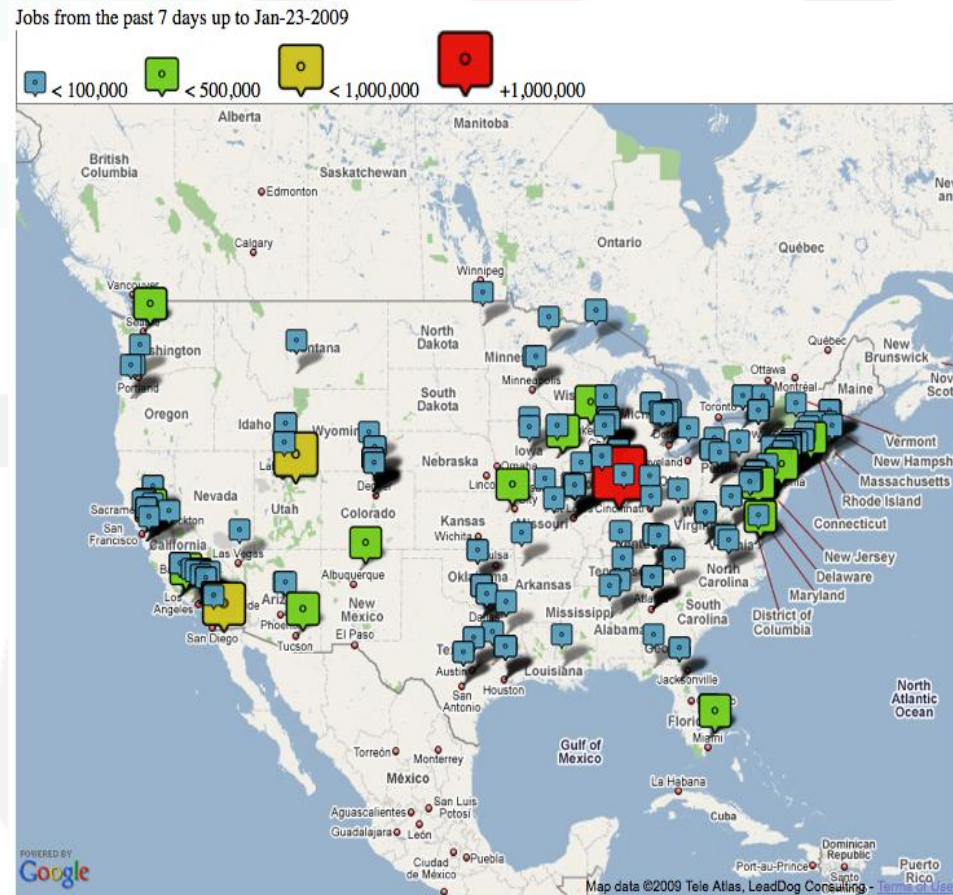
Allocations Updates

- Simplified and accelerated access for new users
  - allow (require) new users to create TGUP login
    - will allow us to eliminate POPS login, postal-mailed user packets
  - integration with Shibboleth
  - challenging and must be done to satisfy authentication requirements of RPs, other grids, security orgs.
- Continued improvements to POPS
  - reduced data entry by PIs for renewals, supplements
  - support for required submission documents
- Improved integration of POPS and TGUP for users

Advanced User Support

- Advanced Support for TeraGrid Applications (ASTA)
  - maintain ~25 ASTA projects throughout the year
  - contributions to Science Highlights
- Advanced Support for Projects (ASP)
  - "foundation work" to allow use of scientific codes and applications on the HPC resources.
  - projects with potential to benefit a number (10+) of users in a domain science or TeraGrid users in general
- Advanced Support for EOT (ASEOT)
  - content development and presentation; workshop, panel and BOF participation
  - outreach to new CI programs: DataNet, iPlant, etc.

