June 26, 2012

XSEDE and OSG

Kim Dillman

kadillma@purdue.edu

Alain Roy

roy@cs.wisc.edu

What is OSG?

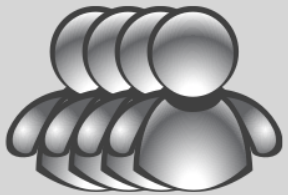
- A multi-disciplinary partnership to federate local, regional, community and national cyberinfrastructures to help share computing and storage resources of research and academic communities at all scales
- We provide common services and support for more than 100 resource providers and scientific institutions using a distributed fabric of high-throughput computational services
- We do not own computing resources but instead provide software and services to users and resource providers to enable the effective use and sharing of their resources based on the principles of Distributed High Throughput Computer (DHTC)

OSG's Community Focus

Communities...

(LHC, LIGO, SBGrid, Campuses...)

... have **users**, application developers, site administrators, researchers, students etc:



... may use custom **services** in addition to OSG's:

Job management
Data management
...

...may provide diverse **resources** via one or more autonomous sites:

Compute Interface	Storage Interface
Site Fabric	System Status Accounting Security

OSG

Consultation Services
- Community
- Security
- Architecture
- Education/Training

Operational Services
- Front-line support
- Security
- Info/monitoring
- WLCG Interaction
- IT Services

Software Services
- Software Distribution
- Software Evaluations
- Support

There is a sharing of software, operational services, and knowledge between the communities and OSG in each of these areas.

Where is OSG?

Open Science Grid

XSEDE

How big is OSG?

From http://display.grid.iu.edu/

OSG delivered across 110 sites

In the last 24 Hours

439,000	Jobs
1,824,000	CPU Hours
2,344,000	Transfers
1,001	TB Transferred

In the last 30 Days

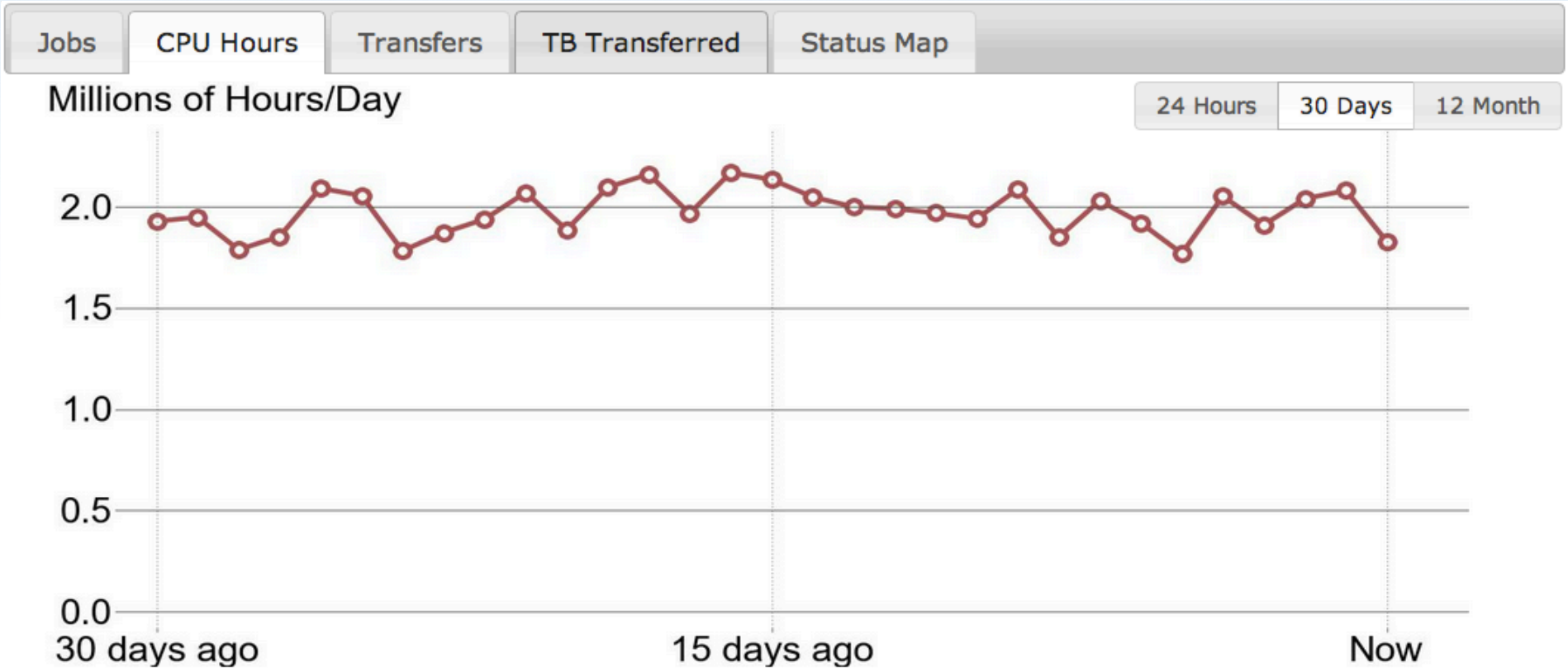
14,097,000	Jobs
61,271,000	CPU Hours
47,775,000	Transfers
27,242	TB Transferred

In the last Year

220,399,000	Jobs
649,521,000	CPU Hours
599,875,000	Transfers
274,284	TB Transferred

How big is OSG?

