

July 27, 2018

# High Performance Computing and Big

## Environment (XSEDE)

Anirban Jana, [anirban@psc.edu](mailto:anirban@psc.edu)  
Pittsburgh Supercomputing Center

# XSEDE

Extreme Science and Engineering  
Discovery Environment

# Outline

- What Is Supercomputing?
- Overview of XSEDE Resources for Computational Economics
  - Computational Resources
  - Data Storage & Transfer Resources
  - Software Resources
- How to Access and Use XSEDE Resources

# What is a supercomputer?

A bunch of “commodity” components + enhancements

- **Processors:** essentially the same as your Mac or PC at home, arranged on “blades” (motherboards) or “nodes”; multiple blades are 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 (“parallel filesystem”)



## Different flavors of supercomputing

- “High Performance Computing” (HPC)
  - Single application running on many, many processors
  - Focused on floating-point operations per second (FLOPS)
    - A good laptop system today: ~50 GigaFLOPS
    - HPC resources like those on XSEDE: Fairly easily achieve TeraFLOPS, and with enough effort, even PetaFLOPS
- “High Throughput Computing” (HTC)
  - Running many (smaller) applications at once
- “Data-intensive Computing” (Big Data)
  - Focused on data storage amount (Terabytes to Petabytes of disk) and access speed (I/O operations per second or IOPS)

# Data-intensive (Big Data) 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 filesystems (RAM disk)
  - Software, e.g.:
    - MapReduce (Large-scale distributed data-analytics, fromGoogle)
    - Hadoop (open implementation of MapReduce)
    - Spark ( a high performance data analytics engine)



**Watson wins 'Jeopardy!' finale**

<http://latimesblogs.latimes.com/showtracker/2011/02/watson-jeopardy-finale-man-vs-machine-showdown.html>

# Machine Learning

- Systems that automatically improve in performing some task with data
- Emphasis back on FLOPs, but data handling is still important
- Strategies
  - Hardware
    - Accelerators (GPUs, FPGAs)
    - Solid-state drives (SSDs)
    - Memory drives (RAM disk)
  - Algorithms/Software
    - Traditional Machine Learning
      - Regression (Linear/Logistic/Lasso/Ridge), Support Vector Machines, Decision Trees, Random Forests, ...
      - scikit-learn (Python), Apache Spark MLlib (Python, R, Java, Scala), MATLAB Statistics and Machine Learning Toolbox, Weka3 (Java) , ...
    - Deep Learning / Neural Networks
      - Convolutions Neural Net (CNN), Recurrent Neural Net (RNN), Generative Adversarial Networks (GAN)
      - Tensorflow, Keras, Torch, Caffe/Caffe2, Apache MXNet, LBANN, ...

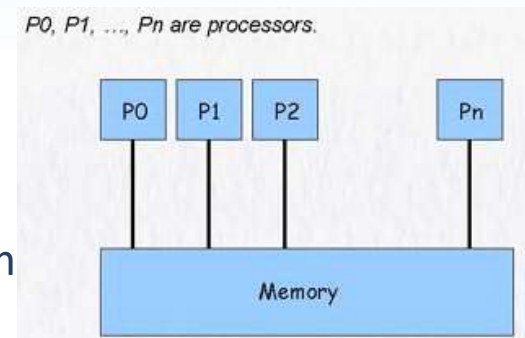
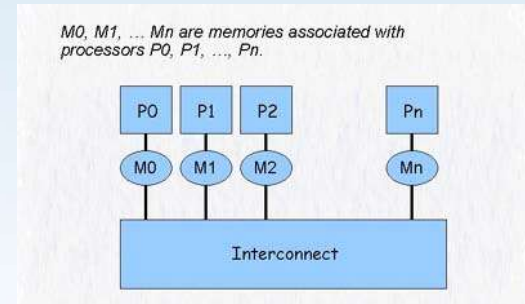


**Carnegie Mellon Reveals Inner Workings of Victorious Poker AI** :Libratus AI Defeated Top Pros in 20 Days of Poker Play, <https://www.scs.cmu.edu/news/carnegie-mellon-reveals-inner-workings-victorious-poker-ai>

Used PSC's Bridges!!!

# Distributed Memory vs Shared memory

- Distributed Memory Computing
  - Each processing element controls its own memory
  - If a processing element needs data from the memory of another, that data has to be explicitly exchanged from the latter to the former through the interconnect
  - MPI
- Shared Memory
  - SMP: Symmetric multiprocessing
  - Every processing element can access all of the mem
  - A single blade or node is a small SMP
  - Large SMPs bring together many nodes into a single memory space
  - Fast interconnect between nodes
  - OpenMP



# Pros and Cons of Distributed vs Shared Memory Computing

- **Distributed memory**
  - Pros
    - Can scale to very high processor counts with MPI, 10K-100K cores
    - Ubiquitous
  - Cons
    - Difficult to program if you have to do it yourself (for complex cases)
- **Shared memory**
  - Pros
    - Can be done easily using OpenMP directives or packages that use OpenMP
    - Directly parallelize existing serial code (parallelize existing loops)
  - Cons
    - Usually a limited amount of parallelism, ~100 cores



# High Performance Computing (cont.)

- Heterogeneous: Incorporate specialized processing elements into traditional HPC system, e.g.:
  - Graphics processing unit (GPU)
    - Only works (well) for some problems
    - Used to be hard to program (CUDA), but now getting easier (OpenACC, OpenCL)
    - Cost-effective performance (high performance per dollar)
  - Field-programmable gate array (FPGA)
    - Only works (well) for some problems
    - Requires experts to program
    - Potentially amazing performance, low power consumption

