



XSEDE

Extreme Science and Engineering
Discovery Environment

DATA SAMPLE PART 01:
Wij proberen voorwaarden te creëren die leidbaar en voorbereidbaar maken voor de toekomst. We hebben de mogelijkheid om verschillende scenario's te ontwerpen en te evalueren, zodat we kunnen zien hoe verschillende factoren elkaar beïnvloeden. Dit helpt ons om betere beslissingen te nemen.

DATA SAMPLE PART 02:
Investigaties van de wereld zijn belangrijk voor ons leven. Ze helpen ons om betere keuzes te maken en ons te helpen om de wereld een betere plek te maken. Ze helpen ons om te begrijpen hoe de wereld werkt en hoe we kunnen ervan gebruik maken.

DATA SAMPLE PART 03:
Deze pagina toont voorbeelden van verschillende scenario's die kunnen worden gebruikt om te helpen bij het maken van beslissingen. De scenario's zijn gebaseerd op verschillende factoren en kunnen worden aangepast aan verschillende situaties.

DATA SAMPLE PART 04:
Dit rapport bevat voorbeelden van verschillende scenario's die kunnen worden gebruikt om te helpen bij het maken van beslissingen. De scenario's zijn gebaseerd op verschillende factoren en kunnen worden aangepast aan verschillende situaties.

Bruce Loftis
bruce.loftis@colorado.edu

Agenda

- ❑ A Little XSEDE History
- ❑ Current Computational Resources
- ❑ Allocations
 - How Can I Play ??
- ❑ Other XSEDE Activities and Resources
 - Support, Software, Science Gateways, Training
- ❑ Blue Waters
 - Not really part of XSEDE, but a cousin-of-sorts



HPC Funded by the National Science Foundation

- ❑ **1985 – HPC Centers Program**
 - NCSA, SDSC, PSC, Cornell, JvNC
- ❑ **1996 – PACI Program**
 - Partnership for Advanced Computational Infrastructure
 - NCSA and SDSC
- ❑ **2001 – TeraGrid**
- ❑ **2011 – XSEDE – 16 partners (and counting)**
- ❑ **2016 – XSEDE-02 – renewal underway**

- ❑ **Resources are free – proposals are peer-reviewed**



What Is XSEDE ???

- ❑ Extreme Science and Engineering Discovery Environment
(no one ever remembers this)
- ❑ Funded by NSF
 - ~\$120M over 5 years
 - Not for hardware
 - XSEDE-02 will be funded soon
- ❑ Rah-Rah: *Most powerful integrated advanced digital resources and services in the world*
- ❑ Includes:
 - computing, visualization, storage, data collections, software, tools, advanced support services, training ...



XSEDE

XSEDE Computing Resources

Stampede	TACC	Large Dell cluster with Xeon Phi – capability computing
Maverick	TACC	Visualization and GPU programming
Wrangler	TACC	Large-scale data management – flash memory – large databases – 10 pb replicated at Indiana U
Gordon	SDSC	Data intensive applications with flash memory and SSDs
Comet	SDSC	Small-or-mid-sized problems
Bridges	PSC	Large memory – new diverse communities – easy-to-use – Databases – Webserver nodes – Data transfer nodes
JetStream	Indiana U	Virtual machine on cloud resources
Darter	NICS	Mid-sized Cray for capability computing
	Open Science Grid	High-throughput computing
Mason	Indiana U	HPC with genome assembly software
	Storage - Support	



Experimental Resources

Summit at CU

- MRI Award by NSF to CU and CSU
- Intel Xeon Phi Knights Landing nodes will be allocated by XSEDE
- Probably available late summer 2016



XSEDE

Types of Allocations

XSEDE

- Peer-Reviewed Proposal
- Startup
 - Up to 100,000 cpu-hours on some systems
 - Just an abstract
- Education – Classroom
- Small award thru Campus Champion
- Extended Collaborative Support
- Storage



XSEDE

XSEDE Allocations

Who Is Eligible to Submit Proposals ??

