File Systems/ Data Transfers

George S. Markomanolis

HPC Engineer

Oak Ridge National Laboratory

Introduction to Summit Workshop

20 May 2019

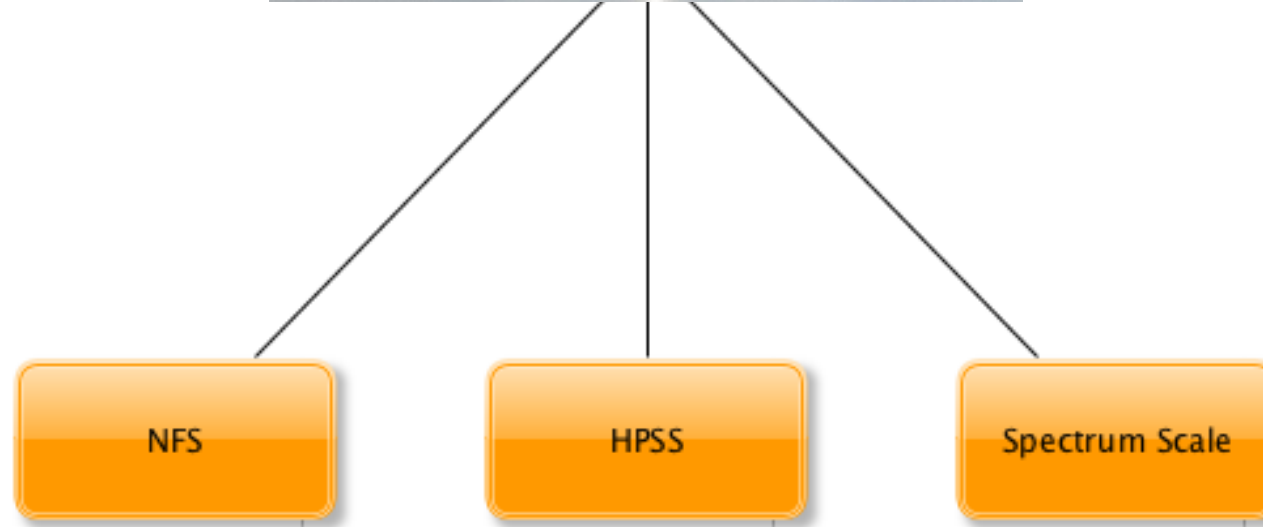OAK RIDGE
National Laboratory

U.S. DEPARTMENT OF
ENERGY

Outline

- **File Systems**
  - Available file systems and options for archiving

- **Data Transfer**
  - Transfer your files between Titan and Summit

OAK RIDGE
National Laboratory

Summit and filesystems

NFS

HPSS

Spectrum Scale

OAK RIDGE
National Laboratory

Open slide master to edit

NFS

- User home: /ccs/home/$USER

- Project home: /ccs/proj/[projid]

- **Long-term** storage for your general data under home or related to project under proj

- **Build** your code in /tmp/$USER it is faster and **install** in /ccs/proj/[projid]

- There is provided a **backup**

- User home and project home are accessible read-only from the Summit compute nodes

- **Not purged**

- **Quota** of 50GB

- User home is user-centric

Open slide master to edit

NFS (cont.)

- Check quota on user home

> quota -Qs

Disk quotas for user gmarkoma (uid 14850):

Filesystem blocks quota limit grace files quota limit grace

nccs-svm1.lb.ccs.ornl.gov:/nccs/home2

3237M 51200M 51200M   49161 4295m 4295m

OAK RIDGE
National Laboratory

NFS (continue)

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

- Answer: snapshots
  - Go to the .snapshot folder (ls will not show this folder):
  - cd .snapshot

ls -l

drwx------ 27 gmarkoma gmarkoma 4096 Nov 21 16:51 daily.2018-11-23_0010

drwx------ 27 gmarkoma gmarkoma 4096 Nov 21 16:51 daily.2018-11-24_0010

…

OAK RIDGE
National Laboratory

HPSS

- User archive: /home/$USER

- Project archive: /proj/[projid]

- **Long-term** storage for large amount of general data under home or related to project under proj.

- **Quota** of 2 TB and 100 TB for user and project archive respectively. If any of the used files during htar is bigger than 68 GB size, then it will fail, similar if there are more than 1 million files per archive

- **Not purged**

- User archive is user-centric

OAK RIDGE
National Laboratory

HPSS (cont.)

