

WVU RESEARCH COMPUTING INTRODUCTION

Introduction to WVU's Research Computing Services



WHO ARE WE?

- Division of Information Technology Services
- Funded through WVU Research Corporation
- Provide centralized HPC systems and support staff to help researchers compute scientific data
- Support 30 individual research groups across campus and ~150 active users



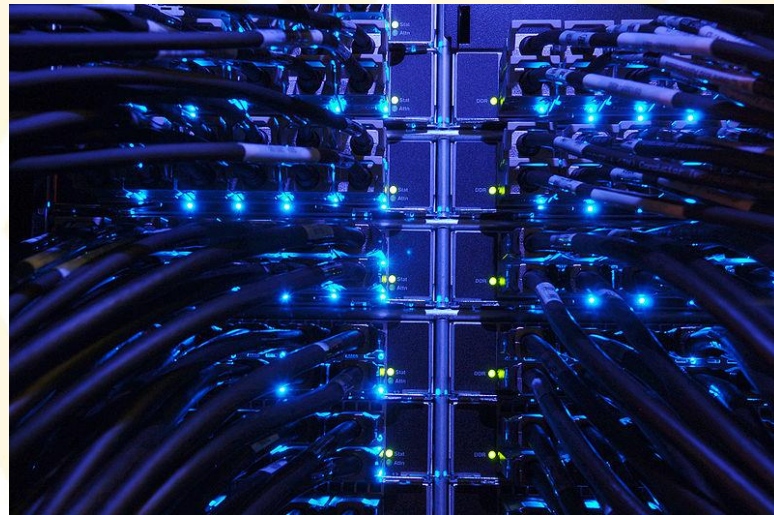
MOUNTAINEER

- 32 Compute Nodes
 - Dual 6 core Intel Xeon (Westmere) processors (12 cores)
 - 48 Gbytes of memory (4 GB per core)
 - 10 Gbps Ethernet Interconnect
- 50 Terabytes of shared storage
- Open community access



SPRUCE KNOB

- 165 Compute Nodes/3,064 Cores
- Current node offerings
 - Small Memory Nodes
(24 Intel Xeon Broadwell Cores/64 GB Ram)
 - Medium Memory Nodes
(24 Intel Xeon Broadwell Cores/128 GB Ram)
 - Large Memory Modes
(24 Intel Xeon Broadwell Cores/512 GB Ram)
 - GPU memory nodes
(24 Intel Xeon Broadwell Cores/64 GB Ram/
NVIDIA k20x GPU)
- 54 Gbps Infiniband Interconnect
- 17 GPU Accelerator cards



SPRUCE KNOB – HYBRID MODEL

Condo Model

- Faculty pay for any nodes that they want in the cluster
- Faculty investors (and their research teams) have priority access to their nodes
- Any Spruce HPC users can use idle faculty owned nodes for up to 4 hours (standby queue)

Community Model

- 51 Nodes/848 Cores are generally available to the HPC community (WV)
- Fair-share scheduling, not subject to owner preemption



HPC SOFTWARE

- MOAB Schedule Suite
- Intel Studio XE (Fortran, C, MPI, etc)
- Matlab
- Module avail command
- Discipline specific software
- Researchers self-installed software



LOCAL STORAGE

- DDN GRIDScaler System
- 1 PB Raw Storage
- ~ 400 TB scratch/transfer space
- Expandable from 2 PB to 8 PB
- GPFS Parallel Filesystem
- Direct Infiniband Connection



RESEARCH DATA DEPOT

Central Storage Repository

- Centrally Managed
- NSF/NIH Data Management Requirements
- Low cost storage solution for researchers
- Expandable Solution
- Multiple protocol support
 - CIFS (Windows Shares)
 - Globus Online
 - SFTP/SCP
- Connected to Science DMZ
- Coming Soon!

DDN GS7K System



ARCHIVING RESEARCH DATA

Google Drive for Education

- Unlimited Storage
- Max file size limit 5TB
- Easy file sharing
- But how do I get it there?
 - See shortly ...



HOW DO I TRANSFER MY DATA?

