March 6, 2014

Campus Bridging: What is it and how do we do it?

Rich Knepper

rich@iu.edu

XSEDE

Extreme Science and Engineering Discovery Environment

NSF

XSEDE Campus Bridging Mission

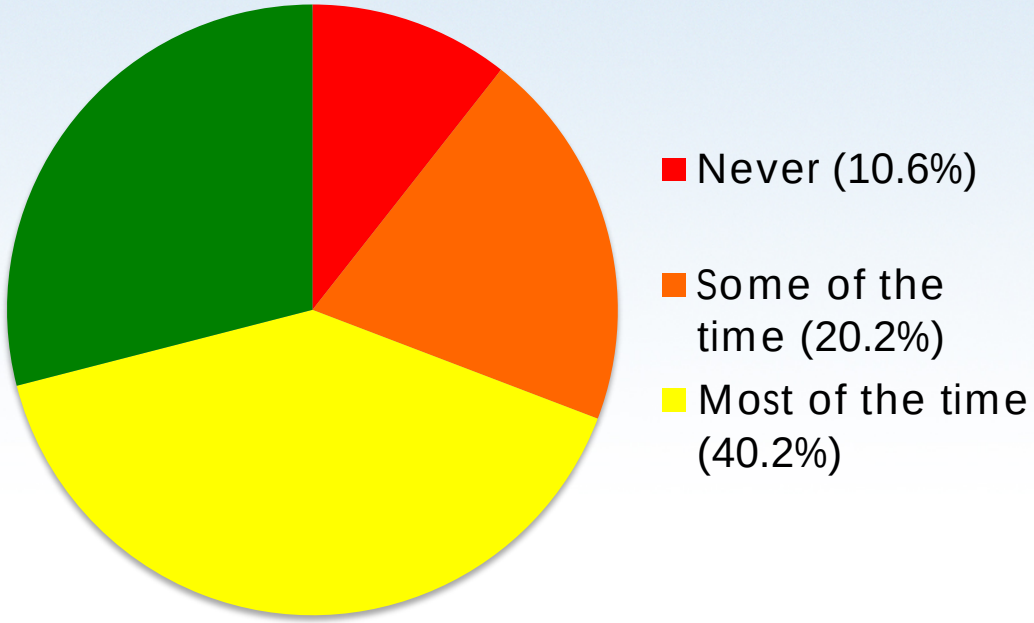
- To make it possible for researchers, educators, and students to access, from their labs, remote cyberinfrastructure resources individually or in concert, as simply as if they were peripherals attached to their own personal laptop or computer;

- To disseminate software, training materials, and information that enables the US research community to leverage the nation's aggregate cyberinfrastructure to maximize US innovation and global competitiveness.

What is Campus Bridging?

- "Bridging" the gap between local researcher cyberinfrastructure, campus CI, and national CI resources
  - Hardware
  - Software
  - Support
  - Outreach

Adequacy of research CI



Responses to asking if researchers had sufficient access to cyberinfrastructure resources – survey sent to 5,000 researchers selected randomly from 34,623 researchers funded by NSF as Principal Investigators 2005-2009; results based on 1,028 responses