- Check HPSS quota:

```
> showusage -s hpss
HPSS Storage in GB:
                    Project Totals
   Project              Storage                Storage
_____|_____|_____
   stf007       |       46868.90         |       0.00
```

OAK RIDGE
National Laboratory
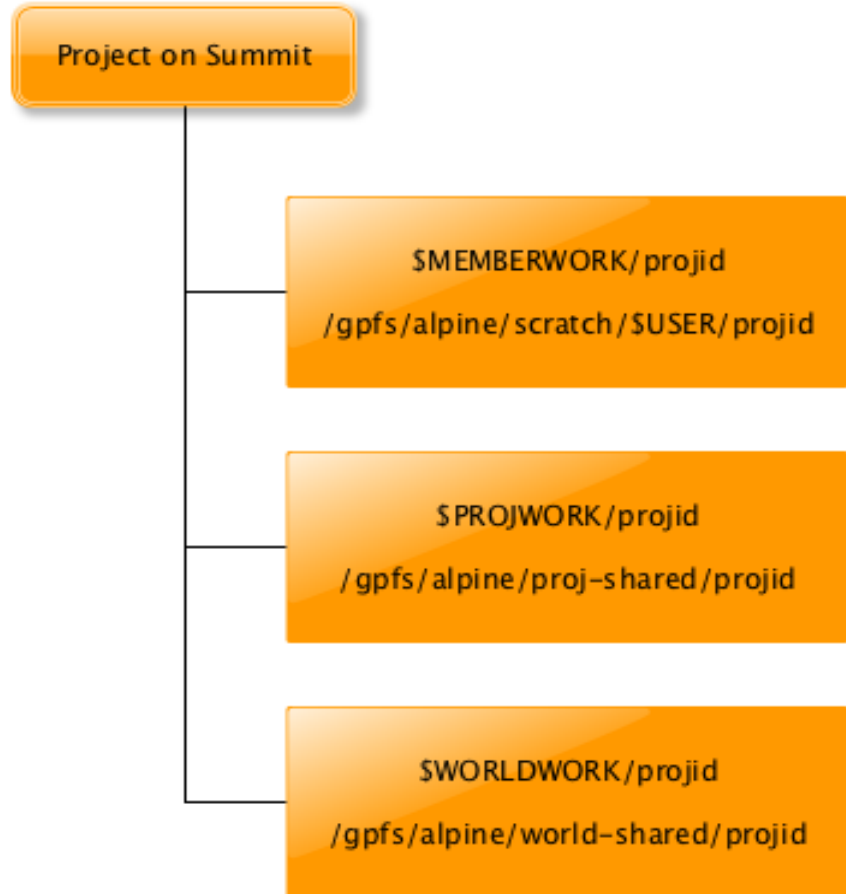
Spider III - Alpine

- Alpine, is a Spectrum Scale (ex-GPFS) file system of 250 PB of used space, which is mounted on Summit and Data Transfer Nodes (DTN) with maximum performance of 2.5 TB/s for sequential I/O and 2.2 TB/s for random I/O

- Largest GPFS file system installation

- Up to 2.6 million accesses per second of 32 KB small files

- It is constituted by 154 Network Shared Disk (NSD) servers

- It is a shared resource among users, supporting File Per Process (FPP), Single Shared File (SSF) and any of their combination

OAK RIDGE
National Laboratory

Alpine (cont.)

Project on Summit

$MEMBERWORK/projid

/gpfs/alpine/scratch/$USER/projid

$PROJWORK/projid

/gpfs/alpine/proj-shared/projid

$WORLDWORK/projid

/gpfs/alpine/world-shared/projid

- 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

- **No backup**
- **Quota 50 TB**
- **Purged after 90 days**

OAK RIDGE
National Laboratory

Open slide master to edit

Storage policy

Name	Path	Type	Permissions	Backups	Purged	Quota	Mounted on Compute nodes
User Home	$HOME	NFS	User Set	yes	no	50GB	Read-only
User Archive	/home/$USER	HPSS	User Set	no	no	2TB	No
Project Home	/ccs/proj/[projid]	NFS	770	yes	no	50GB	Read-only
Member Work	/gpfs/alpine/scratch/[userid]/[projid]/	Spectrum Scale	700	no	90 days	50TB	Yes
Project Work	/gpfs/alpine/proj-shared/[projid]	Spectrum Scale	770	no	90 days	50TB	Yes
World Work	/gpfs/alpine/world-shared/[projid]	Spectrum Scale	775	no	90 days	50TB	Yes
Project Archive	/proj/[projid]	HPSS	770	no	no	100TB	No

