Getting Started with XSEDE

Dan Stanzione
Welcome to XSEDE!

• XSEDE is an exciting cyberinfrastructure, providing large scale computing, data, and visualization resources. XSEDE is the evolution of the NSF TeraGrid.

• Today’s session is a general overview of XSEDE for new XSEDE/TeraGrid users; it is not going to teach you computational science or programming!

• Use the webcast chat window to ask questions!
Outline

- What is XSEDE?
- How do I get started?
- XSEDE User Portal
- User Responsibility & Security
- Applying for an Allocation
- Accessing Resources
- Managing Data and File Transfers
- Other Resources
- Picking the right resource for your job.
- Running Jobs
- Managing Your Software Environment
- Getting Help
- Next Steps
What is XSEDE?

• The Extreme Science and Engineering Discovery Environment (XSEDE): The most powerful integrated advanced digital resources and services in the world. Funded by NSF.

• A single virtual system that scientists can use to interactively share computing resources, data, and expertise.

• XSEDE lowers technological barriers to the access and use of computing resources. Using XSEDE, researchers can establish private, secure environments that have all the resources, services, and collaboration support they need to be productive.

• 9 supercomputers, 3 visualization systems, and 9 storage systems provided by 16 partner institutions (Service Providers or SPs)
  – NCSA, NICS, PSC, TACC, SDSC, Univ. of Virginia, Shodor Education Foundation, SURA, Univ. of Chicago, Indiana U., Purdue, Cornell, Ohio State, UC-Berkeley, Rice, and NCAR
How do I get started using XSEDE

• To get started using XSEDE a researcher needs to:
  – Apply for an Allocation, or
  – Get added to an existing allocation

• To do either of these things, you should start with the **XSEDE User Portal**
XSEDE User Portal (XUP)

• Web-based single point of contact that provides:
  – Continually updated information about your accounts.
  – Access to your XSEDE accounts and allocated resources:
    – 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.org Home Page

• From here, you can create a web account at any time!
User Responsibilities and Security

• The first time you login to the Portal, at the beginning of each allocation term, you will be asked to accept the User Responsibilities form:
  – Explains acceptable use to protect shared resources and intellectual property. Acknowledgment in publications, etc.
  – You are responsible for your account: Do not share accounts
  – User is responsible for protecting the passwords:
    • Includes not sharing passwords, not writing passwords down where they can be easily found, and not using tools which expose passwords on the network
    • This includes private keys: make sure they are password-protected.

• Appropriate Behavior
  – Protecting computing, closing SSH terminals when done, logging out of the User Portal when done, etc.

• Report Suspicious Activity.

• If you have any suspicion that your account or personal computer has been compromised send email to help@xsede.org or call 24/7 1-866-907-2383 immediately.
Getting an Allocation

• If you do not yet have an allocation, you can use the portal to acquire one.

• If you are a first time investigator, request a startup allocation.
Creating an Allocation

Submission Home

User: Barth, William

Research Requests:

1. New
   CCR100034 Submission Status: Approved Active Status: Active
   Title: Developing Advanced MPI capabilities for Petascale Applications
   Submitted On: Jul 16, 2010 Award Start Date: Oct 01, 2010 Award End Date: Mar 30, 2012
   Actions: [View | Transfer | Supplemental]

Staff Requests:

1. New
   STA110019 Submission Status: Approved Active Status: Active
   Title: XSEDE SP TACC
   Submitted On: Sep 19, 2011 Award Start Date: Sep 20, 2011 Award End Date: Sep 20, 2012
   Actions: [View | Transfer | Supplemental]

New Request: Select An Appropriate Category For A New Request

- Research: Next submission period will be 12/15/2011 - 01/15/2012
- These Research requests will be for allocation period 04/01/2012 - 03/31/2013
- Startup: For investigators new to XSEDE. Accepted anytime.
- Educational: Specifically for classroom instruction and training courses. Accepted anytime.
- Campus Champions
Once your allocation is approved:

• The PI (principle investigator), Co-PI, or Allocation Manager can add users to an existing allocation through the portal.
  – XSEDE UserPortal:My XSEDE->Add/Remove User

• Takes the portal name of the user you want to add/remove.

• Accounts at certain Service Providers need to be activated before they can be accessed.
Accessing XSEDE Resources

• Several methods are possible:
  – Direct login access
  – Single Sign On (SSO) through portal
  – SSO between resources
  – Through Science Gateways

• Your choice of method may vary with:
  – How many resources you use
  – How much you want to automate file transfers, job submission, etc.
Accessing Resources(2)

• SSO is the default method; you’ll need to file a ticket to request a direct access password to the machine.

• Direct access:
  – Use a secure shell (ssh) client.
  – From Linux or Mac terminal window:
    • ssh –l <username> <machinename>
    • E.g: ssh –l dstanzi ranger.tacc.utexas.edu
  – From Windows:
    • Download one of many ssh clients
    • Free ones include “putty”
    • Most campuses have a site license for a fancier one.
Single Sign On

- Single Sign-On (SSO) allows you to use just one username and password (your User Portal one) to log into every digital service on which you have an account.
- The easiest way to use SSO is via the XSEDE User Portal, but you can also use SSO via a desktop client or with an X.509 certificate.
- After you authenticate using SSO with your User Portal username and password, you will be recognized by all XSEDE services on which you have an account, without having to enter your login information again for each resource.
SSO thru user portal

- 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 access.
A Java Applet will talk... you may be asked permission to allow it to run.

