OAK RIDGE National Laboratory

OLCF Storage Areas and Data Transfers

Suzanne Parete-Koon NCCS HPC Engineer
2-15-23

ORNL is managed by UT-Battelle LLC for the US Department of Energy

U.S. DEPARTMENT OF ENERGY

Alpine (gpfs)

NFS (user/proj home)

HPSS (archive)

DTNs (Data Transfer Nodes)

Orion (Lustre)

A Storage Area for every Activity

User Centric

- **User Home: (NFS)** Long-term data for routine access that is unrelated to a project. Read-only from compute nodes.
- **Member Work: (Orion/Alpine)** Short-term user data for fast batch-job access. Purged.
- **Member Archive: (HPSS)** Long-term project data for archival access that is not shared with other project members.

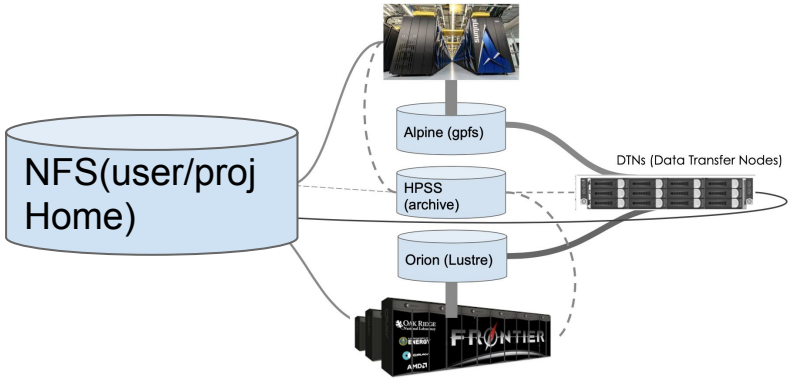
Project Centric

- **Project Home (NFS) :** Long-term project data for routine access that's shared with other project members. Read-only from compute nodes.
- **Project Work: (Orion/Alpine)** Short-term project data for fast, batch-job access that's shared with other project members. Purged.
- **Project Archive: (HPSS)** Long-term project data for archival access that's shared with other project members.

Areas for sharing between projects

- **World Work: (Orion/Alpine)** Short-term project data for fast, batch-job access that's shared with users outside your project. Purged.
- **World Archive:(HPSS)** Long-term project data for archival access that's shared with users outside your project.

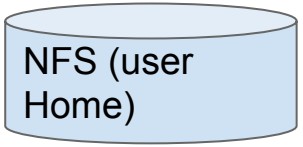
NFS Network File System



- User home: /ccs/home/$USER
  - User home is user-centric
- Project home: /ccs/proj/[projid]
  - Project-centric
- **Long-term** storage for your general data under home or related to project under proj
- User Home and Project Home will be accessible read-only from the Frontier compute nodes
- **Not purged**
- **Quota** of 50GB
- There is an automated **backup**

NFS Backups

I deleted a file from my NFS, how do I recover it?



Answer: snapshots

Go to the .snapshot folder (ls will not show this folder):

```
[Summit ~]$ cd $HOME/.snapshot
[summit .snapshot]$ ls -l
total 2048
drwxr-xr-x 232 suzanne users 61440 Feb  2 14:04 daily.2023-02-03_0010
drwxr-xr-x 232 suzanne users 61440 Feb  7 13:09 hourly.2023-02-08_1605
drwxr-xr-x 232 suzanne users 61440 Feb  2 14:04 weekly.2023-02-05_0015
```

ORION

Orion will be the largest and fastest single file POSIX namespace file system in the world.

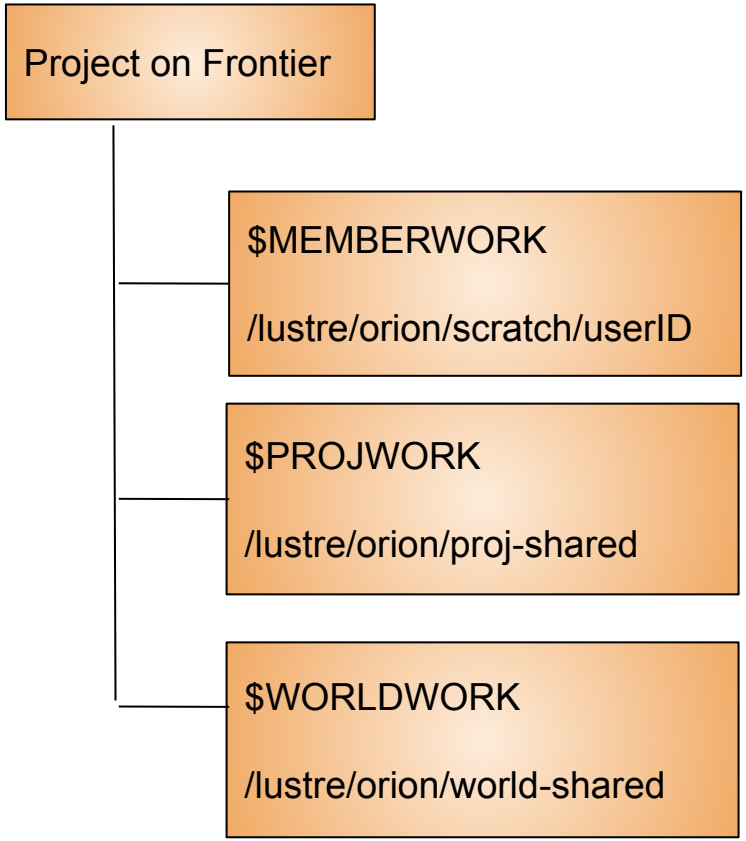
- Orion is a Lustre filesystem
- Flash-based performance tier of 5,400 nonvolatile memory express (NVMe) devices providing 11.5 petabytes (PB) of capacity at peak read-write speeds of 10 TB/s
- A hard-disk-based capacity tier of 679 PB at peak read speeds of 5.5 TB/s and peak write speeds of 4.6 TB/s
- flash-based metadata tier of 480 NVMe devices providing an additional capacity of 10 PB.

ORION

Orion is a Lustre filesystem

- Basic Lustre, in addition to other servers and components, is composed of Objects Storage Targets (OSTs) on which the data for files is stored. A file may be "striped" over multiple OSTs
- Striping provides the ability to store files that are larger than the space available on any single OST and allows a larger I/O bandwidth than could be managed by a single OST
- Orion will have multiple performance tiers for storing different sizes of data, so the concept of striping is even more complex that what is described above.
- While users may control striping, OLCF has built tools to help automatically choose the most efficient striping pattern for most files.

