SG - OSG

Improving Campus Research CI Through Leveraging and Integration:

Developing a SURAgrid-OSG Collaboration

John McGee, RENCI/OSG Engagement Coordinator

John-Paul Robinson, UAB/SURAgrid Governance Committee Co-Chair

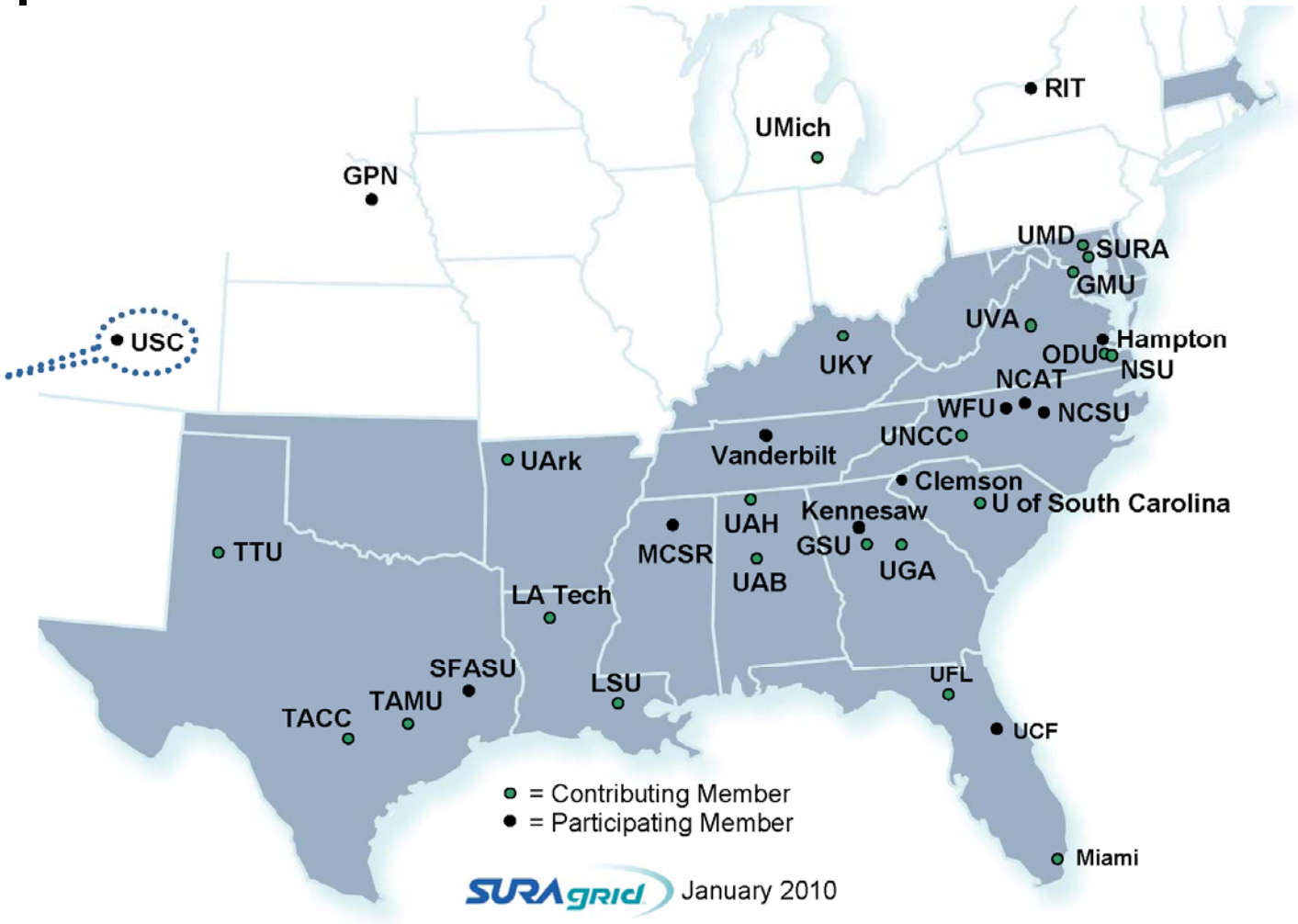
Agenda

- Collaboration History
- SURAgrid Motivation
- OSG Motivation
- Future Directions