- Preferred Method:
Globus Online
- Alternative Methods:
 - scp
 - rsync
 - sftp



GLOBUS ONLINE

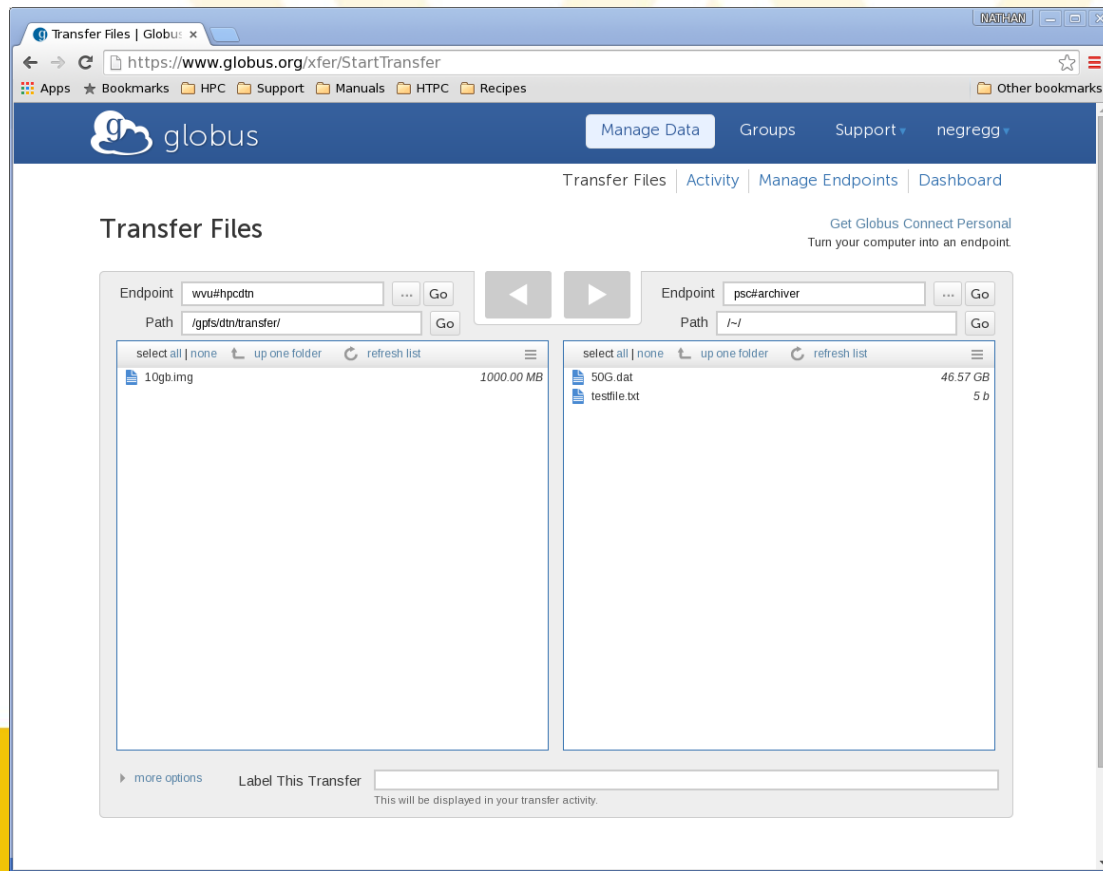
- File Transfers
 - Fast, reliable, secure
 - Restartable
 - Google Drive Connector (Coming Soon!)
- File Sharing
 - Self-managed/no admins needed
 - Custom groups
- Data Publication
 - Publish datasets in searchable web-interface
 - Helps meet NSF/NIH Data Management Plans



WHY GLOBUS ONLINE?



GLOBUS – INTUITIVE WEB INTERFACE



The screenshot displays the Globus Transfer Files web interface in a browser window. The address bar shows the URL <https://www.globus.org/xfer/StartTransfer>. The interface includes a navigation bar with the Globus logo, a 'Manage Data' button, and links for 'Groups', 'Support', and a user profile 'negregg'. Below the navigation bar, there are tabs for 'Transfer Files', 'Activity', 'Manage Endpoints', and 'Dashboard'. The main content area is titled 'Transfer Files' and includes a link to 'Get Globus Connect Personal' with the text 'Turn your computer into an endpoint.'.

