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

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   Ohio Supercomputer Center
      SURA   Cornell
      Indiana Univ   Purdue
      Univ of Chicago   Rice
      Berkeley   NCAR
      Shodor   Jülich Supercomputing Centre

XSEDE supports a breadth of research

- Earthquake Science and Civil Engineering
- 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
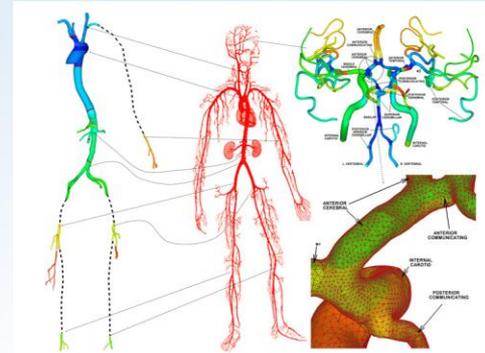
Three-dimensional model of major vessels and bifurcations of the human arterial tree reconstructed with gOREK from a set of computed tomography (CT), digital subtraction angiography CT and magnetic resonance angiography images.
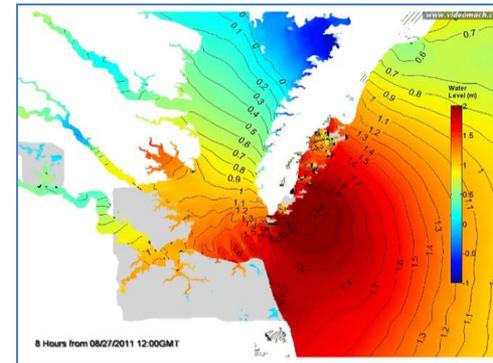
A snapshot of an animation for water level prediction including the wind-wave signature.

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 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 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.

XSEDE

Compute Resources

Kraken @ NICS
- 1.2 PF Cray XT5

Stampede @ TACC
- 7+PF Intel's new MIC technology

Gordon @ SDSC
- 341 TF Appro Distributed SMP cluster

Lonestar (4) @ TACC
- 302 TF Dell Cluster

Trestles @ SDSC
- 100 TF Appro Cluster

Mason
- 3.8 TF HP DL530

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

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

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

Visualization and Storage Resources

- Visualization
  - Longhorn @ TACC

- Storage
  - Archival Storage
  - Standalone Storage
  - Resource filesystem storage

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

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

ECSS Support Services

- Support for
  - Research Teams
  - Community Capabilities
  - Novel and Innovative Projects

ECSS-1-Support for Research Teams

- Optimization
  - Profiling
  - Scaling to higher core count
  - Improving IO
  - Porting to GPUs
  - Finding better solvers (what's better often depends on the degree of parallelization)
- Visualization
- Workflows

How do you get Extended Collaborative Support?

- You have to ask for it, -it's an allocated resource.

- You can always ask for it, i.e. midstream or even as part of a startup request.

- Lasts up to a year.

The 5 Questions

1. *What do you want to accomplish with the help of advanced support? Have you already done any work on this aspect of your software?*

2. *How would the success of this advanced support benefit your project?*

3. *Which member(s) of your team would collaborate with XSEDE advanced support staff?*

4. *Have you had significant interaction on previous projects related to your current proposal or discussed your advanced support needs with any XSEDE staff? If so, please indicate with whom.*

5. *Have you received TeraGrid/XSEDE advanced support in the past? If so, please indicate the time period and how the support you received then relates to the support you request now.*

ECSS-2- Community Capabilities

- Deploying, hardening, and optimizing useful software systems
    - Collaborate with the developers of widely used community codes and tools, including PetaApps, SDCI, STCI, SI2, MREFC and industrial partners
    - E.g. Develop *asynchronous* replica exchange mechanism to enable dynamic scheduling of resources and adaptive control of replica parameters to understand important aspects of the physics of protein-ligand recognition

ECSS-3- Novel and Innovative Projects

- Pro-actively reaching out to communities new to advanced computing, e.g.
  - social science
  - social network analysis
  - language processing
  - genomics
  - digital humanities
- Look for pioneers, work with them to craft the project, and then help them begin to execute it.

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."
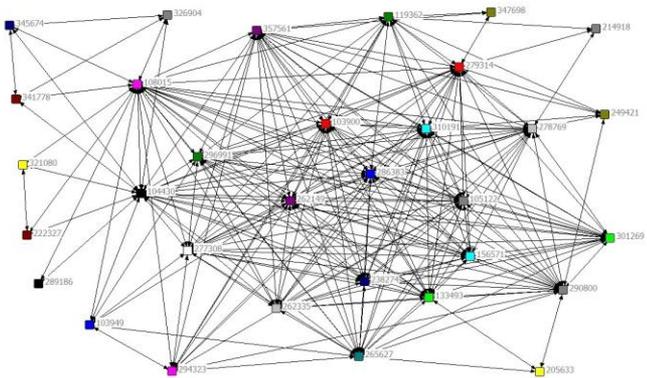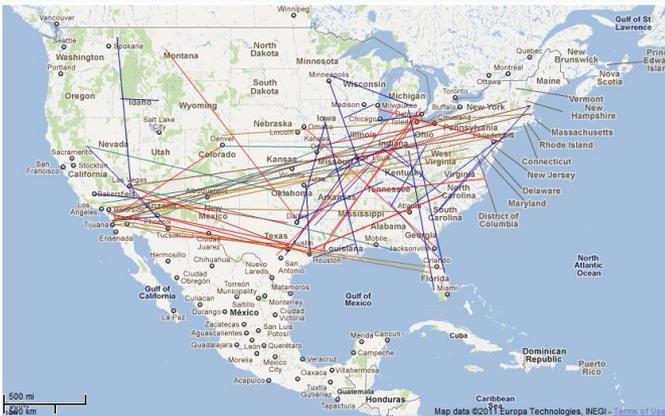
Large Genome Assemblies