ORION



MEMBERWORK:

- Short-term storage of user data related to the project but not shared

PROJWORK:

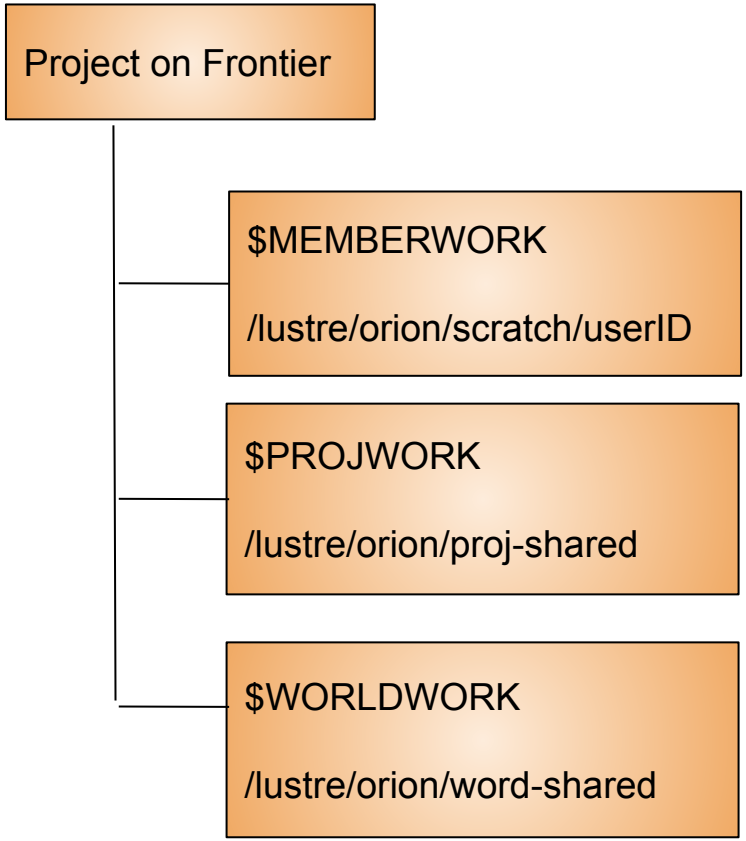
- Short-term storage of project data shared among the members of the project

WORLDWORK:

- Short-term storage of project data shared with OLCF users outside the project

Note:These aliases on Andes and the DTNs will point to Alpine until further notice.

ORION



SHORT TERM STORAGE

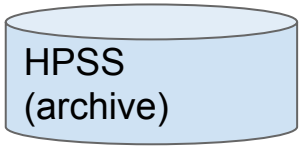
- **No backup**
- **Purged after 90 days**

Orion Storage Policy Prediction

We will update you when Orion is in production

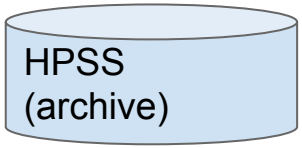
Area	Path	Permission	Backups	Purged	On Compute Nodes
Member Work	/lustre/orion/[projid]/scratch/[userid]	700	No	90 days	Read/Write
Project Work	/lustre/orion/[projid]/proj-shared	770	No	90 days	Read/Write
World Work	/lustre/orion/[projid]/world-shared	775	No	90 days	Read/Write

HPSS



- **Long-term** storage for large amounts of general data related to your project
- Access by htar and hsi from login nodes and DTNs, access by Globus using the “OLCF HPSS” Globus endpoint.
- HPSS is optimized for large files. Ideally, we recommend sending archives 768 GB or larger to HPSS.
  - If any of the individual files included in an htar are bigger than 68 GB size, then htar will fail, if there are more than 1 million files per archive, htar will fail
- Not purged

HPSS



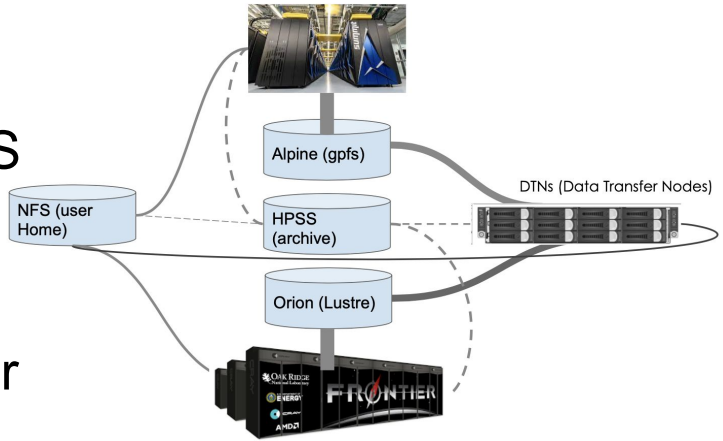
Area	Path	Type	Permissions	Quota	Backup	Purged	On Compute Nodes
Member Archive	/hpss/prod/[projid]/users/$USER	HPSS	700	100 TB	No	No	No
Project Archive	/hpss/prod/[projid]/proj-shared	HPSS	770	100 TB	No	No	No
World Archive	/hpss/prod/[projid]/world-shared	HPSS	775	100 TB	No	No	No

Data Transfer

- When Orion comes up, Frontier will not mount Alpine
- Summit will not mount Orion

There are a few ways you can move data between Alpine and Orion:

- We recommend that you use Globus and the DTNs as first choice (fastest)
- However, if you are already archiving restart files or initial data on HPSS, HPSS may be the most convenient path
- You can use the DTN or logins nodes to move small files from Alpine through User Home, but it will be slow.



Data Transfer Nodes

- The Data Transfer Nodes (DTNs) are hosts specifically designed to provide optimized data transfer between OLCF systems and systems outside of the OLCF network.
- Perform well on local-area transfers as well as the wide-area data transfers for which they are tuned.
- Access
  - ssh <username>dtn.ccs.ornl.gov
  - Globus endpoint OLCF DTN

Globus

- Globus is a fast and reliable way to move files.
- It has a convenient Web-interface at globus.org that you log into with a username and password.
- Transfers are done by activating “endpoints”
  - Endpoints are portals where data can be moved using the Globus transfer
  - Activating the OLCF Globus endpoints is done using your OLCF User name and Token Code
  - Endpoints stay activated for hours or days so you don’t need to enter your credentials for each transfer.
- Has a command-line Interface
  - https://docs.globus.org/cli/
  - https://docs.globus.org/cli/quickstart/

Globus

A few Globus Endpoints have been established for OLCF resources.