Advanced Support for TeraGrid Applications (ASTA)

- Users request ASTA as part of quarterly TRAC
  - reviewed by TRAC and recommendation score provided
  - scores taken into account to select ASTA projects
  - other criterion
    - match of AUS staff expertise for ASTA
    - RP site(s) where PI has allocation
    - ASTA work plan
    - interest of PI, PI team

- Additional mechanisms for ASTA in Sept 2008:
  - Startup and Supplemental ASTA
  - scope of a regular User Services activity elevated to ASTA
  - PI requests ASTA at the next TRAC (support continues uninterrupted)

Number of ASTAs Started per Quarter - Trend
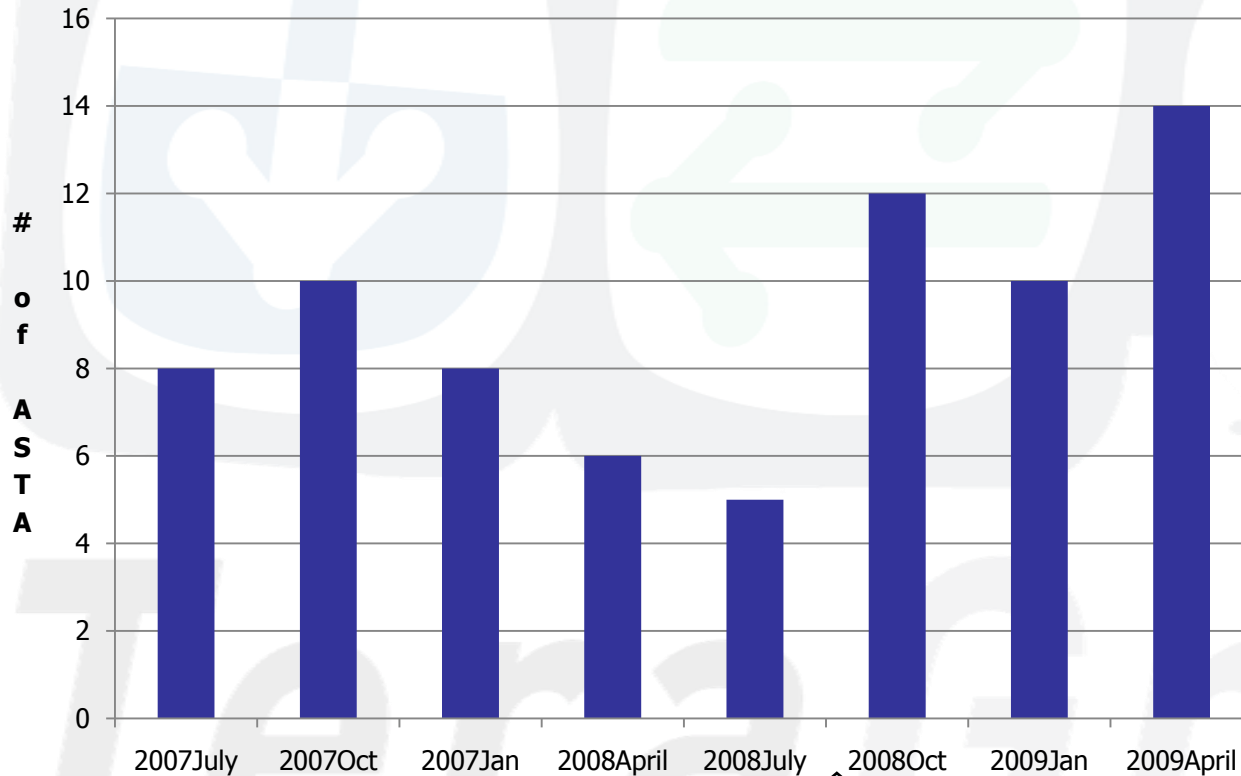
> Currently total of ~40 active ASTAs from last three TRACs, Startup/Supplemental, continuing from 2008

TeraGrid-wide AUS started

Outreach Highlights

- ## Campus Champions
  - grew to 30 institutions in less than a year (currently 39 institutions); in discussion with another ~30 campuses
  - attend User Services working group calls and offer recommendations for improved user services that have been adopted by TeraGrid

