July 23, 2012

Accessing Free, Large-scale Computation and Data Resources for Economics Through the eXtreme Science and Engineering Discovery Environment (XSEDE)

Philip Blood
Pittsburgh Supercomputing Center

Outline

- What Is Supercomputing?
- Overview of XSEDE Resources for Computational Economics
  - Computational Resources
  - Data Storage & Transfer Resources
  - Software Resources
- How to Access XSEDE Resources
- Example: Monte Carlo test of consistency of OLS on a supercomputer

What is a supercomputer?

A bunch of “commodity” components + enhancements

- **Processors**: essentially the same as your Mac or PC at home, arranged as “blades” (motherboards), or nodes, mounted on “racks”
- **Network**: *very* fast connection between blades (you don’t have this at home
- **Software**: Generally some flavor of UNIX, usually **Linux**, optimized for using a lot of processors together
- **Storage**: similar to what you would put into your home computer (traditional hard drives and SSDs), but *a lot* of them configured to work together in parallel for performance

Different flavors of supercomputing

- "High Performance Computing" (HPC)
  - Focused on floating-point operations per second (FLOPS)
    - Fast desktop system ~36 GigaFLOPS
  - Running 1 application across many cores
- "High Throughput Computing" (HTC)
  - Run many (smaller) applications at once
- "Data-intensive Computing"
  - Focused on I/O operations per second (IOPS)

High Performance Computing

- Distributed Memory
  - "Capability" computing
    - MPP: Massively Parallel Processing
    - biggest machines, fastest networks
    - Largest parallel calculations: 10000-100000 cores
  - "Capacity" Computing
    - Cluster
    - Slower network, but still fast
    - Smaller parallel calculations: 2-1000 cores
- Shared Memory
  - SMP: Symmetric multiprocessing
  - A single blade or node is a small SMP
  - Large SMPs bring together many nodes into a single memory space
  - Fast interconnect between nodes
  - Smaller parallel calculations: usually 2-1000 cores

High Performance Computing (cont.)

- Heteregeneous: Incorporate specialized processing elements into traditional HPC system, e.g.:
  - Graphics processing unit (GPU)
    - Only works (well) for some problems
    - Harder to program (but getting easier: CUDA, OpenCL)
    - Potentially great performance for cheap
  - Field-programmable gate array (FPGA)
    - Only works (well) for some problems
    - Requires experts to program
    - Potentially amazing performance, low power

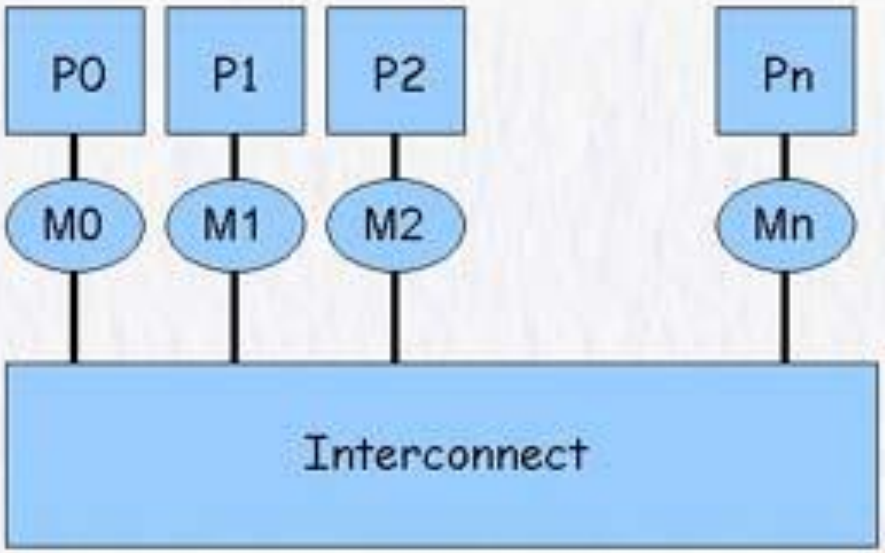
Data-intensive Computing

- "New" field of supercomputing
- Brought on by data deluge in fields like genomics, astronomy
- Data movement and manipulation dominate computation
- Strategies
  - Hardware
    - Solid-state drives (SSDs)
    - Memory drives (RAM disk)
  - Software, e.g.:
    - MapReduce (Google)
    - Hadoop (open implementation of MapReduce)

http://latimesblogs.latimes.com/.a/6a00d8341c630a53ef014e5f48074c970c-popup

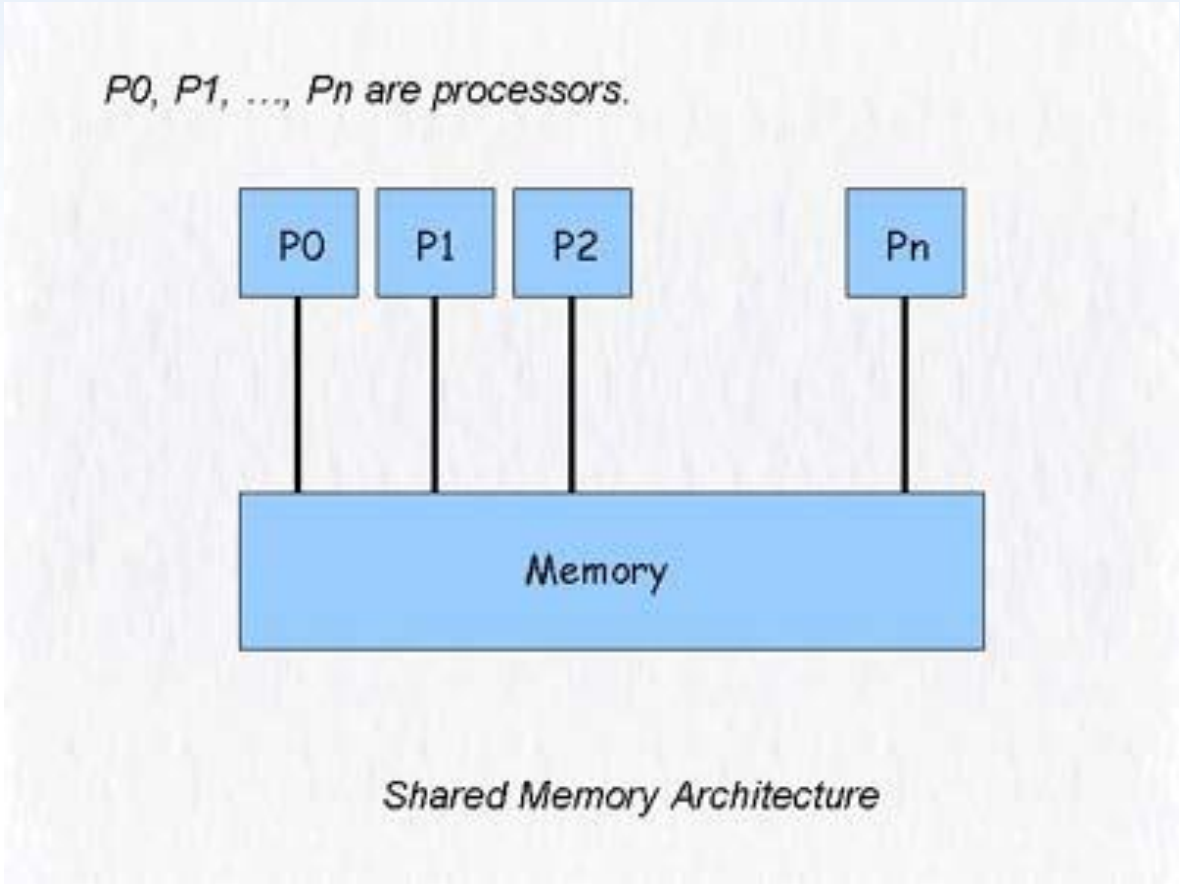
Different types of parallel platforms: Distributed Memory

M0, M1, … Mn are memories associated with processors P0, P1, …, Pn.