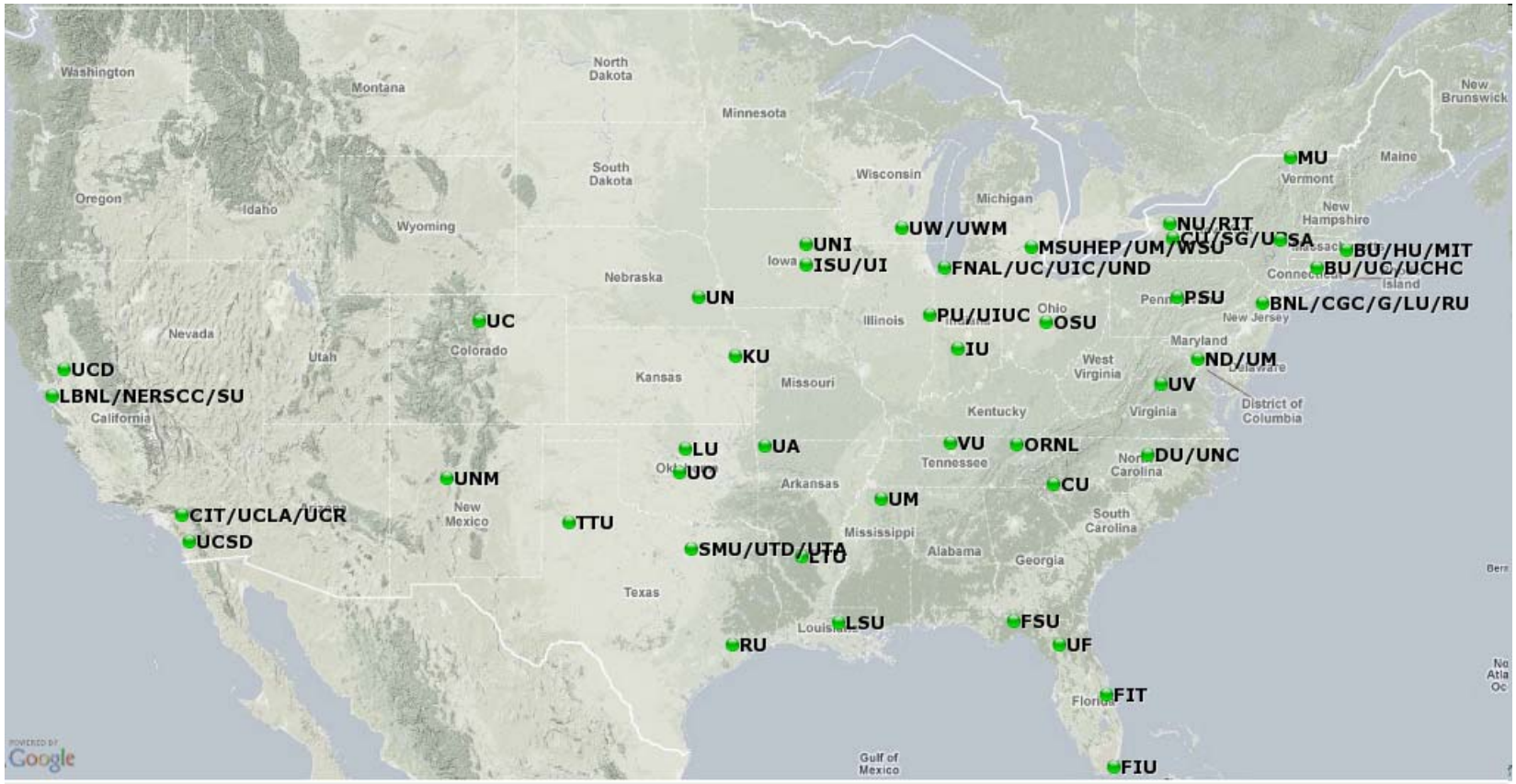
SURAgrid

Open Science Grid

SURAgrid


RIT
UMich
GPN
UMD
SURA
GMU
UVA
Hampton
ODU
NSU
NCAT
USC
UKY
WFU
NCSU
UNCC
Vanderbilt
UArk
Clemson
U of South Carolina
Kennesaw
UAH
GSU
UGA
MCSR
UAB
TTU
LA Tech
SFASU
LSU
UFL
TAMU
TACC
UCF
= Contributing Member
= Participating Member
Miami
SURAgrid
January 2010