OAK RIDGE
National Laboratory
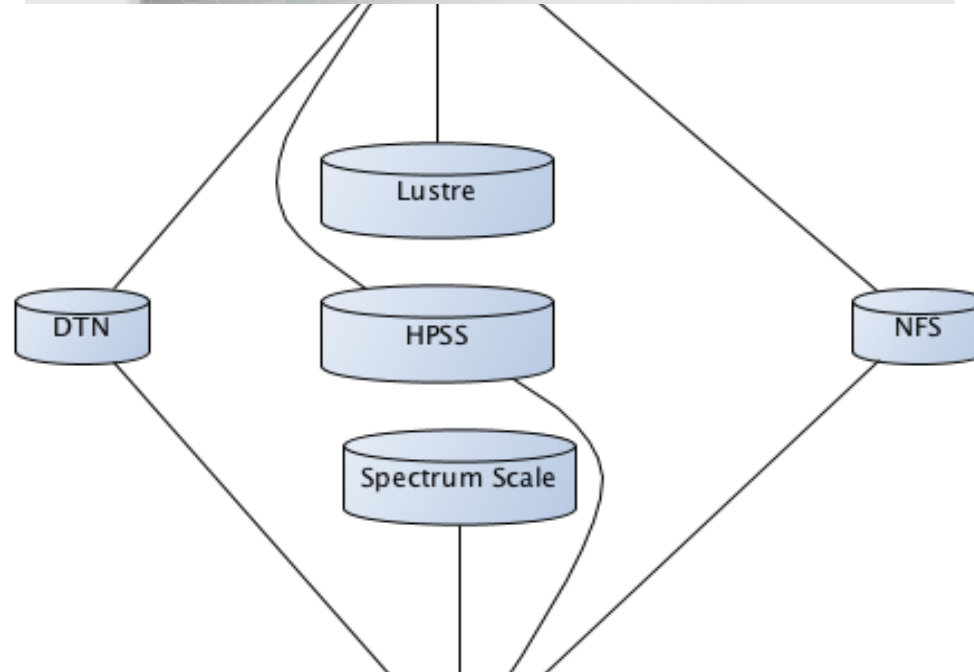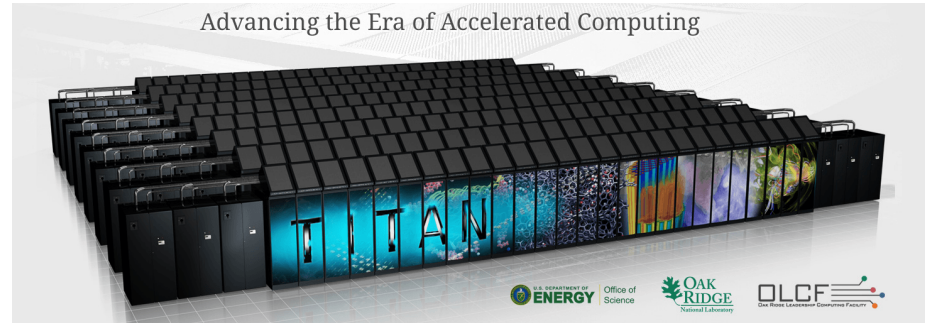
Data Transfer

Data Transfer Nodes (DTN) improve the performance by reducing the load on the login and service nodes of the HPC facilities. Moreover, transfer data outside the HPC facility.

OAK RIDGE
National Laboratory

Data Transfer (cont.)

- When you log-in to Summit you would like to have access to your old files (if you are already user of OLCF HPC facilities)

- There are many ways to transfer files but in many cases we propose Globus

- We will mention all the approaches and some performance results.

OAK RIDGE
National Laboratory

Data Transfer (cont.)

- Using home NFS

- If the data size is less than 50 GB and there is enough free space in your home directory is through home.

> titan> cp –r data $HOME
>
> summit> cp –r $HOME/data .

- It is simple, but is it fast?

OAK RIDGE
National Laboratory

Data Transfer (cont.)

