THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Glen MacLachlan, Jason Hurlburt, Marco Suarez, Adam Wong, Zhen Ni, Hadi Mohammadi

HPC Workshop 1

What we're covering:

- Logging in
- Navigating the shell
- Modules, environment variables and .profile
- Quotas
- Purges
- File transfer and management:
  - scp
  - Globus
  - Lustre vs NFS filesystems
  - Lustre striping & inodes

Colonial One

HPC Workshop - HPC @ GWU

Current Specs:

- Dell C8220 cluster, 213 node
- 53x GPU nodes, 2x NVIDIA K20 GPUs
- 1x 2TB Node, Quad 12-Core 3.0GHz Xeon E7-8857v2 CPUs
- 159x CPU nodes, 2x 2.6GHz 8-core Xeon CPUs, 64/128/256GB of RAM

Totals:

- 3,228 (2,592) Intel Xeon CPU cores
- 264,576 NVIDIA CUDA cores
- over 27 TB of RAM
- Mellanox FDR Infiniband fabric
- Two primary filesystems
- 262 TB NFS fileserver for /home and /groups
- 262 TB Lustre filesystem for high-speed scratch
- 268 TB Dell Compellent for archival

HPC Workshop - HPC @ GWU

Colonial One:

- Serves over 900 users in nearly 150 research groups
- Runs 24/7, 365 days a year
- Processes > 2,000 jobs every day
- User demand is 91% of capacity
- Open to entire GW community
- 129 open proposals for funding reference Colonial One

HPC Workshop - HPC @ GWU

Jobs Run on Colonial One:

- Study structure of subatomic particles
- Large-scale molecular dynamics simulations
- Network analysis
- Drug design for cancer therapy
- Protein engineering for immune response against bacteria and viruses including HIV/AIDS
- fMRI analyses of injured brains
- Genomic sequencing
- Phylogenetic mapping of evolutionary traits
- Satellite imagery
- Population and census dynamics

HPC - Logging In

Requirements:

- SSH Client
- Colonial One account

Log into Colonial One:

ssh username@login.colonialone.gwu.edu

Use your NetID and password!

HPC - Navigating the shell

Pathname

- A path through the directory system
- *pwd* – shows current path
- Absolute vs. Relative path

/ - the forward slash

- Represents the very bottom (root) of the file system
- acts as a divider in between directories on the file system

HPC - Navigating the Shell

- pwd: Print Working Directory, shows you where you are
- . versus .. : Your current directory versus the directory one level above
- The ~ character: Shortcut your home directory
- ls: list current path contents
- ls –la: list all details of the current path in long form
- cd: change directory
    - cd /absolute/path
    - cd path/relative/to/where/I/am

HPC - Modules

Modules load an environment so a program can run correctly.

Module commands:

- module list
- module avail
- module load
- module unload
- module spider

Environment variables

Environment variables are a set of dynamic named values that can affect the way running processes will behave on a computer. They are part of the environment in which a process runs.

Environment commands:

- printenv
- printenv Variable_Name
- echo $Variable_Name
- export Variable_Name=Value

HPC - Shell Configuration Files

- .bashrc: Runs when logging into a BASH session. Local to the BASH shell
  - You can enter the same commands inside .bashrc_profile as you can inside .profile
  - change your prompt: export PS1='[\u@\h:\w]\$ '
    - [hurlburj@login4:~]$
- Other shells have similar names: .cshrc (C shell), .ksh (Korn shell)

HPC - Quotas

Home and Group Quotas

- Soft quotas are in place now
- Home quota: /home/username - default 25GB
- Group quota: /groups/groupname - default 250GB
- Check quota: type "quotareport" at the shell

Colonial One is not meant for archival data. Please remove data from old jobs once you finish your project.

HPC - Purge

What data is purged?

- Home and Group shares are not purged
- The high speed lustre file system IS purged every month
- Lustre is to be used for scratch space while running jobs

When is data purged?

- At the beginning of every month

HPC - Purge Policy

Lustre Purge Policy Coming into Effect 3/1/2017

1. **Frequency:** A purge will be conducted on the first day of every month (starting on **3/1/2017**). In the past, purges have been scheduled based on how close lustre utilization was to capacity. In the updated procedure, a purge will be conducted irrespective of lustre utilization. Again, a purge will be conducted on the first of every month even if the 1st falls on a weekend or holiday.

2. **File Access Time:** All files whose access time is greater than 60 days will be subject to purging. NOTE: updating access times with the sole intent of circumventing purging of files may result in disciplinary action including account suspension.

3. **File Size:** Files will be subject to purging regardless of the size they occupy on disk.