- OLCF DTN:
  - Provides access to User/Project Home areas as well as the Alpine filesystem and will provide access to the Orion filesystem
- OLCF HPSS
  - Provides access to the HPSS

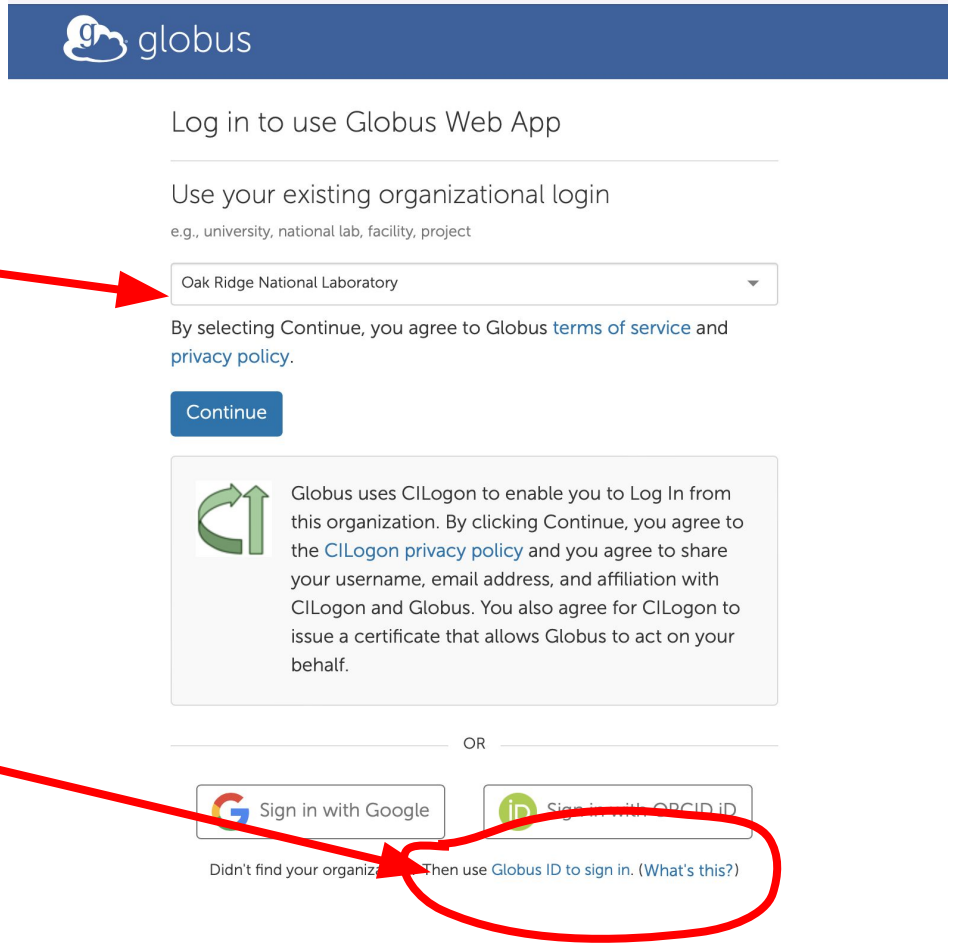
By utilizing these endpoints you can transfer data between OLCF systems and you can use them with an external endpoint to move data outside of OLCF.

