The purpose of this tutorial is to supplement the standard XSEDE new user tutorial with overviews of how to use supercomputers and clusters. Campus Champions who work closely with new users are also included as presenters.

Prerequisites: participants should bring a laptop and have one of the following: Python 3.x, R, Julia, or MATLAB. Participants should have an understanding of basic scripting and programming concepts.

Speakers: Srijith Rajamohan

ARCC Tutorial: Enabling and Advancing Research Computing on Campuses

This workshop is the first of a series to provide a forum for leaders, technology developers, and resources to share, learn, and advance the use of high performance computing. Attendees will learn about the evolutionary aspects of research computing centers and the challenges they face.

Speakers: Robert Gardner, Mats Rynge

Pandas: a high-level library for data analysis

Pandas is a high-level open-source library that provides data analysis tools for Python. It provides an easy-to-use, yet powerful data structures and data analysis tools. Pandas can be used for data cleaning, data manipulation, and data processing.

Speakers: Srijith Rajamohan

3. Download sample data from here

Level: Introductory

1. Computer, mouse with scroll wheel (tablets are not sufficient for this tutorial)

Packaging with Python's distutils

Speakers: Robert Gardner, Robert Deleon

The development of high-performance packages for Python can be challenging due to the large number of dependencies and the complexity of the build process. Packaging tools like distutils can help simplify this process.

Speakers: Robert Gardner, Robert Deleon

Building recipes. After installation, attendees will then be able to login to the installed clusters to gain interactive access. Packages will be built for common components, including scientific computing libraries.

Speakers: Robert Gardner, Robert Deleon

HPC-centric package builds for a variety of common components in an effort to minimize duplication and provide tools that are portable and can be used across different platforms.

Speakers: Robert Gardner, Robert Deleon

MVAPICH2 is an opensource high-performance MPI library that provides a flexible and scalable platform for parallel computing. Attendees will learn how to use MVAPICH2 and its associated MPI-T tools for performance analysis.

Speakers: Robert Gardner, Robert Deleon

Participants will learn how to use the TAU Performance System with MPI and OpenMP and use the MPI-T interface from the MVAPICH2 library on the Stampede system at TACC and on the VM. This will help to prepare participants for further into creating scripted launch of additional resources as needed. Beyond that, elastic computing techniques will be presented in live-demo mode, and each section will feature an exercise so participants can try out the tools on their own.

Speakers: Robert Gardner, Robert Deleon

Earthquake hazard estimation requires systematic investigation of past records as well as fundamental understanding of the physics of seismic source processes. This tutorial will focus on the use of high-performance computing to analyze earthquake hazard.

Speakers: Robert Deleon

Continued advances in high-performance computing architectures constantly move the computational and data processing requirements of scientific research. Researchers are continually faced with the challenge of efficiently using these resources, which often require expertise in programming and configuration. This tutorial will focus on best practices for managing these resources and the tools used to do so.

Speakers: Mehmet Belgin, Blake Fleischer, Ruizhu Huang, Fang Liu, Weijia Xu

Blue Waters engendered by the extremely large jobs that it typically executed. In this tutorial, we will explain how attendees can assist others, as well as briefly describe projects that are being currently carried out.

Speakers: Patrick Schmitz, UC Berkeley

The Advisory Committee of the Social, Behavioral, and Economics Program, National Science Foundation, 2001-2011.

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation

ARCC Conference on Workforce Diversity & Evaluation
4:00pm – 12:30pm – 4:00pm –

Cybersecurity Center of Excellence (CCoE) will be on hand to lead the discussion and answer questions.

Neural network training microbenchmarks, the standard Xeon-based nodes performed up to four times faster of microbenchmarks of linear algebra kernels and machine learning functionality that includes clustering and task for research and data support staff. Purdue Research Computing staff have developed a web-based system.

This session is appropriate for principal investigators, program officers, IT professionals in research and higher education, and graduate students.

9:30-10:30 am — Workshop on Trustworthy Scientific Cyberinfrastructure (Part 1)

US, Canadian, European, and other large-scope research computing jurisdictions.

