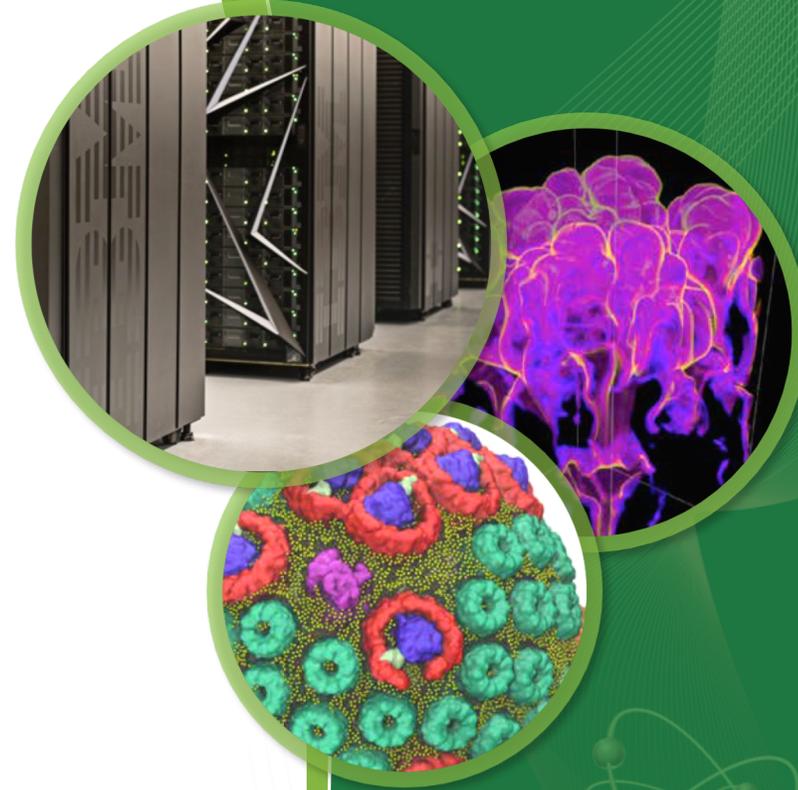


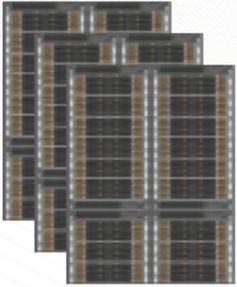
File Systems & Data Transfers Overview

Tom Papatheodore
Oak Ridge Leadership Computing Facility
Oak Ridge National Laboratory