Globus example

- Go to https://www.globus.org and log in

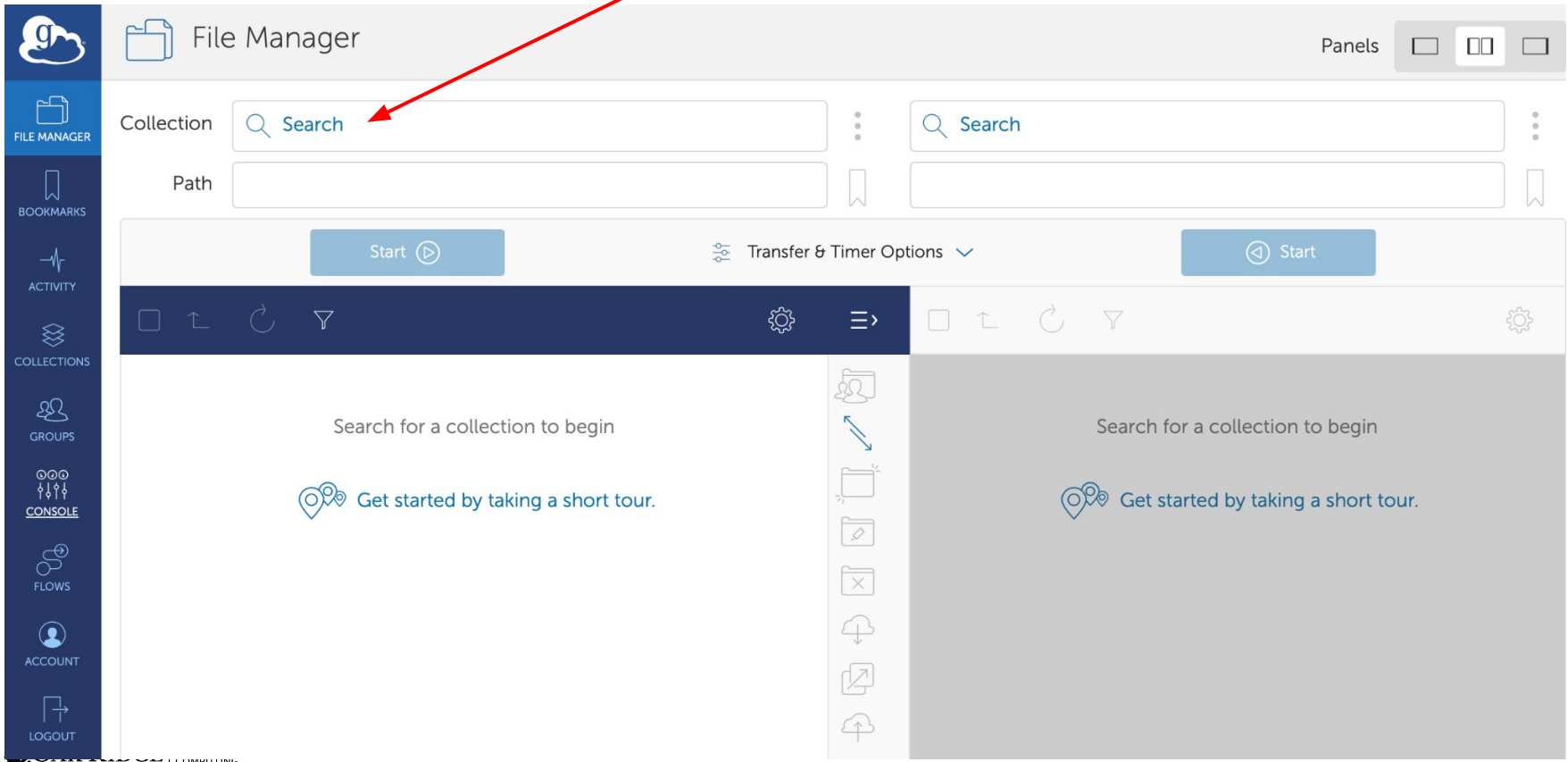
Globus example

- Select the organization that you belong to
- If you don't work for ORNL, do not select ORNL
- If your organization is not in the list, create a Globus account