The interface shows two transfer endpoints side-by-side, each with a list of files and folders:

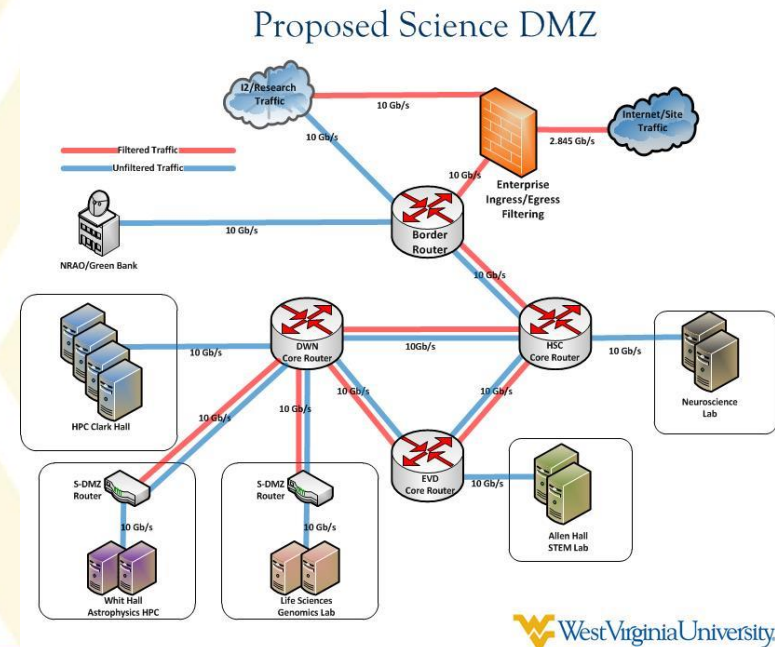
- Left Endpoint:** Endpoint is `www#hpcdn`, Path is `/gifs/dtn/transfer/`. The file list contains one item: `10gb.img` with a size of `1000.00 MB`.
- Right Endpoint:** Endpoint is `psc#archiver`, Path is `/~/`. The file list contains two items: `50G.dat` with a size of `46.57 GB`, and `testfile.txt` with a size of `5 b`.

At the bottom of the interface, there is a section for 'more options' and a 'Label This Transfer' field with a placeholder text: 'This will be displayed in your transfer activity.'



SCIENCE DMZ

- Science/Research Dedicated Dual 10 Gb Link
- Soon connected to:
 - HPC
 - Dr. Spirou Neuroscience Lab @ HSC
 - Astrophysics HPC @ White Hall
 - Genomics Core Lab @ LSB
 - College Education and Human Services
 - Greenbank
 - Outside endpoints
- Monitored with perfSONAR



RESEARCH NETWORKING

- WVU-Pittsburgh Internet2 connection = 10Gbps
- WVU research network to Pittsburgh shared with DOE - NETL
- WVU – member of Internet2
- Internet2 100 Gbps nationwide network
- 3ROX implemented a 100 Gbps connection to I2



ADDITIONAL INFORMATION

- Main Website
 - <http://hpc.wvu.edu>
- Wiki
 - <http://wiki.hpc.wvu.edu>



QUESTIONS?

- Main Contact: <https://helpdesk.hpc.wvu.edu> or helpdesk@hpc.wvu.edu
- Contact Nathan Gregg (Research Computing Manager/HPC Systems Administrator)
 - Email: nathan.gregg@mail.wvu.edu
 - Phone: 304-293-0963
- Contact Guillermo Franco (Research Computing Software Developer)
 - Email: gufranco@mail.wvu.edu
- Contact Brice Knotts (ITS Enterprise Infrastructure -- Executive Director)
 - Email: baknotts@mail.wvu.edu



February 14, 2017

Overview of XSEDE Resources and System

LCSEE Graduate Seminar – February 13, 2017

XSEDE

Extreme Science and Engineering
Discovery Environment



First a little history

- A long history of supercomputing in scientific research
- In the “olden days”...
 - many supercomputers at different institutions
 - very siloed
 - independent, unique administrative procedure
 - little or no coordination
 - many supported by the National Science Foundation

First a little history

- In 2001 NSF recognized the need for a more coordinated approach to scientific computing in the U.S.
- Announced a competitive solicitation to build a large scale, coordinated network of supercomputers and their respective organizations
- Awarded in 2004