- Using HPSS

- Send one folder to HPSS and retrieve it from the destination. There is significant higher data size limit

```
titan> htar -cvf transfer_test.tar transfer_test/*

HTAR: a   transfer_test/data0.txt
HTAR: a   transfer_test/data10.txt
…
HTAR: a   /tmp/HTAR_CF_CHK_8183_1543522594
HTAR Create complete for transfer_test.tar. 23,068,684,800 bytes
written for 22 member files, max threads: 3 Transfer time: 186.324
seconds (123.809 MB/s) wallclock/user/sys: 186.521 30.654 105.275
seconds

HTAR: HTAR SUCCESSFUL

summit> htar -xvf transfer_test.tar
```
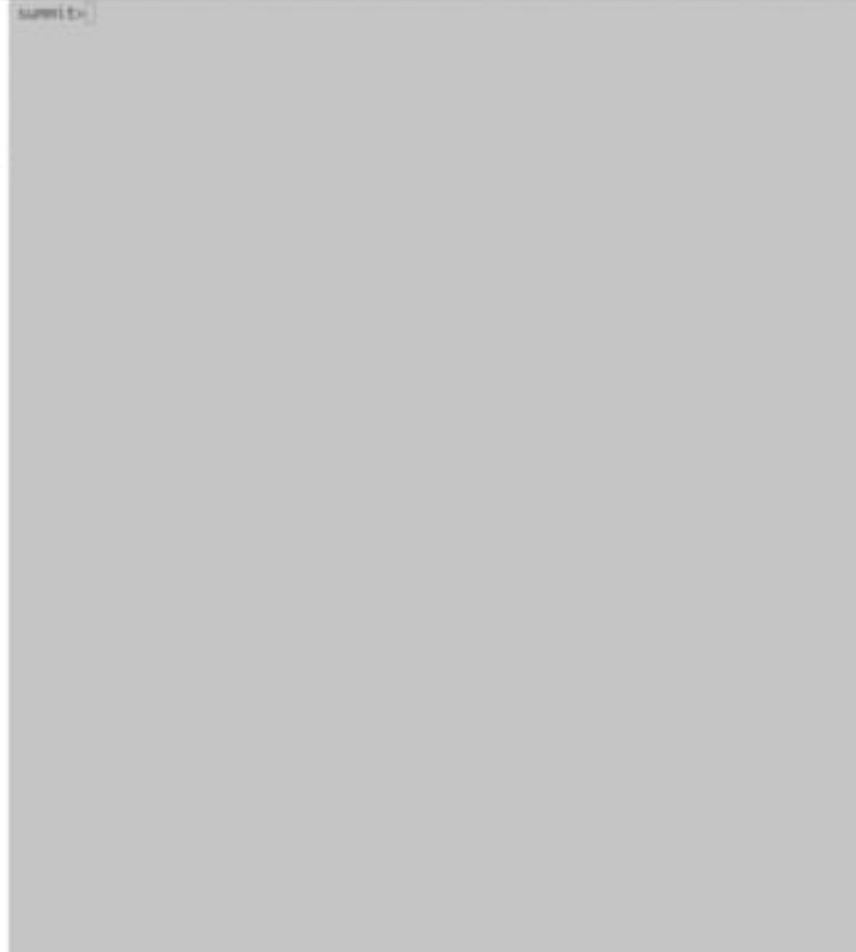
OAK RIDGE
National Laboratory

Open slide master to edit

Transferring files through NFS and HPSS

```
titan>ls -l
total 1024016
-rw-r--r-- 1 gmarkoma gmarkoma 1048576000 Dec  1 09:52 data.txt
titan>htar -cvf small_transfer.tar data.txt
HTAR: a   data.txt
HTAR: a   /tmp/HTAR_CF_CHK_52054_1543676913
HTAR Create complete for small_transfer.tar. 1,048,578,048 bytes written for 1 member
 files, max threads: 3 Transfer time: 8.469 seconds (123.811 MB/s) wallclock/user/sys
: 8.562 1.308 4.796 seconds
HTAR: HTAR SUCCESSFUL
titan>
```

```
summit>
```

https://vimeo.com/304829936

OAK RIDGE
National Laboratory

Open slide master to edit

Globus

- Globus transfers fast, parallel and reliable files between two endpoints

- Endpoints are different locations where data can be moved using the Globus transfer

- Visit www.globus.org to register and/or login

- You can find the **OLCF DTN** endpoint.