HPC - File Transfer with SCP

SCP - Secure Copy

- Usage:
  - scp ***from*** […] ***to***
  - scp <sourcefile> <destfile>
  - scp host:<sourcefile> <destfile>
  - scp user@host:<sourcefile> <destfile>
- Syntax is like cp
  - -r flag to recursively copy directories
  - man scp for more options

HPC - File Transfer with Globus

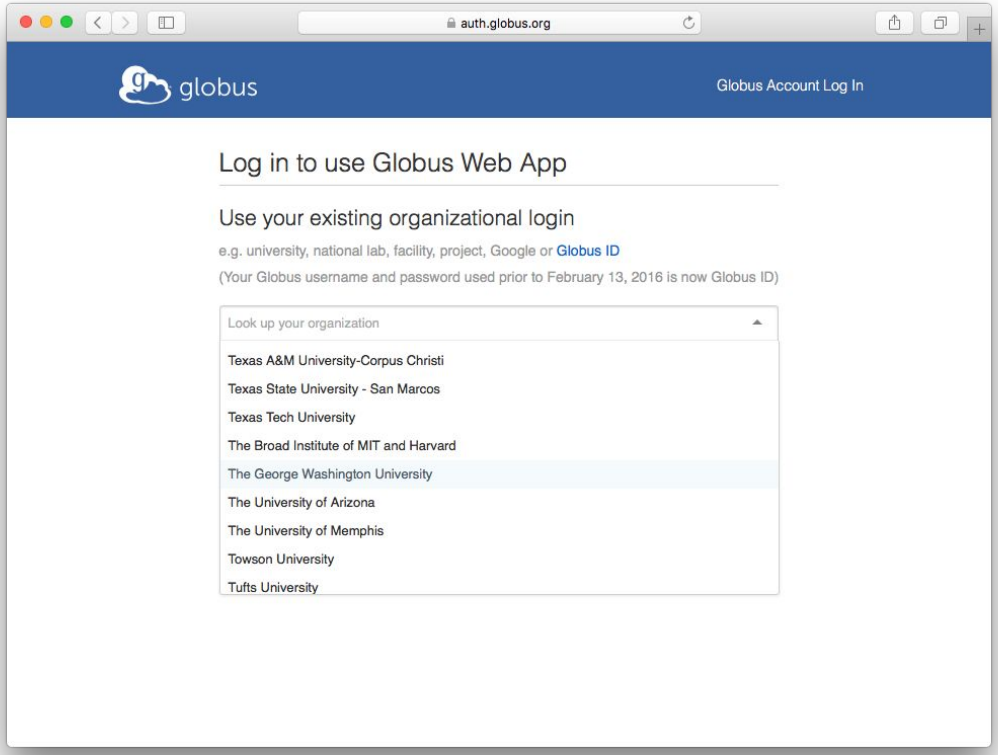
Globus is the industry standard for transferring large amounts of science and engineering research data between datacenters and endpoints.

Key points are:

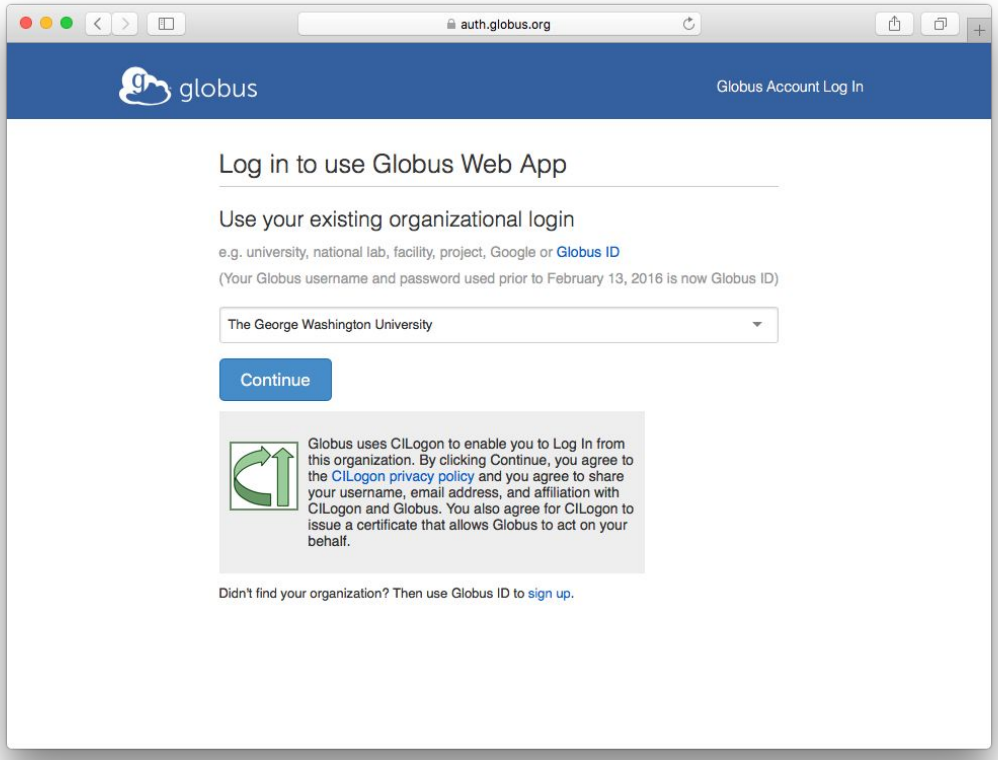
- Built on GridFTP technology.
- Data Transfers can be encrypted in flight (not encrypted by default).
- Transfer run in the background and can be interrupted and restarted, even if a file is partially transmitted.
- Free for individuals (institutes must pay to use the service).
- Globus is used literally everywhere by everyone transferring data in the HPC world. AWS, National Labs, National Supercomputing Centers, Universities, and even GWU!