From http://display.grid.iu.edu/

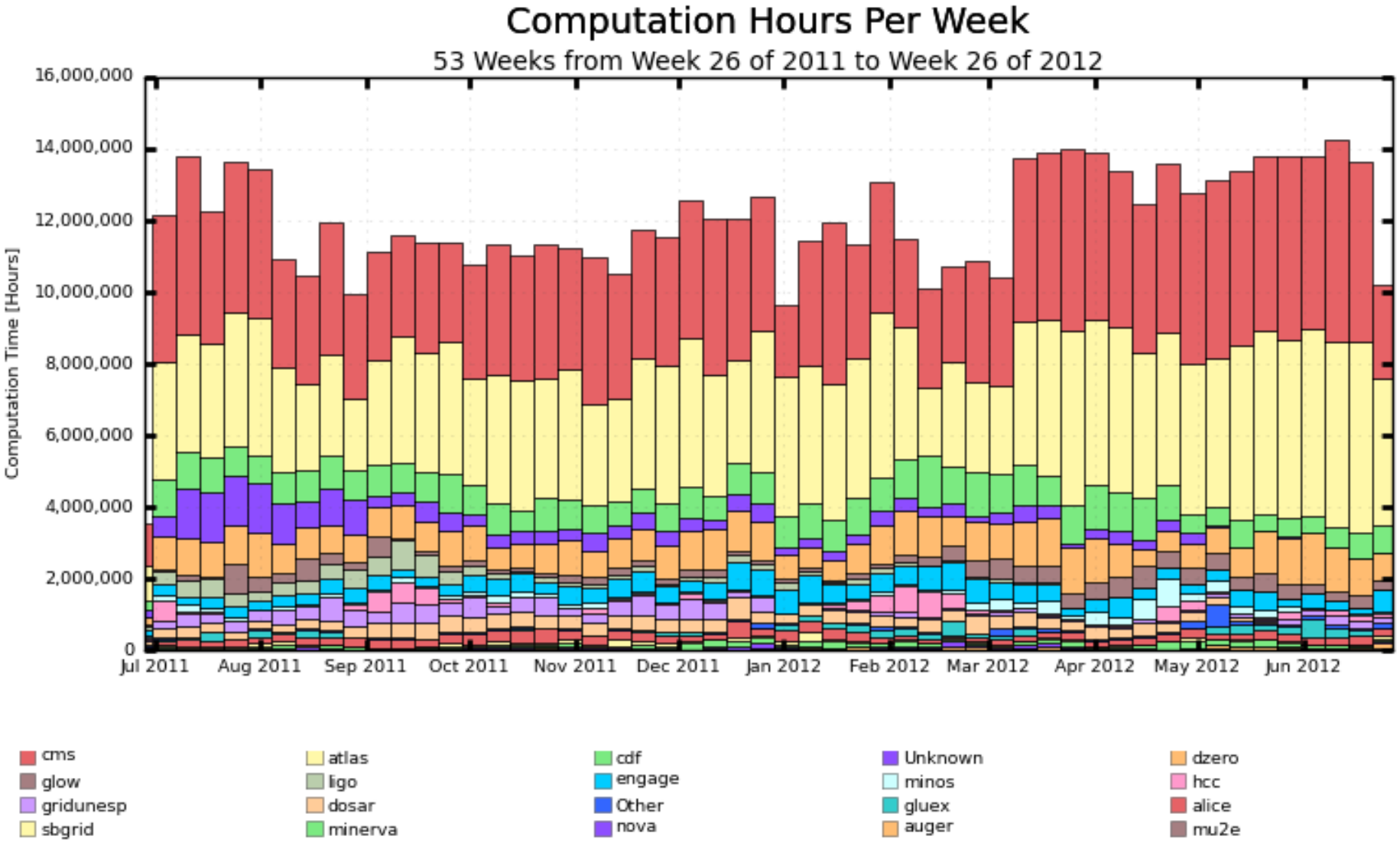


CPU hours spent on an OSG resource are reported to the central accounting system. The above graph shows the number of CPU hours per day. A total of 61,271,000 CPU hours were spent.