July 27, 2018



# XSEDE

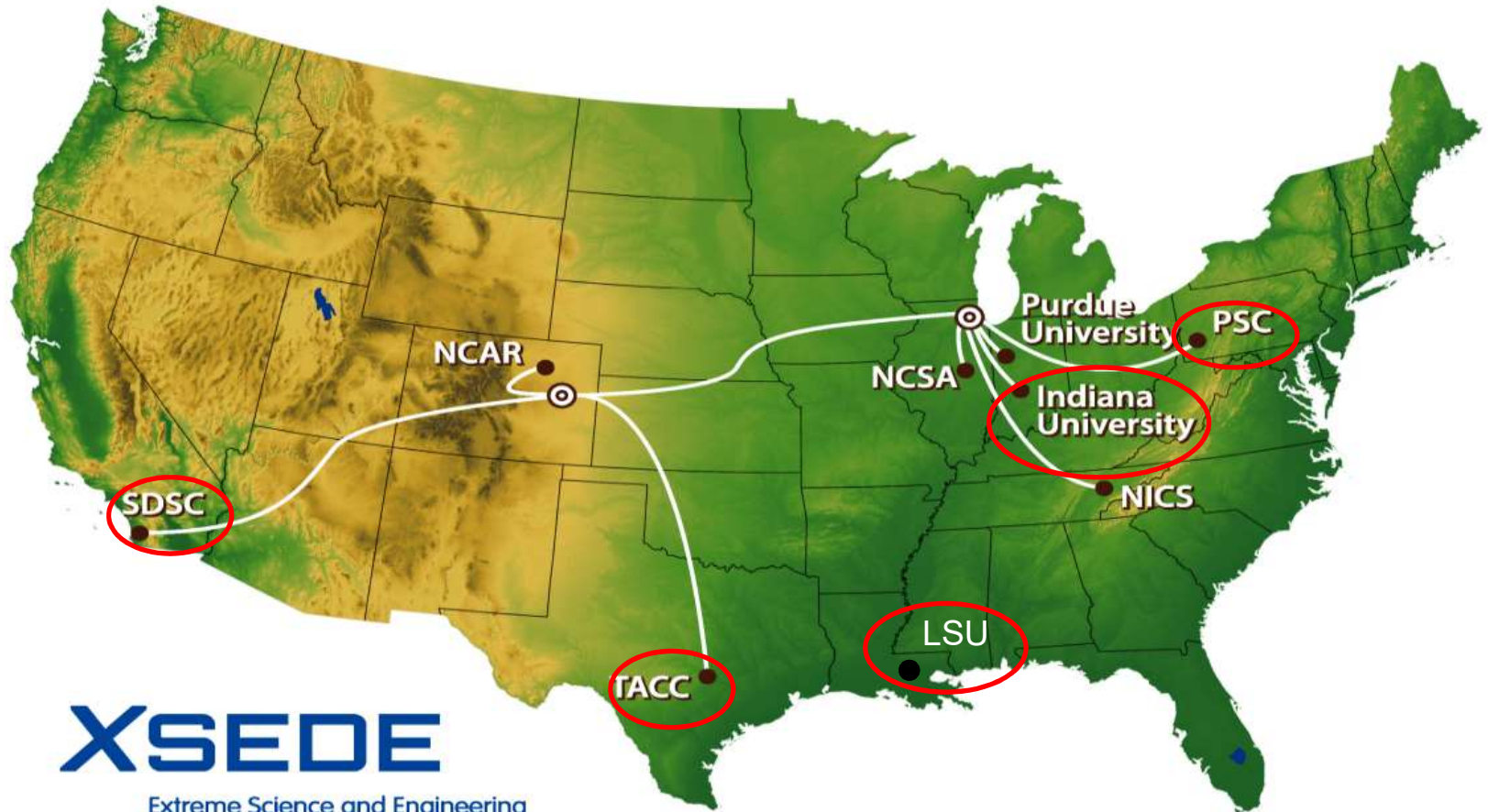
Extreme Science and Engineering  
Discovery Environment

# What is XSEDE ?

---

- eXtreme Science and Engineering Discovery Environment
- An integrated set of leading-edge computational, networking, data, software, and support resources to facilitate science.
- People around the world use these resources and services — things like supercomputers, collections of data and new tools — to improve our planet.
- Funded by the National Science Foundation (NSF)
- Free for open research

# XSEDE Consists of Many Partners



**XSEDE**

Extreme Science and Engineering  
Discovery Environment

- NCSA @ Illinois is the lead organization
- Current major resource providers circled in red
- Many other partners providing various management and training services

**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)
  - Researchers at for-profit institutions doing open research
- **Additional Users:**
  - PI's collaborators (including foreign), students; they need to be added to the PIs XSEDE grant

# What resources are available?

<https://portal.xsede.org/resource-monitor#vis-resources>

- Compute
  - Massively Parallel (100,000 cores)
  - Shared Memory (upto 12 TB)
  - High Throughput Computing
  - Cloud computing
- Special purpose: heterogeneous, data-intensive web hosting
- 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 Service (ECSS), Science Gateway Community Initiative (SGCI)

The screenshot displays the XSEDE User Portal resource monitor. It features a navigation bar with links for MY XSEDE, RESOURCE, DOCUMENTATION, ALLOCATIONS, TRAINING, USER FORUMS, HELP, and ECSS. Below the navigation bar, there are several resource categories:










- Compute Resources:** A table listing various compute nodes with columns for Name, Status, CPU%, Peak F/Flops, Utilization, Running Jobs, Queued Jobs, and Other Jobs. Nodes include Stampede2, Corral, XStream, SuperMIO, Bridges GPU, Bridges Regular Memory, Bridges Large Memory, Corral GPU, and Waverley.
- Visualization Resources:** A table with columns for Name, CPU%, Peak F/Flops, Running Jobs, Queued Jobs, and Other Jobs.
- Storage Resources:** A table with columns for Name and File Space (TB). Storage options include TACC Long-term tape Archival Storage, TACC Long-term Storage (Waverley Storage), PSC Storage (Bridges Python), SONG Medium-term disk storage (Data Drive), and IU/TACC Storage (Jelutaw Storage).
- HTC and Cloud Resources:** A table with columns for Name, Status, CPU%, and Peak F/Flops. Resources include Jetstream and Open Science Grid.

# What resources are available?

## ⚙️ Compute Resources

Name	Status	CPUs	Peak TFlops
Stampede2   User Guide	✓ Healthy	368280	12800.0
Comet   User Guide	✓ Healthy	46752	2000.0
XStream   User Guide	✓ Healthy	1300	1001.7
SuperMIC   User Guide	✓ Healthy	7200	925.0
Bridges GPU   User Guide	✓ Healthy	1344	894.6
Bridges Regular Memory   User Guide	✓ Healthy	21056	894.6
Bridges Large Memory   User Guide	✓ Healthy	160	894.6
Comet GPU   User Guide	✓ Healthy	1728	104.0
Wrangler   User Guide	✓ Healthy	2304	62.0

## 📁 Storage Resources

Name	File Space (TB)
TACC Long-term tape Archival Storage (Ranch)   User Guide	61440.0
TACC Long-term Storage (Wrangler Storage)   User Guide	10000.0
PSC Storage (Bridges Pylon)   User Guide	10000.0
SDSC Medium-term disk storage (Data Oasis)   User Guide	4000.0
IU/TACC Storage (Jetstream Storage) 	960.0

## 🌐 HTC and Cloud Resources

Name	Status	CPUs	Peak TFlops
Jetstream   User Guide	✓ Healthy	15360	516.1
Open Science Grid   User Guide	✓ Healthy	60000	50.0

# 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

- Ranch @ TACC
  - 61.44 PB tape archive
- Wrangler @ IU/TACC
  - 0.5 PB high speed flash + 10PB disk
- Bridges Pylon @ PSC
  - 10 PB disk
- Data Oasis @ SDSC
  - 7 PB disk (Lustre)
- Jetstream storage @ IU/TACC
  - 0.96 PB



# Transferring data to, from and between XSEDE resources

- scp, gsiscp, rsync: command line tools
- 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/>

# What software is available?

- Comprehensive software search feature for existing software: <https://portal.xsede.org/software#/>
- Existing software is mostly open-source but some are commercial
- In some cases, resource providers provides licenses for commercial software, but for others you have to bring your own license (BYOL)
- We can also install software that is not currently available
- Send questions/requests to [help@xsede.org](mailto:help@xsede.org)

# Software for Economics

- Matlab
- R
- Python:
  - SciPy
  - NumPy
- C/C++/Fortran +MPI/OpenMP
- VMs
- Containers (e.g. Singularity)
- Not all available software officially deployed, so may not all appear in “Comprehensive” search
- First 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/software#>

## Software

Find software available on XSEDE Service Provider sites. You can view by Resource, Site or Software type and then search for name, version, URL and more. To view details about a software package, click on the software name to see available versions. For more details click on the version to find out more about the software including how to access the software package.