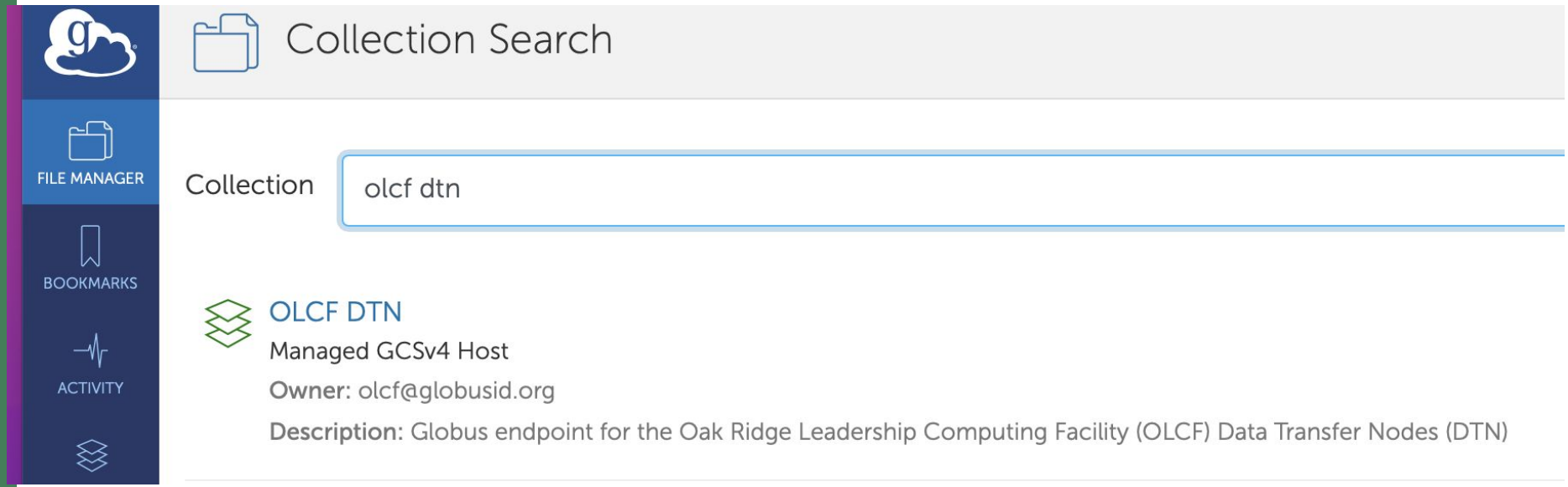
Globus example

- Search for the endpoint OLCF DTN



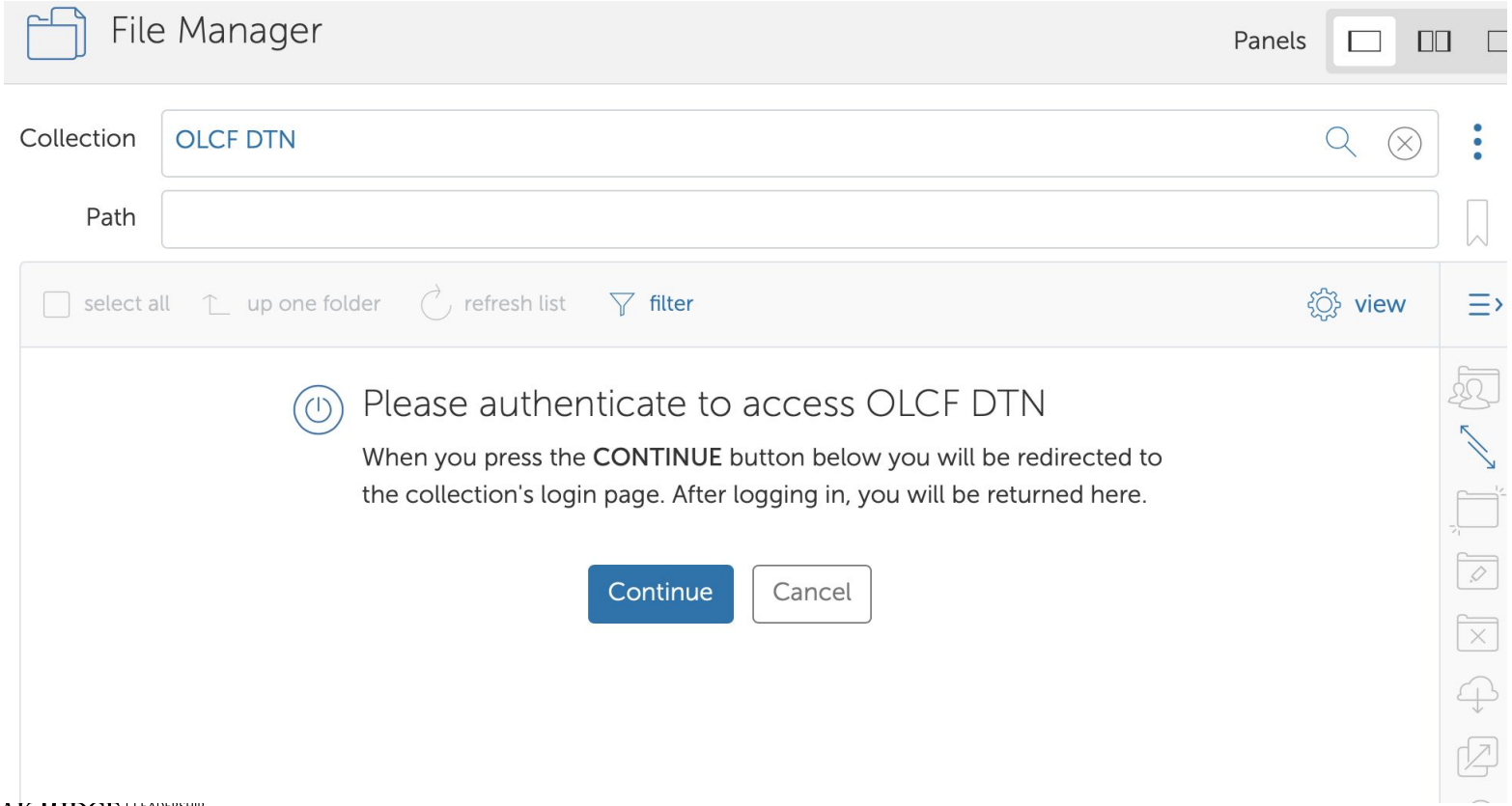
Globus example

- Search for the endpoint OLCF DTN



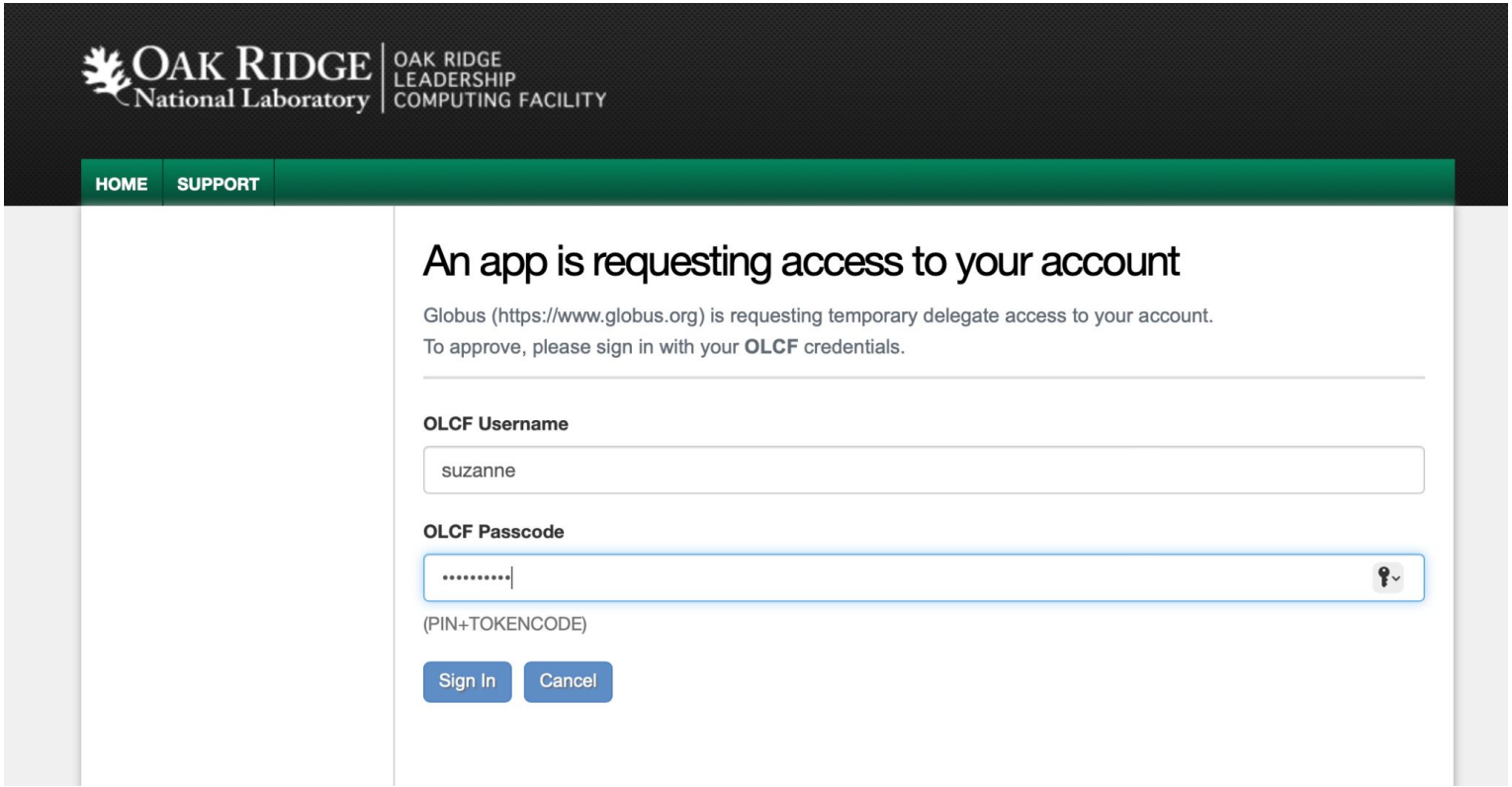