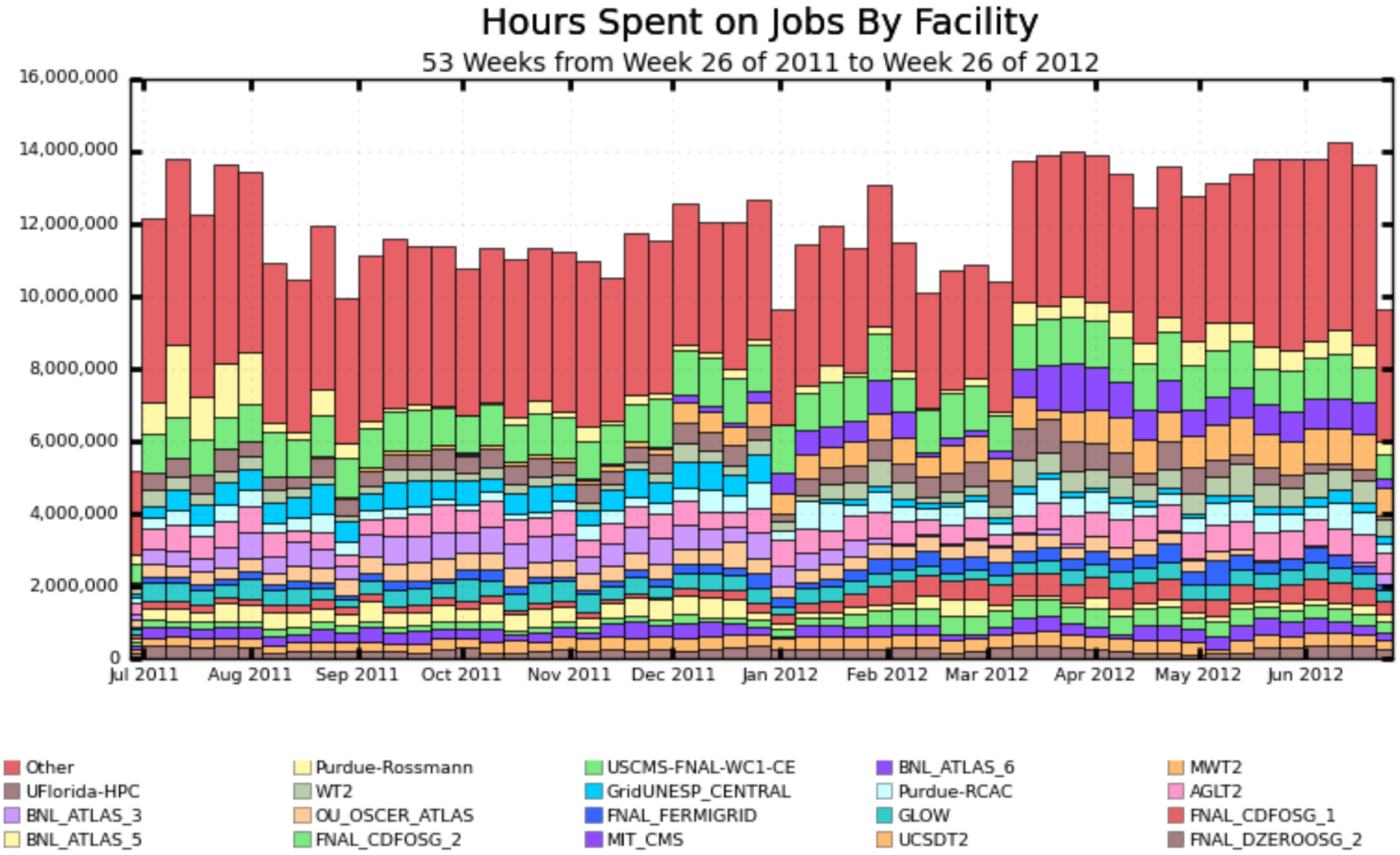
Open Science Grid

XSEDE

Who uses OSG?


Computation Hours Per Week
53 Weeks from Week 26 of 2011 to Week 26 of 2012
Computation Time [Hours]
16,000,000
14,000,000
12,000,000
10,000,000
8,000,000
6,000,000
4,000,000
2,000,000
0
Jul 2011
Aug 2011
Sep 2011
Oct 2011
Nov 2011
Dec 2011
Jan 2012
Feb 2012
Mar 2012
Apr 2012
May 2012
Jun 2012
cms
atlas
cdf
Unknown
dzero
glow
ligo
engage
minos
hcc
gridunesp
dosar
Other
gluex
alice
sbgrid
minerva
nova
auger
mu2e
Maximum: 14,281,694 Hours, Minimum: 3,520,880 Hours, Average: 11,930,397 Hours, Current: 10,192,162 Hours

Who contributes to OSG?



How do you use OSG?

Step 1: Make sure you can use HTC

- Do you have a problem that needs HTC?
- Do you have local computing resources?
  - Use them first
  - You get local support
  - You gain experience with HTC
- When you need more resources, OSG is a good option to grow

Open Science Grid

XSEDE

How do you use OSG?

Step 2: You must be in a Virtual Organization (VO)

- A VO is:
  - A collection of people and/or resources with a common purpose
  - Usually a scientific collaboration
    People + their computing & storage
- Which VO?
  - About 60 VOs in OSG: you might already be affiliated
  - Make your own for your collaboration
  - Join the “OSG” VO: miscellaneous

Open Science Grid

XSEDE

How do you use OSG?

Step 3: Get access to glidein VO Frontend

- OSG runs a glidein factory, but you need a submission point (i.e. condor_schedd) Called a *VO Frontend* (see Igor's talk)
- Larger VOs set up their own frontend
- Smaller VOs can use an OSG supplied frontend
  - We're in a transition period right now
  - This will probably be hosted at UCSD in the near future
- Sites need to authorize your VO

How do you contribute to OSG?

- Usually contribution is based on VO needs
  - Scientific collaboration wants to share within the collaboration: OSG provides the tools
  - Other VOs can utilize unused capacity
- Some sites are simply generous and share
- Install and maintain the OSG software on:
  - Front end job submission computer
  - Storage interface computer (optional)
  - Authorization service (optional)
- Everyone is welcome, but it takes effort to do it (25% of a person's effort?)

Autonomy

- OSG provides:
  - Software
  - Support
  - Central Services
- Sites and VOs are autonomous:
  - They make decisions about their sites
  - They decide when to install, upgrade
  - They make operational decisions
- This is HTC: if we mandate less, more people will join and we will have access to more resources

What is XSEDE?

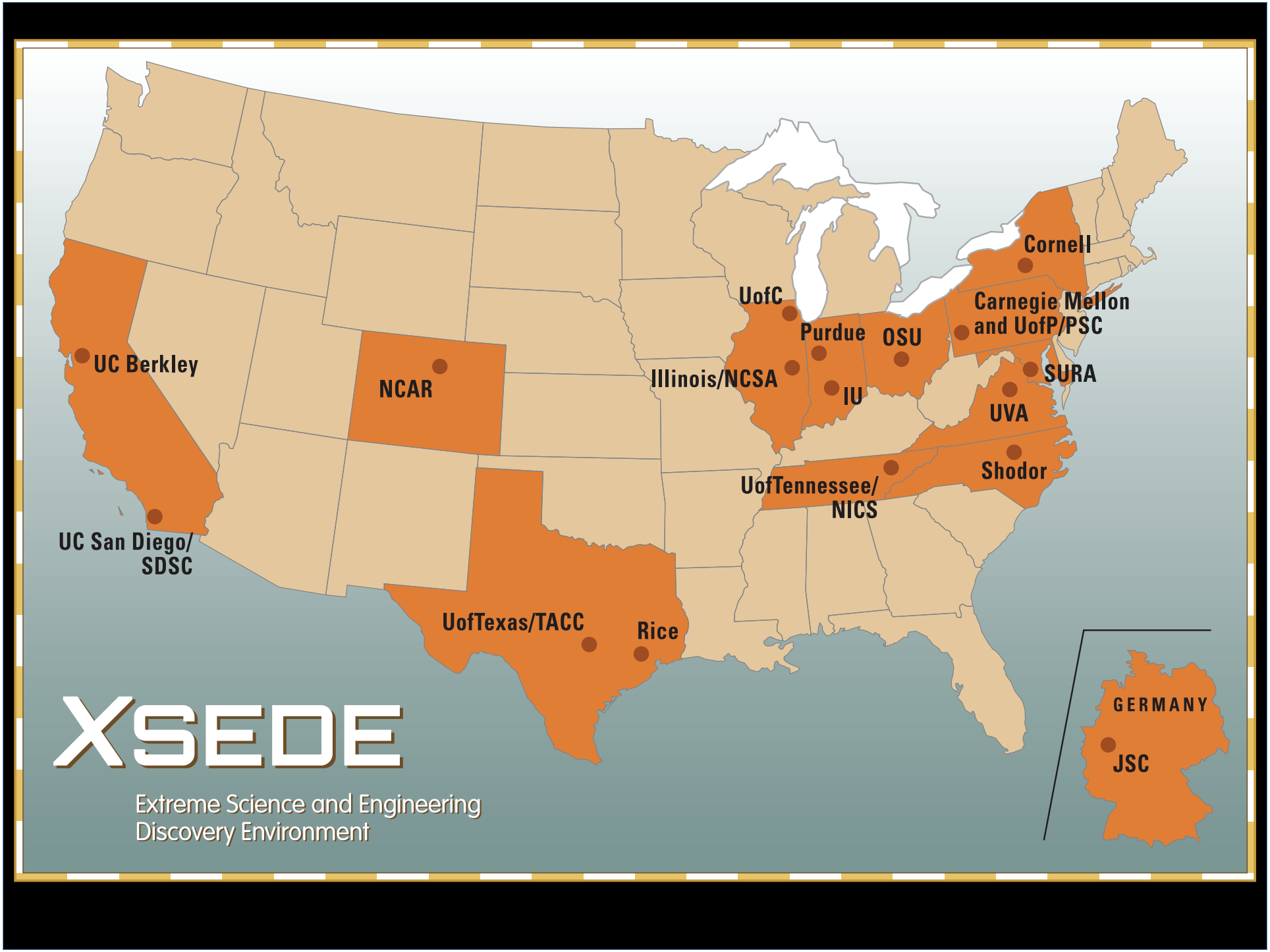
XSEDE is a comprehensive set of advanced, heterogeneous high-end digital services, integrated into a general-purpose infrastructure.