Are you looking for software that is accessible via a science gateway? [Visit the Science Gateways Application List](#)

Help us gauge interest in potential future installations: if there's a software package you'd find useful, [submit a ticket](#) to let us know.

Clear Search

Matlab

Collapse All

View by:

Resources

Bridges

matlab

Science Category/Domain

Comet

matlab

Sites

Stampede2

matlab

mcr

SuperMIC

matlab

Wrangler

matlab

mcr

# Using MATLAB in parallel

- Job level parallelism – launch multiple MATLAB jobs using a (SLURM) script
- Inherently multithreaded functions
- Parallel Computing Toolbox - within a node (e.g., upto 352 cores on Bridges Extreme Shared Memory node)
  - parfor (parallel for loops)
  - spmd (Single Program Multiple Data – distributed arrays, MPI-like)
  - GPU-based
- MATLAB Distributed Computing Server (MDCS) –scale across nodes, available on Comet at SDSC
  - parfor
  - spmd
  - GPU-based

# 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**

Ref: <http://www.nics.tennessee.edu/computing-resources/nautilus/software?&software=r>

# Using R in parallel

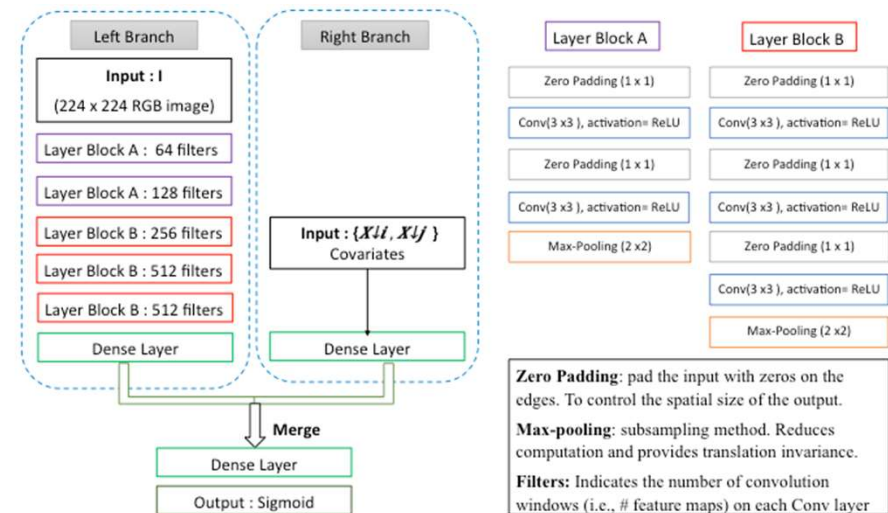
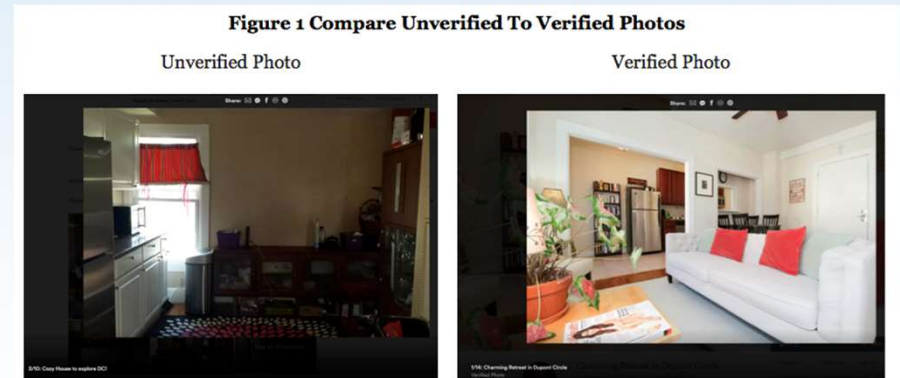
- Parallel libraries available
  - threaded
    - compile with threaded MKL library (ask a sysadmin)
    - **pnmath**: uses OpenMP versions of standard math libraries
  - distributed
    - **multicore** – creates new processes to execute functions on different cores
    - **RMPI** – facilitates using message passing within R
  - **foreach**: allows you parallelize loops using simple directives, using various methods “under the hood” that you can choose: **RMPI**, **multicore**

Ref: <http://www.nics.tennessee.edu/computing-resources/nautilus/software?&software=r>

# Investigating Economic Impacts of Images and Natural Language in E-commerce

Dokyun Lee, CMU Tepper School of Business

- Security and uncertain quality create challenges for sharing economies
  - Lee et al. studied the impact of high-quality, verified photos for Airbnb hosts
  - 17,000 properties over 4 months
  - Used *Bridges'* GPU nodes
- Security and uncertain quality create challenges for sharing economies
  - ✓ Difference-in-Difference (DD) analysis showed that on average, rooms with verified photos are booked 9% more often
  - ✓ Separating effects of photo verification from photo quality and room reviews indicates that high photo quality results in \$2,455 of additional yearly earnings
  - ✓ They found asymmetric spillover effects: on the neighborhood level, there appears to be higher overall demand if more rooms have verified photos





July 27, 2018



**XSEDE**

Extreme Science and Engineering  
Discovery Environment

# 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

The XSEDE logo is displayed in a stylized, white, sans-serif font against a dark blue background with a grid pattern. The background of the entire slide features a space-themed image with planets and a bright light source.

XSEDE

# 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.

The XSEDE logo is displayed in a bold, white, sans-serif font against a dark blue background with a grid pattern. The background of the entire slide features a space-themed image with planets and a bright light source.

XSEDE

# XSEDE User Portal

https://portal.xsede.org/my-xsede#/guest

Apps New Tab

## XSEDE USER PORTAL

Extreme Science and Engineering  
Discovery Environment

Search XSEDE...