Globus example

- Activate the OLCF DTN endpoint with you OLCF credentials



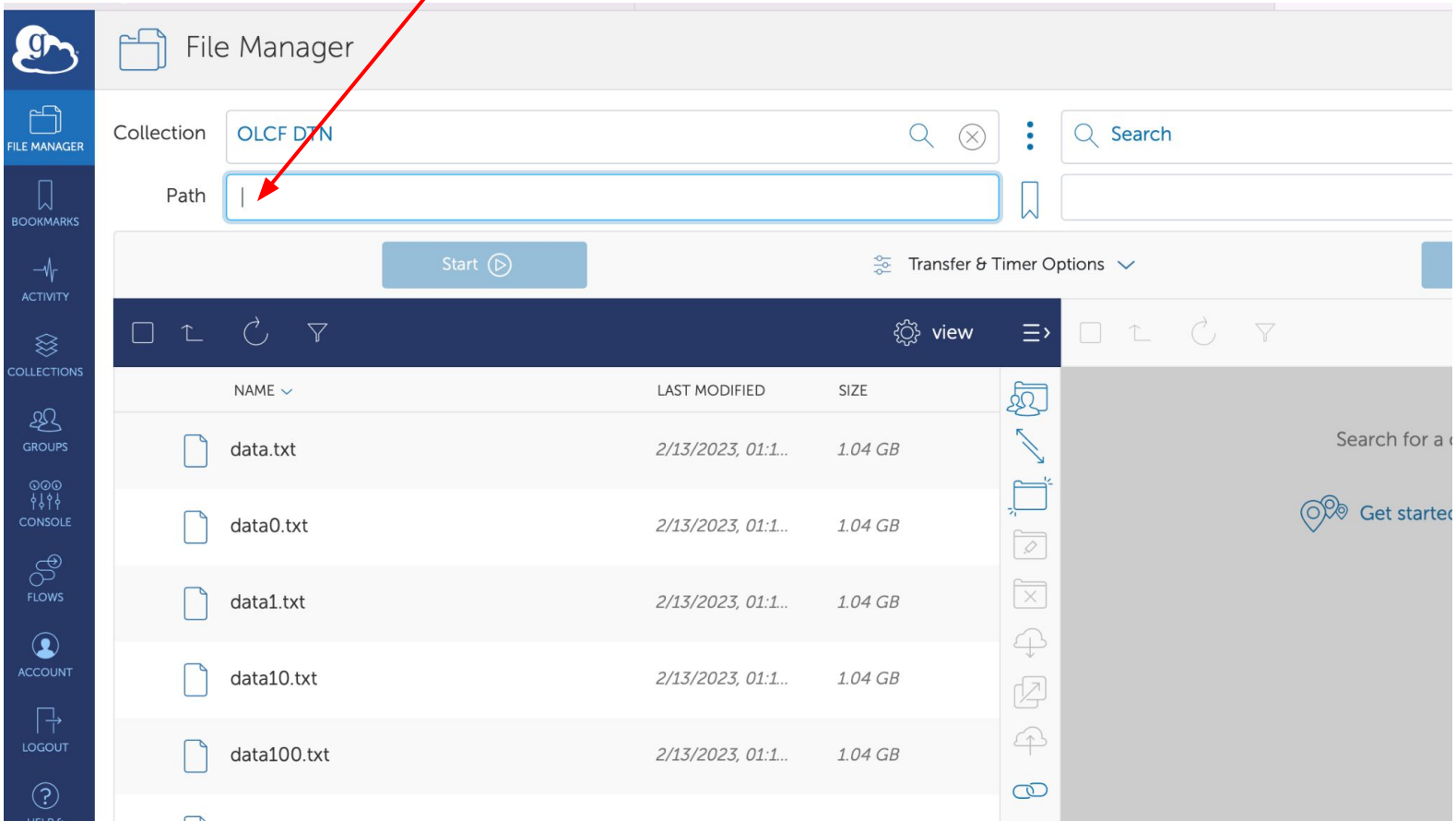
Globus example

- Activate the OLCF DTN endpoint with you OLCF credentials