June 1, 2018



File Systems & Data Storage



GPFS File System

250 PB storage

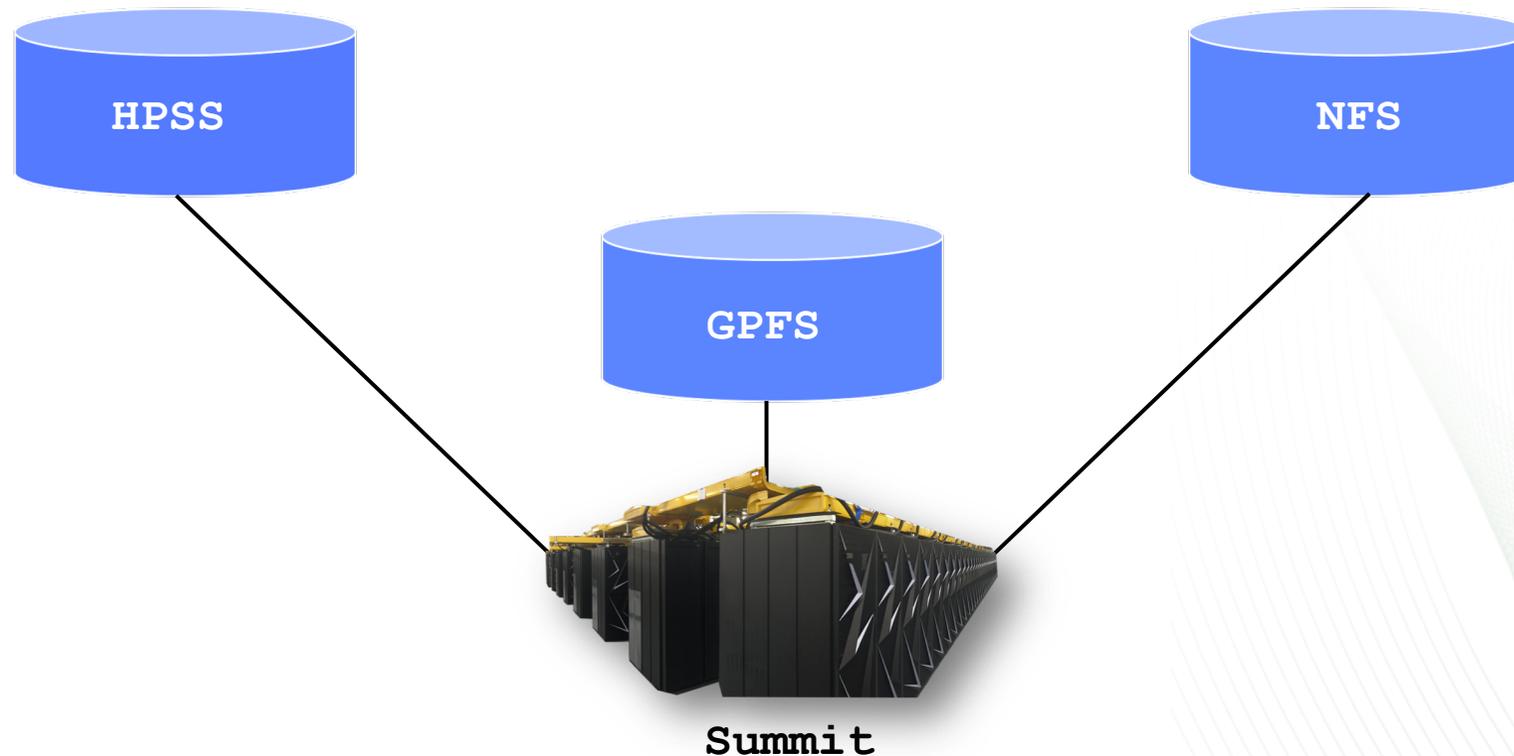
2.5 TB/s read, 2.5 TB/s write

(**2.5 TB/s sequential and 2.2 TB/s random I/O)

- When Summit is accepted, it will mount a 250 PB GPFS file system, Alpine, for production I/O.
- Currently, Summit mounts a 3 PB GPFS file system, called AlpineTDS, for testing and development.
 - Because AlpineTDS is still being tested/developed, it is important to move data off this file system frequently to ensure your data is not lost.
 - It is also important to regularly move data off AlpineTDS because the (relatively) small amount of storage is shared among all early Summit Users.

File Systems & Data Storage

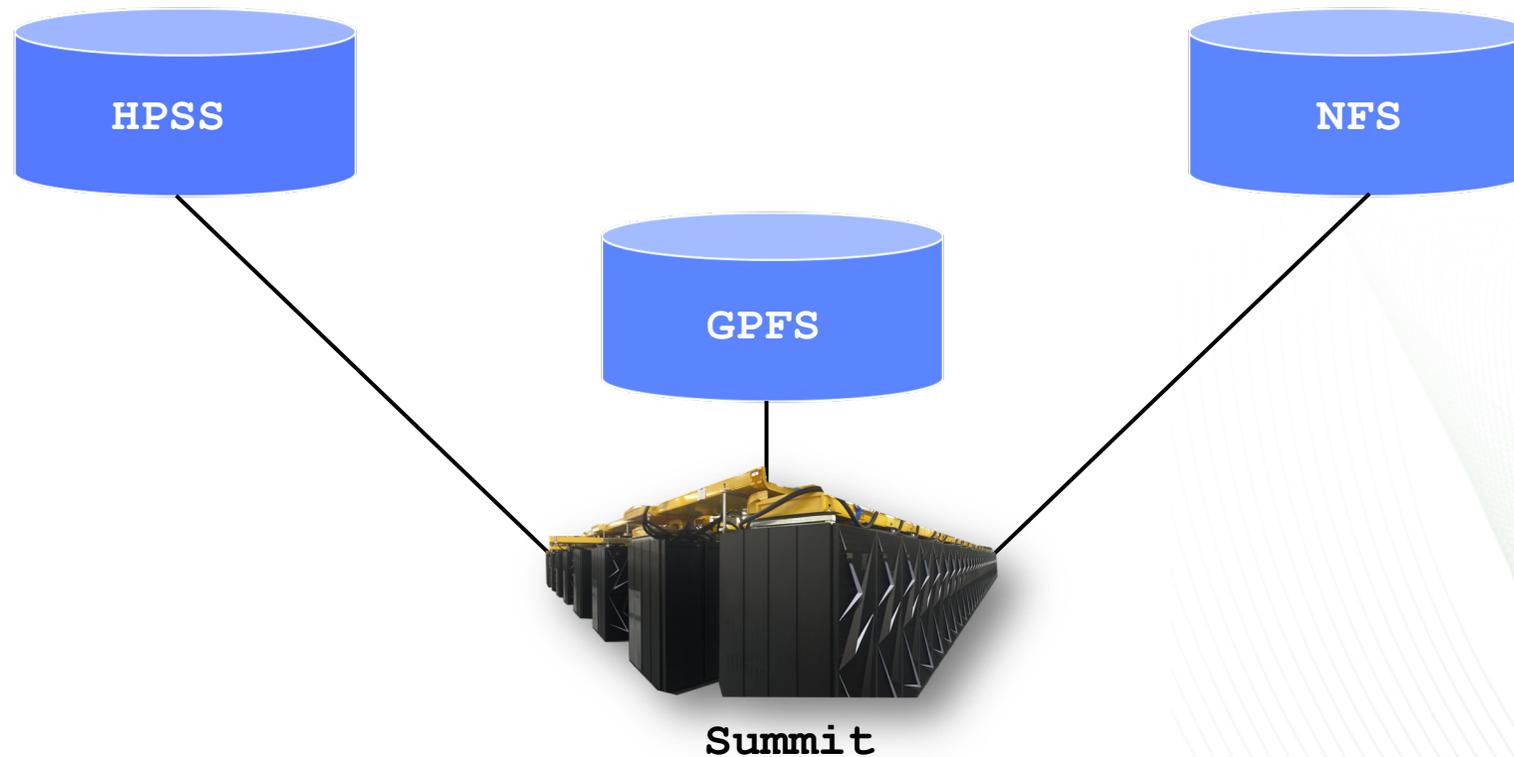
Summit also mounts OLCF's Network File System (NFS) and High Performance Storage System (HPSS).