HPC - File Transfer with Globus

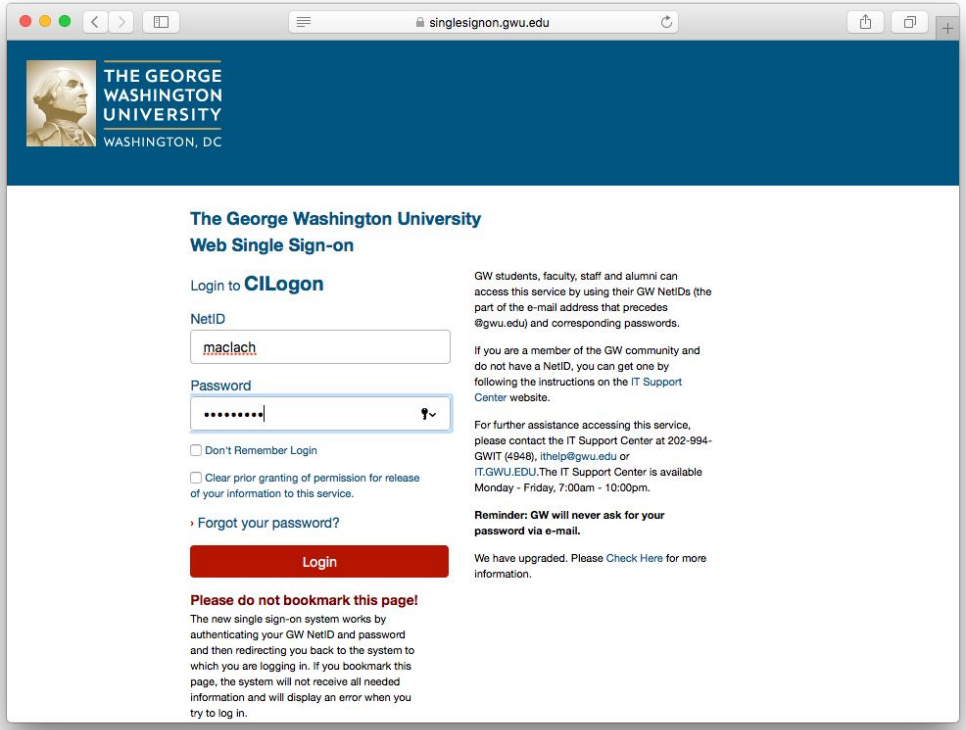
HPC - File Transfer with Globus



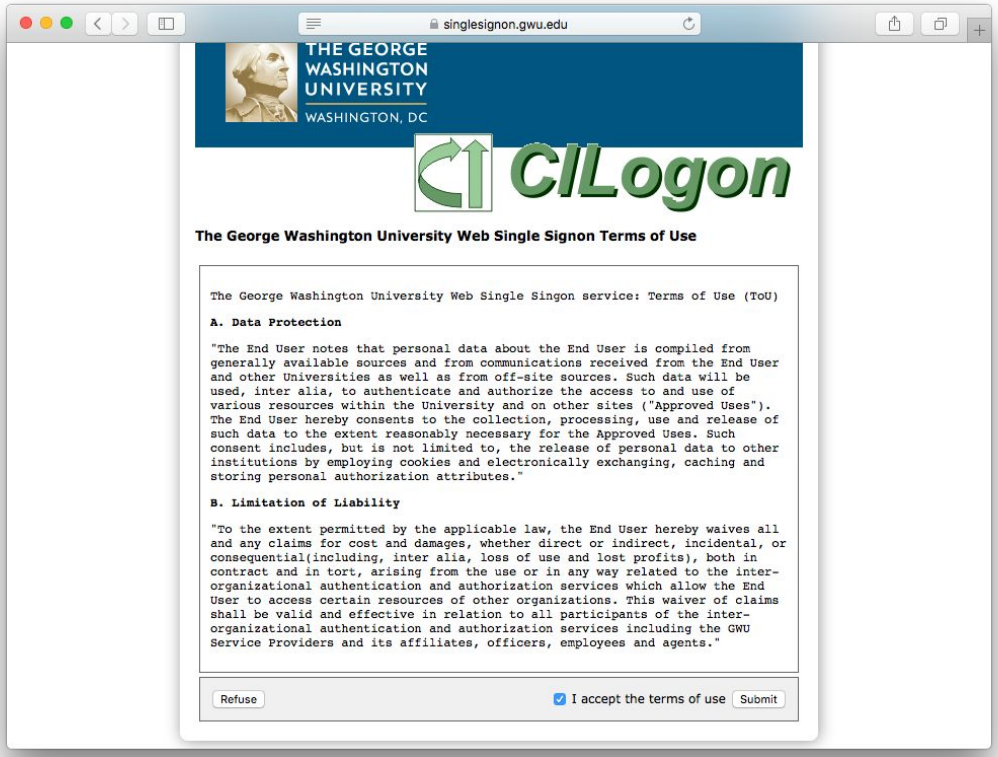
HPC - File Transfer with Globus



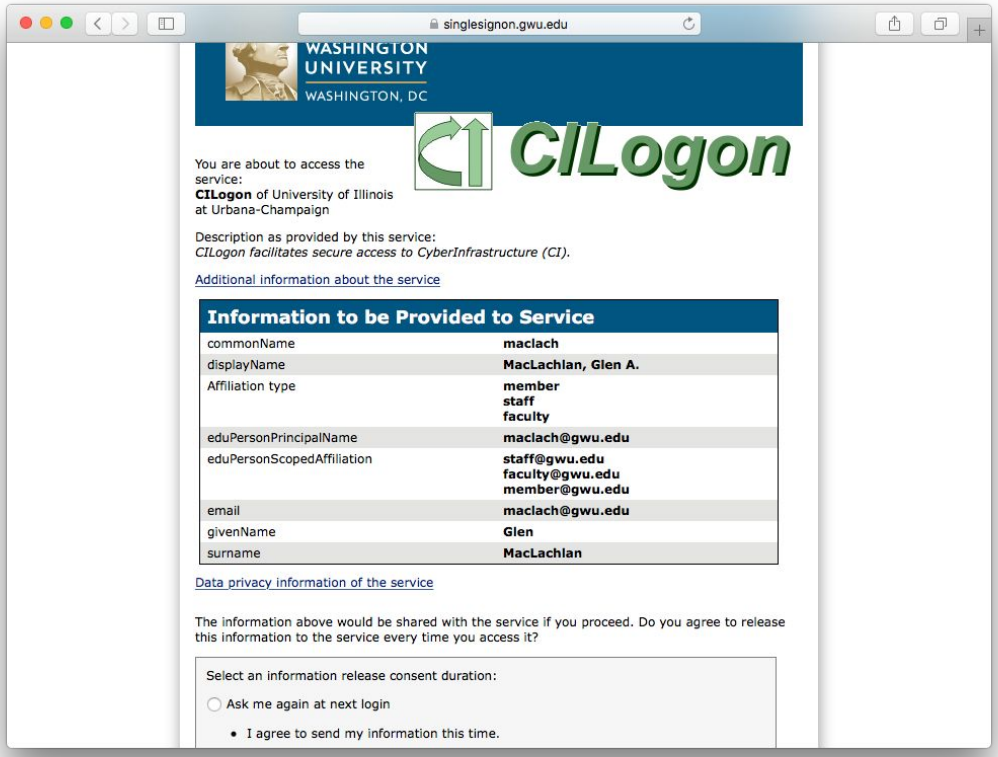
HPC - File Transfer with Globus