SIGN IN

MY XSEDE RESOURCES DOCUMENTATION ALLOCATIONS TRAINING USER FORUMS HELP ECSS ABOUT

Summary Allocations/Usage Accounts Jobs Profile Publications Tickets Change Password Add User Community Accounts SSH Terminal

### Get Started on XSEDE

Sign In

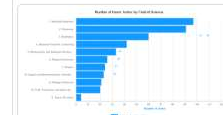
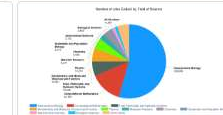
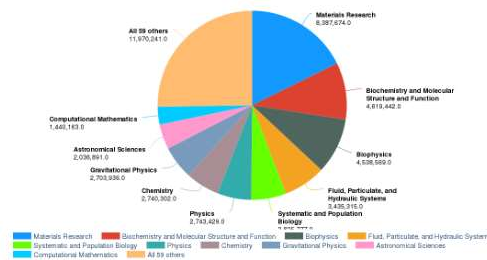
Create Account

### Quick Links

- System Monitor
- Allocations
- User News
- Scheduled Downtimes
- Software Search
- Documentation
- Training
- Help Desk
- ECSS
- User Forums
- Science Gateways

### In The Past 7 Days

XDSUs Charged: Total: by Field of Science



View Gallery

### New to XSEDE? ▾

### XSEDE Compute Resources

Detail View

Name	Status	Load	Jobs
Stampede2   UT Austin	Healthy	92%	R: 973 Q: 635 O: 174
Comet   SDSC	Healthy	79%	R: 2433 Q: 162 O: 54

https://portal.xsede.org/c/portal/login

# XSEDE

# XSEDE User Portal

← → ↻ [https://portal.xsede.org/my-xsede?p\\_p\\_id=58&p\\_p\\_lifecycle=0&p\\_p\\_state=maximized&p\\_p\\_mode=view&saveLastPath=0&\\_58\\_struts\\_action=%2Flogin%2Flogin](https://portal.xsede.org/my-xsede?p_p_id=58&p_p_lifecycle=0&p_p_state=maximized&p_p_mode=view&saveLastPath=0&_58_struts_action=%2Flogin%2Flogin) 🔍 ☆ ⓘ ⋮

Apps 📄 New Tab

## XSEDE | USER PORTAL

Extreme Science and Engineering  
Discovery Environment

Search XSEDE... 🔍

SIGN IN ✉

MY XSEDE RESOURCES DOCUMENTATION ALLOCATIONS TRAINING USER FORUMS HELP ECSS ABOUT

Summary Allocations/Usage Accounts Jobs Profile Publications Tickets Change Password Add User Community Accounts SSH Terminal

### Enter the Portal

USER NAME

PASSWORD

Sign In

REMEMBER ME

Other Sign In Options

[CREATE ACCOUNT](#) [VERIFY ACCOUNT](#) [FORGOT PASSWORD](#) [FORGOT USERNAME](#)

