XSEDE: An Advanced and Integrated Set of Digital Resources for Science and Engineering

Linda Akli, SURA
Assistant Director, Training, Education & Outreach & XSEDE Underrepresented Community Engagement
What is XSEDE?

- Foundation for a national CI ecosystem
  - comprehensive suite of advanced digital services that federates with other high-end facilities and campus-based resources
- Unprecedented integration of diverse digital resources
  - innovative, open architecture making possible the continuous addition of new technology capabilities and services
XSEDE Team

- World-class leadership from CI centers with deep experience: partnership led by NCSA, NICS, PSC, TACC and SDSC
- Partners who strongly complement these CI centers with expertise in science, engineering, technology and education

<table>
<thead>
<tr>
<th>U of Virginia</th>
<th>Ohio Supercomputer Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>SURA</td>
<td>Cornell</td>
</tr>
<tr>
<td>Indiana University</td>
<td>Purdue</td>
</tr>
<tr>
<td>University of Chicago</td>
<td>Rice</td>
</tr>
<tr>
<td>Berkeley</td>
<td>NCAR</td>
</tr>
<tr>
<td>Shodor</td>
<td>Jülich Supercomputing Centre</td>
</tr>
</tbody>
</table>
XSEDE Vision and Mission

• Vision
  – XSEDE aspires to be the place to go to access digital research services.

• Mission
  – Accelerate scientific discovery by enhancing the productivity of researchers, engineers, and scholars by deepening and extending the use of XSEDE’s ecosystem of advanced digital, services and by advancing and sustaining the XSEDE advanced digital infrastructure.
Why XSEDE?
XSEDE Supports a Breadth of Research

- Earthquake Science
- Molecular Dynamics
- Nanotechnology
- Plant Science
- Storm Modeling
- Epidemiology
- Particle Physics
- Economic Analysis of Phone Network Patterns
- Large Scale Video Analytics (LSVA) Decision Making Theory
- Library Collection Analysis

Three-dimensional model of major vessels and bifurcations of the human arterial tree reconstructed with gOREK from a set of computed tomography (CT), digital subtraction angiography CT and magnetic resonance angiography images.

A snapshot of an animation for water level prediction including the wind-wave signature.
XSEDE Compute Resources

Stampede @ TACC
- 10 PFLOPS (PF) Dell Linux Cluster based on 6400+ Dell PowerEdge server nodes, each outfitted with 2 Intel Xeon E5 (Sandy Bridge) processors and an Intel Xeon Phi Coprocessor (MIC Architecture)

Gordon @ SDSC
- Flash-based supercomputer designed for data-intensive applications

Darter @ NICS
- Cray XC30 system providing both high scalability and sustained performance with a peak performance of 250 Tflops

Greenfield
- 360 cores and 18TB of memory in three nodes: two HP DL580s and an HP SuperDome X. Hosts a large number of bioinformatics tools

Mason
- A large memory computer cluster configured to support data-intensive, high-performance computing tasks using genome assembly software

Super Mic @LSU
- Equipped with Intel's Xeon Phi technology. Cluster consists of 380 compute nodes.
New Resources

Wrangler
Data Analytics System combines database services, flash storage and long-term replicated storage, and an analytics server. IRODS Data Management, HADOOP Service Reservations, and Database instances.

Comet
Features the next generation Intel "Haswell" processors with AVX2 and hosts a variety of tools including Amber, GAUSSIAN, GROMACS, LAMMPS, NAMD, and VisIt.

Coming in 2016
Featuring interactive on-demand access, tools for gateway building, and virtualization.

Jetstream
A self-provisioned, scalable science and engineering cloud environment.
XSEDE Visualization and Data Resources

Visualization

**Maverick@ TACC**
- HP/NVIDIA cluster
- 132 TB memory
- VisIt
- ParaView
- Interactive Data Language

**Visualization Portal**
- Remote, interactive, web-based visualization
- iPython / Jupyter Notebook integration
- R Studio Integration

Storage

- **Resource file system storage:** All compute/visualization allocations include access to limited disk and scratch space on the compute/visualization resource file systems to accomplish project goals.

- **Archival Storage:** Archival storage on XSEDE systems is used for large-scale persistent storage requested in conjunction with compute and visualization resources.

