Introduction to FREE National Resources for Scientific Computing

Dana Brunson
Oklahoma State University
High Performance Computing Center

Jeff Pummill
University of Arkansas
High Peformance Computing Center

High Performance Computing Center

Oklahoma Supercomputing Symposium
October 6, 2010

Arkansas High Performance Computing Center

Questions

- What free national resources are available?
- Who can use them?
- What hardware, software, and training materials do they offer and how do I find them?
- Do they have advanced user support?
- Where can I find more information?

Resource Providers

- **DoD Supercomputing Resource Centers (DRSCs)**

  http://www.hpcmo.hpc.mil/cms2/index.php/hpccenters

- **DoE Innovative and Novel Computational Impact on Theory and Experiment (INCITE)**

  http://www.er.doe.gov/ascr/incite/

- **NASA Advanced Supercomputing Division (NAS)**

  http://www.nas.nasa.gov

- **Open Science Grid (OSG)**

  http://www.opensciencegrid.org/

- **TeraGrid**

  https://www.teragrid.org/

Department of Defense
Supercomputing Resource Centers

Department of Defense
Supercomputing Resource Centers (DSRCs)

DoD DRSCs

- DoD Projects requiring HPC are eligible.
- Variety of clusters and shared memory systems. Consolidated hardware list: http://www.afrl.hpc.mil/consolidated/hardware.php
- Many programming and analysis codes available. Consolidated software list: http://www.afrl.hpc.mil/consolidated/softwareALL.php
- User support services begins with Consolidated Customer Assistance Center (CCAC) and can continue with more specialized user assistance.
- More information at: http://www.hpcmo.hpc.mil/cms2/index.php/hpccenters

Department of Energy

Innovative and Novel Computational Impact on Theory and Experiment (INCITE)

Department of Energy

Innovative and Novel Computational Impact on Theory and Experiment (INCITE)

- Open to all researchers and research organizations—academic, governmental, and industrial, through peer review.

- Intent is to serve very large scale computationally intensive projects.

- Argonne National Lab and Oak Ridge National Lab host computational resources.

- More information:
http://hpc.science.doe.gov/allocations/incite/faq.do

Intrepid IBM Blue Gene/P at ANL

Jaguar Cray XT at ORNL

Jaguar
224,162 Compute Cores
#1 on Top500 – Nov '09

NASA Advanced Supercomputing Division (NAS)

NASA
Advanced Supercomputing Division (NAS)

- NASA projects are supported.

- NAS' current high-end computing environment includes three supercomputers, and a 25-petabyte mass storage system for long-term data storage.

- NAS has pioneered many technologies and techniques that have become standards for integrating supercomputers into a production environment.
http://www.nas.nasa.gov/About/legacy.html

Pleiades at Ames Research Center

Open Science Grid (OSG)

Open Science Grid (OSG)

• OSG brings together computing and storage resources from campuses and research communities into a common, shared grid infrastructure over research networks via a common set of middleware.

• OSG offers participating research communities low-threshold access to more resources than they could afford individually, via a combination of dedicated, scheduled and opportunistic alternatives.

• OSG is a consortium of software, service and resource providers and researchers, from universities, national laboratories and computing centers across the U.S., who together build and operate the OSG project. The project is funded by the NSF and DOE, and provides staff for managing various aspects of the OSG.

• OSG provides training through hands-on workshops and focused engagement with the community, helping new users to run applications on the infrastructure and resource owners to make their compute and storage resources accessible to the grid.