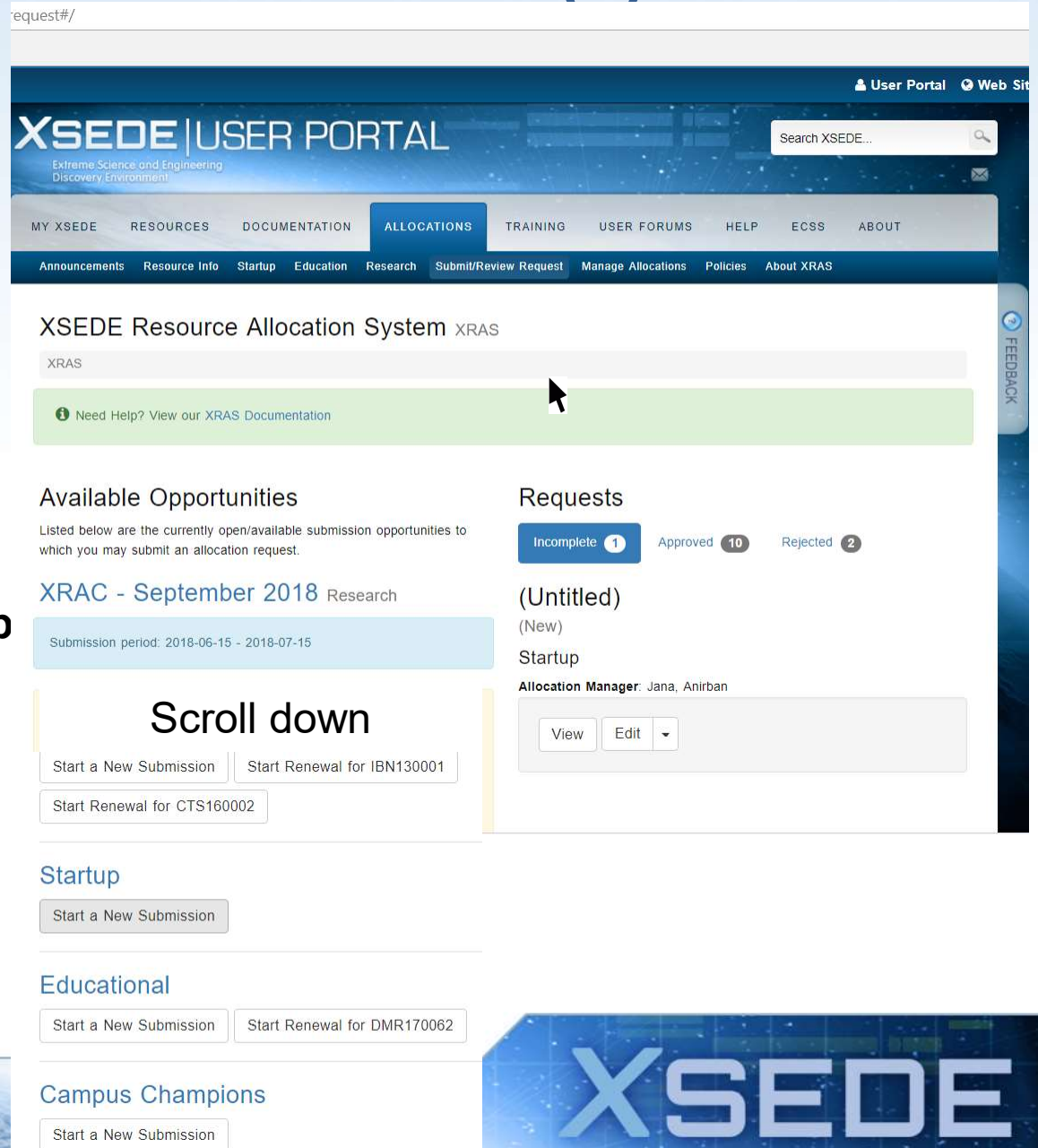
FEEDBACK

MY XSEDE	RESOURCES	DOCUMENTATION	ALLOCATIONS	TRAINING	USER FORUMS	HELP	ECSS
<a href="#">Summary</a>	<a href="#">Systems Monitor</a>	<a href="#">Get Started</a>	<a href="#">Announcements</a>	<a href="#">Overview</a>	<a href="#">Forums</a>	<a href="#">Overview</a>	<a href="#">ECSS Overview</a>
<a href="#">Allocations/Usage</a>	<a href="#">Storage</a>	<a href="#">Manage Data</a>	<a href="#">Resource Info</a>	<a href="#">Course Catalog</a>		<a href="#">Help Desk</a>	<a href="#">ECSS Projects</a>
<a href="#">Accounts</a>	<a href="#">Remote</a>	<a href="#">User Guides</a>	<a href="#">Startup</a>	<a href="#">Course Calendar</a>		<a href="#">Security Incident</a>	<a href="#">ECSS Symposium</a>
<a href="#">Jobs</a>	<a href="#">Visualization</a>	<a href="#">Community</a>	<a href="#">Education</a>	<a href="#">Online Training</a>			<a href="#">ECSS Workflows</a>
<a href="#">Profile</a>	<a href="#">Software</a>	<a href="#">Codes</a>	<a href="#">Research</a>				<a href="#">ECSS</a>
<a href="#">Publications</a>	<a href="#">Queue Prediction</a>	<a href="#">News</a>	<a href="#">Submit/Review</a>				<a href="#">Justification</a>

# XSEDE

# Request an XSEDE allocation (1)

- Go to **Allocations** tab
- Click on **Submit/Review Request** heading
- If you are new to XSEDE, you should submit a startup allocation by clicking **“Start a New Submission”** under **Startup**



The screenshot shows the XSEDE User Portal interface. The top navigation bar includes 'User Portal' and 'Web Site' links. The main navigation menu has tabs for 'MY XSEDE', 'RESOURCES', 'DOCUMENTATION', 'ALLOCATIONS', 'TRAINING', 'USER FORUMS', 'HELP', 'ECSS', and 'ABOUT'. The 'ALLOCATIONS' tab is selected. Below the navigation, there is a search bar and a 'FEEDBACK' button. The main content area is titled 'XSEDE Resource Allocation System XRAS'. A green banner with an information icon says 'Need Help? View our XRAS Documentation'. Below this, there are sections for 'Available Opportunities' and 'Requests'. The 'Available Opportunities' section lists 'XRAC - September 2018 Research' with a submission period of 2018-06-15 to 2018-07-15. The 'Requests' section shows a summary of 'Incomplete' (1), 'Approved' (10), and 'Rejected' (2) requests. A specific request '(Untitled) (New) Startup' is shown, managed by 'Jana, Anirban', with 'View' and 'Edit' buttons. Below these sections, there are 'Startup', 'Educational', and 'Campus Champions' sections, each with a 'Start a New Submission' button. A 'Scroll down' instruction is placed above the 'Startup' section.

# 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

## XRAC - September 2018 Research

Start a New Submission

Start Renewal for IBN130001

Start Renewal for CTS160002

## Startup

Start a New Submission

## Educational

Start a New Submission

Start Renewal for DMR170062

## Campus Champions

Start a New Submission

# 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

Editing Request: [Untitled Request] Advanced

New Submission for Startup

PERSONNEL TITLE/ABSTRACT RESOURCES DOCUMENTS GRANTS PUBLICATIONS SUBMIT

Personnel / Roles [Help](#)

By default you are added to the request as an **Allocation Manager**. You can update your role to be PI or Co-PI, if applicable. Add other people this request by looking them up by their XSEDE Portal Username. Then you can select the person's role for the request.

**Note:** Each person to be added to the request must already have a portal account and must provide you with their XSEDE Portal Username. For security reasons, we do not permit you to search to find portal usernames; you must provide a valid username. After providing a username, the person's full name, email address, institution, and phone number are shown for confirmation.

PI, Co-PIs, and Allocation Managers

Enter username to lookup Find & Add to Request

Person	Role *
Anirban Jana <anirban@psc.edu> Pittsburgh Supercomputing Center 412-268-4960	Allocation Manager

Users (optional)

If you already have a list of users that should be added to this project, should it be awarded, you can add them now. **Each person that is to be added as a user must have an existing XSEDE User Portal account.** You will need to know their username in order to add them.

Please make sure the user has successfully logged in to the XSEDE User Portal and accepted the Acceptable Use Policy.

Add Users

Next



# Request an XSEDE allocation (4)

Compute	Storage	Advanced Services
<input type="checkbox"/> Bridges PSC Regular Memory	<input type="checkbox"/> Bridges Pylon PSC Storage	<input type="checkbox"/> SGCi The Science Gateways Community Institute
<input type="checkbox"/> Bridges GPU PSC GPU	<input type="checkbox"/> Data Oasis SDSC Medium-term disk storage	<input type="checkbox"/> XSEDE Extended Collaborative Support
<input type="checkbox"/> Bridges Large PSC Large Memory	<input type="checkbox"/> Jetstream Storage IU/TACC Storage	
<input type="checkbox"/> Comet SDSC Dell Cluster with Intel Haswell Processors	<input type="checkbox"/> Ranch TACC Long-term tape Archival Storage	
<input type="checkbox"/> Comet GPU SDSC Comet GPU Nodes	<input type="checkbox"/> Wrangler Storage TACC Long-term Storage	
<input type="checkbox"/> Jetstream IU/TACC		
<input type="checkbox"/> OSG Open Science Grid		
<input type="checkbox"/> Stampede2 TACC Dell/Intel Knights Landing, Skylake System		
<input type="checkbox"/> Wrangler TACC Data Analytics System		
<input type="checkbox"/> XStream Stanford Cray CS-Storm K80		

- Choose resources of interest and enter “amount” requested for each
  - For Compute resources, enter requested SUs or Service Units ( the meaning of an SU depends on the specific type of resource, and explained in the next step)
  - For Storage resources, enter requested amount in GBs
  - For Advanced Services, enter Yes/No
  - Sum of SUs on all machines cannot exceed 200,000 SUs (???)
- Upload **CV** for the PI and submit request

# Request an XSEDE allocation (5)

XSEDE Extended Collaborative Support

## Compute

Bridges  
PSC Regular Memory

Bridges GPU  
PSC GPU

Bridges Large  
PSC Large Memory

Comet  
SDSC Dell Cluster with Intel Haswell Processors

Comet GPU  
SDSC Comet GPU Nodes

Jetstream  
IU/TACC

OSG  
Open Science Grid



## Storage

Bridges Pylon  
PSC Storage

Data Oasis  
SDSC Medium-term disk storage

### Bridges PSC Regular Memory

1 Service Unit = 1 Core Hour

In order to request this resource, you must also request one of the following resources:

- [Bridges Pylon PSC Storage](#)

Amount requested \*

SUs

### Bridges GPU PSC GPU

K80 nodes: 1 GPU-hour = 1 SU. P100 nodes: 1 GPU-hour = 2.5 SUs.