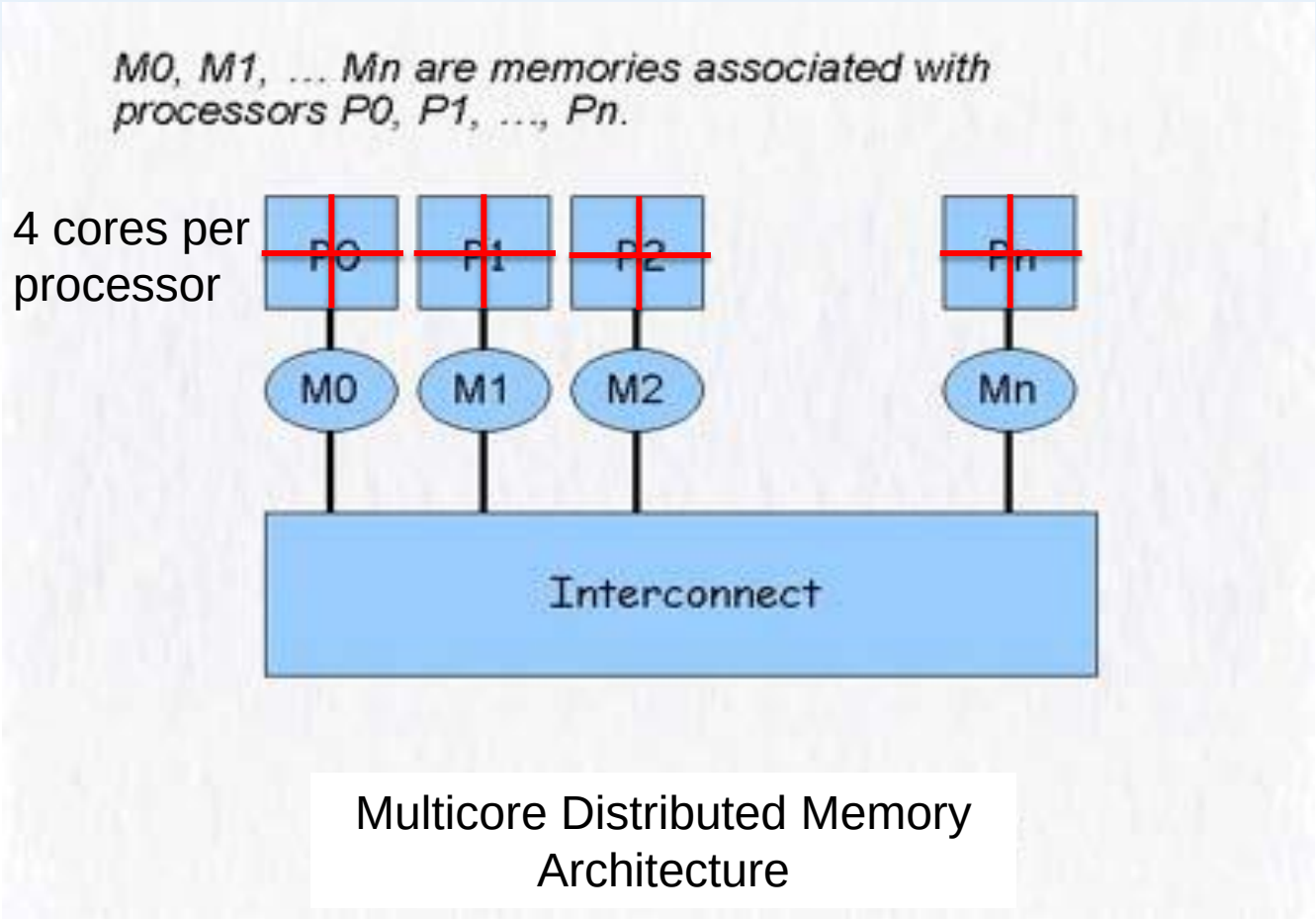


Distributed Memory Architecture

Different types of parallel platforms: Shared Memory



Multicore: a hybrid between distributed and shared memory

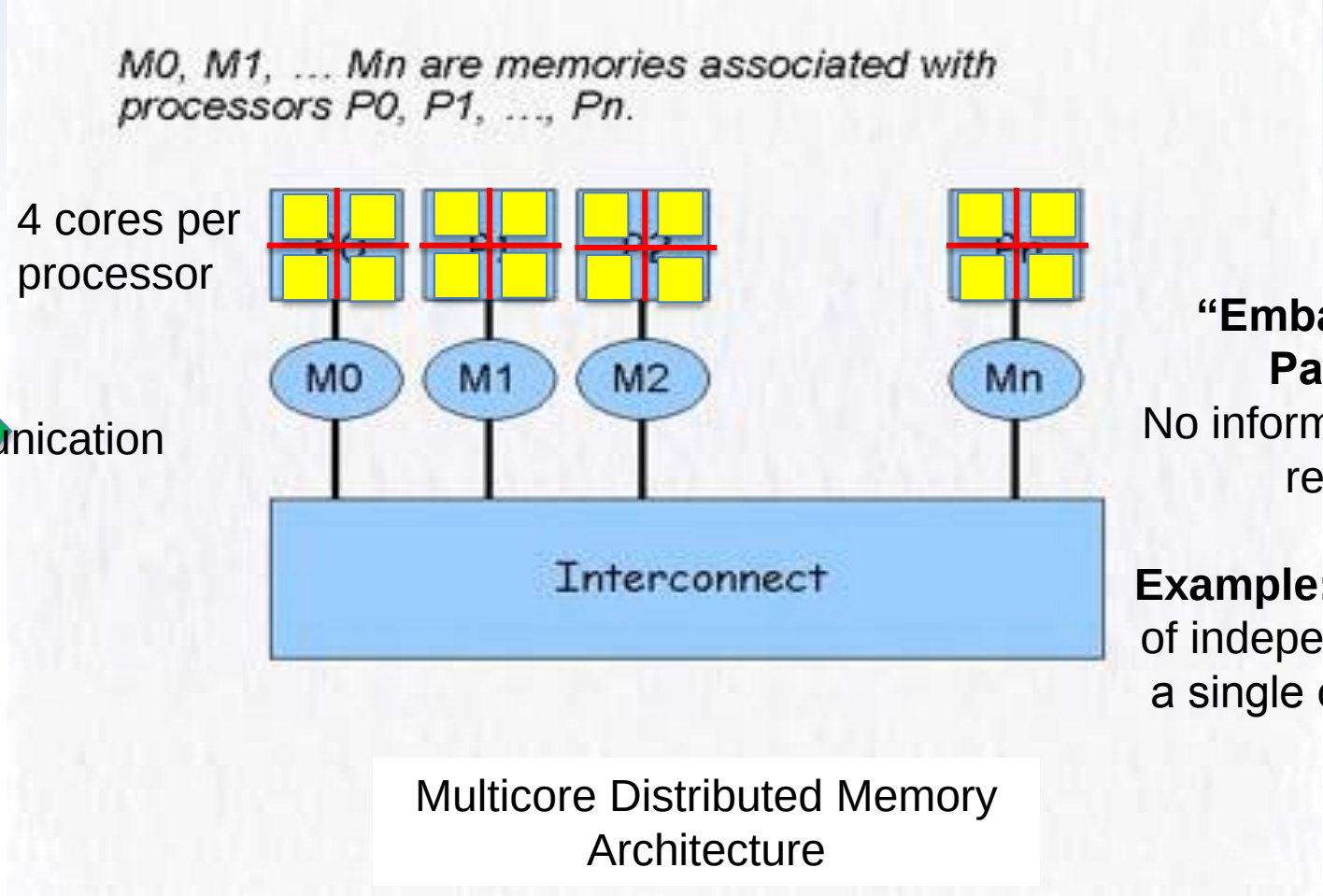


Multicore Distributed Memory Architecture

How to solve problems in parallel: Independent Jobs

Work

Communication



Multicore Distributed Memory Architecture

"Embarassingly Parallel": No information sharing required

Example: Run a bunch of independent jobs on a single core ("serial")

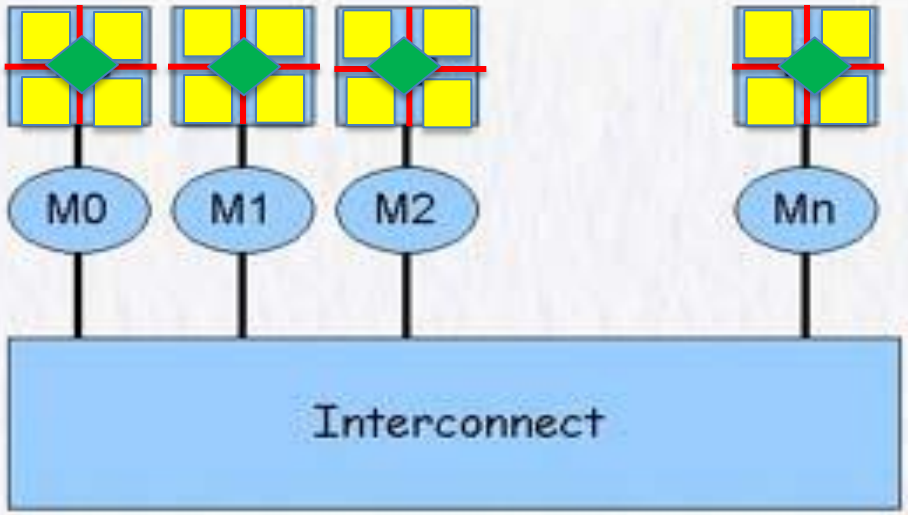
How to solve problems in parallel: Shared memory programming

Work

Communication

M0, M1, … Mn are memories associated with processors P0, P1, …, Pn.

4 cores per processor



Multicore Distributed Memory Architecture

"Shared Memory Parallelism": Threads do work separately; share information through common variables in memory

Example: OpenMP – easily split loops between multiple threads

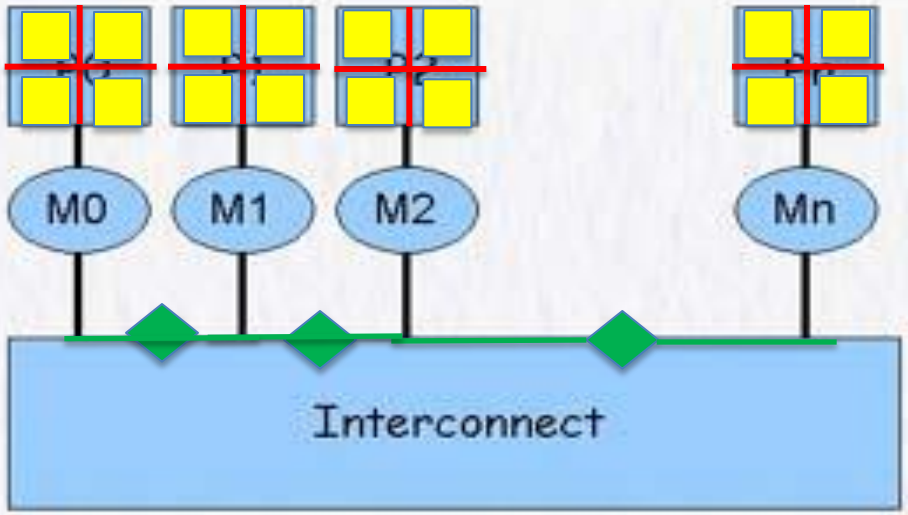
How to solve problems in parallel: Distributed memory programming

Work

Communication

M0, M1, … Mn are memories associated with processors P0, P1, …, Pn.