First a little history

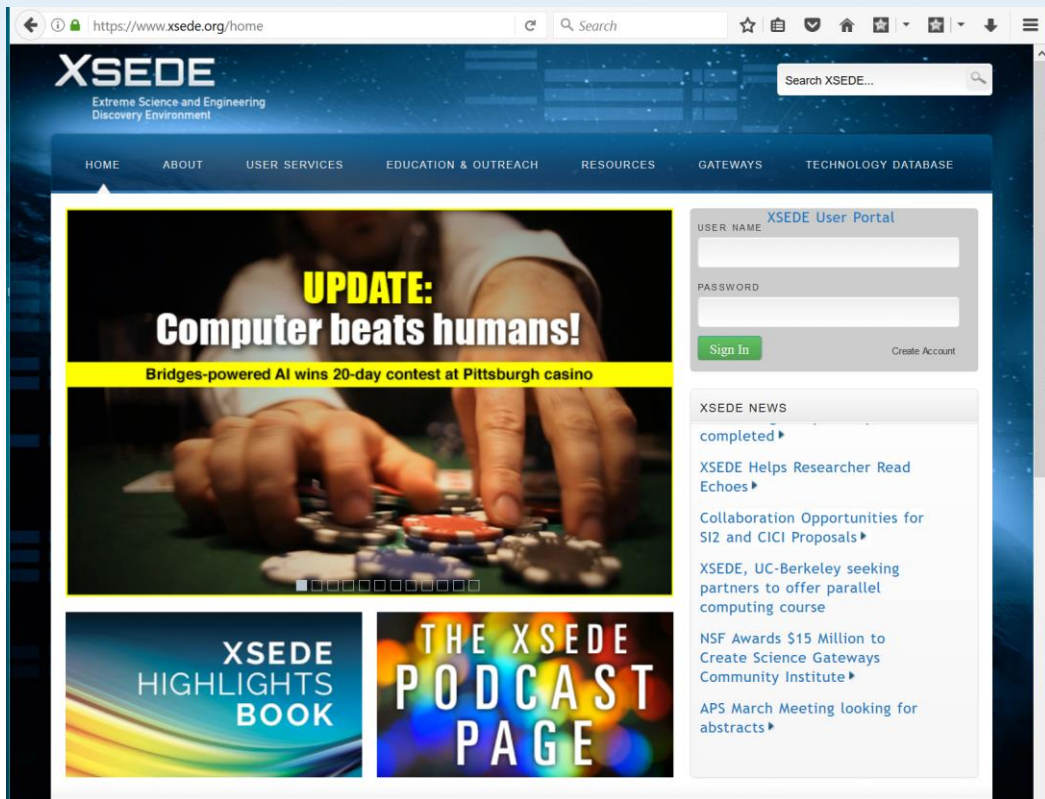
- Awarded in 2004
- Became known as the Teragrid
- Teragrid was a coordinated consortium of scientific computing resources – including computational and storage resources
- Teragrid operated from 2004 to 2011

First a little history

- In July 2011 Teragrid “transitioned” to XSEDE
- A new consortium – many original members
- A new focus



www.xsede.org



The screenshot shows the XSEDE website homepage. The browser address bar displays "https://www.xsede.org/home". The XSEDE logo is at the top left, with the tagline "Extreme Science and Engineering Discovery Environment". A search bar is located at the top right. A navigation menu includes links for HOME, ABOUT, USER SERVICES, EDUCATION & OUTREACH, RESOURCES, GATEWAYS, and TECHNOLOGY DATABASE. The main content area features a large banner with the text "UPDATE: Computer beats humans!" and "Bridges-powered AI wins 20-day contest at Pittsburgh casino". Below the banner are two tiles: "XSEDE HIGHLIGHTS BOOK" and "THE XSEDE PODCAST PAGE". On the right side, there is a "XSEDE User Portal" with fields for USER NAME and PASSWORD, and buttons for "Sign In" and "Create Account". Below the portal is a "XSEDE NEWS" section with several news items, each with a right-pointing arrow.

https://www.xsede.org/home

XSEDE
Extreme Science and Engineering
Discovery Environment

Search XSEDE...

HOME ABOUT USER SERVICES EDUCATION & OUTREACH RESOURCES GATEWAYS TECHNOLOGY DATABASE

UPDATE:
Computer beats humans!
Bridges-powered AI wins 20-day contest at Pittsburgh casino

XSEDE HIGHLIGHTS BOOK

THE XSEDE PODCAST PAGE

XSEDE User Portal

USER NAME
PASSWORD
Sign In Create Account

XSEDE NEWS

completed ▶

XSEDE Helps Researcher Read Echoes ▶

Collaboration Opportunities for SI2 and CICI Proposals ▶

XSEDE, UC-Berkeley seeking partners to offer parallel computing course

NSF Awards \$15 Million to Create Science Gateways Community Institute ▶

APS March Meeting looking for abstracts ▶



What is XSEDE?

