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

Linda Akli, SURA
Assistant Director, Training, Education & Outreach

XSEDE
Extreme Science and Engineering Discovery Environment

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 Vision

The eXtreme Science and Engineering Discovery Environment (XSEDE):

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

and thus

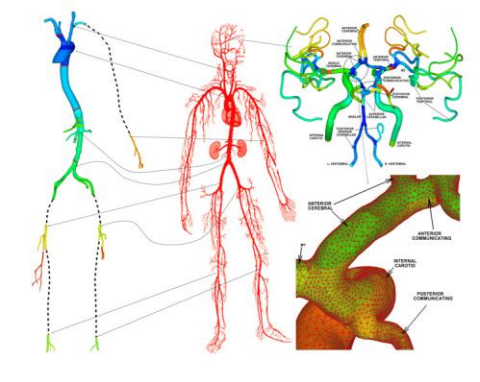
- facilitates scientific discovery while enabling transformational science/engineering and innovative educational programs

XSEDE Team

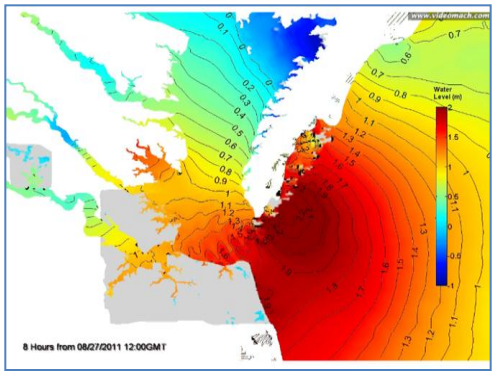
- World-class leadership from CI centers with deep experience: partnership led by NCSA, NICS, PSC, TACC and SDSC
  - PI: John Towns, NCSA/Univ of Illinois
  - Co-PIs: Jay Boisseau, TACC/Univ of Texas Austin
    Greg Peterson, NICS/Univ of Tenn-Knoxville
    Ralph Roskies, PSC/CMU
    Nancy Wilkins-Diehr, SDSC/UC-San Diego
- Partners who strongly complement these CI centers with expertise in science, engineering, technology and education
  - Univ of Virginia
    SURA
    Indiana Univ
    Univ of Chicago
    Berkeley
    Shodor
    Ohio Supercomputer Center
    Cornell
    Purdue
    Rice
    NCAR
    Jülich Supercomputing Centre

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.

Why would you use XSEDE?

Some Jobs Just Need Bigger Tools

How Can XSEDE Help Meet Challenges?

- Massively parallel clusters
- Large shared-memory nodes
- Parallel file systems
- Fast networking
- Very large databases
- Efficient data movement tools
- Consulting services provided by experts (expertise on OpenMP, workflow, science gateway, scientific database, visualization and more)

XSEDE offers huge variety of resources

- Leading-edge distributed memory systems
- Very large shared memory systems
- High throughput systems, including now OSG
- Visualization engines
- Accelerators like GPUs

Many scientific problems have components that call for use of more than one architecture.

Current XSEDE Compute Resources


Stampede @ TACC
- 6 PFLOPS (PF) Dell Cluster w/ GPUs and Xeon PHIs


Gordon @ SDSC
- 341 TF Appro Distributed SMP cluster


Lonestar (4) @ TACC
- 302 TF Dell Cluster


Trestles @ SDSC
- 100TF Appro Cluster


Steele @ Purdue
- 60 TF Dell Cluster


Blacklight @ PSC
- 37 TF SGI UV (2 x 16TB shared memory SMP)