- If you want to save data on HPSS, you could use the **OLCF HPSS** endpoint.

OAK RIDGE
National Laboratory

Globus(cont.)

Globus demo, transfer from Titan to Summit

https://vimeo.com/304829453

OAK RIDGE
National Laboratory

Performance Results

- Study case: Transfer data from Atlas to Alpine with 3 approaches. Copy the files through NFS, use HPSS, or use Globus

Type	Home NFS*	HPSS*	Globus*	HPSS**	HPSS** from DTN
	Time in seconds to finish the transfer				
Transfer 22 files of 1GB each	323	270	10	227	20
Transfer 1 file of 22 GB	308	301	80	345	26
Transfer 4 files of 1GB each	69	53	9	39.2	4.6

- Globus is the most efficient approach to transfer files when you have many files, however, HPSS tools from DTN, can be more efficient in some cases.   There are available some traditional tools such as scp, rsync

* 29 November 2018   ** 16 May 2019

OAK RIDGE
National Laboratory

Open slide master to edit

DTN

- As long as we have both Atlas and Alpine on DTN, we use the following variables (GPFS variables are not active yet)

OAK RIDGE
National Laboratory

Open slide master to edit
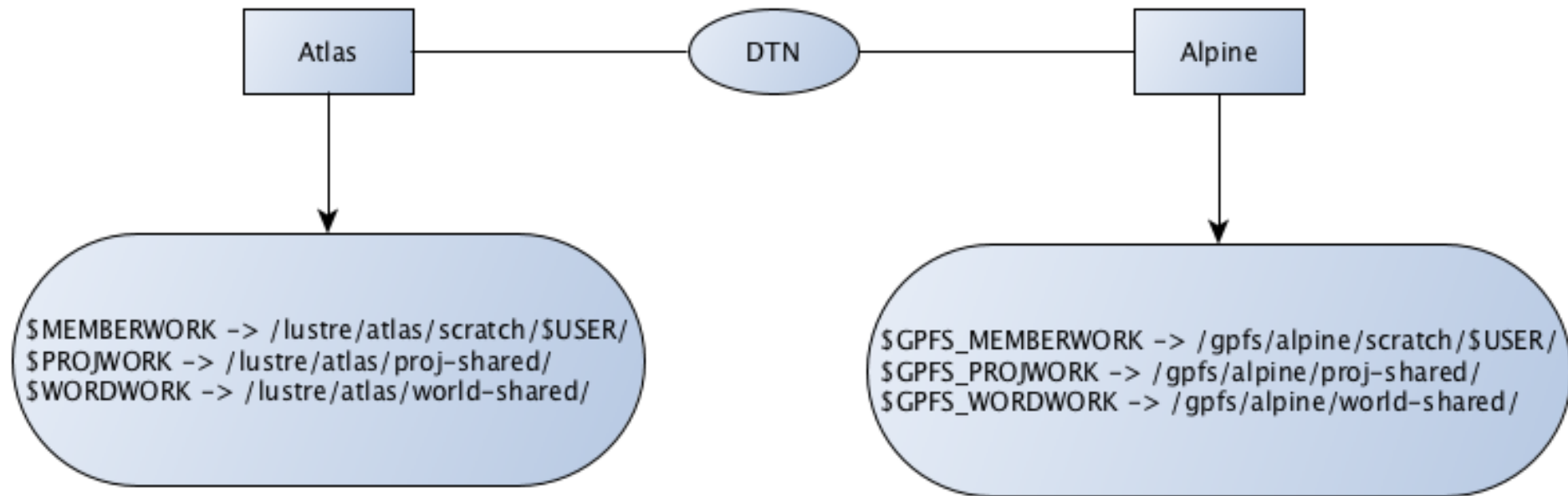
Conclusions – File systems/Data transfer

- Use NFS for installing your libraries (long-term storage)

- There are many approaches to transfer files, some tools are more efficient depending on the number of files and the file size.

- Use HPSS for large files that you don't plan to use soon and to backup soon to expire projects with important data

- Start transferring your files to Summit as soon as you have access. Atlas filesystem will be decommissioned mid of August 2019!

- Do not forget the storage policy!

OAK RIDGE
National Laboratory

Acknowledgement

This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

OAK RIDGE
National Laboratory

Open slide master to edit

Thank you!

Questions?

OAK RIDGE
National Laboratory