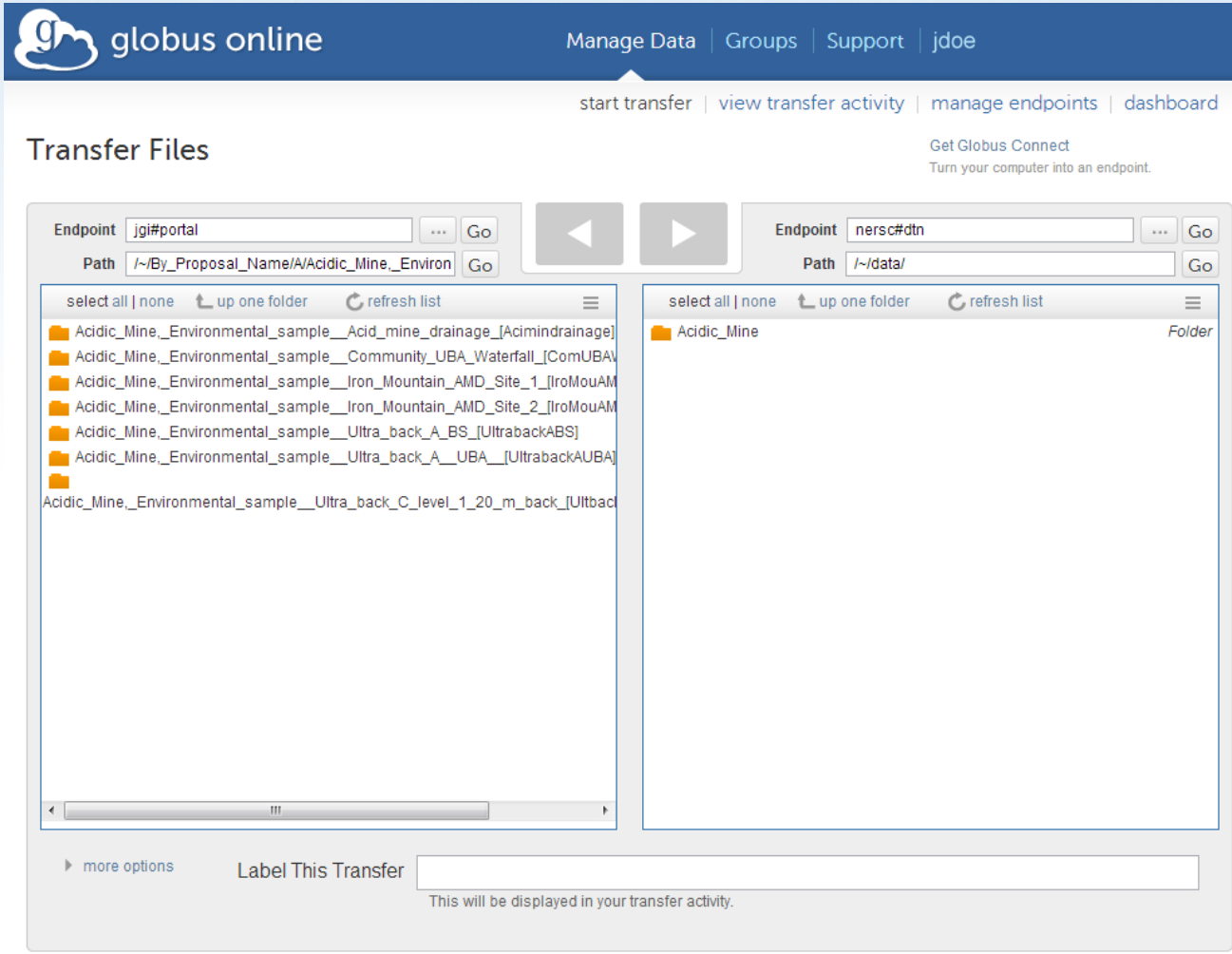
Stewart, C.A., D.S. Katz, D.L. Hart, D. Lantrip, D.S. McCaulay and R.L. Moore. Technical Report: Survey of cyberinfrastructure needs and interests of NSF-funded principal investigators. 2011. hdl.handle.net/2022/9917

Accessing and Moving data

- Because who doesn't have challenges moving data around?
- Our focus is on tools that move data from campus to XSEDE and back
- Many of the tools we are supporting and developing can be used more generally to move data between any two points

Globus Online

- Now in production as an XSEDE tool
- Simple interface for moving files back and forth
- Can be configured to connect between multiple computers and larger resources to transfer data



Screenshot courtesy of http://genome.jgi.doe.gov/pages/portal_apps.jsf

Globus Online 2

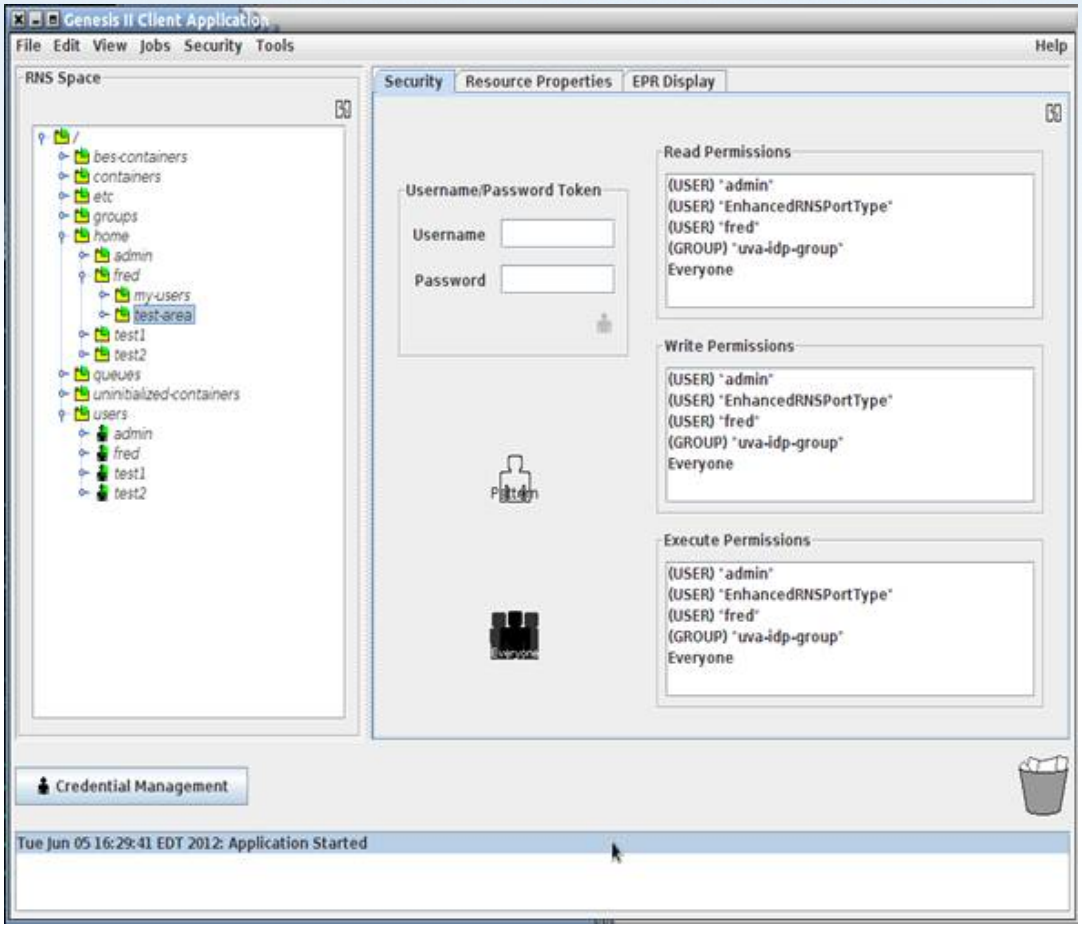
- Set a transfer and start it, walk away, and get a notification when the transfer is complete
- During Q3 2013, GlobusOnline allowed researchers to move over 630TB of data onto the XSEDE network
  - That's just ONE direction!