Amount requested \*

GPU Hours

Comments

### Bridges Large PSC Large Memory

1 Service Unit = 1 TB Memory Hours

Amount requested \*

Memory Hours

Comments

### Bridges Pylon PSC Storage

Amount requested \*

GB

Comments

## Request an XSEDE allocation (6)

- You'll get an email confirming request
- Startup requests usually take 1-2 weeks to be granted
- To start using XSEDE, you must apply for a startup. Later, if you want a larger allocation, you need to write a proper proposal
- See here for more info:  
<https://portal.xsede.org/allocations/research>

# 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 on XSEDE systems for one year
  - That is the equivalent of **8333 days (22.8 years)** of processing time on a single CPU 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

The XSEDE logo is displayed in a bold, white, sans-serif font against a dark blue background with a grid pattern. The background of the entire slide features a space-themed image with planets and a bright light source.

XSEDE

# Adding Users

- Each prospective user must get an XSEDE portal account
- Once you receive the allocation you can add other users: [portal.xsede.org](https://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

**XSEDE USER PORTAL**  
Extreme Science and Engineering  
Discovery Environment

Search XSEDE...

MY XSEDE RESOURCES DOCUMENTATION ALLOCATIONS TRAINING USER FORUMS HELP ECSS ABOUT

Summary Allocations/Usage **Accounts** Jobs Profile Publications Tickets Change Password Add User Community Accounts SSH Terminal

**XSEDE Single Sign on Login Hub**  
You can SSH into any XSEDE system with your PORTAL username and PORTAL password from the convenience of your desktop.

XSEDE recommends you use the XSEDE Single Sign on Login Hub to login to XSEDE resources with your local username and password. Use a local SSH client on your desktop to SSH to login.xsede.org with your portal username and password then easily `gsi-ssh` to any XSEDE system you have an account on with no additional username or passwords. For more information please visit the [XSEDE Single Sign on Login Hub](#) documentation page.

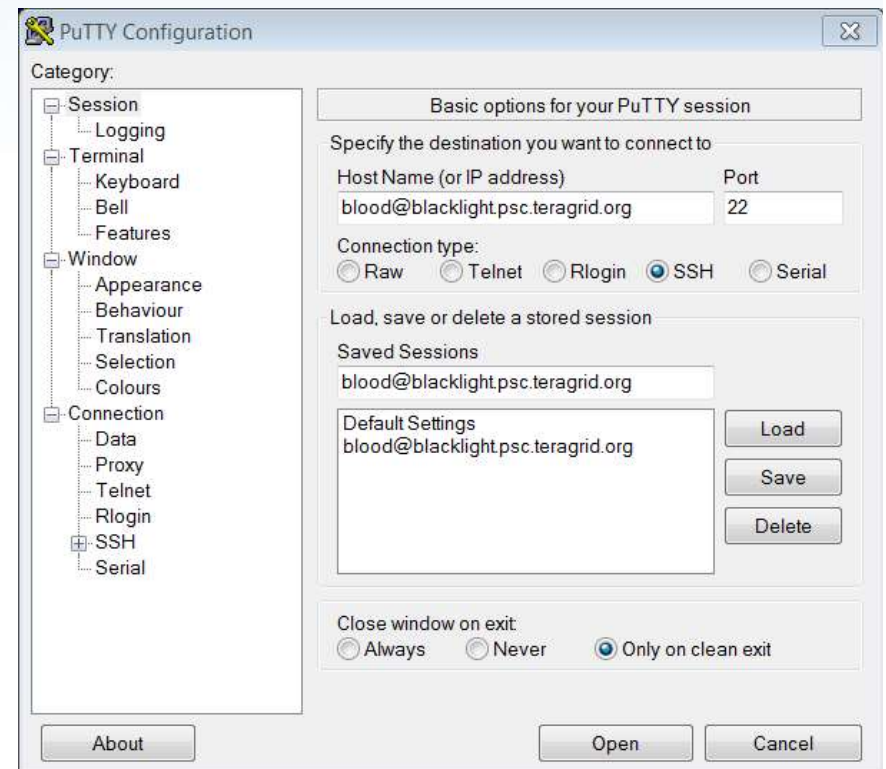
RESOURCE NAME	GSI-SSH LOGIN HOST	INSTITUTION	LOCAL USERNAME
Mason	mason.iu.xsede.org	IU	
SuperMIC	supermic.cct-lsu.xsede.org	LSU CCT	janirban
bridges-gpu.psc.xsede	bridges-gpu.psc.xsede	PSC	anirban
OSG	submit-1.osg.xsede.org	OSG	janirban
stampede2.tacc.xsede	stampede2.tacc.xsede	TACC	janirban
comet-gpu.sdsc.xsede	comet-gpu.sdsc.xsede	SDSC	janirban
XStream	xstream.stanford.xsede.org	Stanford U	xs-janirban
Bridges	bridges.psc.xsede.org	PSC	anirban
Bridges Large	bridges.psc.xsede.org	PSC	anirban
Comet	comet.sdsc.xsede.org	SDSC	janirban
Jetstream	jetstream.tacc.xsede.org	TACC	janirban
rmacc-summit.colorado.xsede	rmacc-summit.colorado.xsede	CU Boulder	
Wrangler	wrangler.tacc.xsede.org	TACC	janirban

- Make sure you are logged into the XSEDE User Portal
- Go to 'My XSEDE' tab
- Go to the 'Accounts' link
- Follow the instructions on the XSEDE Single Sign on Login Hub webpage