- Mason
  - 3.8 TF HP Cluster with large memory nodes (2TB/node)

https://www.xsede.org/web/xup/resource-monitor

Current XSEDE Visualization and Data Resources

- Visualization
  - 

Nautilus @ UTK
    - 8.2 TF SGI/NVIDIA SMP
    - 960 TB disk
  - 
Longhorn @ TACC
    - 20.7 TF Dell/NVIDIA cluster
    - 18.7 TB disk

https://www.xsede.org/web/xup/resource-monitor#advanced_vis_systems

- Storage
  - HPSS @ NICS
    - 6.2 PB tape
  - Data Supercell @ PSC
    - 4 PB tape
  - Ranch @ TACC
    - 40 PB tape
  - Data Oasis @ SDSC
    - 4 PB tape

https://www.xsede.org/web/xup/resource-monitor#storage_systems

Special Purpose Resources

Keenland

- Hybrid CPU/GPGPU System

FutureGrid

- Experimental Infrastructure-as-a-Service cloud environment

Open Science Grid

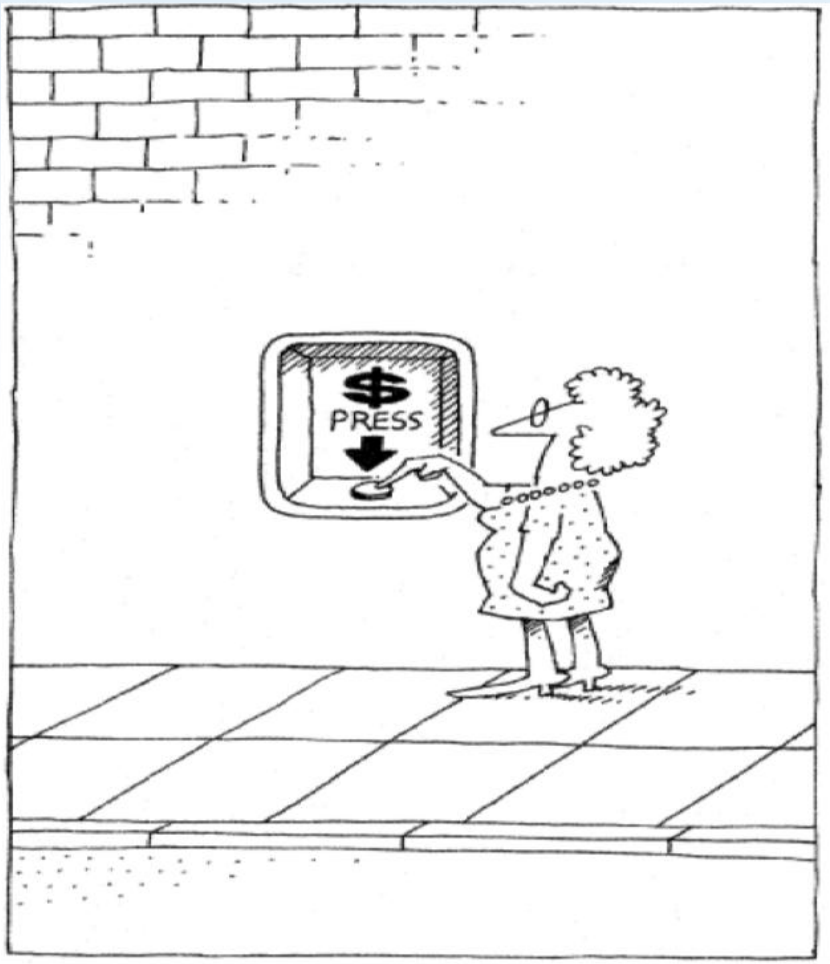
- High Throughput Computing (many jobs that are typically similar and not highly parallel)

https://www.xsede.org/resources/overview

NCSA Blue Waters System

- Funded by the NSF to support very large scale computational science and engineering
- Cray systems
  - 22,640 Cray XE6 nodes - 64 GB of memory per node
  - 3,072 Cray XK7 nodes include NVIDIA processors with 32 GB of memory
  - 26 petabytes of online storage
  - 380 petabytes of tape storage
- Allocations are made via:
  - Applications to the NSF PRAC proposal process
  - Applications to Blue Waters education allocations

Simple Enough



OOPS

XSEDE User Services

XSEDE User Services are grouped into four main areas:

- 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
- User Engagement
  - Includes 'consulting support' to answer questions
  - Also includes user interviews, focus groups, and surveys

XSEDE User Guides and News

- XSEDE provides intro user guides for every XSEDE-allocated system—no matter where it is actually hosted
  - Consistently structured and formatted
  - All available from website and XUP
  - Prepared using expertise of host sites
- XSEDE also provides up-to-date User News about every system, and XSEDE-wide services, available via:
  - Web/portal
  - Email
  - RSS feeds
  - Calendar feeds (for downtimes, training events, etc.)