Open Science Grid

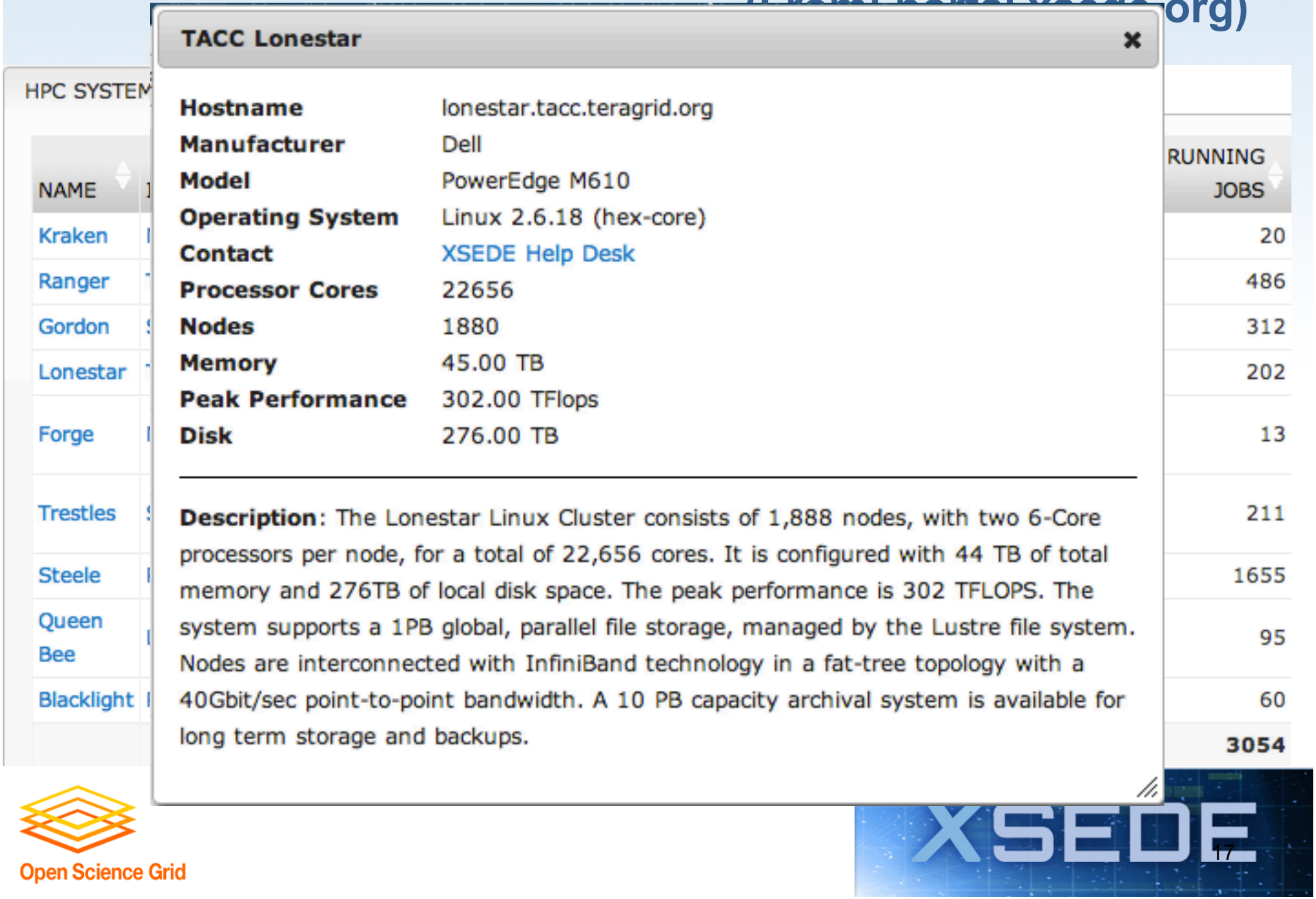
XSEDE

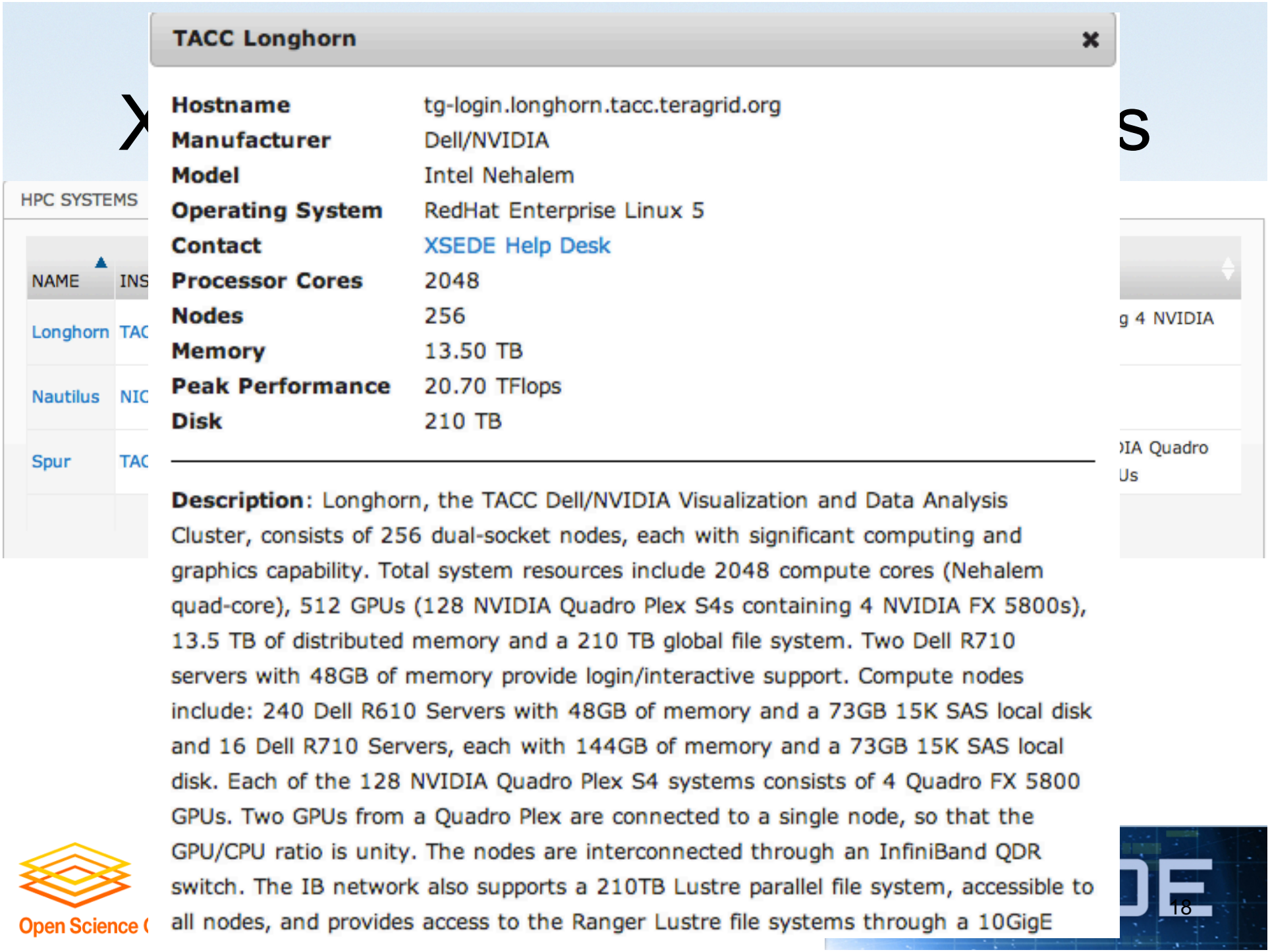
NSF eXtreme Digital (XD) Program

- NSF's Transition from TeraGrid to XD
- High-Performance Computing and Storage Services
- High-Performance Remote Visualization and Data Analysis Services
- eXtreme Digital Resources for Science and Engineering (XSEDE)


Cornell
UofC
Carnegie Mellon
and UofP/PSC
Purdue
OSU
UC Berkley
NCAR
Illinois/NCSA
IU
SURA
UVA
Shodor
UofTennessee/
NICS
UC San Diego/
SDSC
UofTexas/TACC
Rice
GERMANY
JSC
XSEDE
Extreme Science and Engineering
Discovery Environment

(From portal.xsede.org)




TACC Longhorn
Hostname tg-login.longhorn.tacc.teragrid.org
Manufacturer Dell/NVIDIA
Model Intel Nehalem
Operating System RedHat Enterprise Linux 5
Contact XSEDE Help Desk
Processor Cores 2048
Nodes 256
Memory 13.50 TB
Peak Performance 20.70 TFlops
Disk 210 TB
Description: Longhorn, the TACC Dell/NVIDIA Visualization and Data Analysis Cluster, consists of 256 dual-socket nodes, each with significant computing and graphics capability. Total system resources include 2048 compute cores (Nehalem quad-core), 512 GPUs (128 NVIDIA Quadro Plex S4s containing 4 NVIDIA FX 5800s), 