XSEDE

# Access XSEDE with standard SSH client

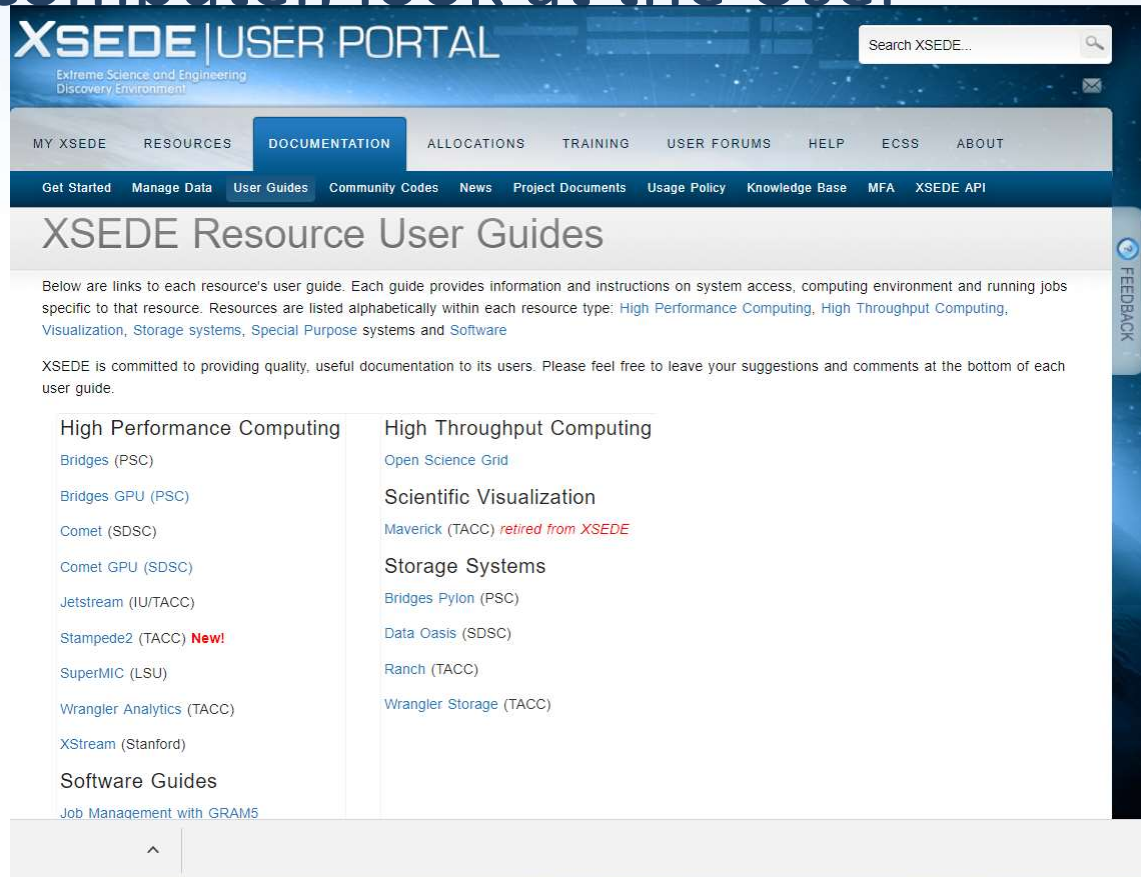
1. Login to [portal.xsede.org](http://portal.xsede.org) and go to “**My XSEDE**” → **Accounts** to find your site-specific username
2. Email [help@xsede.org](mailto:help@xsede.org) to get site-specific password. Or, for Bridges, set your PSC password here:  
<https://apr.psc.edu>
3. Use your site-specific username and password to login to your machine of choice:
  - **Unix/Linux/Mac:** use ‘ssh’ in a terminal window
  - **Windows:** Download and use an SSH client such as 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



The screenshot shows the XSEDE User Portal website. The header includes the XSEDE logo and 'USER PORTAL' text, along with a search bar and navigation links. The main content area is titled 'XSEDE Resource User Guides' and provides a list of links to user guides for various resources, categorized by type.

**XSEDE USER PORTAL**  
Extreme Science and Engineering Discovery Environment

Search XSEDE...

MY XSEDE RESOURCES DOCUMENTATION ALLOCATIONS TRAINING USER FORUMS HELP ECSS ABOUT

Get Started Manage Data User Guides Community Codes News Project Documents Usage Policy Knowledge Base MFA XSEDE API

## XSEDE Resource User Guides

Below are links to each resource's user guide. Each guide provides information and instructions on system access, computing environment and running jobs specific to that resource. Resources are listed alphabetically within each resource type: [High Performance Computing](#), [High Throughput Computing](#), [Visualization](#), [Storage systems](#), [Special Purpose systems](#) and [Software](#)

XSEDE is committed to providing quality, useful documentation to its users. Please feel free to leave your suggestions and comments at the bottom of each user guide.

High Performance Computing	High Throughput Computing
<a href="#">Bridges (PSC)</a>	<a href="#">Open Science Grid</a>
<a href="#">Bridges GPU (PSC)</a>	<b>Scientific Visualization</b>
<a href="#">Comet (SDSC)</a>	<a href="#">Maverick (TACC) <i>retired from XSEDE</i></a>
<a href="#">Comet GPU (SDSC)</a>	<b>Storage Systems</b>
<a href="#">Jetstream (IU/TACC)</a>	<a href="#">Bridges Pylon (PSC)</a>
<a href="#">Stampede2 (TACC) <b>New!</b></a>	<a href="#">Data Oasis (SDSC)</a>
<a href="#">SuperMIC (LSU)</a>	<a href="#">Ranch (TACC)</a>
<a href="#">Wrangler Analytics (TACC)</a>	<a href="#">Wrangler Storage (TACC)</a>
<a href="#">XStream (Stanford)</a>	
<b>Software Guides</b>	
<a href="#">Job Management with GRAM5</a>	

XSEDE



# Know Your Filesystems

Where your data resides on XSEDE and the appropriate storage is your responsibility. Below are some filesystems you may typically encounter on the different resources:

- **\$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. Backed up.
- **\$SCRATCH:** More space, but may be TEMPORARY depending on the machine. Recommended for running jobs. **Backup your files left here! If a purge policy is in place, then the files get deleted (purged) after a certain time. Also not typically automatically backed up,**
- **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.

The XSEDE logo is displayed in a large, white, sans-serif font against a dark blue background with a grid pattern. The background of the entire slide features a space-themed image with planets and a bright light source.

# Create a Batch Script

```
#!/bin/bash
#SBATCH -N 1
#SBATCH -p RM
#SBATCH --ntasks-per-node 28
#SBATCH -t 5:00:00

# echo commands to stdout
set -x

# move to your appropriate pylon5 directory
# this job assumes:
# - all input data is stored in this directory
# - all output should be stored in this directory
cd /pylon5/groupname/username/path-to-directory

# run OpenMP program
export OMP_NUM_THREADS=28
./myopenmp
```

- Use a linux text editor (nano, pico, emacs, vi)
- Example **SLURM** script for running an OpenMP job on Bridges at PSC.
- Special (**#SBATCH**) directives at top, followed by regular linux shell script
- **Actual commands can be specific to each system, but they follow general principles.**
- May need some modification to run on other XSEDE machines

## Batch jobs

Submit the script that you have created:

Batch system should be used to run your job.

Some actual commands can be machine specific, but they follow general principles.

```
sbatch myscript  
queue -u $USER  
scancel myjobid  
man sbatch
```

## Interactive jobs

On many of the resources, you can also get Interactive access to a compute node. The exact command for this is machine specific.