XSEDE Allocations

- XSEDE allocates access/time on powerful, valuable systems providing different capabilities at NO COST TO YOU
  - HPC
  - High throughput computing
  - Remote visualization
  - Data storage
  - Etc.
- Users may request XSEDE staff support to assist with optimization of research codes, visualization, workflows, novel projects, and science gateways
- Single Sign-On allows you to use just one username and password (your User Portal one). You will be recognized by all XSEDE services on which you have an account, without having to enter your login information again for each resource.

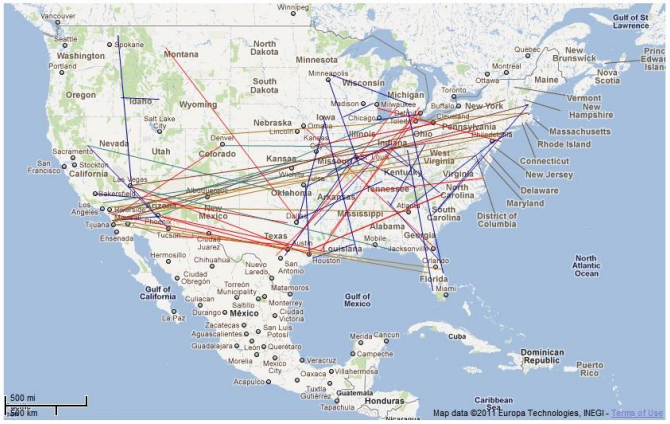
XSEDE offers more in-depth support Extended Collaborative Support Service

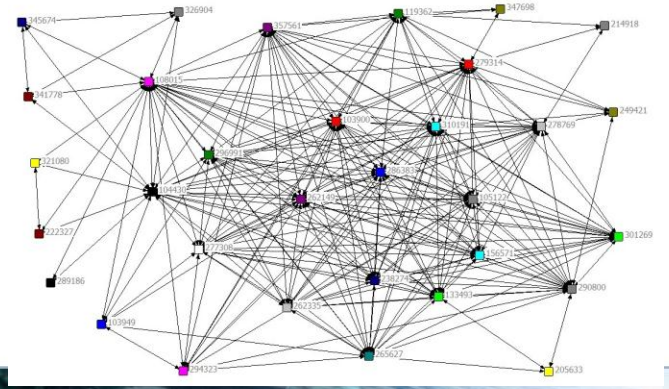
- Support people who understand the discipline as well as the systems (perhaps more than one support person working with a project).
- 37 FTEs, spread over >70 people at more than half a dozen sites.
- Distributed support
  - Easier to find the right expert for the project
  - allows us to cover many more disciplines than if every site had to staff the common applications.
  - support does not have to move with platform change

Virtual Worlds Exploratorium

Marshall Scott Poole, Univ. of Illinois

- Analyze logs (over 10 TB) from several online games





TASK: "The variables of interest for research have to be derived from the game databases through complex analyses that transform the raw log data into terms meaningful for social science research. By taking advantage of massive parallel processing, efficient workflow management and abundant memory availability, access to XSEDE resources could enable us to speed up our research."