Information sharing and collaboration among resource providers be beneficial? Is there desire to form a

Given the importance of collaboration in current scientific practice, user collaboration should be a prime

For example, a chemistry research group may want to run simulations collaboratively, analyze experimental data multiphase flows in the porous media of a reservoir. This solver is realized by the coupling of DPM (discrete

Student-Exhibitor Speed Networking

challenges that these opportunities are not offered at far too many institutions, course curricula are not keeping diverse computational, data-enabled and HPC workforce in all fields of study across all sectors of society. There

and favorite tricks. This will be a place where we will discuss our progress and best practices, articulate

Is there a silver bullet for successful job management in science gateways? Probably not, but both the process

"Technology Audit and Insertion Service" component of the "TeraGrid Phase III: eXtreme Digital Resources for

Supercomputing Center, The University of Texas at Austin Texas Advanced Computing Center, University of

is a uniquely capable resource for empowering new research communities and bringing together HPC and Big

fragments. Among the main features are the minimal technical expertise requirement for the end user and an

GenApp is a universal and extensible tool for rapid deployment of applications. GenApp builds fully functioning

offering resources to local campus users is under constant budgetary pressure. This is partly due to the difficulty

existing algorithms and developing new algorithms and running them on the Comet supercomputer to analyze

Engineering Discovery Environment) and Extended Collaborative Support Service (ECSS). The team is refining

chemistry, biology, and physics earlier in the curriculum is imperative to success in advanced classes and

The structure and dynamics of proteins are an essential part of understanding the molecular foundations of

collaboration to other regional users.

a structured environment for development of graduates with the skills needed to enter the HPC workforce. The

utilizing supercomputing at mid-level and smaller research institutions. Simply having HPC hardware and some

In this presentation we will discuss best practices and methodology for HPC software engineering. We will

required physics analytics very quickly. In this work, we present our approach on applying deep learning for the

means of engaging our current student population to acquire high tech skills in the context of working on relevant

Speakers: Linda Akli, Marques Bland, Susan Fratkin, Kelly Gaither, Rosalia Gomez, Ruby Mendenhall

Research & Development: BOFs and Workshops

this work: 1) the challenges of registering and authenticating users in a distributed environment, and 2) the security requirements of the award agencies, in this evolving future of CyberInfrastructure.

Break

— Matt Rehme, NCAR

— Scott Pearse, Pedro Jimenez, NCAR

Tutorial Half-day

Registration

Imperial 5AB (Level 4)