HPC - File Transfer with Globus



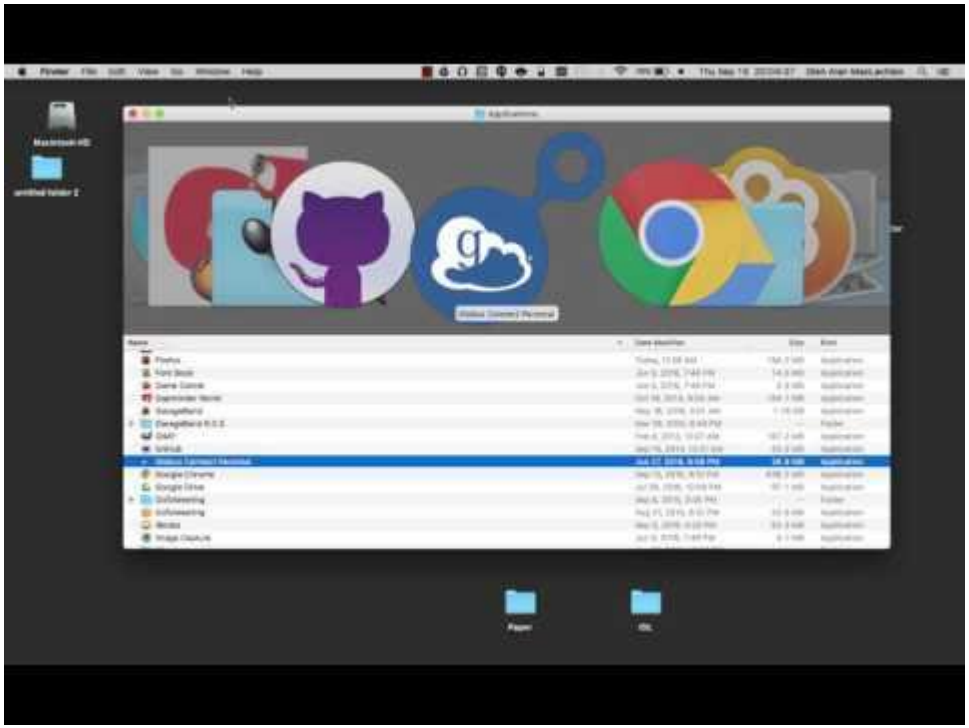
HPC - File Transfer with Globus



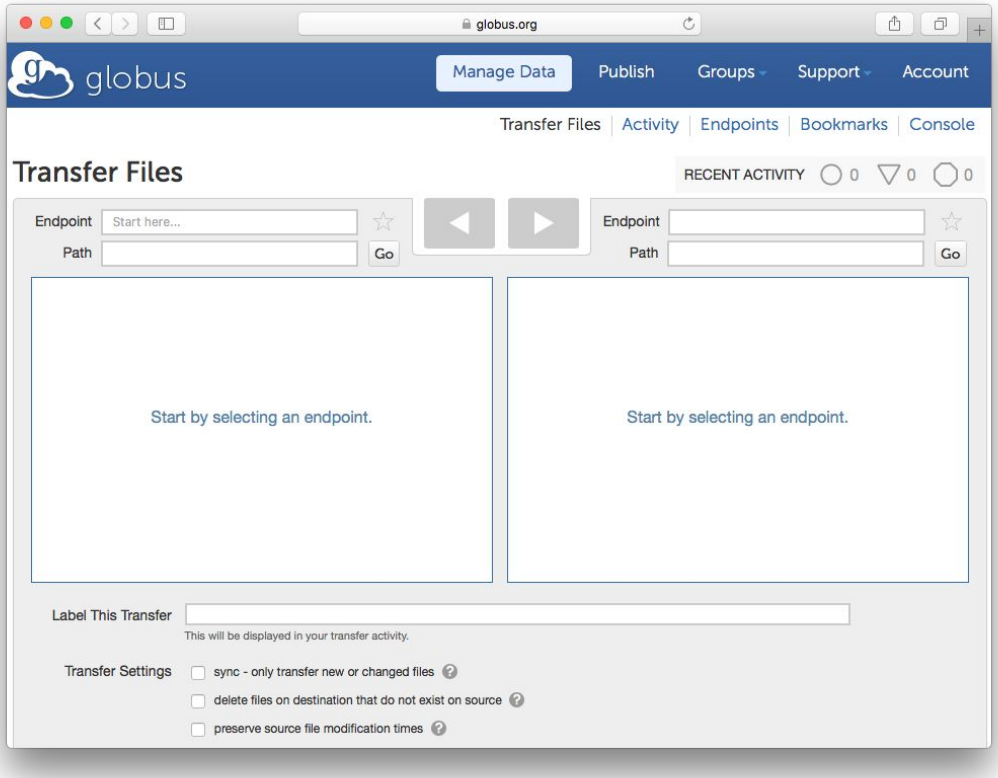
HPC - File Transfer with Globus

Setting up a
personal endpoint…