13.5 TB of distributed memory and a 210 TB global file system. Two Dell R710 servers with 48GB of memory provide login/interactive support. Compute nodes include: 240 Dell R610 Servers with 48GB of memory and a 73GB 15K SAS local disk and 16 Dell R710 Servers, each with 144GB of memory and a 73GB 15K SAS local disk. Each of the 128 NVIDIA Quadro Plex S4 systems consists of 4 Quadro FX 5800 GPUs. Two GPUs from a Quadro Plex are connected to a single node, so that the GPU/CPU ratio is unity. The nodes are interconnected through an InfiniBand QDR switch. The IB network also supports a 210TB Lustre parallel file system, accessible to all nodes, and provides access to the Ranger Lustre file systems through a 10GigE
HPC SYSTEMS
NAME
Longhorn
Nautilus
Spur
Open Science
18

XSEDE Storage Systems

HPC SYSTEMS | ADVANCED VIS SYSTEMS | STORAGE SYSTEMS | HTC SYSTEMS

NAME	INSTITUTION	SYSTEM	ONLINE STORAGE TB	OFFLINE STORAGE TB
Albedo Lustre-WAN	PSC	Lustre	1024	N/A
Data Capacitor	IU	Lustre	535	N/A
Data Replication Service	TACC	iRODS	N/A	1024
HPSS	NICS	HPSS	12	6220
NCSA Tape Storage	NCSA	EMC DiskXtender	250	10000
Ranch	TACC	Sun StorageTek Mass Storage Facility	110	40000
		Total:	**1931**	**57244**