For example, on PSC's Bridges, the following command will give you interactive access to 1 (P100) GPU+14 CPU cores for 2.5 hours:

```
interact --gpu -t 2:30:00
```

# 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 help' <name> provides a short help ( **IMPORTANT!** )
- 'module show' <name> shows exact actions to be taken when loading it

### Example on PSC's Bridges

```
% module load gcc/5.3.0
% which gcc
/opt/packages/gcc/5.3.0/bin/gcc
```

```
% module switch gcc gcc/7.2.0
% which gcc
/opt/packages/gcc/7.2.0/bin/gcc
```

```
% module unload gcc
% which gcc
/usr/lib64/ccache/gcc
```

Ref: <http://modules.sourceforge.net/>

The XSEDE logo is displayed in a large, bold, white font against a dark blue background with a grid pattern. The background of the entire slide features a space-themed image with planets and a bright light source.

## Ask questions, get help:

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

The XSEDE logo is displayed in a stylized, light blue font against a dark blue background with a grid pattern. The background of the entire slide features a space-themed image with planets and a bright light source.

XSEDE

# Extended Collaborative Support Services (ECSS)

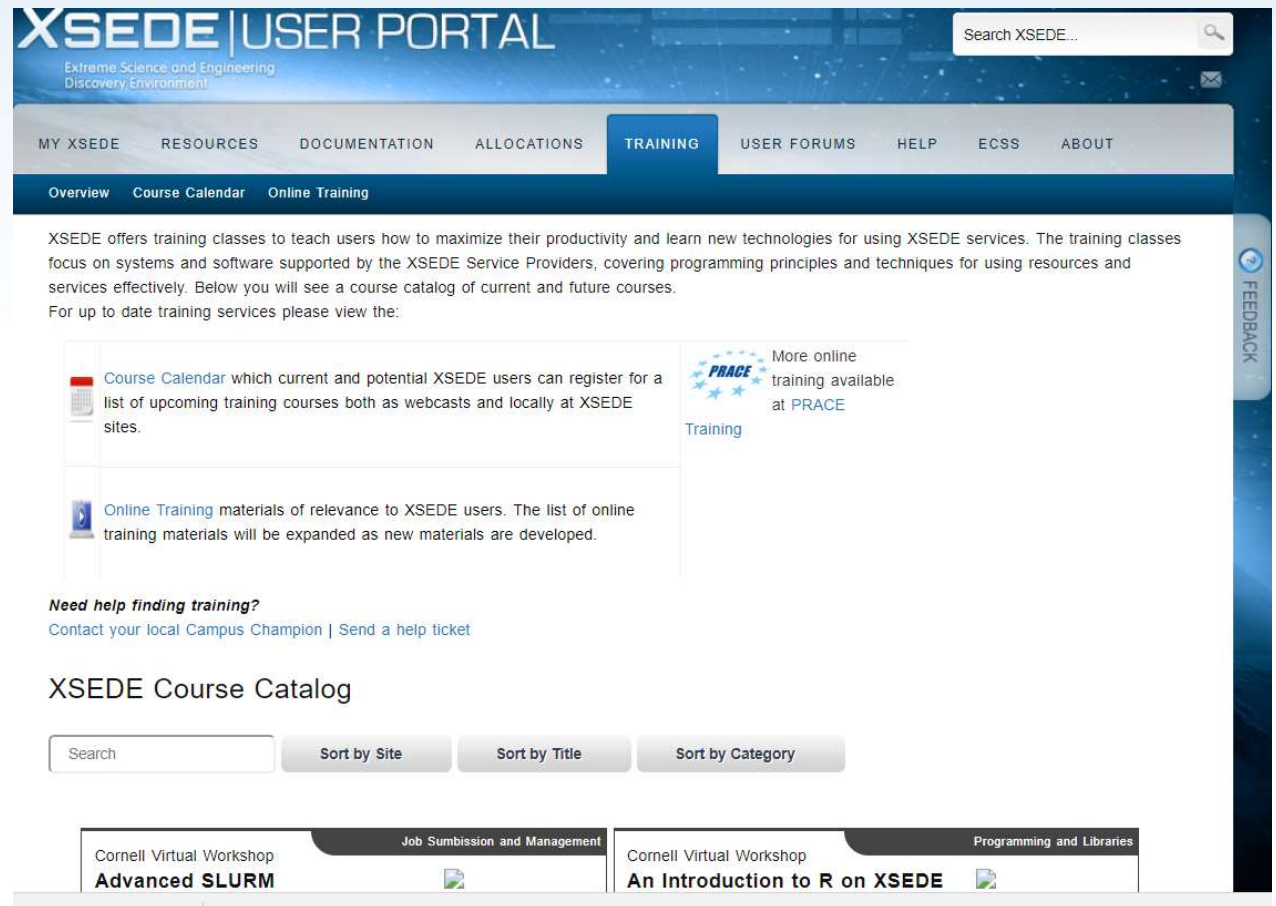
- Collaboration between researchers and XSEDE
- Expertise is available in a wide range of areas:
  - performance analysis and optimization
  - parallelization, optimization
  - gateway and web portal development
  - specialized scientific software.
- Can solicit Advanced Support at any time through the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

The XSEDE logo is displayed in a bold, white, sans-serif font against a dark blue background with a grid pattern. The background of the slide features a space-themed image with planets and a bright light source.

XSEDE

# Need training?

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



The screenshot shows the XSEDE User Portal with the 'TRAINING' tab selected. The page features a search bar, navigation menu, and a main content area with text and links. At the bottom, there is a 'XSEDE Course Catalog' with search and sorting options, and a list of training materials.

**XSEDE USER PORTAL**  
Extreme Science and Engineering  
Discovery Environment

Search XSEDE...

MY XSEDE RESOURCES DOCUMENTATION ALLOCATIONS **TRAINING** USER FORUMS HELP ECSS ABOUT

Overview Course Calendar Online Training

XSEDE offers training classes to teach users how to maximize their productivity and learn new technologies for using XSEDE services. The training classes focus on systems and software supported by the XSEDE Service Providers, covering programming principles and techniques for using resources and services effectively. Below you will see a course catalog of current and future courses. For up to date training services please view the:

- Course Calendar which current and potential XSEDE users can register for a list of upcoming training courses both as webcasts and locally at XSEDE sites.
- Online Training materials of relevance to XSEDE users. The list of online training materials will be expanded as new materials are developed.

More online training available at PRACE Training

**Need help finding training?**  
Contact your local Campus Champion | Send a help ticket

### XSEDE Course Catalog

Search [ ] Sort by Site Sort by Title Sort by Category

Cornell Virtual Workshop <b>Advanced SLURM</b>	Job Submission and Management	Cornell Virtual Workshop <b>An Introduction to R on XSEDE</b>	Programming and Libraries
---	-------------------------------	--	---------------------------





# Campus Champions

- **Campus Champions** are faculty or staff at a particular institution/department 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?
  - To check, go to <https://www.xsede.org/web/campus-champions> and click on “Current Champions” on the right
  - Interested in becoming a campus champion? Scroll down to the bottom of the campus champion webpage for a link to send a request