- 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…"

Today, there are approximately 35 gateways using 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 Community Engagement

- Training
- Education
  - Curriculum development workshops,
  - Working with institutions to offer CS&E Programs.
- Outreach
  - Immersing students in training, internships, and mentoring programs,
  - Engaging new and under-represented communities, and
  - Supporting Campus Champions to assist local users.
- Campus Bridging
  - Facilitating data exchange among campuses and XSEDE

XSEDE

Outreach Services

GOAL: Recruit a large and diverse scientific, academic, and industrial workforce capable of advancing scientific discovery using XSEDE services.

- **Student Engagement**
  - Provide meaningful experiences for undergraduate and graduate students to become engaged in Extreme Digital environments
- **Underrepresented Communities Engagement**
  - Bring XSEDE to new faculty, and students, providing dedicated training and support
- **Campus Champions**
  - Build broader and deeper programs based on TeraGrid foundation

XSEDE

XSEDE UR Engagement Goals

- 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

MSI Campus Engagement

- MSI Engagement
  - Identify established and emerging programs and researchers
  - Expand awareness of XSEDE via campus visits, professional conferences
  - Regional Workshops
  - Targeted deep engagement that connect researchers with XSEDE expertise

XSEDE

Minority Research Community

- Goal
  - Build peer support community and connect with XSEDE services
- Identification/Selection of Participants
  - Invite faculty identified at campus visits and current MSI campus champions, require participation by travel support awardees and add referrals from XSEDE staff and MRC participants
- Activities
  - Monthly Call
  - Listserv
  - XSEDE Conference Sessions
- Discussion Topics
  - Soup to Nuts
- Future
  - Scale and dive deeper
  - Learn from Campus Champions

XSEDE Scholars & Minority Faculty Council

- XSEDE Scholars

  – Create support community, paring students with mentors

  – Provide education and professional development to participants

- Minority Faculty Council

  – External Advisors

  – Ongoing state-of-the-practice focus

XSEDE

XSEDE Campus Bridging Overview

- General Goal: Create a sense of "virtual proximity, " i.e., resources feel like peripherals to laptop computer workstations. (The origin of the name goes back to the days when if you tried to migrate from campus infrastructure to national systems it felt like falling off a cliff… hence the need for bridges)

- Specific Goals:
  - Make XSEDE convenient, transparent, and intuitive at multiple practical scales: personal, departmental and campus computing systems.
  - Simplify integration of non-NSF funded resources with resources enabled and coordinated by XSEDE

Campus Bridging: XSEDE-compatible clusters

- Goal – make it easy for campus cluster managers (including faculty members managing a cluster) to create a cluster that is compatible with XSEDE clusters

- Rocks – a "roll" lets you install a cluster from scratch, set up with open source software and configurations consistent with XSEDE clusters., along with a variety of application software. Job management software and Globus Online included. More information, including a link to download the cluster building software, at https://www.xsede.org/web/xup/knowledge-base/-/kb/document/bdpe

- Already have a cluster that you think works pretty well? Add additional software components as RPMs to your current cluster and gain the same sort of functionality as with the Rocks Roll – you just assemble pieces as you wish, yourself.

XSEDE-compatible cluster - advantages

- Advantages:
  - Automate what you can automate, spend personnel time doing local customizations
  - XSEDE training resources (see https://www.xsede.org/web/xup/overview) can be customized for your local training needs
  - Youtube video (Craig Stewart & Marcus Alfred): http://www.youtube.com/watch?v=E96mIWruATM&feature=youtu.be

XSEDE

Moving data from campus resources to XSEDE, and Back

- Globus Online - production XSEDE data movement service.
  - Convenient user interface – "data movement as a service"
  - Globus Online is one of the most widely used data movement tools in existence, inside and outside of XSEDE
- Coming soon
  - EMS (Execution Management Service) (based on Unicore) – has data movement tools. To be deployed this year
  - General Federated File System (GFFS) (Based on Genesis II) – has many data movement tools. To be deployed this year.

Campus Bridging and Computing Resources

- Make it easier to use XSEDE from your cluster
  - EMS will have capabilities to easily submit jobs from your local cluster to XSEDE
- Capability to create your own 'virtual cluster"
  - You can do this without having any relationship between your virtual cluster – which can be created from clusters in multiple locations – and XSEDE. Just combine forces with your friends for more efficient use of your clusters.
  - Coming later: ability to share cycles and get credits by running XSEDE jobs on your cluster

XSEDE

Campus Bridging Support

- Live human telephone support - call 812-318-2872 9 am to 5 pm eastern time

- Email: send questions and requests to campusbridging@xsede.org.

Student Engagement Opportunities

- Apply for the XSEDE Scholar's Program
- Apply for XSEDE Summer Research Experience
- Sign up for the XSEDE Student Mailing List
- Participate in XSEDE14, July 2014, Atlanta
  - Volunteer
  - Papers
  - Posters

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

XSEDE

Campus Engagement Opportunities

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

Questions? Or Feedback

- XSEDE User Portal – www.xsede.org
- Help – https://www.xsede.org/get-help
- Feedback - https://portal.xsede.org/feedback
- Linda Akli - akli@sura.org

XSEDE

Our reach will forever

exceed our grasp, but,

in stretching our horizon,

we forever improve our world.

XSEDE

Extreme Science and Engineering
Discovery Environment