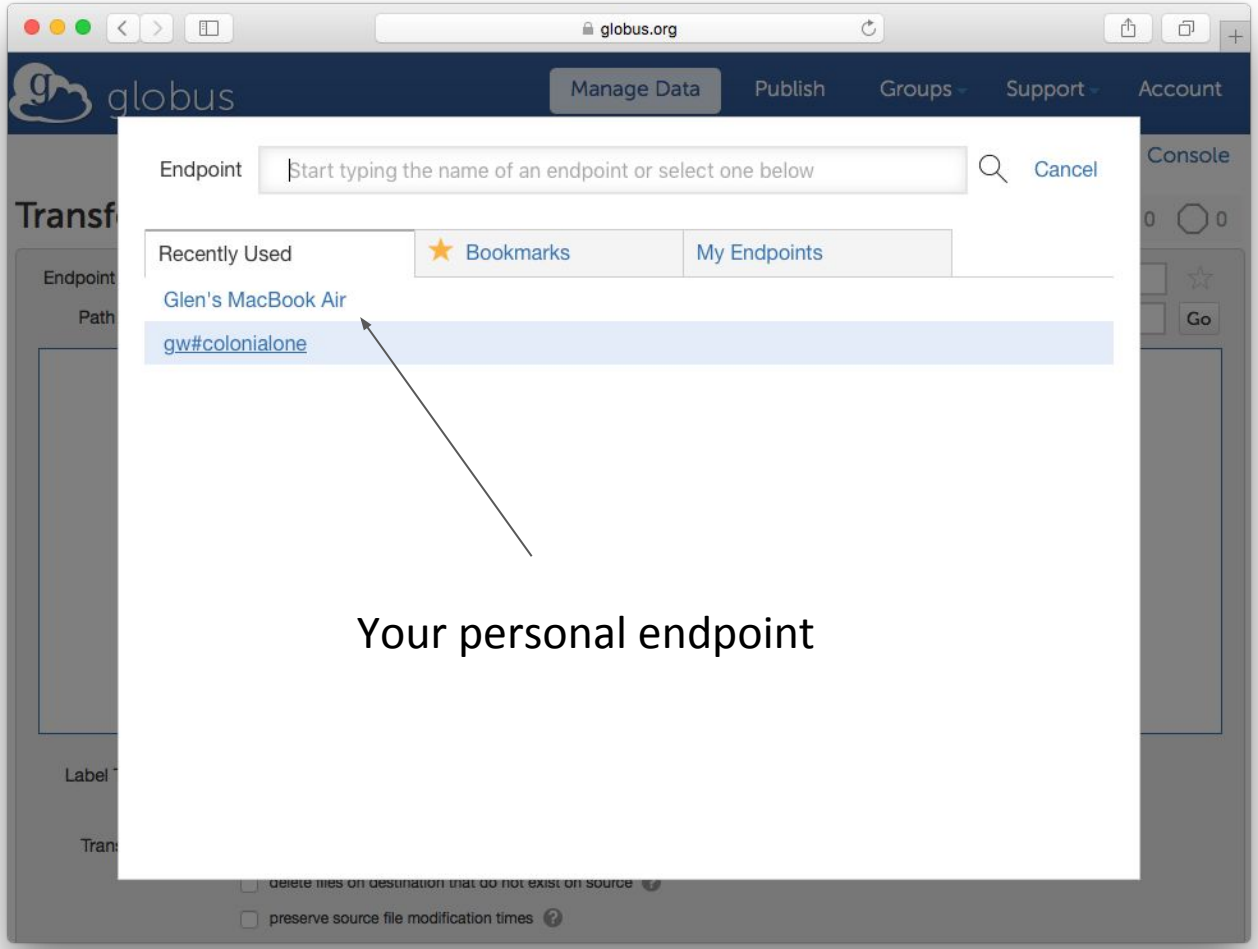
1. Go to EndPoints
2. Add globus personal connect endpoint and name it
3. generate and copy set up key
4. Download installer and install.
5. Open app and paste setup key
6. Go back to webpage and find endpoint