Globus example

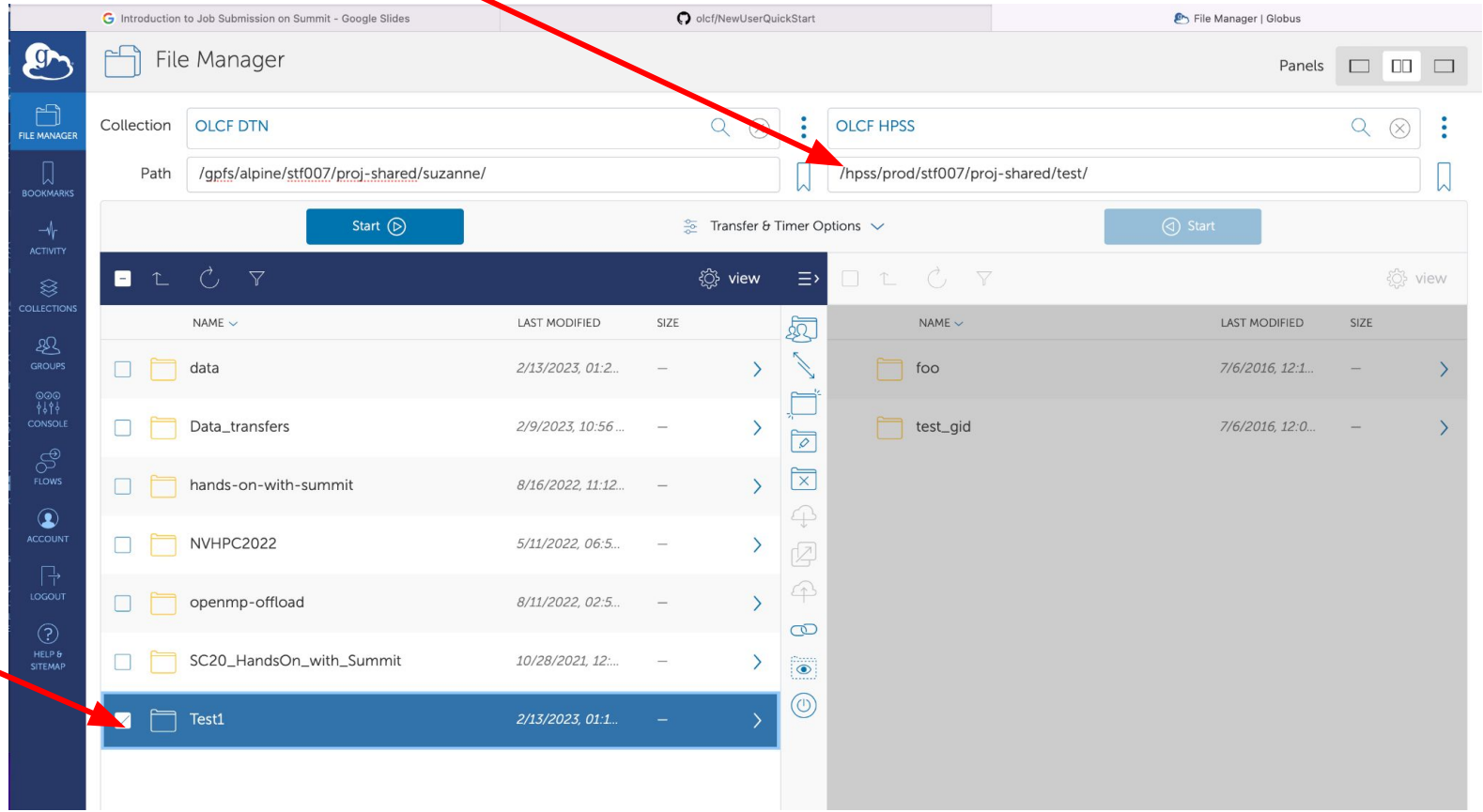
- Enter your path



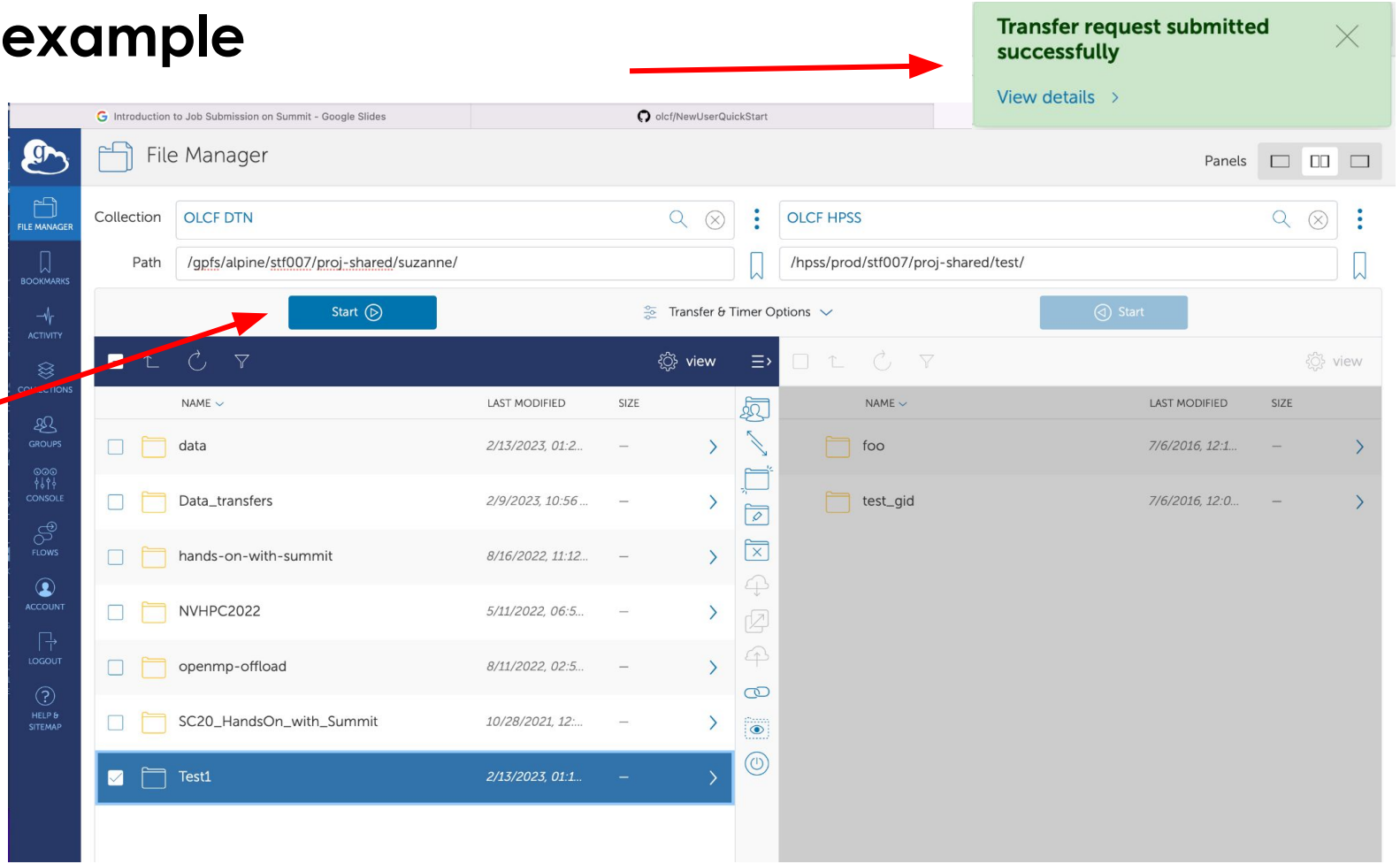
Globus example

Enter the desired path for HPSS

Select the file(s) you want to transfer.