Open Science Grid



Collaboration History

- Leveraging of the Work of Established Grid Initiatives and Standards Groups
- SURAgrid TTU-led Packaged Stack Development based on OSG SW Stack (VDT)

  http://omnius.hpcc.ttu.edu/SURAgrid_wiki/ServerStack
- Grid Technology Cookbook co-sponsors
  http://hv3.phys.lsu.edu:8000/cookbook/gtcb/index.php
- 2009 SURAgrid-OSG Statement of Shared Interest

SURAgrid

Open Science Grid

SURAgrid Services Profile

- Governance Committee
  - Coordinate Member Activities and Identify Needs
  - Prioritize and Align Agenda with Strategic Plan
  - Provide Cross-site Engagement to Develop Operations
- CA and Bridge CA
  - Provide “catch-all” grid credentials
  - Integrate campus CA into trust network
- Server Stack
  - Based on VDT
  - Direct Support for SURAgrid-specific Configuration
- Operations & Support Services
  - Mailing Lists: Announce, Support, Working Group
  - Grid Portal: Basic job submit & Resource Monitoring
  - LDAP-based Member Registry
  - Accounting Data using OGF UR-WG Profile

SURAgrid

Open Science Grid

Successes and Challenges

- Successes
  - Campus Engagement
    - Interest in and Need for Resource Sharing is Strong
  - Model for Engagement
    - Membership and Governance Provide Voice
  - Identifying Campus Requirements
    - Found the “gap” in CI
- Challenges
  - Moving from Local to Remote
  - Matching Users with Compute Cycles
  - Managing Ownership of Core Resources
  - Maintaining Resource Availability
  - Insufficient Service Profile

SURAgrid

Open Science Grid

SURAgrid Motivation

- Success with organizing, community building, engaging campus
  - Successful demonstrations
  - Hands on training, peer learning
- Threshhold of Operations
  - Strong desire to move from test and demo mode to operational mode
- Desire not to reinvent successful model
- Have folks come on board and say “hey this is great” but then have a hard time transitioning to operation

SURAgrid

Open Science Grid

SG-OSG & VOMS

- Domain Science Community Practices
  - OSG has proved successful as an organizing framework for domain science communities
  - Communities typically adopt the OSG tool chain and build their practices around it
  - Scientists “join” the OSG science VO and accept whatever practices are deemed necessary to accomplish that participation

- Campus Practices
  - Potential for a campus to simply adopt a solution like VOMS is limited
  - Integrating campus source systems with VOMS is the best way to allow the site to express its membership information and roles in the context of OSG
  - Current organizing practices lead sites to maintaining the integration independently
  - Thin staffing means integration rarely occurs and campuses remain without the ability to join a larger community of resource sharing

The Open Science Grid

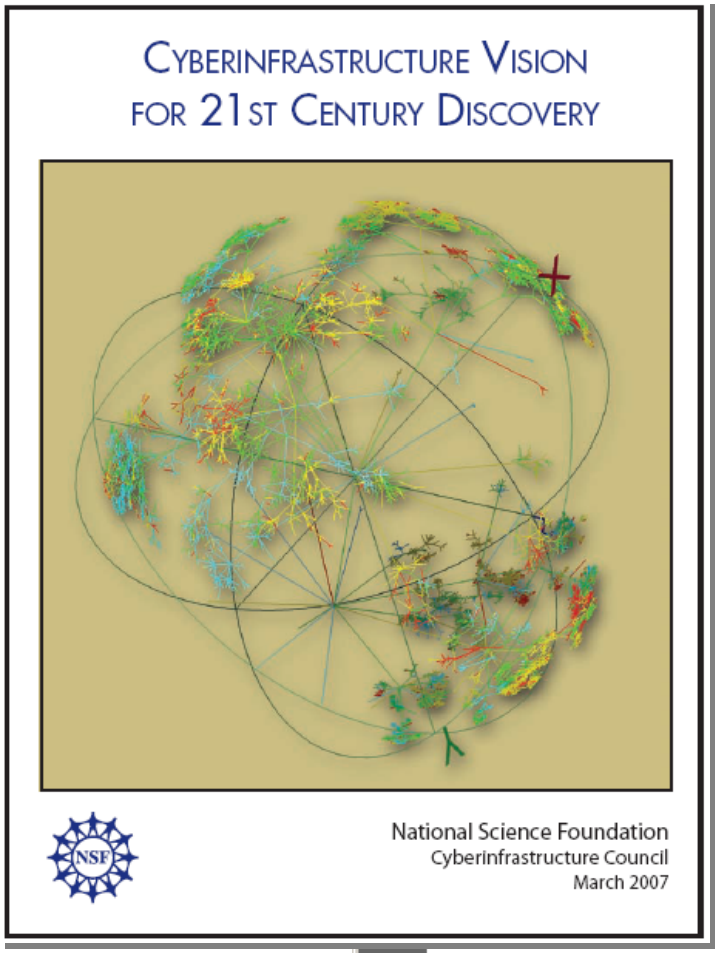
A framework for large scale distributed resource sharing