- A network of scientific computing resources...
 - with coordinated administration
- an organization



XSEDE

- A network of scientific computing resources...
 - High performance computing
 - High throughput computing
 - Visualization
 - Storage
 -



XSEDE

- High performance computing
- A collection of HPC systems at institutions across the country
- Including...
 - Stampede at TACC 9600 Pflops
 - Comet at SDSC 7000 Pflops
 - Bridges regular memory at PSC 895 Pflops
 - Bridges large memory at PSC

•



XSEDE

- High performance computing
 - Stampede at TACC
 - Dell Poweredge cluster with Intel Xeon Phi coprocessors
 - OS: Linux – CentOS
 - Cores: 102400
 - Nodes: 14336
 - Storage: 14 PB
 - SLURM job management system
 - Nvidia Kepler2 GPUs



XSEDE

- High performance computing
 - Bridges – Regular memory at PSC
 - HP cluster with Intel Xeon processors
 - 128 GB Memory per node
 - OS: Linux – CentOS
 - Cores: 21056
 - Nodes: 752
 - Intel Omnipath Interconnect
 - Shared storage system + 8 GB storage per node
 - Many unique features



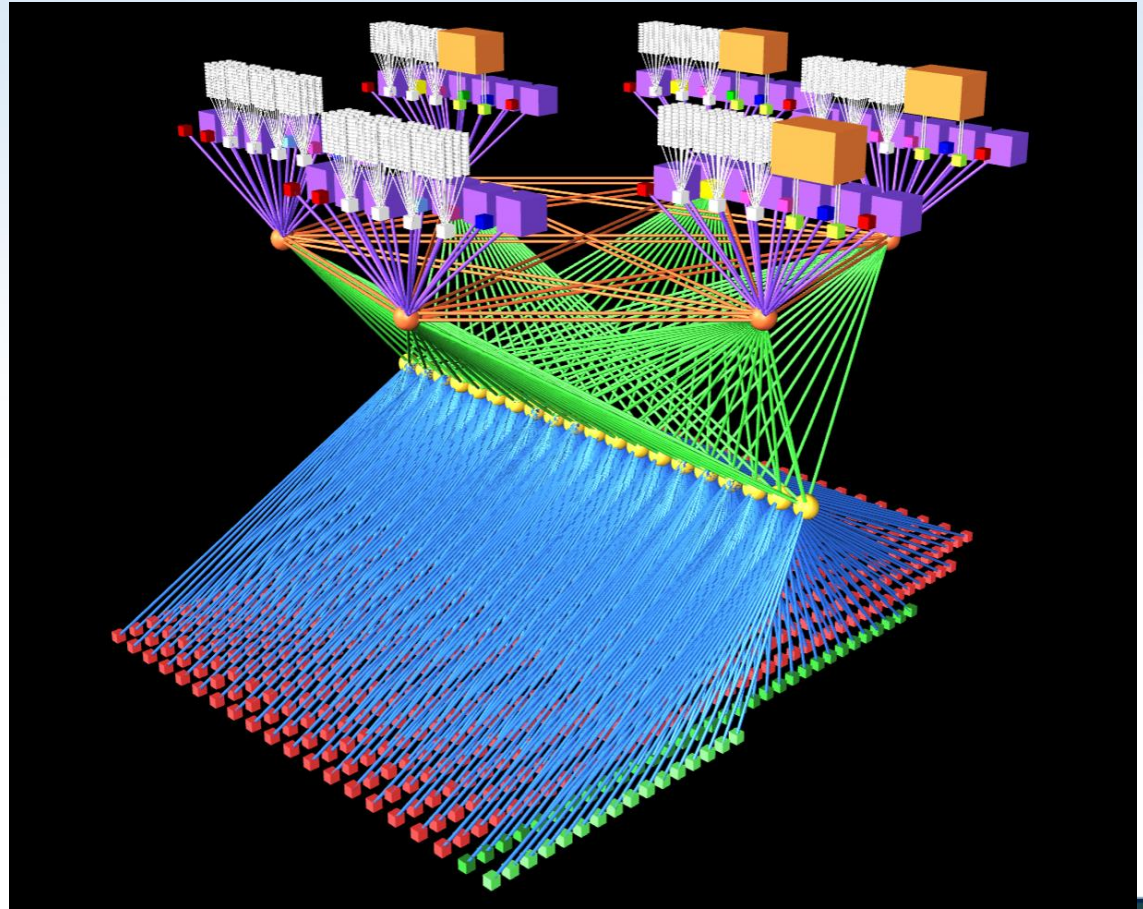
XSEDE

- High performance computing
 - Bridges – Large memory at PSC
 - HP nodes – 4 with 12 TB memory, 42 with 3 TB memory each
 - OS: Linux – CentOS
 - Cores: 160
 - Intel Omnipath Interconnect
 - Pylon - Shared storage system + 8 GB storage per node
 - Many unique features