- ## TeraGrid Pathways
  - awareness sessions conducted at more than 12 professional society meetings
  - over 130 contacts with potential new under-represented users
  - provided User Services/AUS support to 12 new users, mentoring for 7 users, and fellowship awards to 4 research teams
  - worked with Science Gateways teams: Elizabeth City State University (ECSU) on PolarGrid, and Navajo Technical College on GEON

- ## TeraGrid Conference: TG'08 – 350 participants; TG'09 >400 participants, >120 students!

"The help of the TeraGrid champions has been invaluable both in writing and submitting the application of allocation and later, in resolving specific issues with software installation, correspondence with the TeraGrid IT support and a close follow up on my progress", Dr. Nurit Haspel, Rice University

Training

- **Multi-modal delivery for different learning styles and timelines**
  - expanded synchronous content delivery offered
  - on-line tutorials are reaching larger audiences with more topics
  - over 30 new training topics delivered (e.g. Ranger, Kraken)
- **HPC University in production**
  - catalog of training and education resources from NSF, DOE, etc.
  - calendar of HPC events
  - opportunities for educators and students to get involved in HPC

HPC University is working to fill critical gaps in HPC workforce preparation for students and professionals in all fields

Quality Assurance / Common User Environment

- **Quality Assurance Working Group**
  - expediting problem resolution when service failures detected
  - improving the use of the Inca monitoring framework
  - validate common user environment (with CUE WG)
  - develop/propose a more formal process for CTSS software deployment (with Software WG)

- **Common User Environment Working Group**
  - remove barriers to user movement between TeraGrid resources
  - working with RP staff and TG WG to implement:
    - CUE Documentation (CUED)
    - CUE Management System (CUEMS)
    - CUE Build Environment (CUBE)
    - CUE Testing Platform (CUETP)
    - CUE Variable Collection (CUEVC)

Technology Plans: Wide Area Filesystems

- Providing a TeraGrid global file system is a highly requested service
- Global file systems imply a single file system mounted on most TeraGrid resources
  - no solution currently exists for production global file systems
- Wide area file systems give the look and feel of a single file system. Possible with technologies such as GPFS-WAN or Lustre-WAN
  - GPFS-WAN has licensing issues and isn't available for all platforms
  - Lustre-WAN is preferable for both licensing and compatibility reasons
- pNFS is a possible path for global file systems, but is far away from being viable

Lustre-WAN Progress

- Initial production deployment of Indiana's Data Capacitor Lustre-WAN on IU's BigRed, PSC's Pople
- Successful testing and commitment to production on LONI's QueenBee, TACC's Ranger/Lonestar, NCSA's Mercury/Abe, and SDSC's IA64
  - additional sites (NICS, Purdue) will begin testing this year
- Significant work to improve performance and authentication infrastructure
  - work in parallel with production

Technology Plans: Retiring Resources

- **Retiring Resources and TeraGrid Extension**
  - all non-Track 2 resources slated to discontinue as TeraGrid resources after March 31, 2010
  - TeraGrid Extension proposal pending
    - requests continued support for several resources
      - 12 months: Abe, QueenBee, Lonestar, Steele, Quarry, Lincoln
      - 3 months: Pople
  - orderly transition of users to resources with ongoing support

TeraGrid → TeraGrid Extension (?) → XD Transition Planning

- All activity areas have effort reserved for TeraGrid → XD transition effort as appropriate
  - transition issues exist for nearly all areas
    - effort ear-marked to support transition issues
    - more was learned about timing after PY5 IPP was developed
- Start of XD for CMS/AUSS/TEOS deferred for one year (1 April 2011)
  - induced TeraGrid Extension Proposal
    - 12-month funding to support most GIG functions and some non-Track 2 RP resources
  - uncertainty in sequence of events
  - still need to address many changes in TeraGrid going into presumed Extension Period
    - many resources exit Teragrid
- PY5 planning process will *likely* need to address:
  - TeraGrid Extension following PY5
  - necessarily include transition to XD in extension period