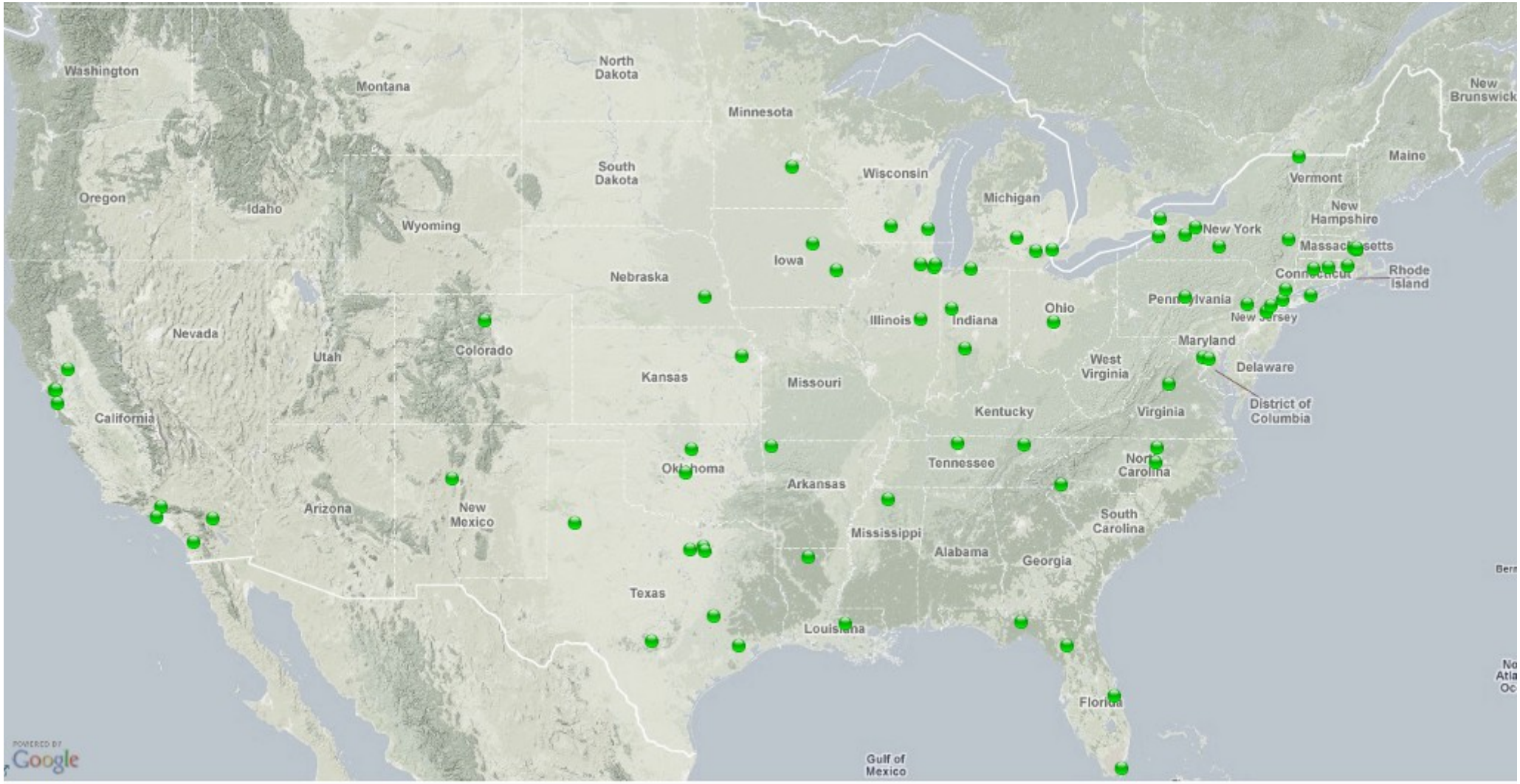
OSG: How it works

• Resource owners register their resource with the OSG.

• Scientific researchers gain access to these resources by registering with one or more Virtual Organizations (VOs).

• The VO administrators Register their VOs with the OSG. All members of the VO who have signed the acceptable use policy (AUP) are allowed to access OSG resources, subject to the policies of the resource owners.

• Each resource and each VO is supported by a designated, and in some cases shared, "Support Center (SC)," determined at registration time.


Washington
Montana
North Dakota
Minnesota
South Dakota
Wisconsin
Michigan
Oregon
Idaho
Wyoming
Nebraska
Iowa
Illinois
Indiana
Ohio
Nevada
Utah
Colorado
Kansas
Missouri
West Virginia
Maryland
Delaware
District of Columbia
Virginia
Kentucky
California
Oklahoma
Tennessee
Arkansas
Arizona
New Mexico
Mississippi
Alabama
Georgia
South Carolina
Texas
Louisiana
Florida
Gulf of Mexico
Vermont
Maine
New Hampshire
New York
Rhode Island
New Jersey
Google

TeraGrid

What is TeraGrid?