- Researchers at American universities (faculty, postdocs)
- Not students (with rare exception)
- American researchers at foreign universities (probably)
- Not Not-American researchers at foreign universities (can be co-PI)
- Researchers at US Government labs
- Researchers at non-profit companies (not common)
- Researchers at US industry (not sure how this works)

Preference During Review

- Lotsa NSF funding



Prepare a Successful XRAS Proposal

❑ Proposal Outline

- Research objectives
- Computational methodology
- Computational plan
- Performance and scaling
- Justification for service units requested
- External funding
- Other available computing resources
- 15 pages



XSEDE

Interesting Allo-Facts

- ❑ ~700 – Research requests per year
- ❑ ~900 – Startup and other requests per year
- ❑ ~4.0B – Service Units (cpu-hours) requested per year
- ❑ ~1.8B – Service Units awarded per year



XSEDE

Science Gateways

❑ Can Leverage Other Investments

- Researchers can focus on challenging science problems rather than challenging infrastructure problems
- Increase access to instruments and technologies
- To simulate and to analyze data
- To encourage workforce development in underserved and other populations
- Increase public awareness of the value of investment in large facilities
- Slice bread – Pack the kids' lunch, ...

❑ <https://www.xsede.org/web/guest/gateways-listing>

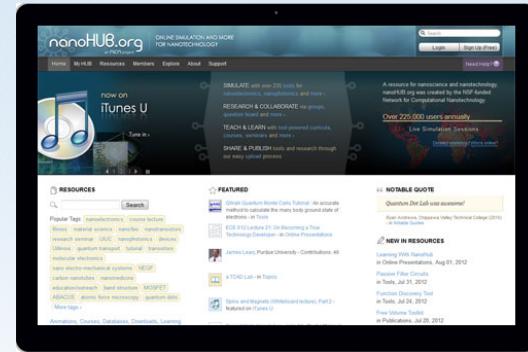


XSEDE

Science Gateways

❑ nanoHUB and HUBzero

- Online simulation for nanotechnology and ...



❑ National Virtual Observatory

- Access to sky surveys



❑ PolarGrid

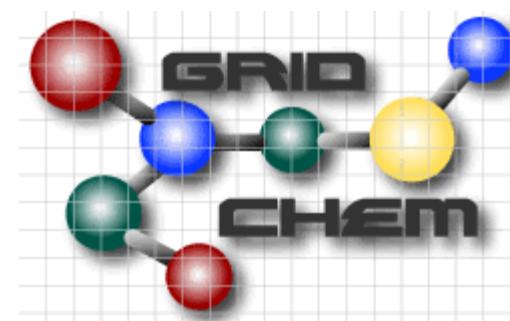
- Study glacial melting with polar ice sheet data

❑ Social Informatics Data Grid

- Data and analysis tools

❑ GridChem

- Access to large-scale chemistry codes



❑ Many more ...



Advanced User Support

Extended Collaborative Support Services (ECSS)

- 40 Computational Scientists available
- Port applications to new resources
- Help with portal and gateway development
- Implement algorithmic enhancements
- Implement parallel math libraries
- Improve scalability of codes to higher processor counts
- Optimize codes to efficiently use specific resources
- Assist with visualization, workflows, data analysis, file transfer, ...



Training – Expand HPC Skills

- **XSEDE Training** (online and in-person)
- **CI-Tutor** (asynchronous)
 - www.citutor.org
 - www.citutor.org/browse.php
 - MPI, OpenMP, Debugging, Visualization, ...
- **Cornell Virtual Workshop** (asynchronous)
 - www.cac.cornell.edu/vw
 - Parallel computing, scripting, visualization, GPU and MIC, performance, data analysis, languages, ...
- **Virtual School for Computational Science & Engineering** (synchronous)
 - www.vscse.org