File Systems & Data Storage

NFS (`/ccs/home/` and `/ccs/proj/`)

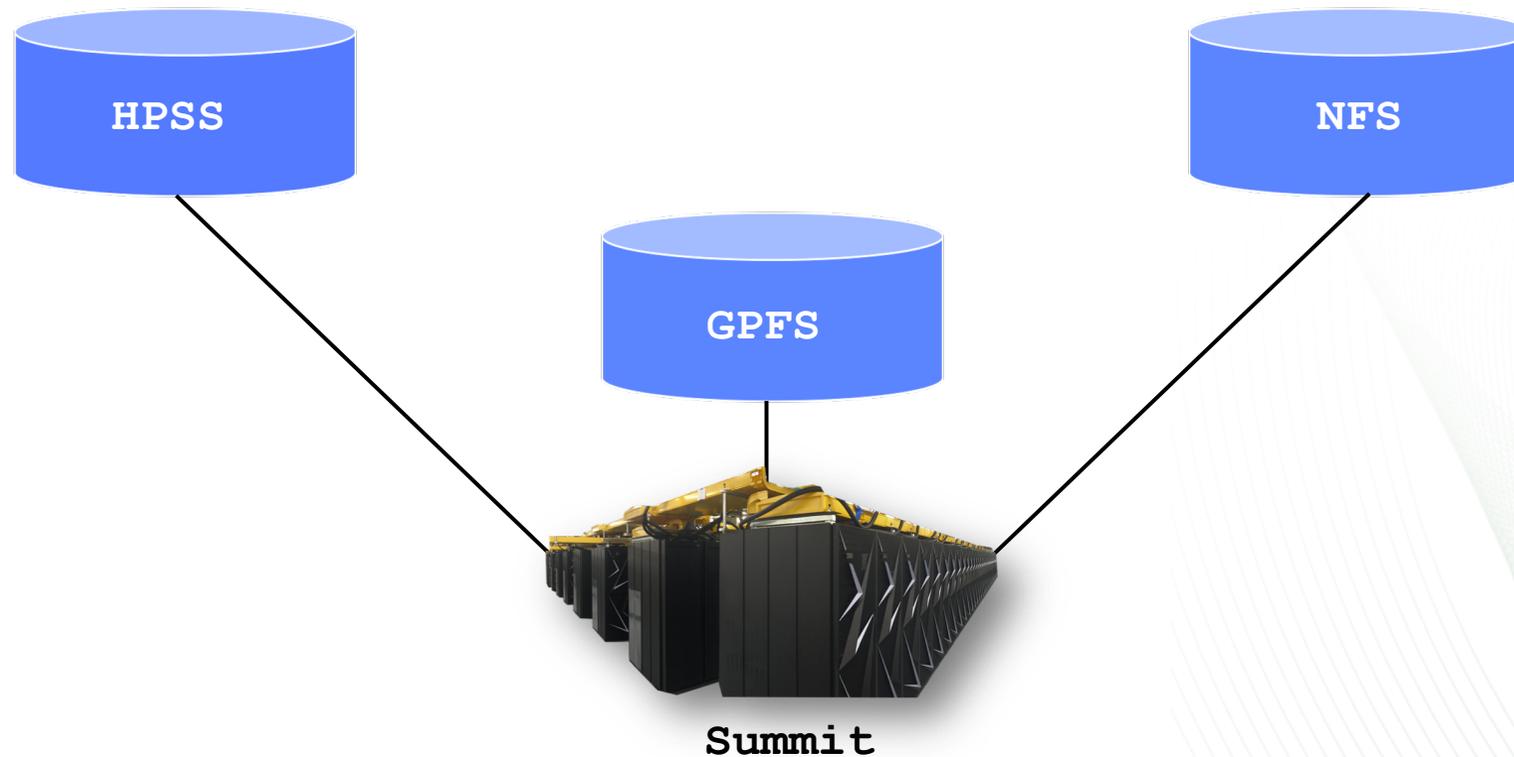
Provides long-term storage for routinely accessed user or project data, respectively.