4 cores per processor



Multicore Distributed Memory Architecture

"Message Passing": Processes send **messages** to each other when they need to communicate

Example: MPI -- Message Passing Interface

Pros and Cons of parallel programming methods

- **Shared memory**
  - Pros
    - Can be done easily using OpenMP directives or packages that use OpenMP
    - Directly parallize existing serial code (parallelize existing loops)
  - Cons
    - Usually a limited amount of parallelism
    - Sharing memory between threads can be perilous
- **Distributed memory**
  - Pros
    - Can scale to very high processor counts with MPI
    - Ubiquitous
  - Cons
    - Difficult to program if you have to do it yourself (for complex cases)

July 23, 2012

What is XSEDE?

Extreme Science and Engineering
Discovery Environment

XSEDE in less than 20 words

XSEDE is a single virtual system that scientists can use to interactively share computing resources, data and expertise.

XSEDE Consists of Many Partners

- eXtreme Science and Engineering Discovery Environment (XSEDE)
- Freely accessible virtual system for open scientific research, funded by NSF
- Provides compute, data, and visualization resources as well as training and support services.
- Integrates resources via centralized services, common software, and fast networks
- Currently composed of 17 Service Providers (SPs) from around the world

XSEDE

Who Can Access XSEDE Resources?

- **XSEDE PI Requirements**: Researchers/educators at U.S. academic/non-profit institutions, (post-docs yes, but not grad students)
- **Additional Users:**

  PI's collaborators (including foreign), students
- How to access: web portal or direct log in – we will explain later in tutorial

Now in my words…

An integrated set of leading-edge computational, networking, data, software, and support resources to facilitate science.

July 23, 2012

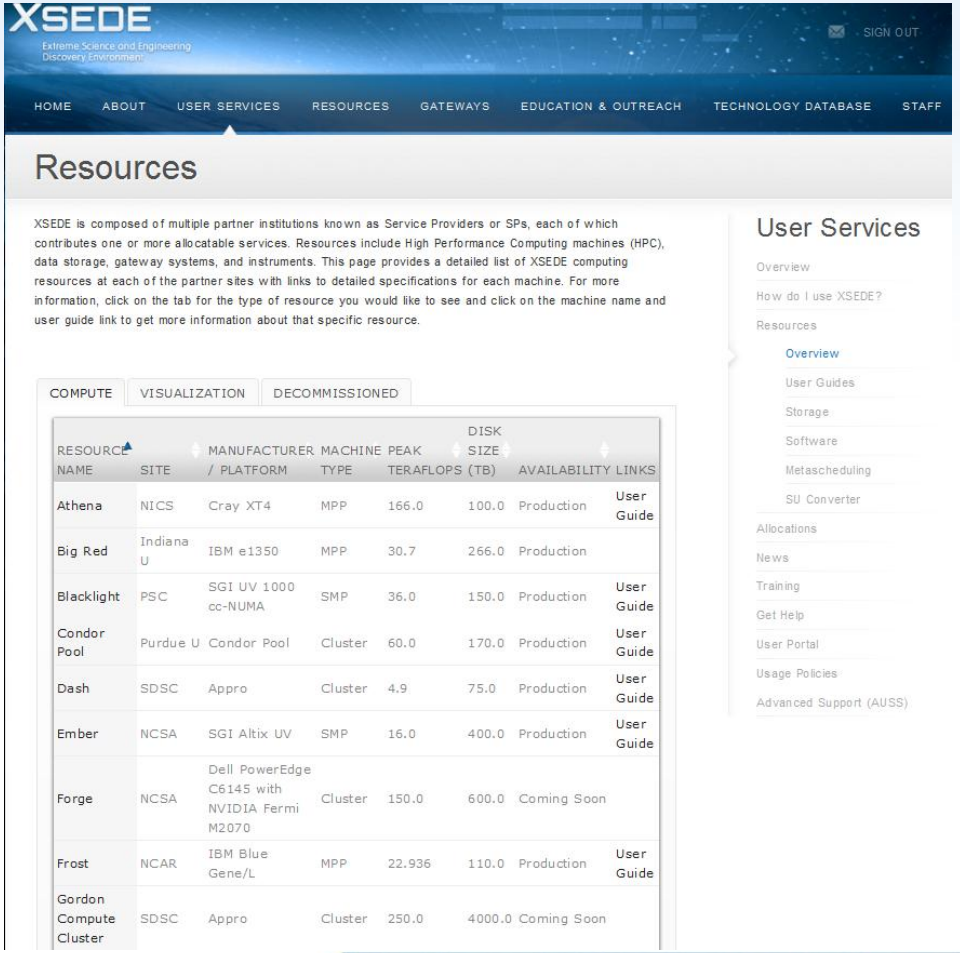
XSEDE Resources for Economics

Computational Resources

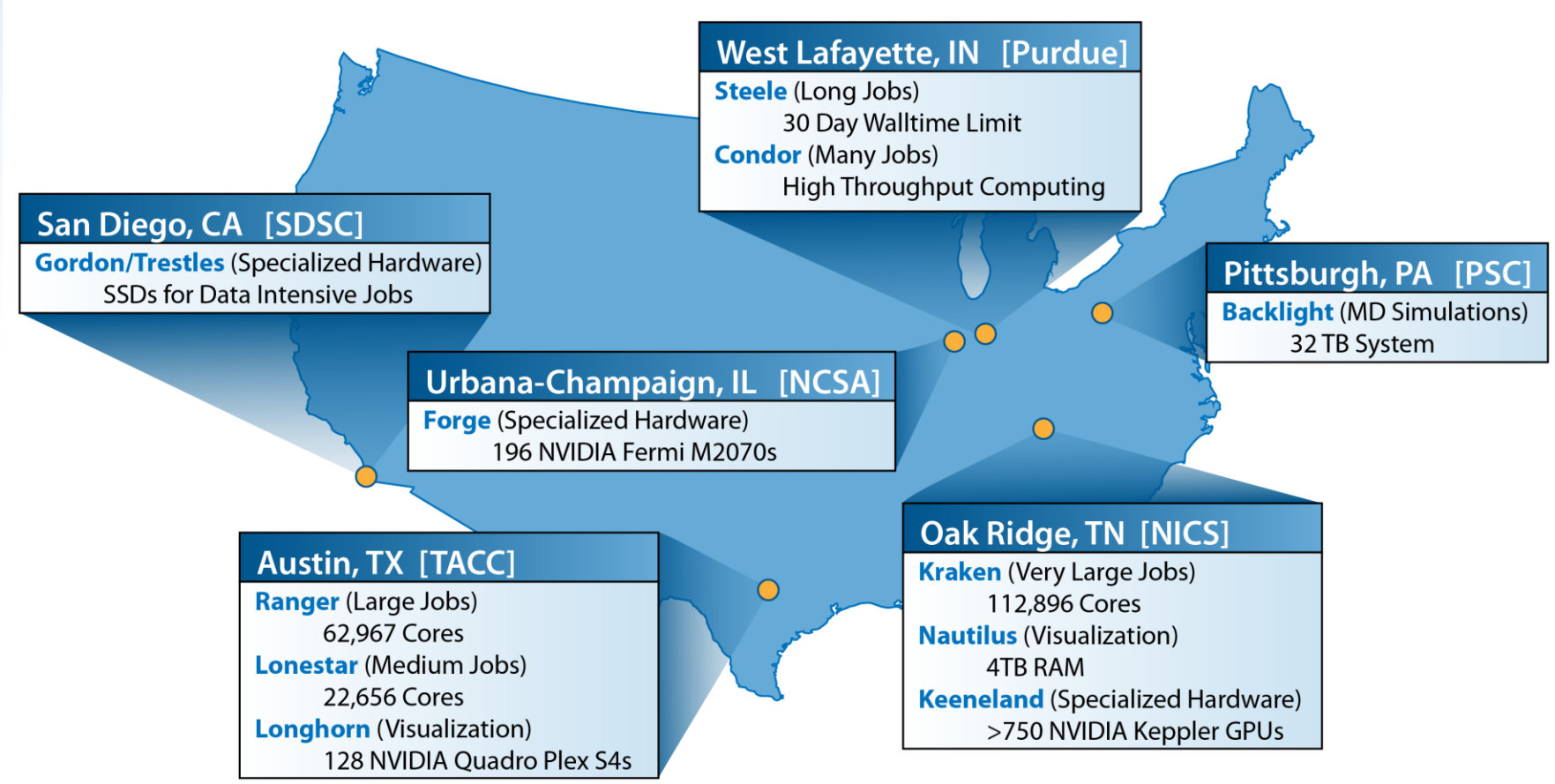
What resources are available?