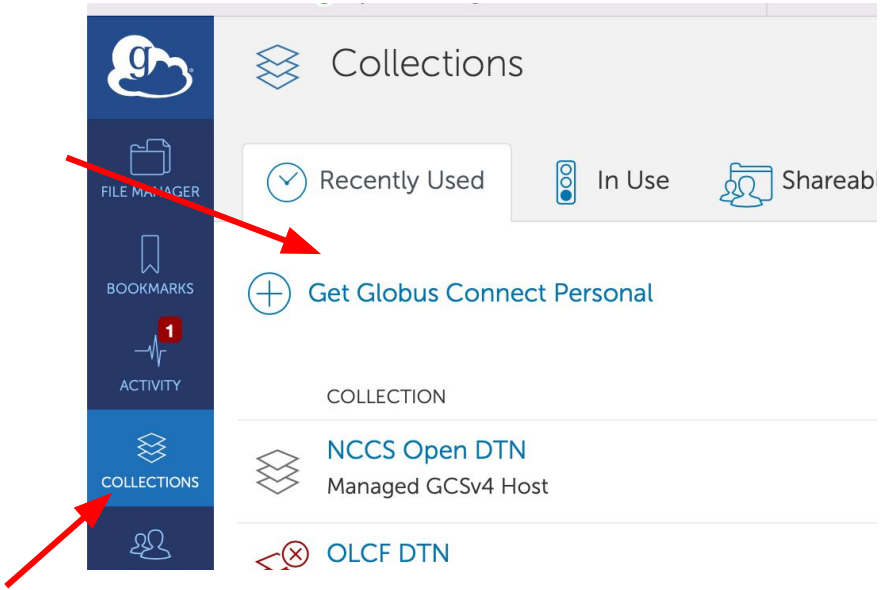


Globus example

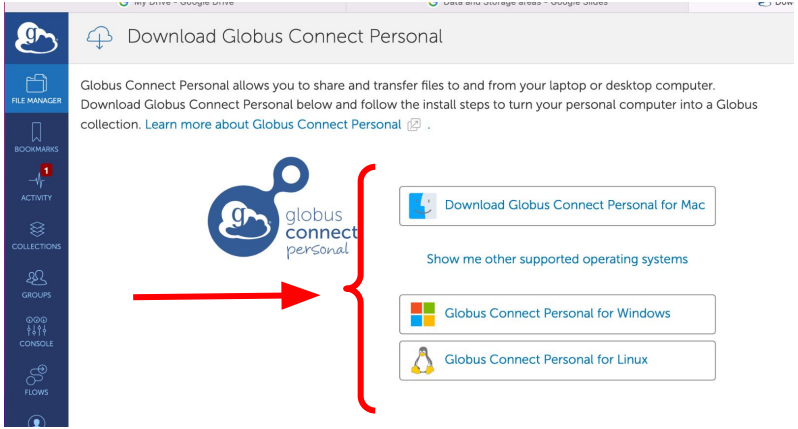


Click start

Globus endpoint for your laptop



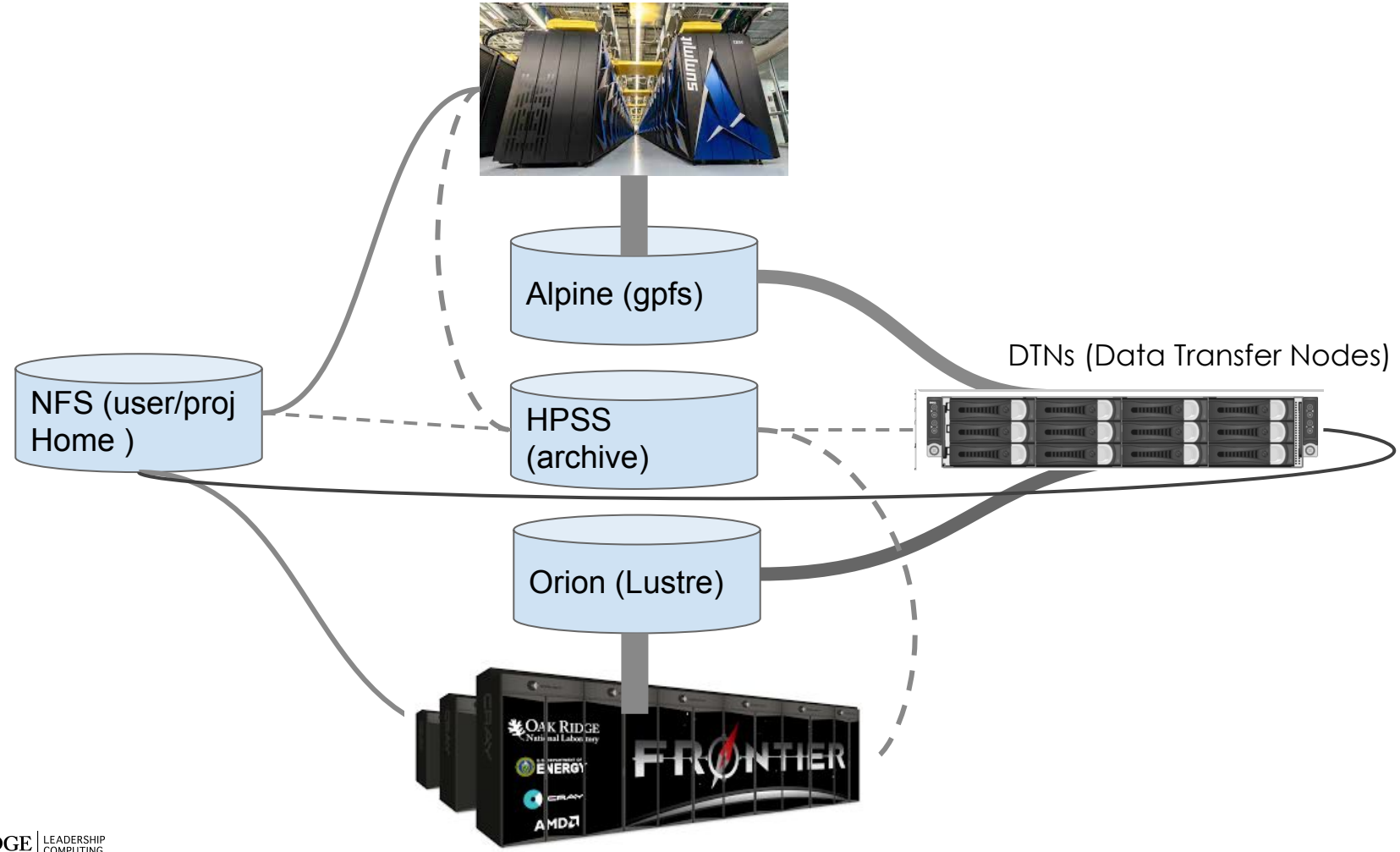
1. Go to Collections
2. Click "Get Globus Connect Personal"



3. Download the version for your machine and follow the given instructions

4. Once installed, globus must be running and your laptop must be open for the transfer to happen

5. Don't expect to see the same transfer speed to/from your laptop as you see when you use endpoint on DTNs


Alpine (gpfs)
DTNs (Data Transfer Nodes)
NFS (user/proj
Home)
HPSS
(archive)
Orion (Lustre)
OAK RIDGE
National Laboratory
ENERGY
AMD
FRONTIER

HPSS: htar example https://docs.olcf.ornl.gov/data/index.html#htar

To move data from Summit/Alpine to the project shared area of HPSS:

```
Summit> htar -cvf /hpss/prod/stf007/proj-shared/Test1.tar Test1

creating HPSS Archive file /hpss/prod/stf007/proj-shared/Test1.tar
HTAR: a  Test1/
. . .
HTAR: a   /tmp/HTAR_CF_CHK_4042346_1676312676
HTAR Create complete for /hpss/prod/stf007/proj-shared/Test1.tar. 10485767168 bytes
written for 10 member files, max threads: 3 Transfer time: 15.901 seconds (659.440
MB/s) wallclock/user/sys: 16.198 6.593 7.431 seconds
HTAR: HTAR SUCCESSFUL
```

To move data from HPSS to Frontier/Orion

```
Frontier> htar -xvf /hpss/prod/stf007/proj-shared/Test1.tar Test1
.  .  .
wallclock/user/sys: 25.243 0.368 4.898 seconds
```

Data Management

OLCF systems generate lots of data very quickly; projects should develop a data strategy *as soon as possible*. (It's easier to fix things with 100 files than with 100,000!)

Some things to consider:

- How are files/directories shared among project members?
  - Where will project members store data?
  - What file attributes (permissions, group, etc.) are needed?
- What happens when someone leaves the project?
- What happens when the project ends?
  - Where does the data need to go?
  - How much data is there, who's moving it, and how long will it take?

Other Command line tools for transfers

SCP and rsync - please use the DTN
https://docs.olcf.ornl.gov/data/index.html#command-line-terminal-tools

HSI another HPSS tools
https://docs.olcf.ornl.gov/data/index.html#hsi

Questions

Thank you for your attention!