File Systems & Data Storage

GPFS (/gpfs/alpinetds)

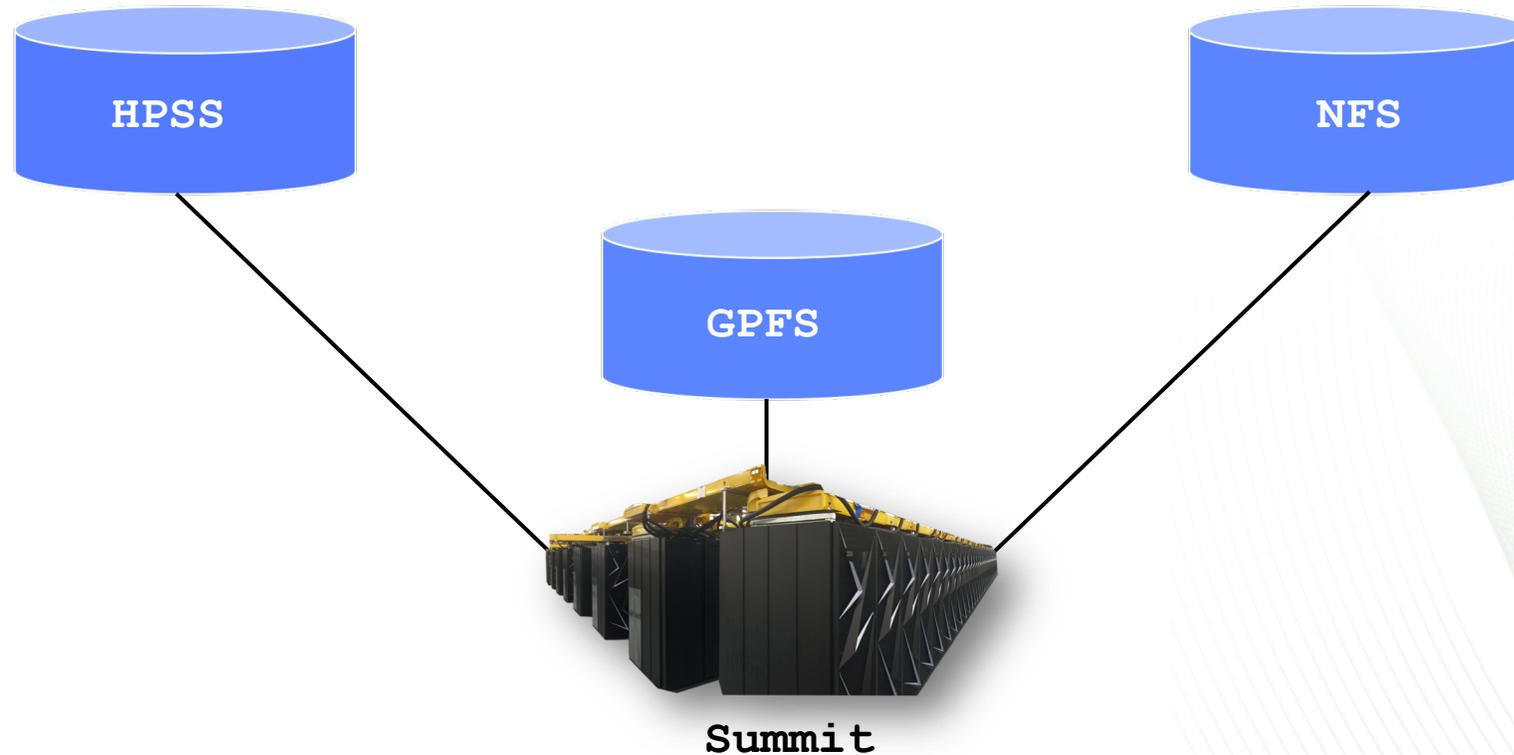
Provides temporary storage for data created on Summit.