- Compute
  - Condor Pool, OSG (HTC)
  - Shared Memory (16 TB)
  - Massively Parallel (100,000 cores)
  - Click on name in portal for info
- Special purpose: heterogeneous, data-intensive web hosting
- Visualization
- Networking and Storage
  - $HOME, $SCRATCH, and archival storage comes with allocation
  - Can also request purely data allocations
- Science Gateways (Web Portals)
- **Human**: Extended Collaborative Support

https://portal.xsede.org/web/guest/resources



XSEDE Resource Map



CONTACT: HELP@XSEDE.ORG

XSEDE

July 23, 2012

XSEDE Resources for Economics

Data Transfer & Storage Resources

XSEDE

Extreme Science and Engineering Discovery Environment

Transferring Data to (and between) XSEDE resources

- scp: (XSEDE uses high performance scp)
  - Just use regular ‘scp’ or ‘gsiscp’ between sites
- Globus Online: web-based optimized file transfer
  - Easy way to use fast gridftp on XSEDE, 100+ MB/s
  - Tracks transfers and automatically restarts, other features
  - https://www.globusonline.org/

XSEDE Data Storage Resources

- All active users with compute allocations get access to:
  - short-term storage
  - long-term storage
- Special 'data allocations' also available

Long-term storage

- Albedo
  - 1 PB Lustre distributed WAN filesystem
- Data Supercell @ PSC
  - 4 PB disk
- HPSS @ NICS
  - 6.2 PB tape
- MSS @ NCSA
  - 10 PB tape
- Ranch @ TACC
  - 50 PB tape

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

July 23, 2012

XSEDE Resources for Economics

Software Resources

What software is available?

- Comprehensive software search feature: https://portal.xsede.org/web/guest/software
- If you have a license for it, we can (usually) run it (somewhere)
- We can install software that is not currently available
- Send questions/requests to help@xsede.org

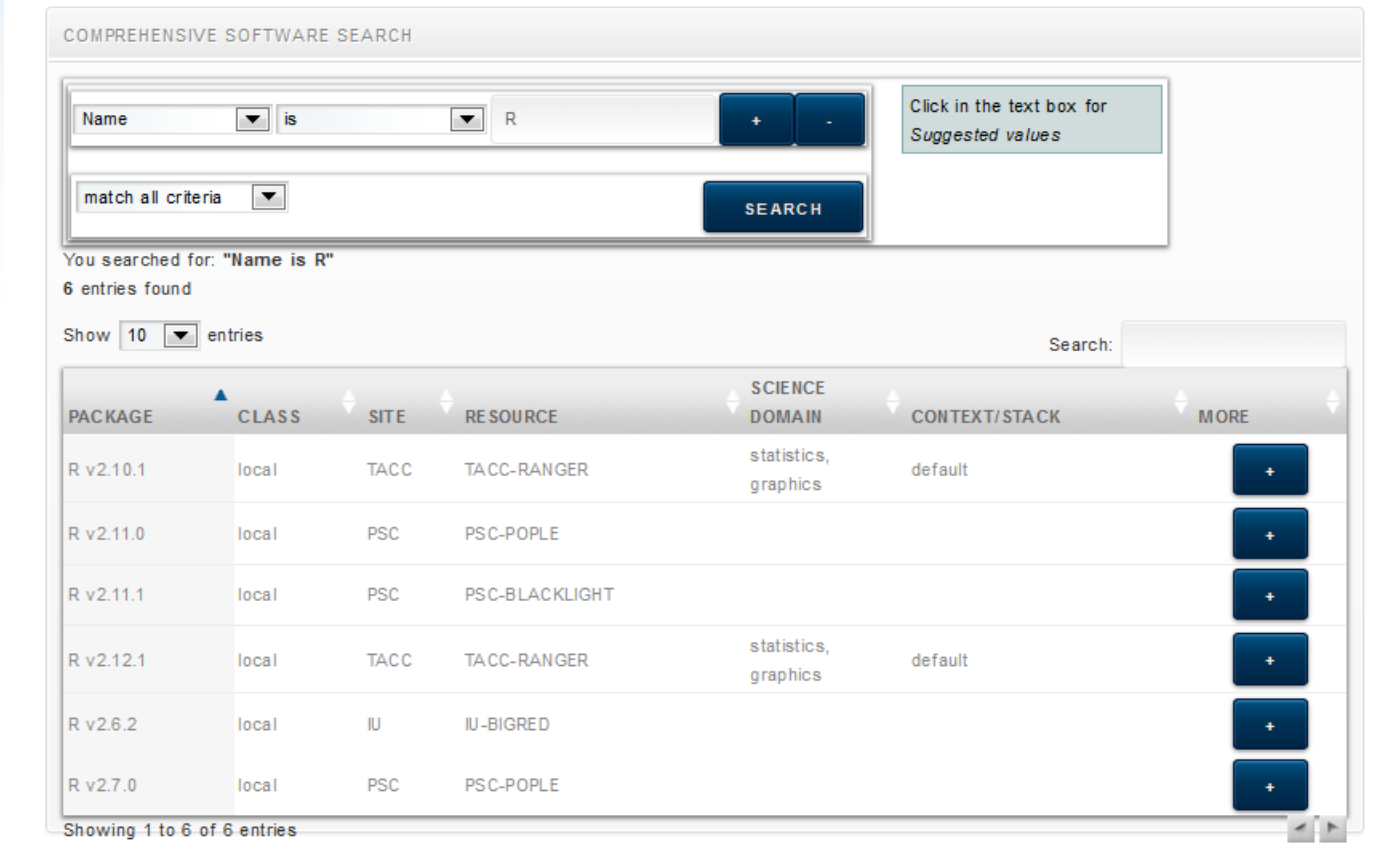
Software for Economics at PSC

- Matlab
  - http://psc.edu/index.php/matlab
- R
- Python:
  - SciPy
  - NumPy
- MPI/OpenMP

- Not all available software officially deployed, so may not all appear in "Comprehensive" search
- Most software can be installed on request
- 1st Determine what computational/data resources are best for you and software needs can usually be addressed (assuming it is open, or you have a license)

Example: Comprehensive software search

https://portal.xsede.org/web/guest/software



Using R in parallel

- Parallel libraries available
  - threaded
    - compile with threaded MKL library (ask a sysadmin)
    - **pnmath**: uses OpenMP versions of standard math libraries
  - distributed
    - **multicore** – creates new processes to execute functions on different cores
    - **RMPI** – facilitates using message passing within R
  - **foreach:** allows you parallelize loops using simple directives, using various methods "under the hood" that you can choose: **RMPI, multicore**
- http://www.nics.tennessee.edu/computing-resources/nautilus/software?&software=r