XSEDE HPC – Bridges

<https://www.psc.edu/index.php/bridges-virtual-tour>



XSEDE

- High throughput computing
 - Open Science Grid
 - 60,000 compute nodes (distributed)
-



XSEDE

- Visualization
 - Maverick – 2640 cores
 - 20 Pbytes of storage
 - designed specifically to do scientific data visualization
 - does visualization remotely
-

XSEDE

- Storage
 - DATA OASIS @ San Diego Supercomputing Center
 - Default/max storage per XRAC: **500GB/50TB**
 - PYLON @ the Pittsburgh Supercomputing Center
 - Default/max storage per XRAC: **500 GB/none**
 - RANCH @ the Texas Advanced Computing Center
 - Default/max per XRAC: **none/1 PB**
 - Wrangler @ the Texas Advanced Computing Center
 - Default/max per XRAC: **TBD**



Software

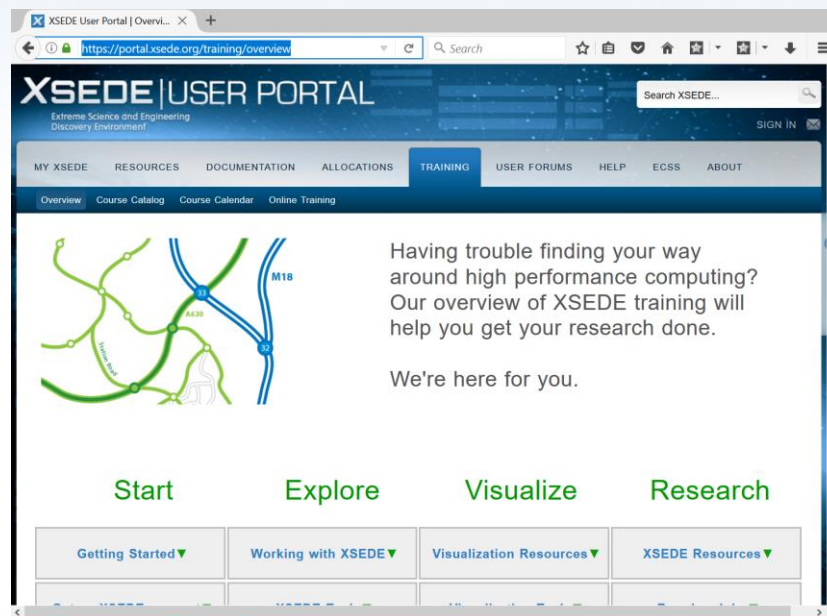
- Too much to tell - too little time
- See –
 - <https://www.xsede.org/software>



XSEDE

Training

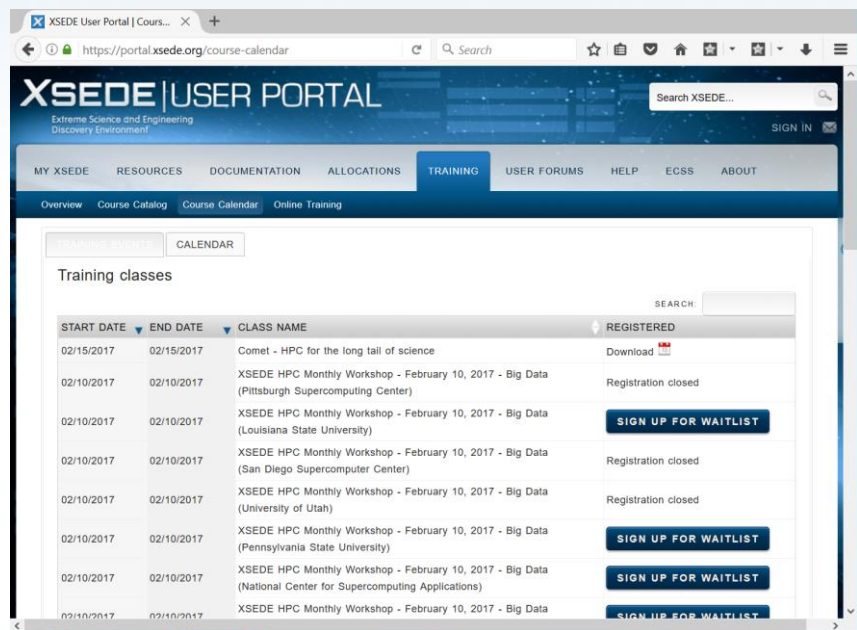
- <https://portal.xsede.org/training/overview>