Open Science Grid

XSEDE

HPC SYSTEMS

NAME	ORY TBYTES
Condor Pool	54.74
	54.74

Purdue Condor Pool

Hostname	tg-condor.purdue.teragrid.org
Manufacturer	Intel
Model	Linux Cluster
Operating System	Unknown
Contact	XSEDE Help Desk
Processor Cores	N/A
Nodes	N/A
Memory	N/A
Peak Performance	N/A
Disk	N/A

Description: The Purdue Condor pools consist of over 4500 CPU of computation: 1645 LINUX/X86_64 CPUs, 1192 LINUX/INTEL (ia32) CPUs, and 1815 WINNT51/INTEL CPUs, as well as a small number of Itanium Linux, Solaris and MacOS X machines. Memory on compute nodes range from 512 MB to 16 GB, and most CPUs run at 3 GHz or better. With a total of approximately 150 TFLOPS available, the Purdue Condor pools can provide large numbers of cycles in a short amount of time. All shared areas and software packages available on Lear are available on Condor. Condor is designed for high-throughput computing, and is excellent for parameter sweeps, Monte Carlo simulation, or most any serial application. Also, some classes of parallel jobs (master-worker) may be run in Condor.

XSEDE Vision

The eXtreme Science and Engineering Discovery Environment (XSEDE):

> enhances the productivity of scientists and engineers by providing them with new and innovative capabilities.

Thus, XSEDE:

> facilitates scientific discovery while enabling transformational science and engineering, and innovative educational programs.

XSEDE fulfills this vision by creating an advanced, capable, and robust cyberinfrastructure supported by the combined expertise of a distributed team of leading CI (cyberinfrastructure) professionals.

Open Science Grid

XSEDE

XSEDE Characteristics:

- XSEDE forms the foundation of a national cyberinfrastructure (CI) ecosystem
  - Its comprehensive suite of advanced **digital services** work with other high-end facilities and campus-based resources
- XSEDE integrates diverse digital resources
  - Its open architecture allows continued addition of new technology capabilities and services

Open Science Grid

XSEDE

XSEDE is about . . . .

- **Increasing productivity**
  - leading to more science
  - making the difference between a feasible project and an impractical one
- **Transformative impact through**
  - active, formal requirements gathering processes to
    - understand the needs of the community
  - new and expanded extended collaborative support that includes
    - external short-term contracting for expertise beyond the current team
    - Novel and Innovative Projects: supports novel science areas, demographic diversity, innovative technologies, science gateway development, data repositories, and campus bridging

Open Science Grid

XSEDE

And…

- National Training and Education and Outreach programs with the scope and scale to:
  - increase diversity of topics, modes of delivery, and reach to new communities and audiences
  - broaden participation among under-represented communities
  - campus bridging for effective use of CI (cyberinfrastructure) resources
    - integrate with campuses through expanded Champions program and additional bridging activities
  - establish academic certificate and degree programs
    - institutional incorporation of CS&E curricula; professional development certificate
    - prepare undergraduates, graduates and future K-12 teachers