July 23, 2012

How to Access XSEDE Resources

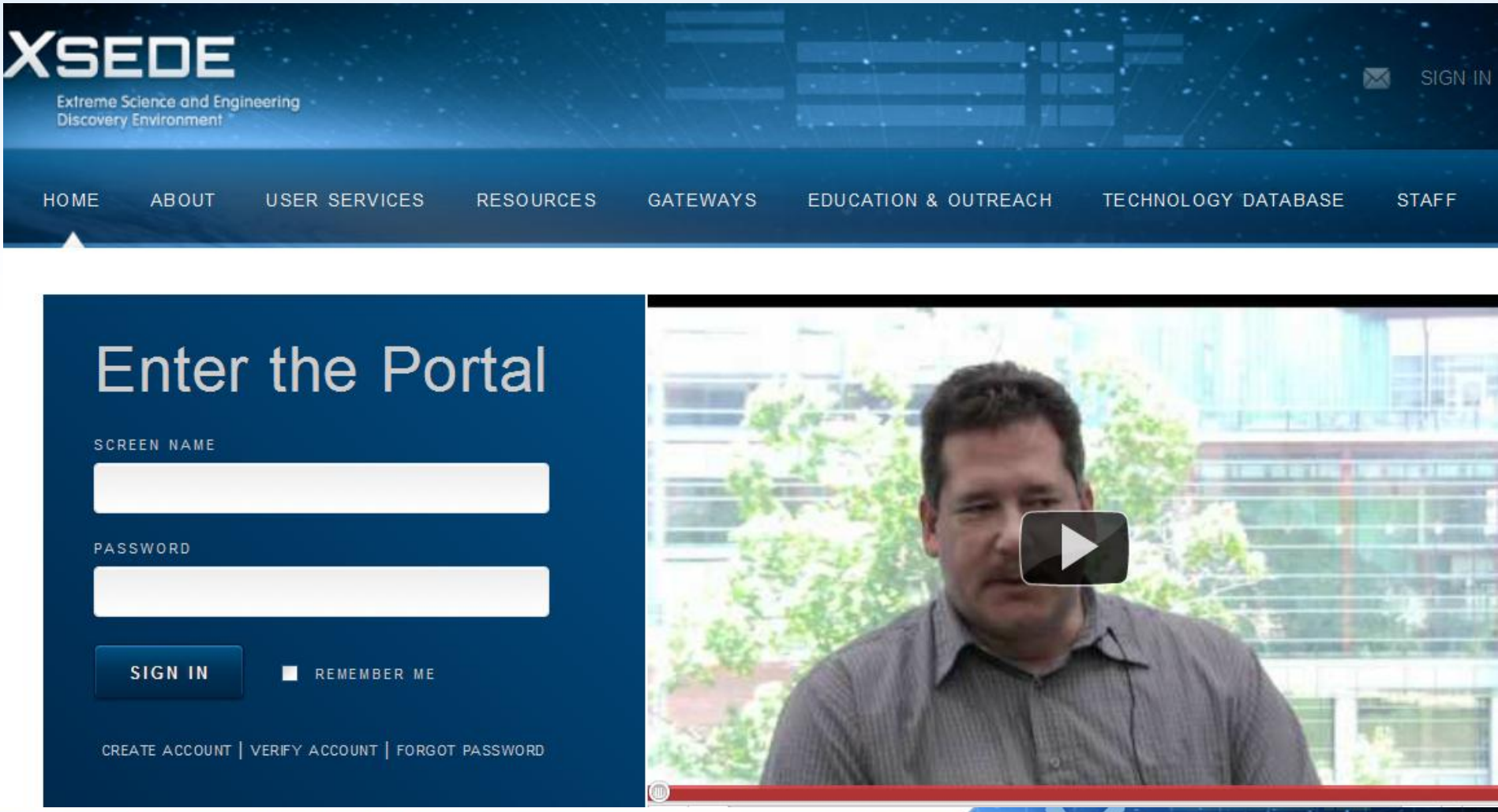
Allocations

- To get started using XSEDE a researcher needs to:
  - Apply for an XSEDE allocation
  - Or request to be added to an existing allocation (Additional User).
  - You do either of these through the XSEDE User Portal.
- In addition to research, can also receive allocations for educational/instructional purposes

XSEDE User Portal

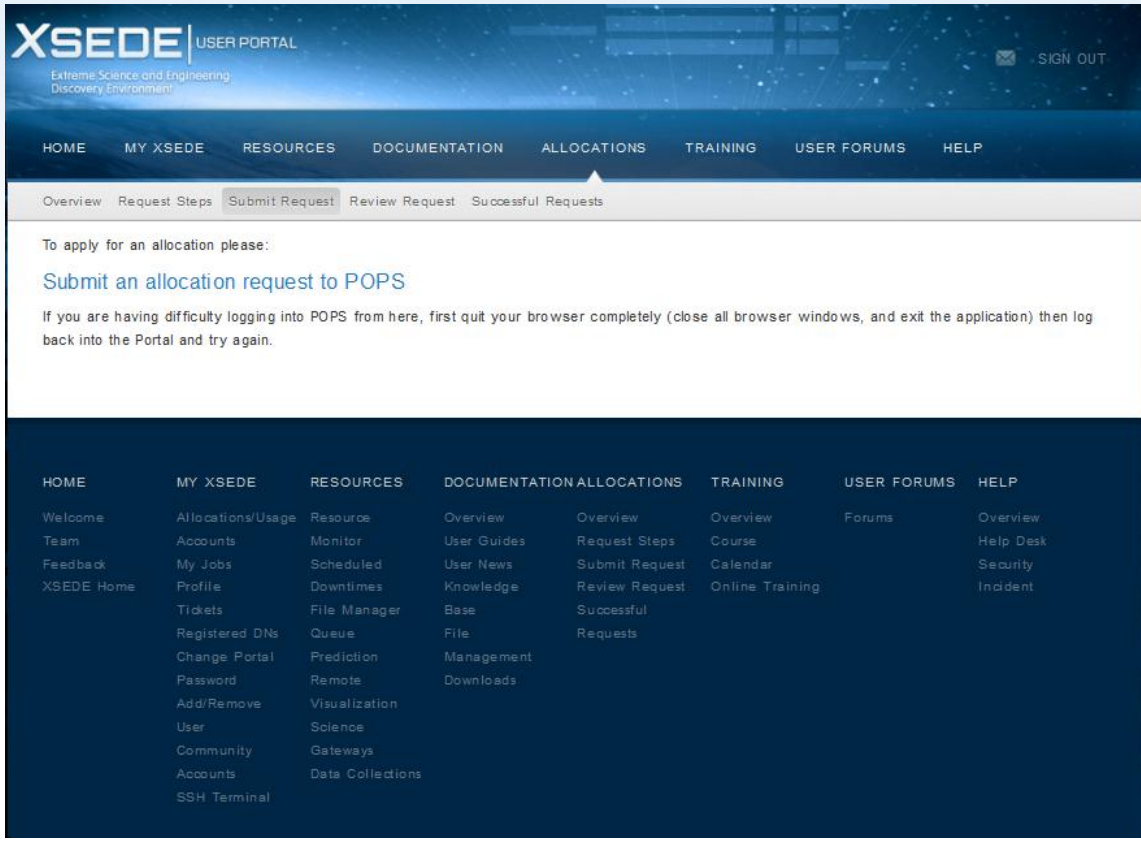
- Web-based single point of contact that provides:
  - Continually updated information about your accounts.
  - Access to your XSEDE accounts and allocated resources: *The Portal provides a single location from which to access XSEDE resources. One can access all accounts on various machines from the Portal.*
  - Interfaces for data management, data collections, and other user tasks and resources
  - Access to the Help Desk.

Create a portal account from the main xsede.org site



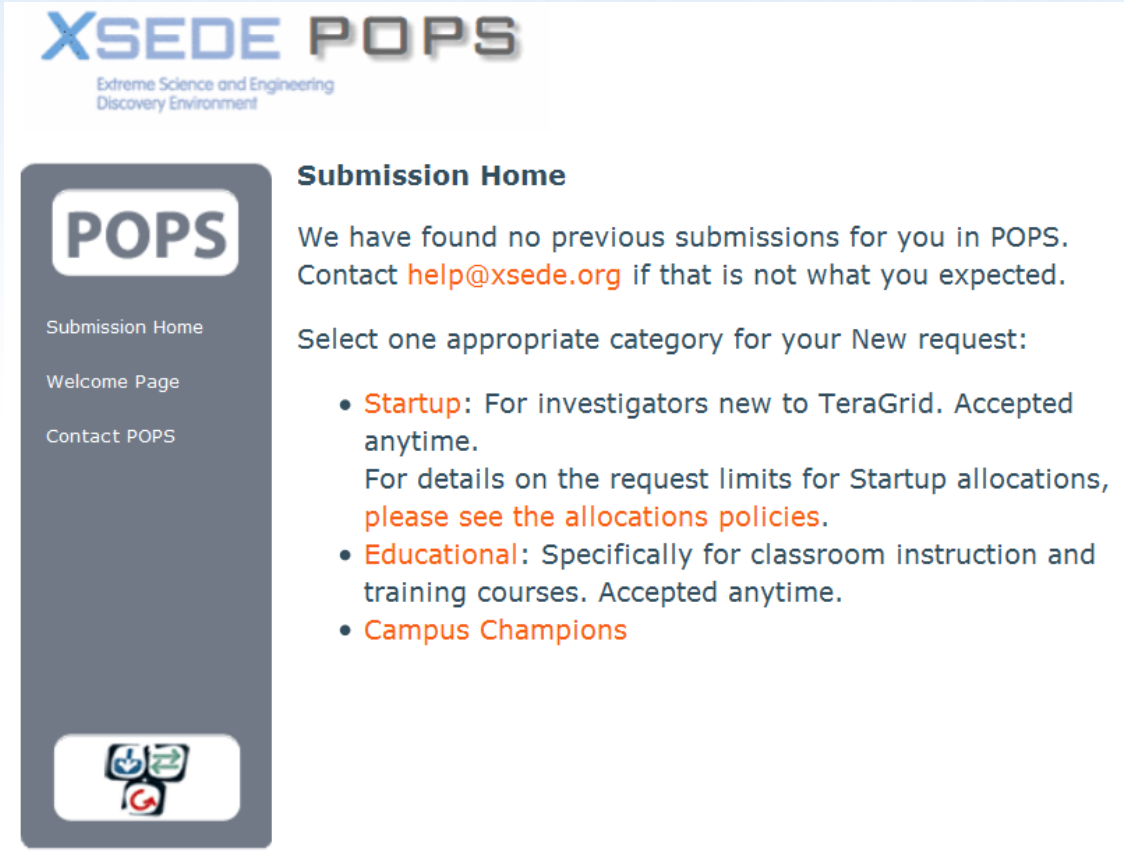
Request an XSEDE allocation (1)

- Go to **Allocations** tab
- Click on **Submit Request** heading
- Click "**Submit an allocation request to POPS**"
- From next page: "**Click to Enter or View a Request**"