Large Genome Assemblies

- ECSS Staff working with leading researchers and code developers
- Largest ever metagenome assembly, using 3.5 TB RAM on PSC Blacklight
- Cold Spring Harbor collaboration to assemble wheat genome (17 Gigabases)

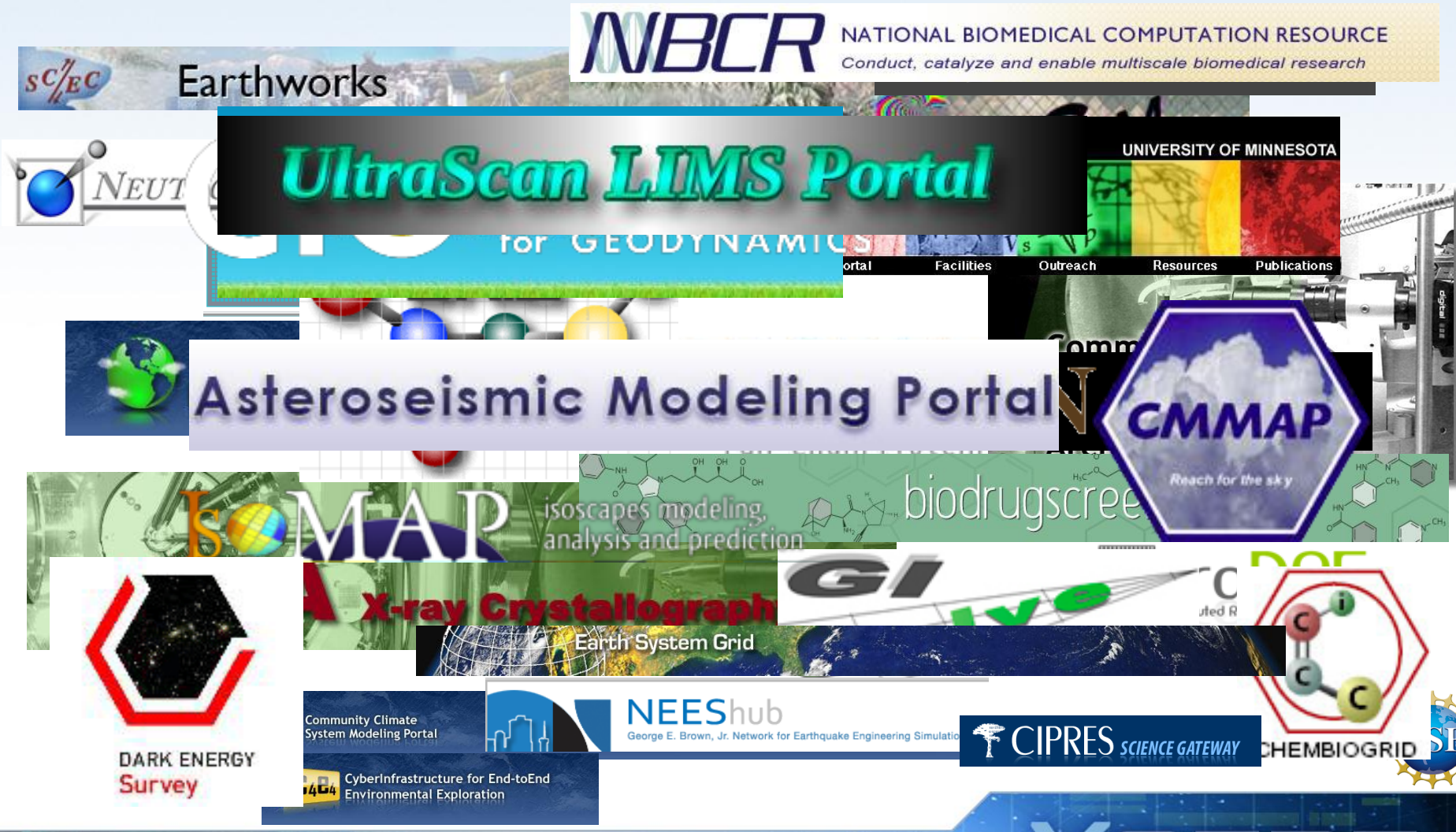
"I wouldn't have been able to do anything on Blacklight without ECSS staff… (consultant) took a real interest and solved a lot of things that were hard for me. He found bugs in the software and got them resolved with the software authors. I'd worked for months and not made that progress. Without his expertise, I might have given up…"