TeraGrid is an open scientific discovery infrastructure combining leadership class resources at eleven partner sites to create an integrated, persistent computational resource.

Using high-performance network connections, the TeraGrid integrates high-performance computers, data resources and tools, and high-end experimental facilities around the country. Currently, TeraGrid resources include more than a petaflop of computing capability and more than 30 petabytes of online and archival data storage, with rapid access and retrieval over high-performance networks. Researchers can also access more than 100 discipline-specific databases. With this combination of resources, the TeraGrid is the world's largest, most comprehensive distributed cyberinfrastructure for open scientific research.


TERAGRID
UC/ANL
PSC
NCSA
PU
NCAR
IU
ORNL
NICS
SDSC
LONI
TACC

TeraGrid – Who is Eligible?

• To qualify for an allocation, the principal investigator (PI) must be a researcher or educator at a U.S. academic or non-profit research institution.

• A qualified advisor may apply for an allocation for his or her class, but a high school, undergraduate or graduate student may not be a PI.

• A postdoctoral researcher can also be a PI. (After receiving an allocation, PIs can request that students be given accounts to use the allocation.)

In general, TeraGrid follows the guidelines described in the current NSF Grant Proposal Guide. However, investigators with support from any funding source, not just NSF, are encouraged to apply. If your institution is not a university or a two- or four-year college, special rules may apply.

Contact help@teragrid.org or your local Campus Champion for details.

Kraken
98,928 Compute Cores
#3 on Top500 – Nov '09


Ranger
62,976 Compute Cores
#9 on Top500 – Nov '09

and in 2011...

...IBM Blue Waters!

With more than 300,000 compute cores, Blue Waters will achieve peak performance of approximately 10 petaflops (10 quadrillion calculations every second) and will deliver sustained performance of at least 1 petaflop on a range of real-world science and engineering applications.

Blue Waters will have:
- a peak memory bandwidth of nearly 5 petabytes/second
- more than 1 petabyte of memory
- more than 18 petabytes of disk storage
- more than 500 petabytes of archival storage.

The base system and computing environment will be enhanced through a multi-year collaboration among NCSA, the University of Illinois, IBM, and members of the Great Lakes Consortium for Petascale Computation. The enhanced environment will increase the productivity of application developers, system administrators, and researchers by providing an integrated toolset to use Blue Waters and analyze and control its behavior.

Advanced Support for TeraGrid Applications (ASTA)

Advanced Support for TeraGrid Applications (ASTA) provides collaboration between Advanced User Support (AUS) staff and users of TeraGrid resources. The objective of the program is to enhance the effectiveness and productivity of scientists and engineers. As a part of the ASTA program, guided by the allocation process, one or multiple AUS staff will join the principle investigator's (PI's) team to collaborate for up to a year, working with users' applications.

AUS staff from TeraGrid resource provider sites have expertise and experience in many areas of high performance computing, domain sciences, data visualization and analysis, data management, and grid computing. Collaborative work can include any of the following:

-Porting applications to new resources
-Providing help for portal and gateway development
-Implementing algorithmic enhancements
-Implementing parallel math libraries
-Improving the scalability of codes to higher processor counts
-Optimizing codes to efficiently utilize specific resources
-Assisting with visualization, workflow, and data analysis/transfer
-Developing a Science Gateway

Teragrid Science Gateways

- Science Gateway is an integrated, community-developed set of tools, applications, and data.
- Access is usually via a portal or a suite of applications, often in a graphical user interface, that is further customized to meet the needs of a targeted community.
- Most existing gateways are available for use by anyone.
- Existing science gateways include resources for astronomy, chemistry, earthquake mitigation, geophysics, global atmospheric research, biology and neuroscience, cognitive science, molecular biology, physics and seismology, among others.
- Current list: https://www.teragrid.org/web/science-gateways/gateway_list

Questions?