XSEDE

Training

- <https://portal.xsede.org/course-calendar>



The screenshot shows the XSEDE User Portal interface. The header includes the XSEDE logo and navigation links: MY XSEDE, RESOURCES, DOCUMENTATION, ALLOCATIONS, TRAINING (selected), USER FORUMS, HELP, ECSS, and ABOUT. Below the header, there are tabs for Overview, Course Catalog, Course Calendar (selected), and Online Training. The main content area is titled 'Training classes' and features a table of upcoming workshops. The table has columns for START DATE, END DATE, CLASS NAME, and REGISTERED. The first row shows a workshop on February 10, 2017, at the Pittsburgh Supercomputing Center, with a 'Download' button. The second row shows a workshop at Louisiana State University with a 'SIGN UP FOR WAITLIST' button. The third row shows a workshop at the San Diego Supercomputer Center with a 'Registration closed' status. The fourth row shows a workshop at the University of Utah with a 'Registration closed' status. The fifth row shows a workshop at Pennsylvania State University with a 'SIGN UP FOR WAITLIST' button. The sixth row shows a workshop at the National Center for Supercomputing Applications with a 'SIGN UP FOR WAITLIST' button. The seventh row shows a workshop at the University of Utah with a 'SIGN UP FOR WAITLIST' button.

START DATE	END DATE	CLASS NAME	REGISTERED
02/15/2017	02/15/2017	Cornet - HPC for the long tail of science	Download
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data (Pittsburgh Supercomputing Center)	Registration closed
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data (Louisiana State University)	SIGN UP FOR WAITLIST
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data (San Diego Supercomputer Center)	Registration closed
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data (University of Utah)	Registration closed
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data (Pennsylvania State University)	SIGN UP FOR WAITLIST
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data (National Center for Supercomputing Applications)	SIGN UP FOR WAITLIST
02/10/2017	02/10/2017	XSEDE HPC Monthly Workshop - February 10, 2017 - Big Data	SIGN UP FOR WAITLIST



XSEDE

Resource Allocations – how do you get to use these resources?

- Resources Types –
 - Computational Resources
 - Visualization Resources
 - Storage Resources
- <https://portal.xsede.org/allocations-overview>



XSEDE

Resource Allocations

- Allocation Types
 - **Trial** – very limited 1000 SUs, 6 months, COMET
 - **Startup** – limited, getting started, benchmarks, available to faculty/post-doc & K-12 teachers
 - **Educational** – for training, workshops, classroom instruction and classroom projects, limited
 - **Research** – large scale resources, awarded based on research proposals, scientific merit+, competitive, proposal submissions quarterly, awarded quarterly, good for one year



XSEDE

Campus Champions

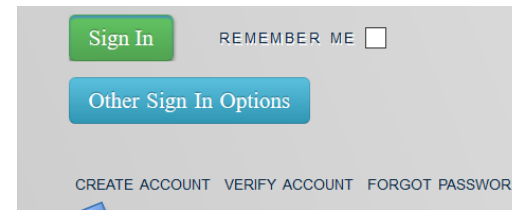
- Support and assistance in using XSEDE resources, getting accounts, etc.
- Liaison between you and XSEDE
- @ WVU
 - Nathan Gregg
 - Nathan.Gregg@mail.wvu.edu
 - Don McLaughlin
 - Don.McLaughlin@mail.wvu.edu



XSEDE

Final comments

- XSEDE Resource Allocations are available to
 - Faculty and Post-docs
 - not graduate students and undergraduate students
 - However, students can be on project teams and be given access to faculty project allocations
 - But, **everyone** must have a portal account




Thank you

Any Questions?





XSEDE



Our reach will forever
exceed our grasp, but,
in stretching our horizon,
we forever improve our world.

XSEDE

Extreme Science and Engineering
Discovery Environment