addressing the technology, policy, and social requirements of sharing

OSG is a consortium of software, service and resource providers and researchers, from universities, national laboratories and computing centers across the U.S., who together build and operate the OSG project. The project is funded by the NSF and DOE, and provides staff for managing various aspects of the OSG.

Brings petascale computing and storage resources into a uniform grid computing environment

Integrates computing and storage resources from over 80 sites in the U.S. and beyond



CI_Vision_March07.pdf - Adobe Acrobat Professional

File Edit View Document Comments Forms Tools Advanced Window Help

itiative. NSF will

elopment and

An effective computing environment designed to meet the computational needs of a range of science and engineering applications will include a variety of computing systems with complementary performance capabilities. NSF will invest in leadership class environments in the 0.5-10 petascale performance range. Strong partnerships involving other federal agencies, universities, industry and state government are also critical to success. NSF will also promote resource sharing between and among academic institutions to optimize the accessibility and use of HPC assets deployed and supported at the campus level. Supporting software services include the provision of intelligent development and problem-solving environments and tools. These tools are designed to provide improvements in ease of use, reusability of modules, and portable performance.

---

[1] A picosecond is 10^{-12} second

[2] A petascale is 10^{15} operations per second with comparable storage and networking capacity

e density for iron oxide (FeO) within the local density approximation, with ors represent the spin density, showing the antiferromagnetic ordering.

8.50 x 11.00 in

OSG Engagement Program

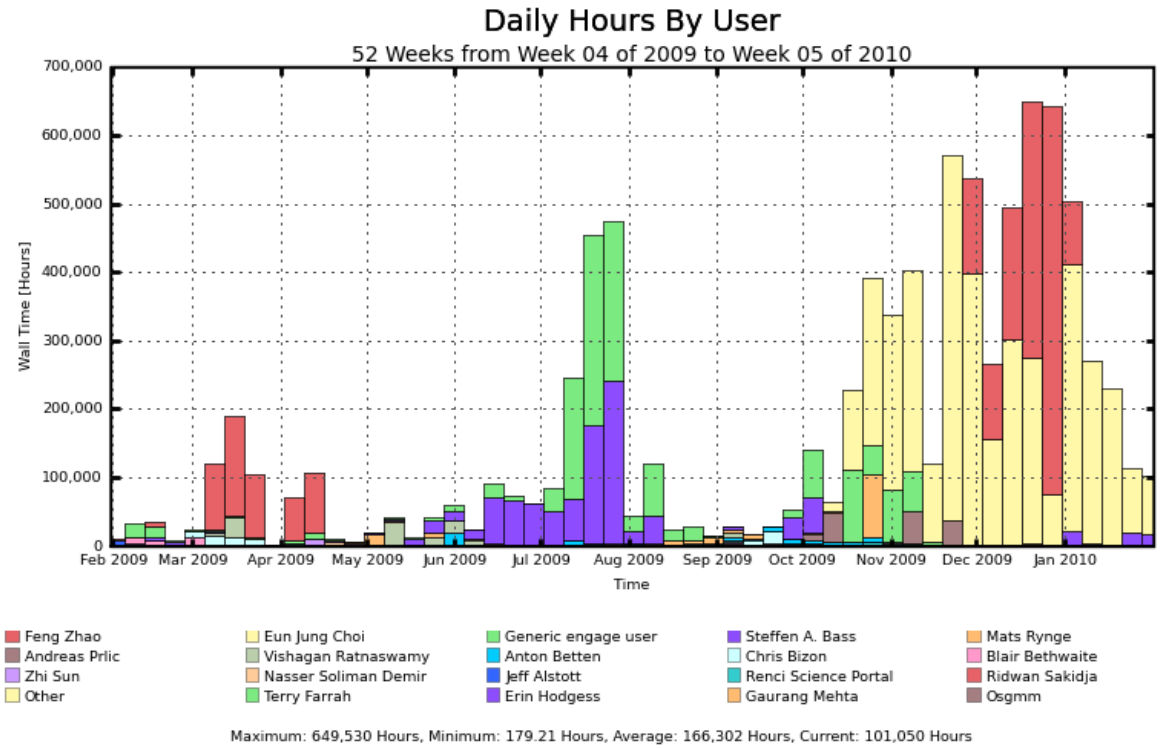
- OSG Satellite Project
  - funded by NSF CI-TEAM
- Engagement Program Objectives
  - Develop new communities of CI enabled researchers by helping new users run scientific codes at OSG scale
  - Help University campuses leverage OSG experience and methodologies to implement campus level resource sharing
  - Use the experience from the two objectives above to drive enhancements and new capabilities into OSG