After the applet starts, a blank terminal window will appear in your web browser.

The window will fill with text indicating that you have been successfully logged into the resource of your choice.

You can now work on this machine, and connect to other machines from this terminal, using the command:

```
gsissh machine-name
```
Starting SSO from an XSEDE Resource

- If you have logged in directly to a machine (with ssh), and you desire SSO access to other XSEDE machines, you will need to start up SSO manually.
- First, generate a certificate using the command:
  - `myproxy-logon -l <username>`
- Where username is your Portal username. You will get a prompt to enter your MyProxy passphrase. *Use your current XSEDE Portal username and password for this step.*
  - Type `man myproxy-logon` to learn more about options to this command.
- Run `grid-proxy-info` to verify a proxy has been established. This will also tell you how long your proxy will last. Now you can access other resources without a password using `gsissh`
Another Access Path: Science Gateways

• There are many sites that give you web-based, domain-specific access to applications running on XSEDE.

  – Collectively, we call them “Science Gateways”
  – View a list of them on the User Portal, in the “Resources” tab.
  – Access methods vary; click on the specific gateway to find out more (dozens available, across many fields!).
File Transfers: Small Files (<2GB)

• To transfer *small files* between XSEDE Resources and/or your own workstation you can use `scp` or `sftp`.

• From Linux or Mac, you can run these commands directly from the terminal.

• From Windows, use your ssh client to do this (putty has free downloads for these tools, too! – just Google “putty sftp”).

• These are easy to use and secure, but provide poor performance for large files.
File Transfer: User Portal

• Log into the XSEDE User Portal

• Select “Resources” tab

• Select “File Manager” tab
  – (now wait for Java Applet to load)
  – May need to allow access for applet to run by clicking OK.

• You will see a list of all machines. This includes:
  – your local machine.
  – XSEDE$Share: 2GB of space to collaborate. Allows you to share files with your collaborators.
Transferring Large Files with User Portal

- For large file transfers, we need to set a few parameters.
- Before clicking on the resource, Right click on the resource you’re going to transfer data from and select Edit. This will bring up the file transfer parameters:
  - Click the checkbox next to “Stripe Transfers” - Click OK
  - Repeat for the other panel using the destination resource
  - Repeat this every time you change Resources
  - Drag and drop the file from source to destination to transfer.
Large Files – Command Line Transfers

• Within XSEDE, you can use the command line to transfer large files with `uberftp` or `globus-url-copy`.

• Example: From PSC “Blacklight” to TACC “Ranger”:

  – Optimized for large files
  – Look here for names of gridftp servers at each site: https://www.teragrid.org/web/user-support/transfer_location

• `speedpage.psc.edu`: provides information on transfer speeds you can expect using `globus-url-copy` with the optimized parameters above.
What is Globus Online?

- Initial implementation of XSEDE User Access Services (XUAS)
- Reliable data movement service
  - High performance: Move terabytes of data in thousands of files
  - Automatic fault recovery
  - Across multiple security domains
- Designed for researchers
  - Easy “fire and forget” file transfers
  - No client software installation
  - New features automatically available
  - Consolidated support and troubleshooting
  - Works with existing GridFTP servers
  - Ability to move files to any machine (even your laptop) with ease

"We have been using Globus Online to move files to a TeraGrid cluster where we analyze and store tens of terabytes of data... I plan to continue using GO to access these resources within XSEDE to easily get my files where they need to go.”
-- University of Washington user

“The service is reliable and easy to use, and I look forward to continuing to use it with XSEDE. I've also used the Globus Connect feature to move files from TeraGrid sites to other machines -- this is a very useful feature which I'm sure XSEDE users will want to take advantage of.”
-- NCSA user
Case Study: Indiana University

- **Fast**: Reduced transfer times
- **Easy**: Fire-and-forget transfers
  - Automated retry
  - No file pre-staging
  - No complex infrastructure
  - Convenient CLI or GUI interfaces

“Globus Online frees up my time to do more creative work than typing `scp` commands or devising scripts to initiate and monitor progress to move many files.”

“I moved 100 7.3 GB files tonight in about 1.5 hours. I am very impressed. I also like the new commands and help system.”

Indiana University researcher moved ~6 TB from Oak Ridge to TACC in 2 days.
How It Works

1. User initiates transfer request
2. Globus Online moves files
3. Globus Online notifies user

Data Source
(XSEDE site or other location)

Data Destination
(XSEDE site or other location)
We strive to make Globus Online broadly accessible...