HPC - File Transfer with Globus

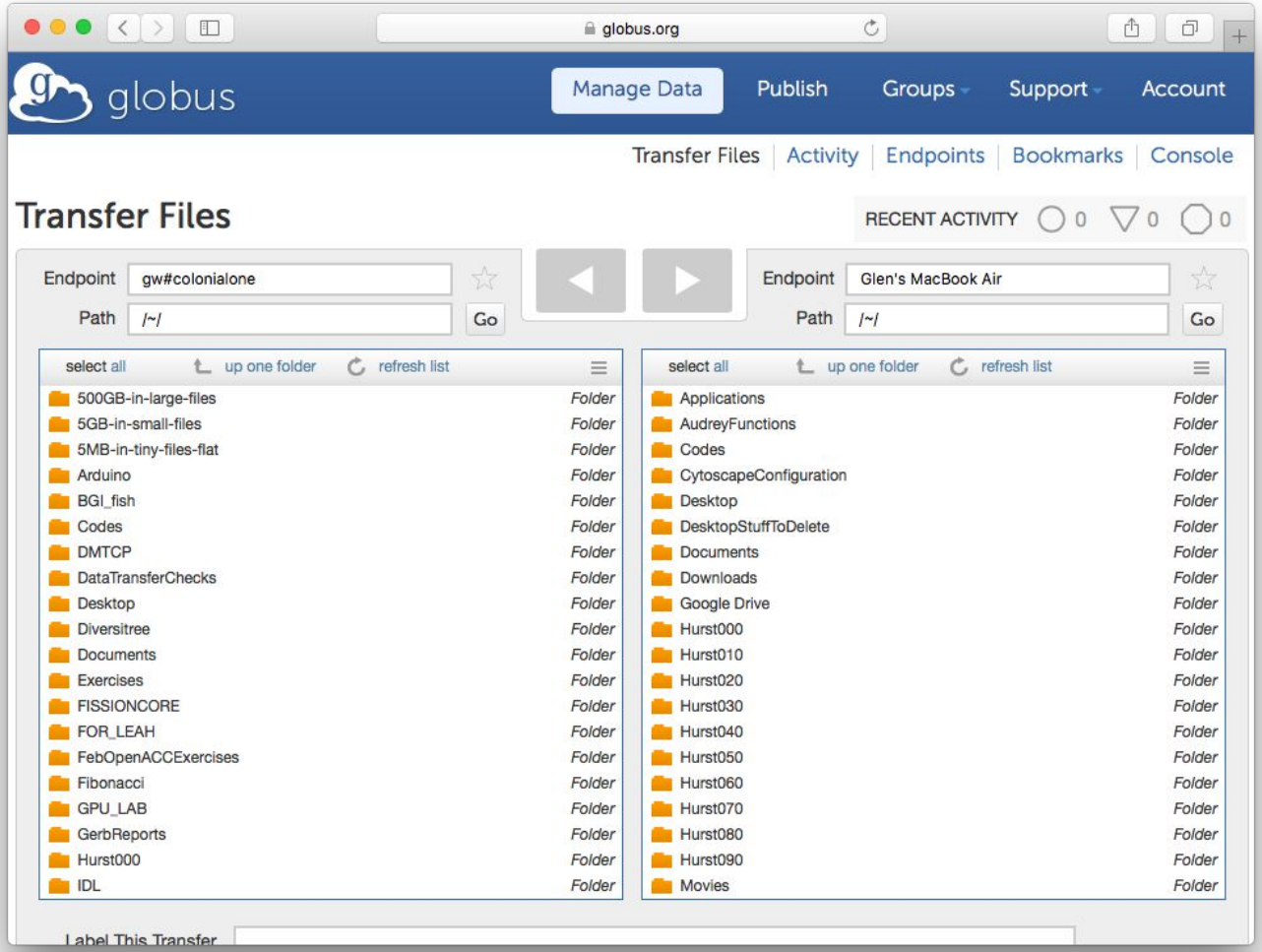


HPC - File Transfer with Globus



Your personal endpoint

HPC - File Transfer with Globus



HPC - Lustre and NFS

Lustre:

- Lustre is a free and open standard for creating a parallel high-speed file system
- It works by “striping” data over several different storage volumes.
- Lustre is a high speed storage system
- Lustre should be used for running jobs
- Is purged monthly

NFS:

- Network File System
- Hosts /home and /group directories
- NFS is slow compared to lustre
- Is not purged

HPC - Lustre

How to use Lustre:

- Using Lustre is one of the simplest things you can do on Colonial One or any cluster. You simply need to read or write to a lustre directory. Nothing else is required!