GFFS (based on the Genesis II software project)

- GFFS allows users to export a directory into a globally federated file system that is visible from any system which is running a GFFS client
- Not quite yet in production

GFFS

GFFS provides a fuse-based filesystem, which sits within user space on your system. It can be mounted by remote systems in XSEDE to access data when it runs your job



Screenshot courtesy http://genesis2.virginia.edu

The basic XSEDE-compatible cluster

- Building a cluster is a lot of work. A lot of that work can be automated
- XSEDE has developed the concept of a "basic XSEDE-compatible cluster software stack"
- Benefits to you:
  - Eases later transition to larger CI
  - Saves investment on development and updates
  - Allows researchers to more easily reuse XSEDE training resources
  - Teaches students interested in HPC to operate within a context similar to XSEDE
  - Focus your effort addressing specific local challenges!

Basic XSEDE-compatible cluster from scratch - Rocks Rolls

- New cluster? Rocks ISOs can automate much of the work of setting up a basic XSEDE-like cluster
- Press the lever, get a cluster
- Maintained and updated by XSEDE Campus Bridging
- Also useful if you just hate your cluster and want to start over from scratch
- See a video on the Rocks Rolls distribution and its value to researchers at: http://www.youtube.com/watch?v=E96mlWruATM

Adding Basic XSEDE-compatible cluster functionality to an existing cluster - YUM Repository

- Existing cluster? XSEDE provides documentation to configure your cluster such that you can use these resources
- Install YUM Repo locally, manage XSEDE packages as if they were part of the OS, just like you would with the base or EPEL repositories
- Add tools that let you leverage XSEDE – like Globus Online – without loosing the effort you have invested in your cluster already

Job management and submission from local clusters

- Two components
  - Unicore 6 Basic Execution Services (BES)
    - Allows for jobs to be submitted as files into a directory
    - Directory is a jobs queue and BES executes job files as jobs in a queue
  - Global Federated File System (GFFS)
    - Standardizes user interface across compute resources
    - GUI component allows drag and drop uploading, user and group management, and POSIX-style permissions management independent of the command line

Who ya gonna call? XSEDE

- Help by telephone - 812-318-2872, by email to campusbridging@xsede.org - 9-5 eastern time
- Email help@xsede.org
- XSEDE Knowledge Base - https://www.xsede.org/web/xup/knowledge-base
  - http://bit.ly/yumrepo
  - http://bit.ly/rocksxsede
- We would be happy to visit your campus to help set up Globus Online or install basic XSEDE-compatible cluster software

And how do we tell if we have done our job?

- Adoption of CB practices and implementation of CB software
- We may well be asking for information from you on a survey . . . If we do, please take time to fill it out
- **And for now: if you do nothing else, go to xsede.org, set up a portal identify, and look around some**

Summary

- Campus bridging aims to create virtual proximity between researchers and resources
- Streamlining the analytical part of research
- Ways we are working to bridge research computing:
  - Basic XSEDE-compatible software stack
  - Full-time dedicated support and consulting
  - GlobusOnline
  - GenesisII

For more information…

- https://www.xsede.org/campus-bridging
- http://cb-repo.iu.xsede.org/xsederepo/
- https://www.xsede.org/globus-online
- http://genesis2.virginia.edu
- campusbridging@xsede.org
- +1 (812) 318-2872 (M-F, 9a-5p Eastern)

- Any questions?


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