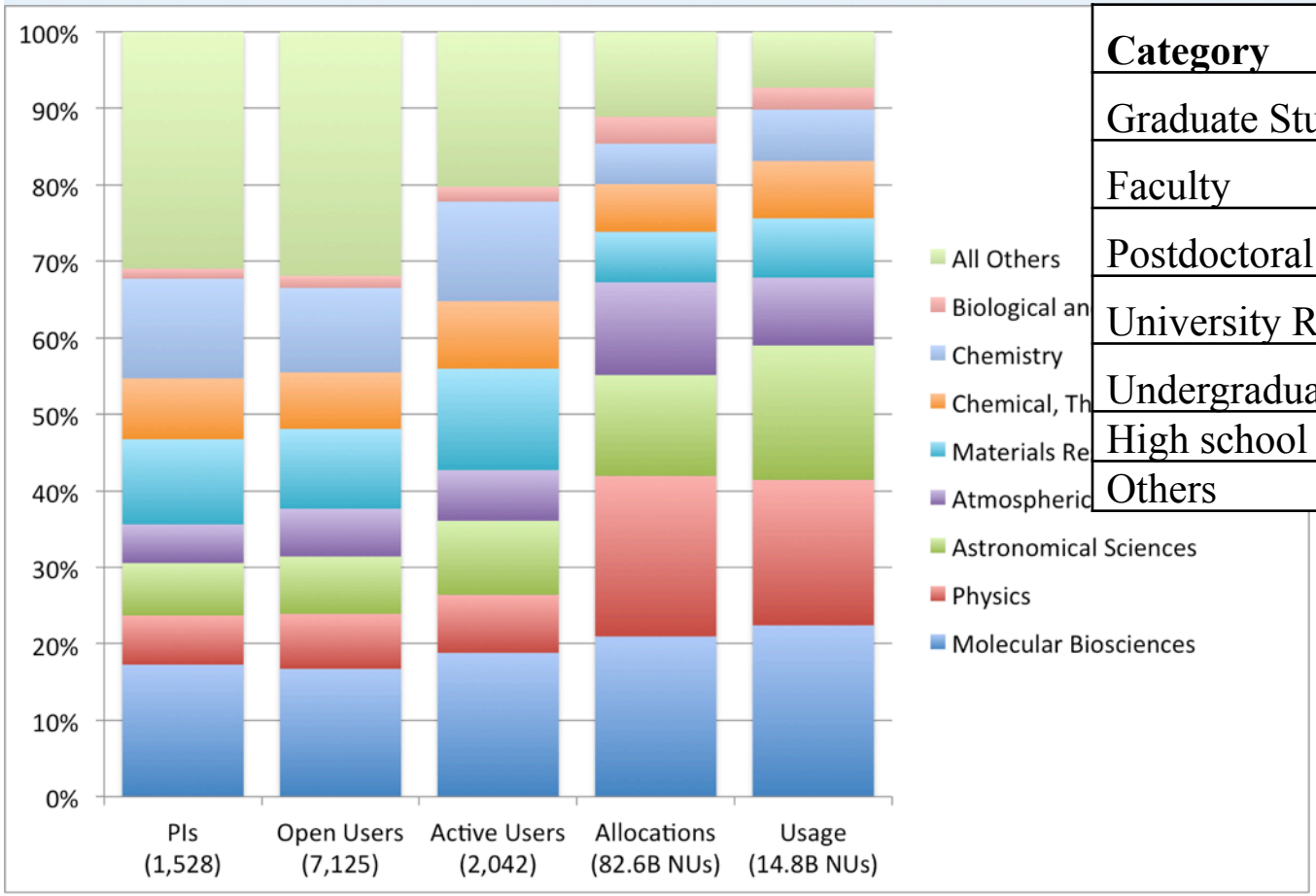
XSEDE Components

- Coordination and Management Service
- Technology Audit and Insertion Service
- Extended Collaborative Support Service
- Training, Education and Outreach Service

Best of all you can use it for FREE!

Who Uses XSEDE?

End of quarter XSEDE open user accounts by user excluding XSEDE staff.

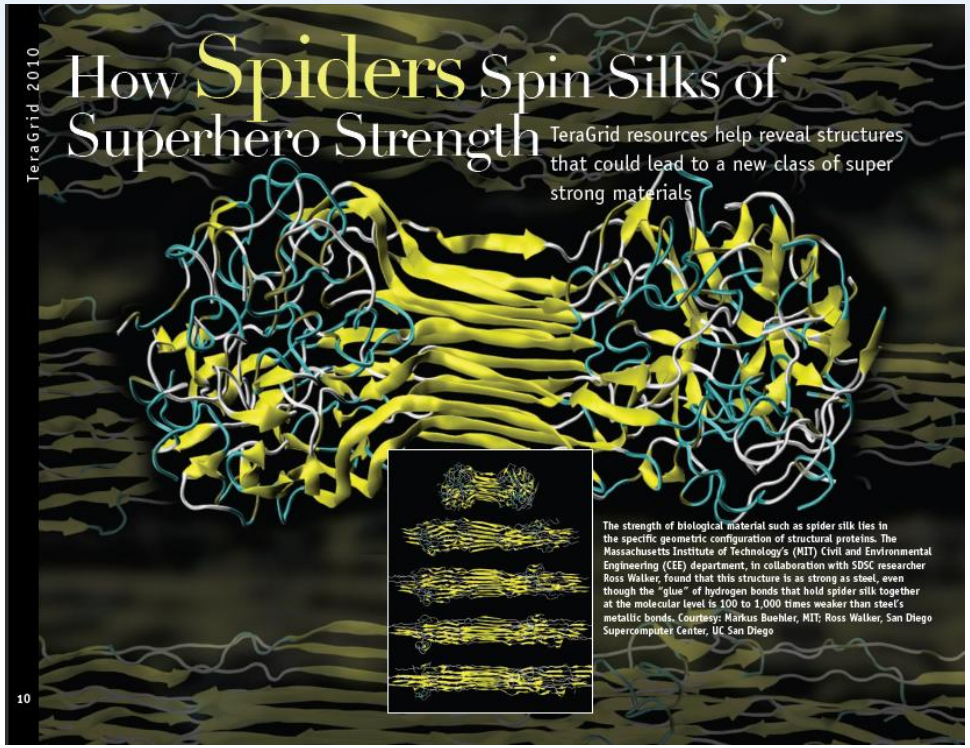


Category	Users
Graduate Student	2,368
Faculty	1,336
Postdoctoral	1,008
University Research Staff	506
Undergraduate Student	505
High school	5
Others	328

Open Science Grid

XSEDE

Science Highlights 1:

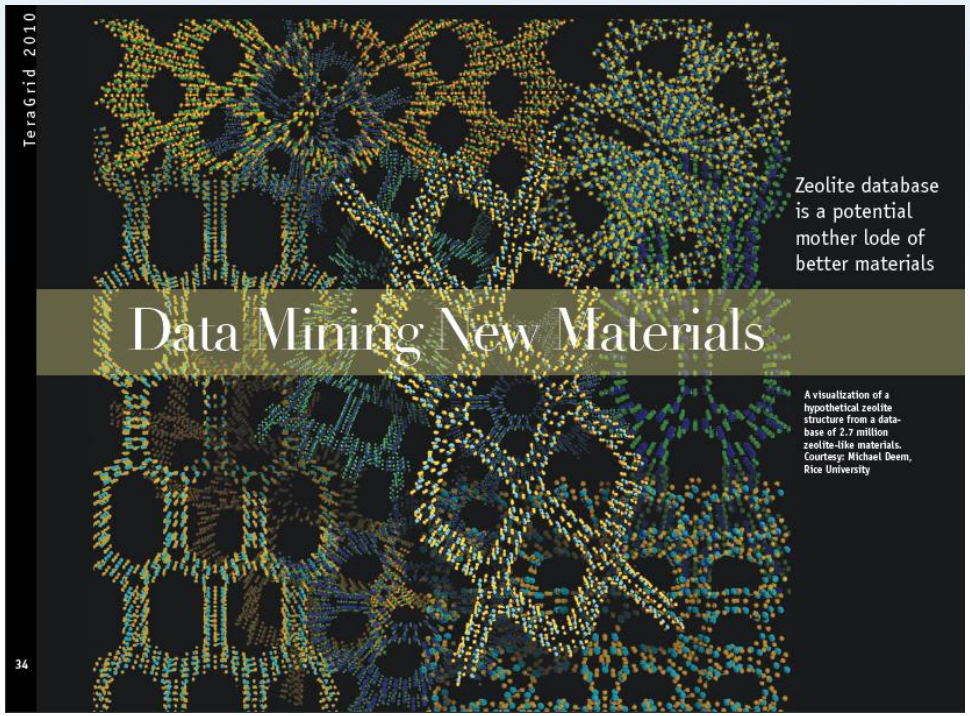


- PI: Markus Buehler
- Institution: MIT
- "We found that the structure of spider silk at the nanoscale can explain why this material is as strong as steel, even though the "glue" of the hydrogen bonds holding spider silk together at the molecular level is 100 to 1,000 times weaker than steel's metallic bonds." says Buehler.