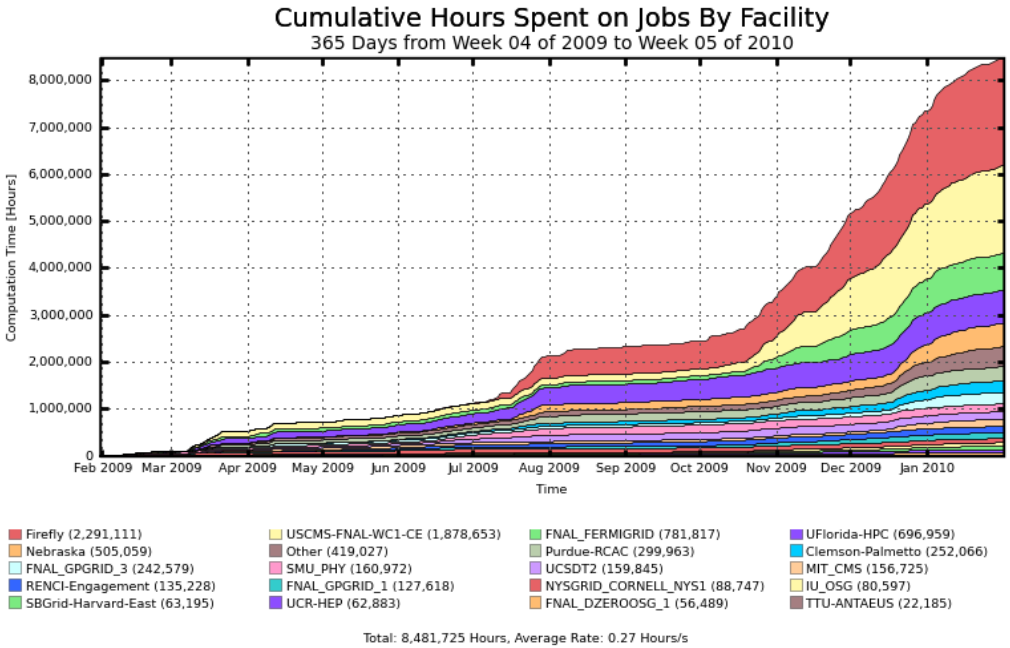
SURAgrid

Open Science Grid

User Engagement

stats for the
OSG Engagement Program
over the previous 12 months





Biochemistry, Genetics, Structural Biology, Materials Science, Nuclear Physics, and more …

running at LHC/HEP scale

OSG Campus Engagement

- Successful when there is campus leadership with a national cyberinfrastructure agenda:
  - Purdue, Clemson, Wisconsin, NY State, Nebraska
  - very difficult otherwise

- Interesting comparison to Campus Shared Cluster model (aka condominium computing)
  - broadly successful resource sharing at the campus level
  - typically initiated by central IT/campus research computing; OSG resource sharing on campus is initiated by domain scientists
  - “safe”: stays within campus borders
  - cluster management is well understood; mgmt of shared resources across multiple resources and owners less so
  - https://twiki.grid.iu.edu/bin/view/CampusGrids/WorkingMeetingFermilab