Areas for Potential OSG-TG "Harmonization"

- **Improving Operations:**
  - Grid clients:
    - common grid client distribution
      - TeraGrid Client Toolkit work sub-awarded to Univ of Wisconsin; possible to build common client toolkit
  - Security:
    - common configuration for security (CAs)
  - Information Services:
    - align information services (class-ads, service registries) so that users don't have to use different configurations even when they have the same client software
    - identify common use cases and requirements
    - consider information services interoperability, shared functionality or implementations
      - enable OSG and TeraGrid users to more easily discover and access capabilities, software, and services across both infrastructures
  - User Support:
    - possible to collaborate in user support activities?
    - cross participation in conference calls? more?

TeraGrid™

Areas for Potential OSG-TG "Harmonization" (2)

- **Interoperability for a larger virtual cyberinfrastructure**
  - Job Gateways:
    - [need to affirm assumption that this is the right way to go…]
    - multiple activities under way; multiple approaches; insufficient coordination
    - Indiana, LONI, NCSA, Purdue
  - Coordinated Applications Support
    - identified set of candidate applications that can benefit from coordinated use of TG and OSG resources
    - develop OSG-TG joint teams to work with applications
    - learn from these activities to inform interoperability needs
    - SCEC, GridChem, Robetta
  - Allocations, Account/Identity Management, Accounting
    - multiple opportunities here; very different current approaches
    - may provide some of the highest payoff activities both operationally and for the user community

Areas for Potential OSG-TG "Harmonization" (3)

- **Education, Outreach, Training**
  - Some Good Things happening already
    - HPC University, cross participation in training events, iSGTW
  - Campus Outreach
    - seems much of current activity is complimentary; can we define join program of activities for leverage
    - CI Days: what status does this have for each of us?
  - Workforce development
    - initial joint whitepaper focused on student development
      - hopefully will produce something from NSF
    - more coordination of student programs

"Bigger Picture" Collaboration

- **Some groundwork has been done**
  - Pordes, Livny, Towns authored a white paper submitted to NSF/DOE in Jan 2009
  - ongoing discussions with various folks from OSG, TG, NSF, DOE and some other CI providers
  - fruitful discussion during July 30 meeting
    - "shared principles" document emerging

- **national federation of cyberinfrastructures**
  - collection of major CI providers
  - initiate a national (international) discussion of the future path for CI in the US (internationally)
  - establish a coherent framework across all cyberinfrastructure programs
  - discussion of "in the trenches" issues to make things work for researchers

"Shared Principles"

- Enact ways to collaborate at the management and implementation levels
  - cross project teams to interact of defined activities
    - potential funding opportunities?
  - cross involvement in management meetings
    - ex-officios at TG Quarterly Meeting, OSG Exec Council meetings
- Develop a clearer understanding of each other's organization, purpose and scope
  - present organizations to one another
  - cross participation in middle-management layer
  - understand the definition/scope of the communities we serve and understand their intersection
- Develop ways to access resources managed by TG and accessible through OSG via a unified process and single mechanism
  - allocations policies
  - allocations procedures

TeraGrid™

"Shared Principles" (2)

- Explore how education, outreach, training, and external relations activities can be coordinated to maximize leverage and impact
  - campus outreach coordination meeting
    - OSG Engagement, OSG Campus Grid, TG Campus Champions, TG Advanced User Support
  - OSG and TG's External Relations coordination
- Work together to more effectively support the needs of research teams who rely on TG and OSG resources
  - cross-project teams (mentioned earlier)
  - explore joint support for over the next six months
- We want to establish activities that will contribute to the success of the fu-ture XD program and provide continuity as we transition from today's TG - OSG to the future XD – OSG' era.
  - OSG will interact equivalently with all XD proponents

TeraGrid™