Strand 12

Strand 11

Bolden 2

Bolden 6

Bolden 5

Bolden 2
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00am – 12:30pm</td>
<td><strong>Student Modeling Day Presentations</strong>&lt;br&gt;<strong>Moderators:</strong> Katherine Cartell&lt;br&gt;The teams from Student Modeling Day will describe their work to all interested attendees.</td>
<td>Strand 13</td>
</tr>
<tr>
<td>11:00am – 12:30pm</td>
<td><strong>Overcoming Bias in the Workplace — Workshop Part 2</strong>&lt;br&gt;<strong>Speakers:</strong> Kim Stephens&lt;br&gt;Meeting your workforce development goals in HPC will require cultural shifts and transformation of your recruiting and retention strategies. Overcoming bias is a challenge for any organization. Join our workshop with presenter Kim Stephens of IBM to learn how overcome unconscious and conscious bias in the workplace.</td>
<td>Strand 13</td>
</tr>
<tr>
<td>11:30am – 12:00pm</td>
<td><strong>COSMIC2: A Science Gateway for Cryo-Electron Microscopy Structure Determination Using the CIPRES Workbench Framework and Globus Services for Terabyte-sized Data Transfer</strong>&lt;br&gt;<strong>Speakers:</strong> Michael Cianfrisco, Andres Leschziner, Nick Wagner, Mona Wong, Choonhoon Yoon&lt;br&gt;Structural biology is in the midst of a revolution. Instrumentation and software improvements have allowed for the full realization of cryo-electron microscopy (cryo-EM) as a tool capable of determining atomic structures of protein and macromolecular samples. These advances open the door for solving new structures that were previously unattainable, which will soon make cryo-EM a ubiquitous tool for structural biology worldwide, serving both academic and commercial purposes. However, despite its power, new users to cryo-EM face significant obstacles. One major barrier consists of the handling of large datasets (10+ terabytes), whereas new cryo-EM users must learn how to interface with the Linux command line while also dealing with managing and submitting jobs to high performance computing resources. To address this barrier, we are developing the COSMIC2 Science Gateway as an easy, web-based, science gateway to simplify cryo-EM data analysis using a standardized workflow to run on XSEDE’s Extreme Science and Engineering Discovery Environment (XSEDE) supercomputers. This gateway will lower the barrier to high performance computing tools and facilitate the growth of cryo-EM to become a routine tool for structural biology. With the support of XSEDE’s Extended Collaborative Support Services (ECSS) and the Science Gateway Community Institute’s (SGCI) Extended Developer Support (EDS), we have adapted the successful Cyberinfrastructure for Phylogenetic Research (CIPRES) Workbench to the cryo-EM analysis workflow and are in the process of adding Globus Auth and Globus Transfer to enable the transfer of hundreds of gigabytes to several terabytes of data for analysis at the San Diego Supercomputer Center (SDSC).</td>
<td>Strand 12</td>
</tr>
<tr>
<td>11:30am – 12:00pm</td>
<td><strong>Building Bridges - The System Administration Tools and Techniques used to Deploy Bridges</strong>&lt;br&gt;<strong>Speakers:</strong> Richard Underwood&lt;br&gt;High Performance Computing is continually growing in scope and areas of research. To cover these new areas of research, HPC has to become more flexible to handle the wide variety of workloads. As the computing becomes more flexible, the infrastructure becomes more complex to accommodate these new varied workloads. At Pittsburgh Supercomputing Center, a Level 1 XSEDE Service Provider, we faced this exact problem for Bridges, our new NSF-funded configurable computing resource. This talk will cover the technical decisions that were made for Bridges, why we made them, the tools we chose, and what the users gain. We chose OpenStack Ironic for system installation, Openstack for managing virtual machines; Puppet for configuration; Slurm for scheduling, Naemon, Elastic Search/Logstash/Kibana and InfluxDB for monitoring and reporting. This software gives flexibility to the users a wide range of ways to do computing at PSC. Additionally, it gives the ability to maintain an even higher level of monitoring and reporting that changes automatically as the systems change functionality.</td>
<td>Strand 11</td>
</tr>
<tr>
<td>12:00pm – 12:30pm</td>
<td><strong>The Community Software Repository from XSEDE</strong>&lt;br&gt;<strong>Speakers:</strong> Maytal Dahan, Richard Knepper, David Lulka, Lee Liming, John-Paul Navarro, Craig Stewart&lt;br&gt;The Extreme Science and Engineering Discovery Environment (XSEDE) aims to be a connector of cyberinfrastructure (CI) resources, software, and services. By bringing together advanced digital infrastructure, expert support, and training services, XSEDE enables scholars, researchers, and engineers to participate in multidisciplinary collaborations while seamlessly accessing advanced computing resources and sharing data to tackle society’s grand challenges. To realize this vision, XSEDE must both enable and participate in a software ecosystem, and must collectively engage with software developers, integrators, vendors, users, and funding agencies to gather requirements, develop, share, and deploy software tools; and provide software documentation, training, and support. To enable community collaboration and accelerate connecting new resources, software, and service into CI, XSEDE recently introduced the Community Software Repository (CSR): a single interface to the tools and information used by XSEDE to plan, document, and organize its software-related services and features. This paper describes the CSR vision and strategy, current capabilities, future plans, and related XSEDE efforts.</td>
<td>Strand 12</td>
</tr>
<tr>
<td>12:00pm – 12:30pm</td>
<td><strong>OpenMP 4 Fortran Modernization of WSM6 for KNL</strong>&lt;br&gt;<strong>Speakers:</strong> Martin Berzins, Robert Mike Kirby, Aaron Knoll, Timbaowo Aime Judicael Ouermi&lt;br&gt;Parallel code portability in the petascale era requires modifying existing codes to support new architectures with large core counts and SIMD vector units. OpenMP is a well-established and increasingly supported vehicle for portable parallelization. As architectures mature and compiler OpenMP implementations evolve, best practices for code modernization change as well. In this paper, we examine the impact of newer OpenMP features (in particular OMP SIMD) on the Intel Xeon Phi Knights Landing (KNL) architecture, applied in optimizing loops in the single moment 6-class microphysics module (WSM6) in the US Navy’s NEPTUNE code. We find that with functioning OMP SIMD constructs, low thread invocation overhead on KNL and reduced penalty for unaligned access compared to previous architectures, one can leverage OpenMP 4 to achieve reasonable scalability with relatively minor reorganization of a production physics code.</td>
<td>Strand 11</td>
</tr>
<tr>
<td>12:30pm – 2:00pm</td>
<td><strong>Awards Luncheon</strong>&lt;br&gt;We close out PEARC17 with a plated luncheon and presentation of awards to best papers in each track, best student paper and poster, and best visualization showcase entry. Be sure to stick around—it could be you!</td>
<td>Empire CD</td>
</tr>
<tr>
<td>2:00pm – 4:00pm</td>
<td><strong>XSEDE All-Staff Meeting</strong>&lt;br&gt;<strong>Speakers:</strong> Ron Payne</td>
<td>Strand 11</td>
</tr>
<tr>
<td>3:00pm – 5:00pm</td>
<td><strong>XSEDE ECSS Training: Part 2</strong>&lt;br&gt;<strong>Speakers:</strong> Jay Alameda&lt;br&gt;Closed event. Training session for XSEDE ECSS staff.</td>
<td>Strand 11</td>
</tr>
</tbody>
</table>