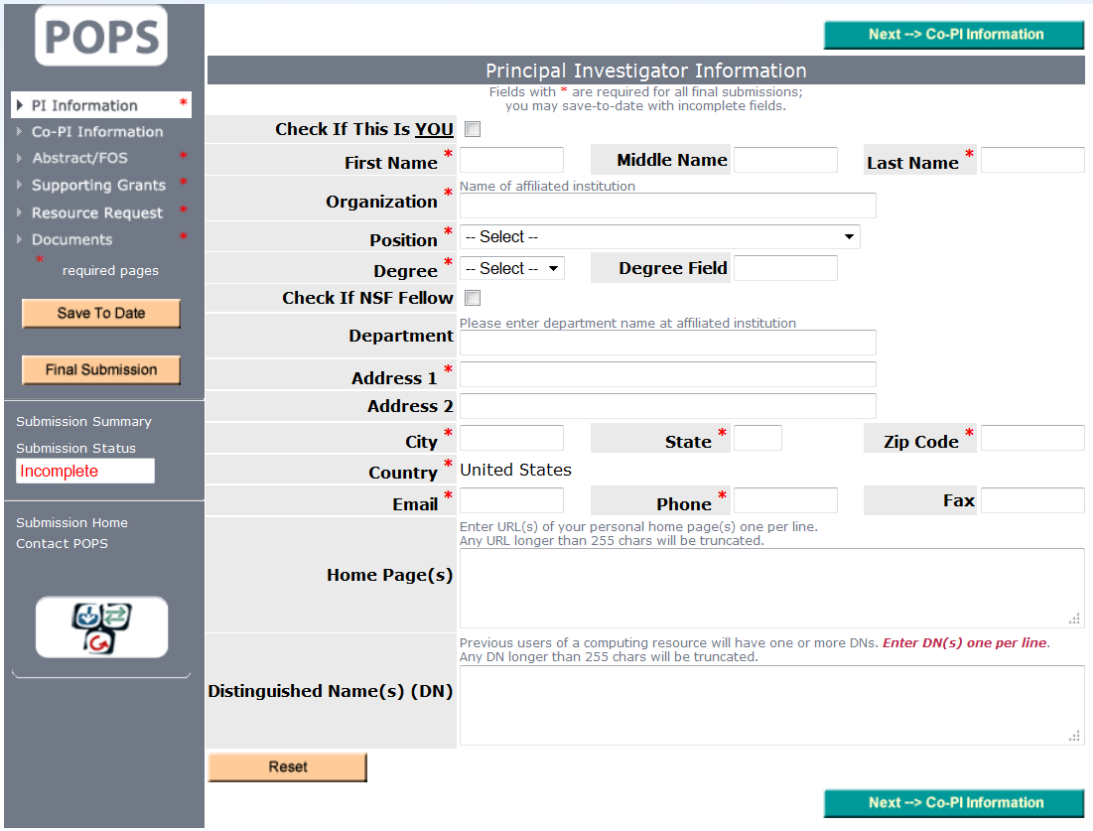
Request an XSEDE allocation (2)

- You can request various allocations types
- **Startup** allocations are good for new users
  - Total limit of 200K SUs
  - Individual resource limit varies, but typically 30K-200K SUs
  - Require only abstract and CV
- **Research** can be requested during quarterly allocation windows (Mar 15-Apr 15, Jun 15-Jul 15, Sep 15-Oct 15, Dec 15-Jan 15)
  - Requires written proposal



Request an XSEDE allocation (3)

- Fill in **PI** information
  - PI can be faculty/staff (including postdocs) at U.S. academic institutions
  - **Cannot** be graduate student
  - Once you get your allocation you can add other users (including foreign collaborators)
- Write Abstract (paragraph or two) explaining what you'd like to do using XSEDE resources



Request an XSEDE allocation (4)

- Choose resources of interest
- Enter requested SUs (SU = one core-hour)
- Check machine-specific startup limits
  - Go here: https://portal.xsede.org/web/guest/resources
  - Click on name of machine
  - Look for “Startup allocation limit”
  - Sum of SUs on all machines cannot exceed 200,000 SUs
- Upload **CV** for the PI and submit request

Request an XSEDE allocation (5)

- You'll get an email confirming request
- Startup requests take 1-2 weeks to be granted
- If you want a larger allocation, you need to write a proposal
- See here for more info: https://portal.xsede.org/web/guest/allocations

You can get a lot of computing for little effort!

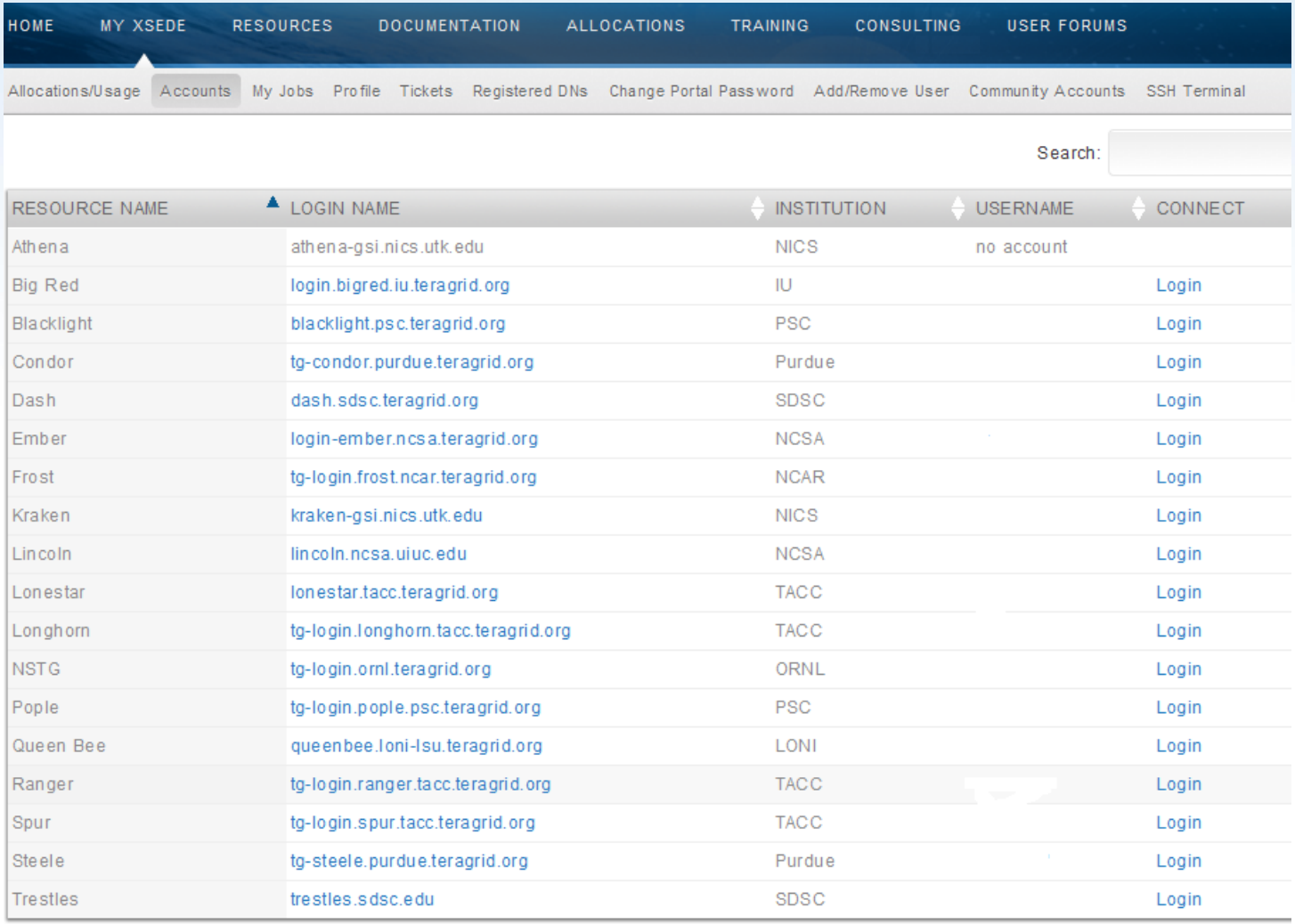
By submitting an abstract, your CV, and filling out a form, you get:

- A Startup allocation
  - Up to 200,000 SUs (core hours) on XSEDE systems for one year
  - That is the equivalent of **8333 days (22.8 years)** of processing time on a single core!
- Access to consulting from XSEDE personnel regarding your computational challenges
- Opportunity to apply for Extended Support
  - Requires answering 5 questions addressing your need for Extended Support
  - Can be done together with your Startup request, or at anytime after that

Adding Users

- Each prospective user must get an XSEDE portal account
- Once you receive the allocation you can add other users: portal.xsede.org→My XSEDE→ Add/Remove user
- Use each person's XSEDE portal username to add them to your allocation