Open Science Grid

XSEDE

Science Highlights 2:

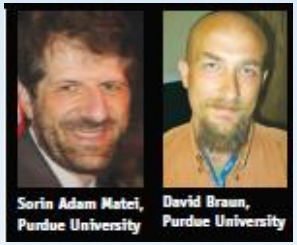


- PI: Michael Deem, David Earl
- Institution: Rice University, University of Pittsburgh
- Identified millions of potentially new zeolites by searching computationally for properly configured, hypothetically stable structures.
- Zeolites are used to make everything from gasoline and asphalt to aquarium filters, laundry detergent and medical-grade oxygen.

Open Science Grid

XSEDE

Science Highlights 3:



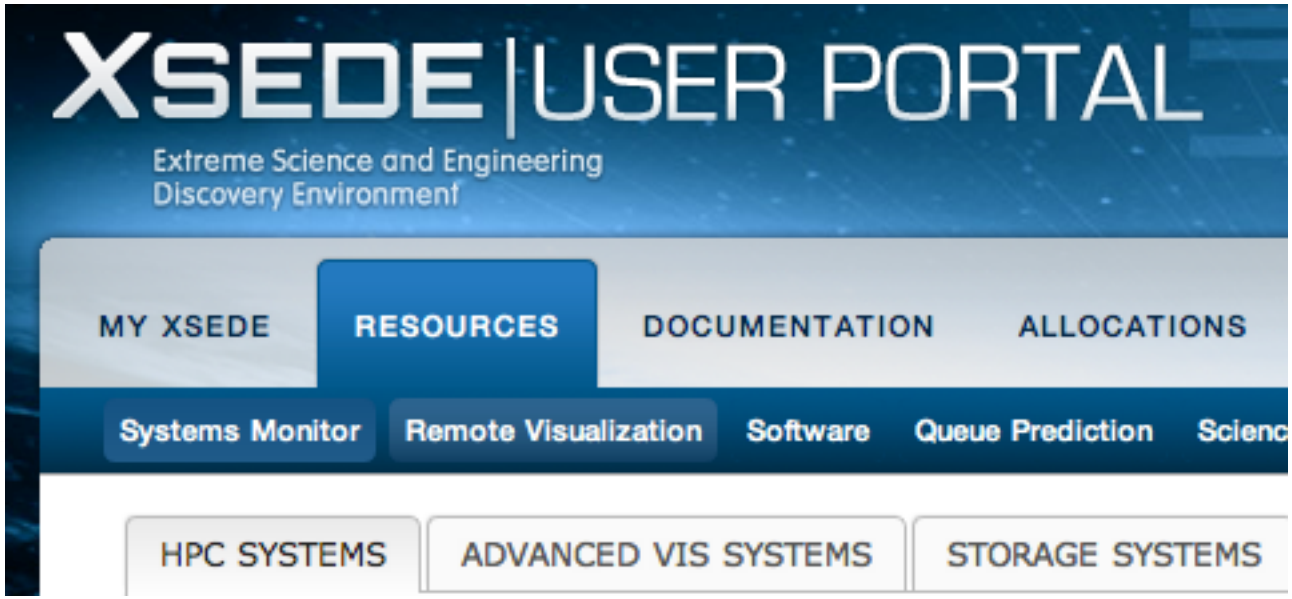
- PI: Sorin Matei, David Braun
- Institution: Purdue University
- Purdue researchers led by Sorin Adam Matei are analyzing the entire collection of articles produced in Wikipedia from 2001-2008, and all their revisions – a computationally demanding task made possible by TeraGrid resources.
- "We looked at how article production is distributed across users' contributions relative to each other over time. The work includes visualizations of patterns to make them easier to discern," says Matei.

Open Science Grid

XSEDE

How do you use XSEDE?
Step 1: The basics

- Setup an XSEDE User Portal Account
  - https://portal.xsede.org/
  - You can do it right now: it's free and easy
- Review available resources and determine which will match the needs of your research
  - Resources -> Systems Monitor



Open Science Grid

How do you use XSEDE?
Step 2: Startup Allocation

- Submit a request (anytime)
- Easy to get—lightly reviewed
- Small: 200K hours
- Use to investigate the resources where your code will run
- Use to benchmarking your code to determine the best resource and how many SUs you will need to complete your research

Open Science Grid

XSEDE

How do you use XSEDE?
Step 3: Research Allocation

- Write and submit a full research allocation request
- Can be submitted quarterly
- Peer-reviewed
- Can be very large
- You should know exactly what resources you need

How do you use XSEDE?
Step 4: Use

- Log into individual resources
  - Can log in via User Portal (built in ssh + authentication)
  - Can download application to login (ssh + authentication)
  - Can request direct ssh access from individual sites
- When you run out:
  - You will be unable to submit more jobs
  - It is possible to request extensions to your allocation

Open Science Grid

XSEDE

Differences between XSEDE and OSG

- Authentication
  - OSG : certificates obtained by user, member of VO
  - XSEDE : User has certificate but never sees it. Transparent access via MyProxy with name/password
- System Access:
  - OSG:
    - Access based on VO membership
    - Jobs submitted remotely (usually glideins)
  - XSEDE
    - Access based on allocations
    - Direct login to resource to submit jobs (via User Portal or gsissh or ssh)

Differences (2)

- Amount of Compute Cycles:
  - OSG : only limited by excess cycles on available resources
  - XSEDE : based on granted allocations (different amounts available for request based on allocation type with research allocation requests peer reviewed)
- Job Types:
  - OSG: serial or single-node parallel
  - XSEDE: serial, shared memory, large parallel, large data
- Other:
  - XSEDE: Can request an allocation of "people" resources (ECSS) to assist with technical aspects of project/code
  - OSG: Can access all systems via a single submit host

Questions? Comments?