- Researchers with no IT background can just move files using the **Web GUI**
- Developers who want to automate workflows can use the **Command Line Interface (CLI)**
- System builders who don’t want to re-engineer file transfer solutions can use the **REST API**
Getting Started (2 easy steps)

1. **Sign up**: Visit [www.globusonline.org](http://www.globusonline.org) to create an account
Getting Started (2 easy steps)

2. Start moving files: Pick your data and where you want to move it, then click to transfer
For More Information

• Visit https://www.globusonline.org/signup to:
  – Get a free account and start moving files
• Visit www.globusonline.org for:
  – XSEDE User Guide
  – Tutorials
  – FAQs
  – Pro Tips
  – Troubleshooting
• Contact support@globusonline.org for:
  – Help getting started
  – Help using the service
The Mobile User Portal

• [https://mobile.xsede.org](https://mobile.xsede.org)
• Allows browsing of all XSEDE systems, file downloading, and third-party transfers.
• It provides several features for mobile users such as one touch file publishing to the user's public folder, simple creation of shared groups for any file/folder, and one click permission management of all XSEDE systems, file downloading, and third-party transfers.
XUP Resource Monitor

- View system information: TFLOPS, memory, today’s load, jobs running in queue. **Status: up or down**: takes you to the news announcements that tells you when the machine is expected to come back up.
User Portal: User Forums

• The User Forums are a great place to ask questions, get help, or discuss ideas about XSEDE.
Running Jobs

• Each system in XSEDE has some local options you will need to know about to run jobs.
• To learn about the specifics of each system, check out the user guides:
  – In the portal, under “Documentation” select “User Guides”
• Pay particular attention to:
  – File Systems
  – Batch job submission
File Systems on XSEDE Resources

• 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; use while you are running your jobs. *Scratch space is temporary; it is not backed up and has limited redundancy, and is periodically purged of old files!*
  – Archival Storage: Long term storage of large amounts of data (often tape); slower access, accessible from all sites.
Batch Jobs

• All XSEDE compute resources use some form of batch scheduler.
• Compute jobs *can not* be run on the login nodes (no faster than a normal workstation!)
• There are several batch systems in use, but all work basically the same way.
  – Request number/type of nodes you need.
  – Specify how long you need to run.
  – Specify where your output files go.
  – Jobs typically described with a job script:
Sample Job Script for Grid Engine on TACC Lonestar:

#!/bin/bash

#$ -N myMPI # Job Name
#$ -j y # Combine stderr and stdout
#$ -o $JOB_NAME.o$JOB_ID # Name of the output file
#$ -pe 12way 24 # Requests 12 tasks/node, 24 cores total
#$ -q normal # Queue name "normal"
#$ -l h_rt=01:30:00 # Run time (hh:mm:ss) - 1.5 hours

ibrun ./a.out # Run the MPI executable named "a.out"
Submitting/manipulating batch jobs

• Submit the script that you have created:
  – Actual commands are machine specific, but they follow general principles.
  – `qsub jobname`
  – `qstat -a`
  – `qstat -u username`
  – `qdel jobid`
  – `man qsub`
Managing Your Environment: Modules

- Allows you to manipulate your environment.
- `module list` shows currently loaded modules.
- `module avail` shows available modules.
- `module show <name>` describes module.

http://modules.sourceforge.net/

% 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
Getting Help

• For continually updated information on your account, allocation, and more, login to the XSEDE User Portal
  – portal.xsede.org -> Help
  – portal.xsede.org -> My XSEDE -> Tickets
  – portal.xsede.org -> Documentation -> Knowledge Base
    • Provides answers to frequently asked questions

• xsede.org -> Resources
  – User support documentation and user guides

• portal.xsede.org -> User Forums
  – Discuss your questions with other users and XSEDE staff.

• portal.xsede.org -> Training
  – Seminars or in-person training sessions. Enroll online.

• Get face-to-face help from XSEDE Campus Champions
• Extended Collaborative Support
Reporting and Tracking Issues

• portal.xsede.org -> Help
  – Help Desk: Submit ticket
  – Security Incidents
    • E.g. your account has been compromised

• portal.xsede.org -> My XSEDE -> Tickets
  – Submit ticket
  – View past tickets (both open and closed)

• Email: help@xsede.org

• Phone: 1-866-907-2383 (24x7x365)
More Resources

• xsede.org -> User Services
• Resources available at each Service Provider
  – User Guides describing each system
    • CPUs, memory, filesystems, etc.
    • Storage facilities
    • Software
• Training
• How to get help
• Extended Collaborative Support
  – The service formally known as Advanced User Support (AUSS)
ECSS can include

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

• Expertise available in a wide range of areas
  – Performance analysis
  – Petascale optimization
  – Gateways and web portals
  – Specialized scientific software

• Can solicit ECSS support at any time though the Allocations tab in the XSEDE User Portal

• Requires written justification and a project plan

• Inquire at help@xsede.org
Questions? Need Help?

- First, try searching the knowledge base or other documentation
- Next, submit a ticket
  - portal.xsede.org -> My XSEDE -> Tickets
- Send email
  - help@xsede.org
- Or call the Help Desk
  - 1-866-907-2383
Need more training?

• portal.xsede.org -> Training
  – Course Calendar
  – On-line training
Which Resource is Right for Me?
Thanks for listening, and welcome to XSEDE!

dan@tacc.utexas.edu