On Colonial One the lustre file system is found here:

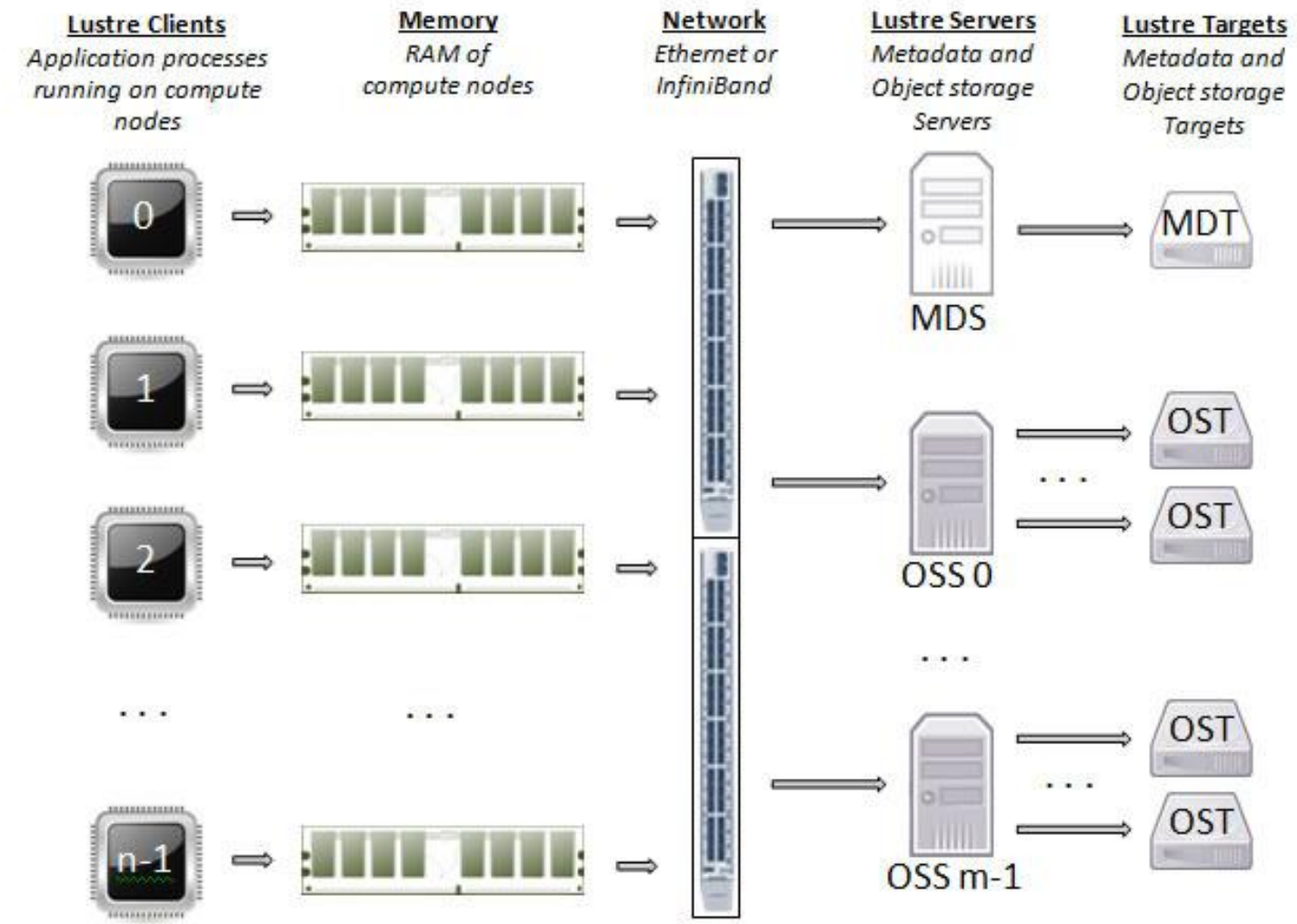
- cd /lustre/groups

HPC - Lustre Overview

Lustre System Components

1. Object Storage Server (OSS) – The OSS handles requests for storing files on one or more local Object Storage Targets.
2. Object Storage Target (OST) -The OST stores file data as “stripes” of files. A single file may be stripped across one or more OSTs. When a file is striped, pieces of the file are stored on more than one OST. The OSS provides information about where and how files are stripped.
3. Metadata Server (MDS) – The MDS is the gateway node that provides access the the Metadata Target nodes, MDTs.
4. Metadata Target (MDT) – The MDT stores metadata such as file attributes like file size, ownership, permissions, access times.
5. Lustre Clients – The nodes that mount the lustre file system. The login and compute nodes on Colonial One are examples of lustre clients.