Accessing XSEDE via portal.xsede.org

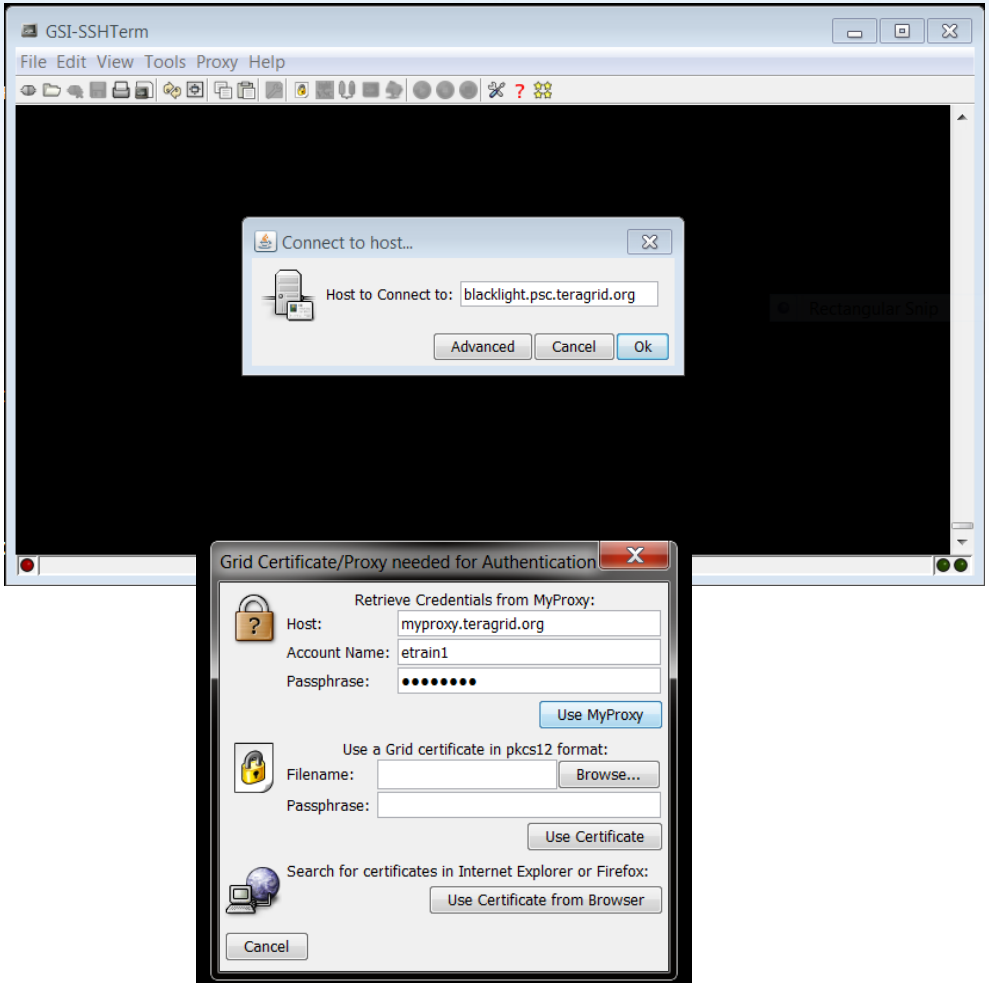


RESOURCE NAME	LOGIN NAME	INSTITUTION	USERNAME	CONNECT
Athena	athena-gsi.nics.utk.edu	NICS	no account	
Big Red	login.bigred.iu.teragrid.org	IU		Login
Blacklight	blacklight.psc.teragrid.org	PSC		Login
Condor	tg-condor.purdue.teragrid.org	Purdue		Login
Dash	dash.sdsc.teragrid.org	SDSC		Login
Ember	login-ember.ncsa.teragrid.org	NCSA		Login
Frost	tg-login.frost.ncar.teragrid.org	NCAR		Login
Kraken	kraken-gsi.nics.utk.edu	NICS		Login
Lincoln	lincoln.ncsa.uiuc.edu	NCSA		Login
Lonestar	lonestar.tacc.teragrid.org	TACC		Login
Longhorn	tg-login.longhorn.tacc.teragrid.org	TACC		Login
NSTG	tg-login.ornl.teragrid.org	ORNL		Login
Pople	tg-login.pople.psc.teragrid.org	PSC		Login
Queen Bee	queenbee.loni-lsu.teragrid.org	LONI		Login
Ranger	tg-login.ranger.tacc.teragrid.org	TACC		Login
Spur	tg-login.spur.tacc.teragrid.org	TACC		Login
Steele	tg-steele.purdue.teragrid.org	Purdue		Login
Trestles	trestles.sdsc.edu	SDSC		Login

- Make sure you are logged into the XSEDE User Portal
- Go to ‘My XSEDE’ tab
- Go to the ‘Accounts’ link
- Resources you have access to will be indicated by a ‘login’ link
- Click on the ‘login’ link of the resource you would like to login to

Access XSEDE with Java SSH client

1. Download **JAVA client** here:
   http://portal.xsede.org/downloads/xsede.jnlp
2. Double click to run
3. Enter hostname of machine you want to run on (blacklight.psc.teragrid.org)
4. Enter XSEDE portal username in “Account Name”
5. Enter XSEDE portal username in “Passphrase”
6. Click “Use MyProxy”



Access XSEDE with standard SSH client

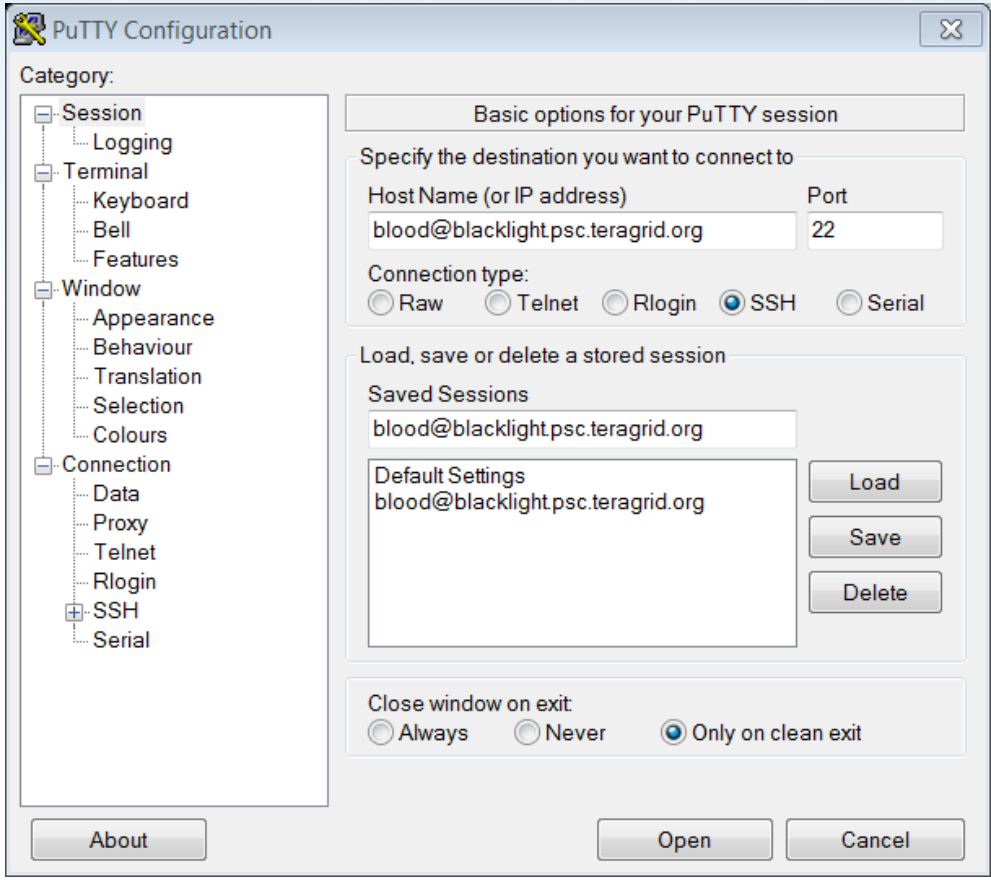
1. Login to **portal.xsede.org** and go to "**My XSEDE**"→**Accounts** to find your site-specific username

2. Email help@xsede.org to get site-specific password. Or, set your PSC password here:

   https://apr.psc.edu

3. Use your site-specific username and password to login to your machine of choice:
   - **Unix/Linux/Mac**: use 'ssh' in a terminal window
   - **Windows**: Download and use Putty

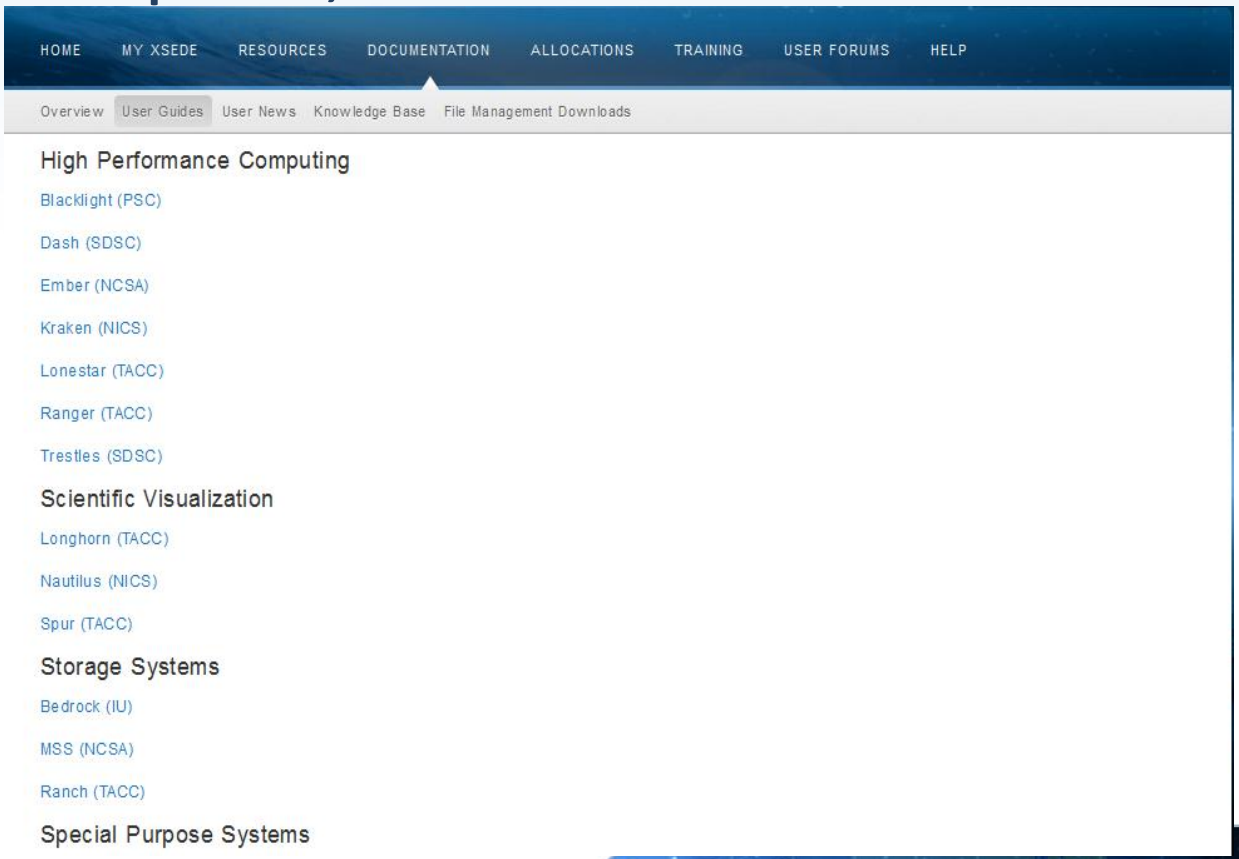
     http://the.earth.li/~sgtatham/putty/latest/x86/putty.exe