- **Stand-alone Storage:** Stand-alone storage allows storage allocations independent of a compute allocation.
Gateways: Democratizing Access

• Almost anyone can investigate scientific questions using high end resources
  – Not just those in high profile research groups

• Gateways allow anyone with a web browser to explore

• Foster new ideas, cross-disciplinary approaches
  – Encourage students to experiment

• Used in production
  – Significant number of papers resulting from gateways, including GridChem, nanoHUB
  – Scientists can focus on challenging science problems rather than challenging infrastructure problems
Science Gateways
XSEDE User Services

• Technical information
  – Always available via web site and XSEDE user portal

• Allocations
  – Request access to XSEDE’s systems

• Training
  – Sign up for classes to learn to use XSEDE resources

• Help Desk/Consultants

• Extended Collaborative Support Services
  – Human resources to help with performance analysis, petascale optimization, efficient use of accelerators, I/O optimization, the development of community gateways and work and data flow systems
XSEDE Training

- XSEDE provides extensive training
  - Covering every major resource
  - From beginner to advanced classes
  - At locations across the country
  - Online via
    - asynchronous technologies
    - Webcasts
- Web-based education credit courses
Community Engagement

- Champions Program
- Campus Bridging
- Education
- Under-represented Community Engagement
- Annual XSEDE Conference
Champions Program

• Campus Champions
  – Representatives to spread information about XSEDE to local faculty, students and staff

• Student Champions
  – Students assist the Campus Champions

• Regional Champions
  – Representatives to spread information about XSEDE to other campuses in the area

• Domain Champions
  – Disciplinary people able to assist others with domain specific HPC questions
Campus Bridging

The goal of campus bridging is to create a sense of “virtual proximity.” Any resource should feel as if it’s just a peripheral to their laptop or workstation. The goal is to make it convenient and intuitive to simultaneously use your personal computing systems, departmental and campus systems (at your campus and others), and national resources liked XSEDE . . . all (almost) transparently and easily.
Education Program

• Development of competencies for undergraduate and graduate computational science programs
  – Assisting campuses with organizing formal certificate programs
  – Sharing instructional materials

• Campus visits to promote computational science
  – Meetings with faculty and administrators
  – Professional development workshops
Underrepresented Community Engagement

• Expand awareness of XSEDE
• Identify programs and researchers who can benefit from XSEDE services
• Enable institutions and faculty to use advanced digital services to increase their research productivity
  – By establishing and growing a thriving collaborative peer support community
  – Through the delivery of training mapped to their needs
  – By connecting researchers with XSEDE services and expertise for targeted deep engagement
• Create scalable and sustainable models and best practices
  – By supporting the establishment of certificate and degree programs and enhanced curriculum
  – By developing and supporting productive campus champions
SAVE THE DATE!

JULY 17-21, 2016
INTERCONTINENTAL MIAMI HOTEL

Miami, one of the most distinct cultural locations in the country, will host the XSEDE16 conference. The themes will be the importance of Diversity, Big Data, Science at Scale, and how they interconnect to deliver the next-generation of science and technology. We look forward to seeing you next year!
XSEDE16 Conference

• Submissions will be accepted for papers, panels, tutorials, BOFs, student programs
• Topics span accelerating discovery, advanced technologies, software, science gateways and portals, and education, outreach and training
• Expect over 600 people from academia, industry, government, and other organizations
• Support for student participation
How Do I Get Started?
Faculty Opportunities

- Use XSEDE Resources for research or teaching
- Participate in Training
- Attend Summer Institutes
- Join the Minority Research Community
- Participate in XSEDE16, July 2016, Miami
Student Opportunities

• XSEDE Scholars – Applications Due Feb 2016
  – Year-long webinars
  – Travel to XSEDE Annual Conference
  – Internships

• Blue Waters Internship – Applications Due Feb 2016
  – 2 week training institute for undergrads and grads
  – year-long computational science problem solving

• Blue Waters Graduate Fellowship
  – similar to NSF Graduate Fellowships
  – year-long engagement

• XSEDE Annual Conference
  – travel support for students to attend the annual Conference
Workshop Agenda

- XSEDE New User Training
- Bioinformatics Databases and Tools
- Statistical and Numerical Tools: Parallel "R"
- Computational Thinking:
- Consulting
- Poster Session
- Participation Acknowledgement
More Information

- XSEDE Website: www.xsede.org
- XSEDE Staff
  - Linda Akli, akli@sura.org
  - Jay Alameda, alameda@illinois.edu (New User Training)
  - Steve Gordon, sgordon@osc.edu (Curriculum)
  - Marcela Madrid, mmadrid@psc.edu (Bioinformatics)
  - Paul Rodriguez, prodriguez@sdsc.edu (“R”)

27
Our reach will forever exceed our grasp, but, in stretching our horizon, we forever improve our world.