HPC - Lustre Overview



HPC - Lustre Striping

Files are split into “chunks” and stored on separate OSTs so they can be read (or written) in parallel and therefore increase I/O bandwidth.

For large files it makes sense to stripe over many OSTs (on Colonial One we have 12 OSTs). However, for small files better performance can be gained by disabling striping.

```
lfs getstripe <dir|filename>
```

```
lfs getstripe --verbose <dir|filename>
```

Stripe size: The size of the chunks in bytes

Stripe count: The number of OSTs to stripe across. -1 means use all OSTs in the filesystem.

Stripe offset: The index of the OST where the first stripe is to be placed. The default is -1, which results in random selection.

HPC - Lustre Striping

You can set stripe settings for either files or directories. If you set striping for a directory, all files in that directory inherit the directory's settings.

lfs setstripe <dir|filename>

lfs setstripe -S stripe_size -c stripe_count -o stripe_offset <dir|filename>

Setting "-c 1" disables striping!

You can pre-create files of zero length and then write to them later.

lfs setstripe -c 1 bigdir

tar cf bigdir.tar bigdir

HPC - Useful Lustre Commands

lfs ls

lfs df -h

lfs df -i

lfs quota -h -v -u maclach

lfs osts <dir|filename>

lfs find /lustre/groups/phys6130_10 -mtime +30 -type f -print

See man lfs for more examples!

HPC - Lustre Striping Exercise

Create a large file with dd using different stripe patterns and see how the timings vary with the number of OSTs. For example,

lfs setstripe -c 1 test_1;dd of=test_1 if=/dev/zero bs=1024k count=100000 iflag=count_bytes

lfs setstripe -c 2 test_2;dd of=test_2 if=/dev/zero bs=1024k count=100000 iflag=count_bytes

lfs setstripe -c 4 test_4;dd of=test_4 if=/dev/zero bs=1024k count=100000 iflag=count_bytes

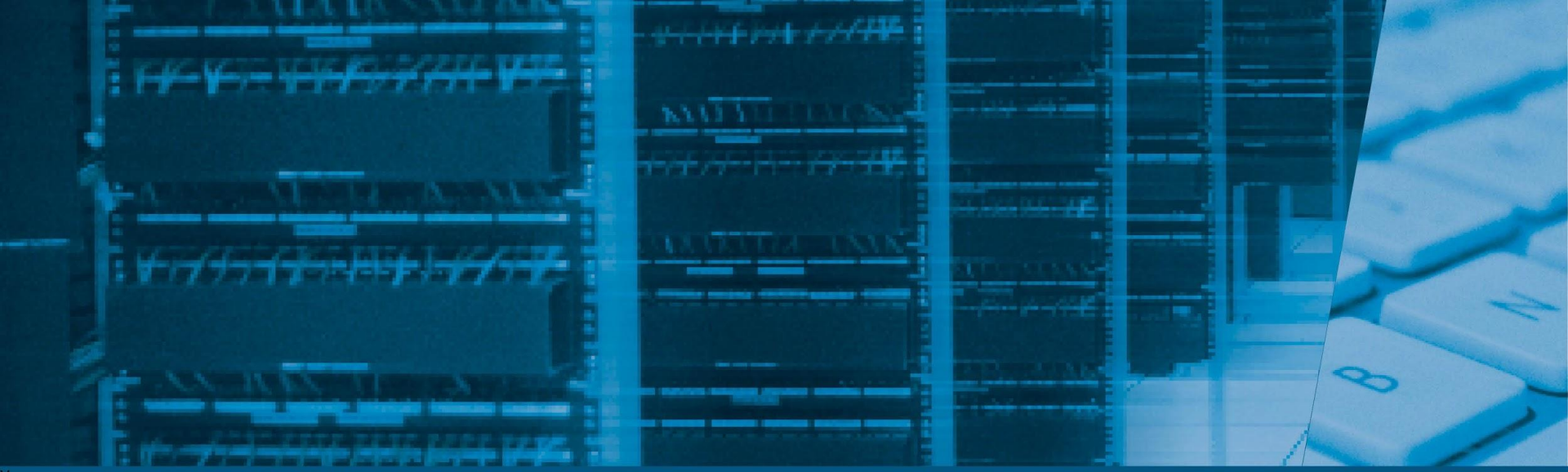
lfs setstripe -c 8 test_8;dd of=test_8 if=/dev/zero bs=1024k count=100000 iflag=count_bytes

lfs setstripe -c 12 test_12;dd of=test_12 if=/dev/zero bs=1024k count=100000 iflag=count_bytes

HPC - Job Scheduler

Colonial One uses SLURM to schedule and prioritize jobs on the cluster.

SLURM (Simple Linux Utility for Resource Management) is a software package for submitting, scheduling, and monitoring jobs on large compute clusters.



For More Information

Colonial One overview:

http://it.gwu.edu/colonialone-high-performance-computing

User documentation:

http://colonialone.gwu.edu

Or send us email:

Colonial One support - hpchelp@gwu.edu