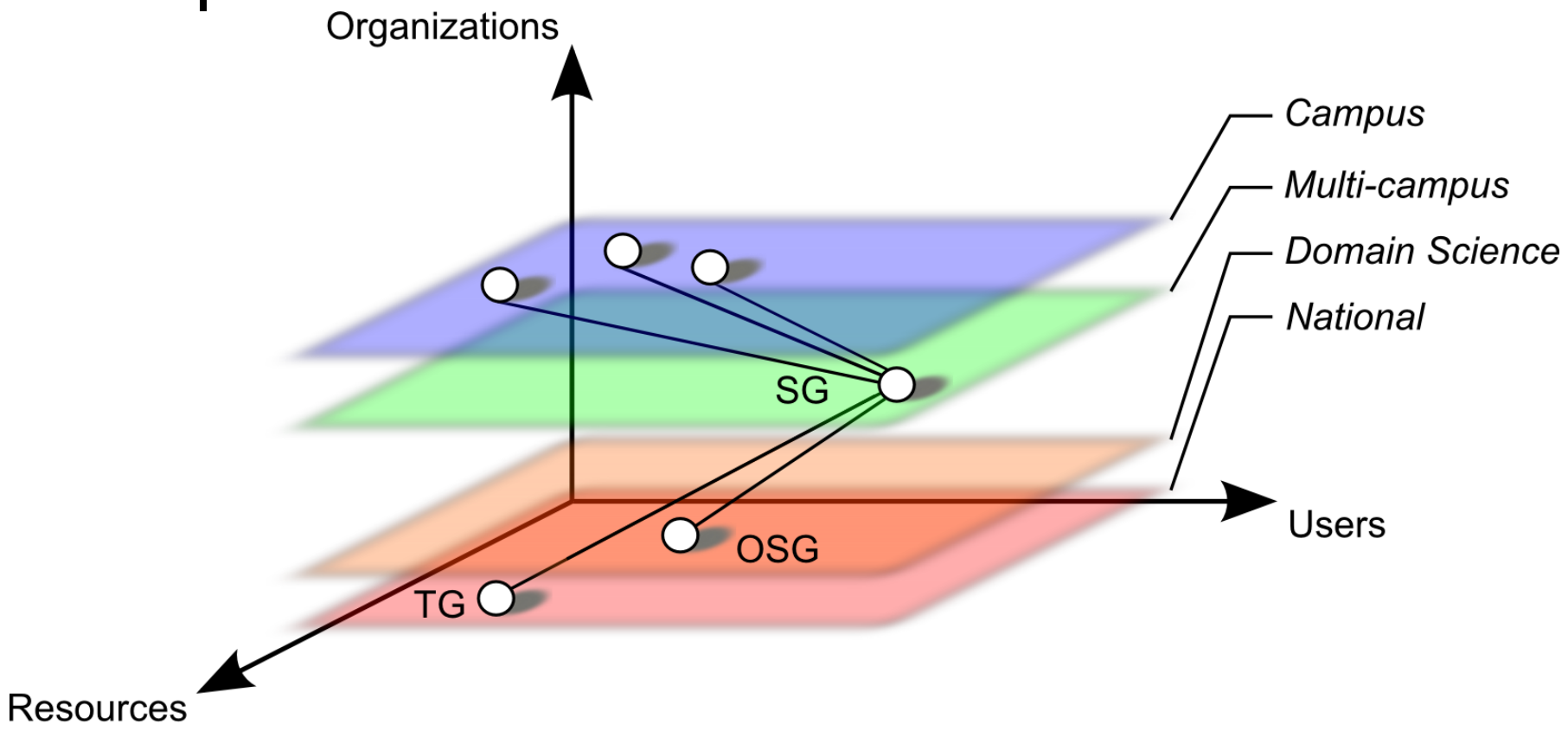
Future Directions

- Enormous unrealized potential
  - current and future campus resources
    - eg Purdue Condor backfill ~ 15% of capacity
    - the original promise of Grid/utility computing
  - cultural and sociological changes that come with a resource sharing ecosystem
  - campus identity and role management leveraged into large multi-campus science efforts
  - aggregated purchasing power for commercial cloud services
  - need for reusable integration and human effort sharing
  - SURAGrid campus expertise + OSG cross organization operational expertise

SURAgrid

Open Science Grid

Coherent CI Landscape



TG=TeraGrid, OSG=Open Science Grid, SG=SURAgrid

Image taken from "Enabling and Sustaining Campus-to-Campus CI" White Paper