Blue Waters

Highest Sustained Performance

PRAC Proposals
Blue Waters

National HPC Systems
XRAS Proposals
XSEDE Virtual Organization

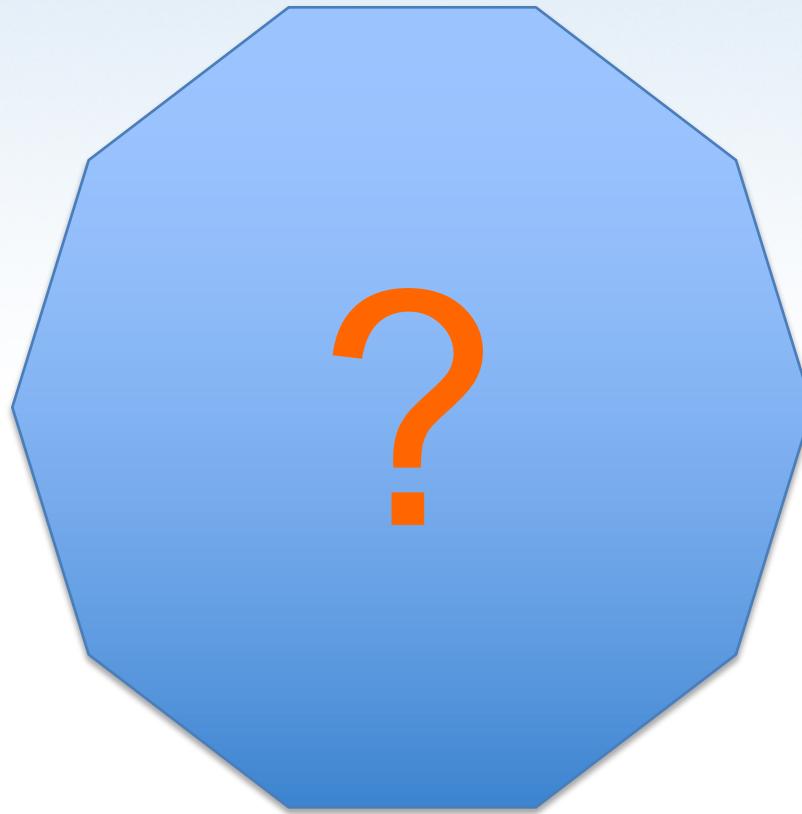
Campus Systems



Blue Waters

- Cost => Many \$\$\$
- Located at NCSA
- Designed for sustained petascale performance on 100,000+ cores for a broad class of problems
- Cray
 - 22,640 AMD XE6 compute nodes => 362,240 cores *plus*
 - 4224 XK compute nodes @ 8 cores + 1 NVIDIA Kepler GPU
 - 7.1 + 6.2 Pflops peak performance
 - 1.4 PBytes + 135 TBytes memory
 - I/O system: 26 Pbytes and >1 TB/sec bandwidth
- Anticipate 10-15 allocations per cycle





bruce.loftis@colorado.edu



Bridges at PSC

- ❑ Designed to support familiar, convenient software and environments for **both traditional and non-traditional HPC users**.
- ❑ Compute nodes with hardware-supported shared memory ranging from 128GB to 2TB per node to support genomics, machine learning, graph analytics and other fields where partitioning data is impractical
- ❑ GPU nodes to accelerate applications as diverse as machine learning, image processing and materials science
- ❑ Database nodes to drive gateways and workflows
- ❑ Webserver nodes to host gateways and provide access to community datasets
- ❑ Data transfer nodes with 10 GigE connections to enable data movement between Bridges and XSEDE, campuses, instruments and other advanced cyberinfrastructure



JetStream at Indiana U

- ❑ <http://jetstream-cloud.org>
- ❑ Cloud-based computation
- ❑ Focus on ease-of-use and broad accessibility
- ❑ Researchers will be able to create virtual machines on the remote resource that look and feel like their lab or home machine – but much larger
- ❑ Run interactively
- ❑ Available spring 2016



Wrangler @ TACC

- ❑ The only Hadoop/HDFS system is on the TACC nodes, so if you need MapReduce or Spark, TACC is where you need to run.
- ❑ TACC nodes have flash storage - IU nodes do not

Stampede at TACC

- ❑ Dell system – Intel Xeon nodes
- ❑ 16 large memory nodes, each with 32 cores and 1 TB of RAM
- ❑ 128 nodes with GPUs
- ❑ 102400 cores
- ❑ 2 Pflops peak performance – 7 on the Xeon Phi
 - your mileage may vary

Comet @ SDSC

- Intel Haswell nodes
- 24 cores per node
- 2 Pflops peak