Running Jobs on XSEDE Resources

To learn about the specifics of how to run on a particular supercomputer, look at the User Guide.

- File Systems
- Batch jobs



Know Your Filesystems

Where your data resides on XSEDE and the appropriate storage is your responsibility. In general, all resources provide:

- **$HOME:** Permanent space, but small. A good choice for building software and working file collections of small to medium sized files, where a medium sized file is less than 50 MB.

- **SCRATCH:** More space, but TEMPORARY. Recommended for running jobs. **Backup your files left here! They get deleted (purged).**

- **Archival (mass) storage:** Long-term storage of large amounts of data, is accessible from all sites, but offers slower access time. Tar files before storing.

Create a Batch Script

```
#!/bin/csh
#PBS -l ncpus=16
#ncpus must be a multiple of 16
#PBS -l walltime=5:00
#PBS -j oe
#PBS -q batch

set echo

ja

#move to my $SCRATCH directory
cd $SCRATCH

#copy executable to $SCRATCH
cp $HOME/mympi .

#run my executable
mpirun -np $PBS_NCPUS ./mympi

ja -chlst
```

#PBS –M blood@psc.edu
#PBS –m n

- Use a linux text editor (nano, pico, emacs, vi)
- Example **PBS** script for running an MPI job on Blacklight at PSC.
- Special (#PBS) directives at top, followed by regular linux shell script
- Actual commands are specific to each system, but they follow general principles.
- Needs to be modified to run on other XSEDE machines.

Batch jobs

Submit the script that you have created:

Batch system should be used to run your job.

Actual commands are machine specific, but they follow general principles.

qsub myscript

qstat –a

qstat -u $USER

qdel myjobid

man qsub

Managing Your Environment on XSEDE resources: **Modules**

- Allows you to manipulate your environment and set environment variables related to a particular application.
- ‘module list’ shows currently loaded modules.
- ‘module avail’ shows available modules.
- ‘module show’ <name> describes module.

```
% module load gcc/3.1.1
% which gcc
/usr/local/gcc/3.1.1/linux/bin/gcc

% module switch gcc/3.1.1 gcc/3.2.0
% which gcc
/usr/local/gcc/3.2.0/linux/bin/gcc

% module unload gcc
% which gcc
gcc not found
```

http://modules.sourceforge.net/

Running a Monte Carlo test of consistency of OLS on a supercomputer

- Log in to blacklight.psc.teragrid.org using your method of choice
  - XSEDE portal
  - Java client (using XSEDE portal username and password)
  - standard ssh client (using PSC-specific username and password)

- get instructions here:
  - http://staff.psc.edu/blood/ICE12/OLSLabICE.pdf

Ask questions, get help:

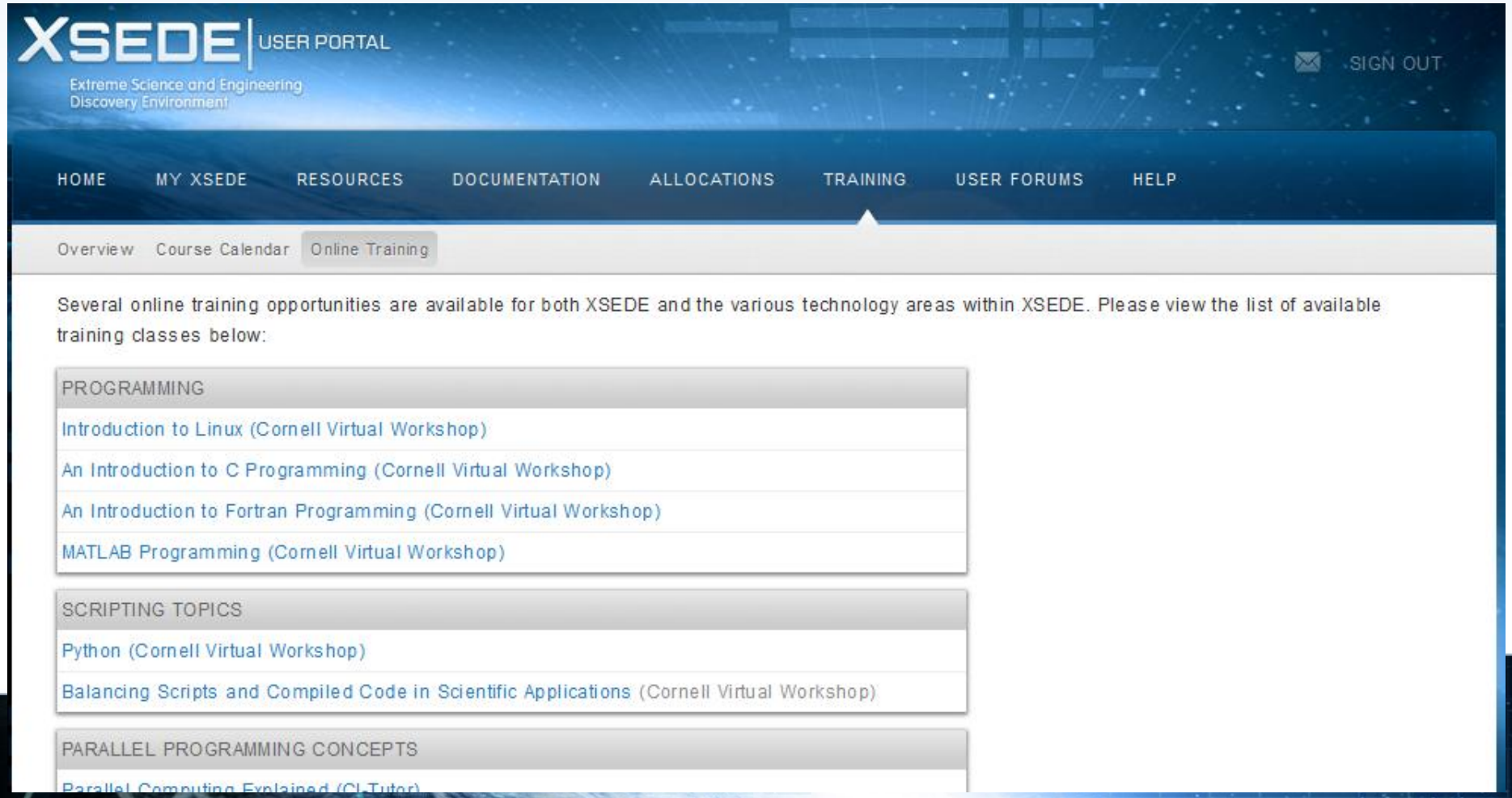
- portal.xsede.org → Help
  - Help Desk: Submit ticket
  - Security Incident, for ex. your account has been compromised.
- portal.xsede.org → My XSEDE → Tickets
  - Submit ticket
  - View past tickets (both open and closed)
- Can also email help@xsede.org or call 1-866-907-2383, at any hour (24/7)

Extended Collaborative Support Services (ECSS)

- Collaboration between researchers and XSEDE
- Expertise is available in a wide range of areas:
  - performance analysis and optimization
  - parallelization, optimization
  - gateway and web portal development
  - specialized scientific software.
- Can solicit Advanced Support at any time through the Allocations tab at the XSEDE User Portal
- Requires written request
- Inquire at help@xsede.org

Need training?

- portal.xsede.org → Training
  - Course Calendar
  - On-line training



Campus Champions

- **Campus Champions** are faculty or staff at a particular institution with interest in helping others move to the “next level” in computing
- Receive specialized training and help from XSEDE
- Is there a Campus Champion on your campus?
  - Check here: https://www.xsede.org/web/guest/current-champions
- Interested in learning more about the program?
  - Email me (blood@psc.edu) or Kay Hunt (kay@purdue.edu)

Questions? Need Help?

Email me:

blood@psc.edu


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