Science Gateways

- Researchers using tools where inputs could be standardized
  - Same executables (no need to recompile)
    - GridChem, CHARMM
- Creating standardized workflows
- Input is streamed data to the web

e.g. LEAD takes radar data, and determines whether the pattern suggests possible formation of a tornado cell, in which case more fine-grained simulation is needed

Today, there are approximately 35 gateways using XSEDE



XSEDE

Gateways democratize access to high end resources

- 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
- But used in production, too
  - Significant number of papers resulting from gateways, including GridChem, nanoHUB
  - Scientists can focus on challenging science problems rather than challenging infrastructure problems

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
- Signing up is **simple--**in the XSEDE User Portal!

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

Community Engagement Activities

- Student Programs
- Under-represented Community Engagement
- Champions Program
- Campus Bridging
- Campus Visits
- Annual XSEDE Conference

Student Programs

- XSEDE Scholars
  - engaging undergraduates and graduates in year-long series of webinars attend annual XSEDE Conference
- XSEDE Summer Research Experience
  - Summer internship with XSEDE staff or user
- Blue Waters Internship
  - 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
- HPC University
  - Lists other student engagement opportunities

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

Campus Champions Program

- Over 150 campuses are members – no cost to join
- Champions receive monthly training and updates
- Champions provided with start-up accounts
- Champions are members of User Services team
- Forum for sharing and interactions
- Access to information on usage by local users
- Registrations for annual XSEDE14 Conference waived
- Community building across campuses

Campus Champions Role

- Raise awareness locally
- Provide training
- Get users started with access quickly
- Represent needs of local community
- Provide feedback to improve services
- Attend annual XSEDE14 conference
- Share their training and education materials
- Build community among all Champions

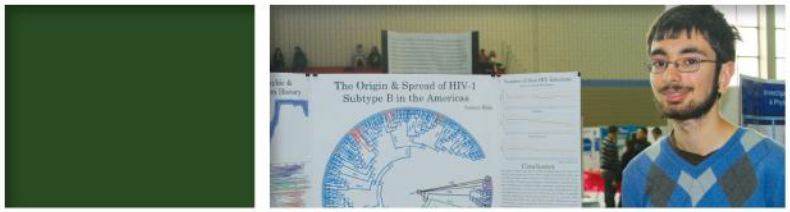
Champions Program

- 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 in general 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.



XSEDE[14]

July 13-18, 2014
Atlanta Marriott Marquis

Registration opens April 14!

www.xsede.org/xsede14

XSEDE14 Conference

- Theme is Engaging Communities
- Submissions being accepted for papers, panels, tutorials, BOFs, student programs
- Topics span accelerating discovery, advanced technologies, software, science gateways and portals, and education, outreach and training
- Over 700 people from academia, industry, government, and other organizations
- Support for Champions and student participation

Faculty Engagement Opportunities

- Create XSEDE Portal Account
- Use XSEDE Resources for research or teaching
- Participate in Training Webinars
- Attend In-Person Training & Summer Institutes
- Be a Campus Champions/Campus Champions Fellows
- Join the Minority Research Community
- Participate in XSEDE14, July 2014, Atlanta

Campus Engagement Opportunities

- Campus Champions
- Campus Bridging
- Education – Computational Science Curriculum, Certificate, or Degrees
- MSI Campus Engagement
- Regional Workshops
- Summer Institutes

Stay Connected

- XSEDE's public web site is www.xsede.org
- Create an XSEDE User Portal signon and receive news and notices
- Training events are announced via the public web site; and registrations via the XSEDE User Portal
- For access to additional training and educational resources www.hpcuniversity.org
- Contact Linda Akli <akli@sura.org> to join the Minority Research Community or schedule a campus visit


Our reach will forever
exceed our grasp, but,
in stretching our horizon,
we forever improve our world.
XSEDE
Extreme Science and Engineering
Discovery Environment