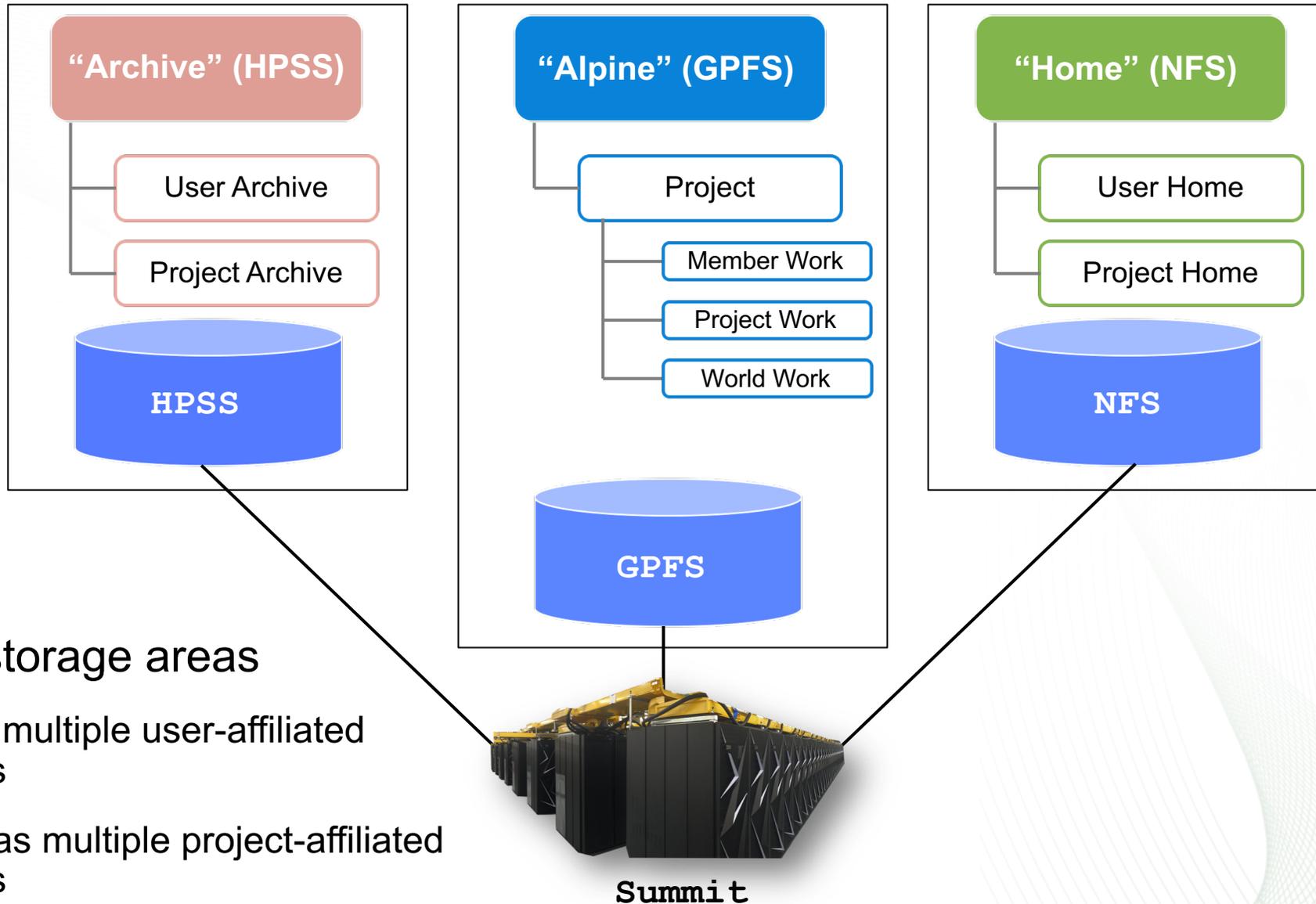
File Systems & Data Storage

HPSS (/home/ and /proj/)

Provides longer-term/archive storage for large amounts of data created on the OLCF compute systems.



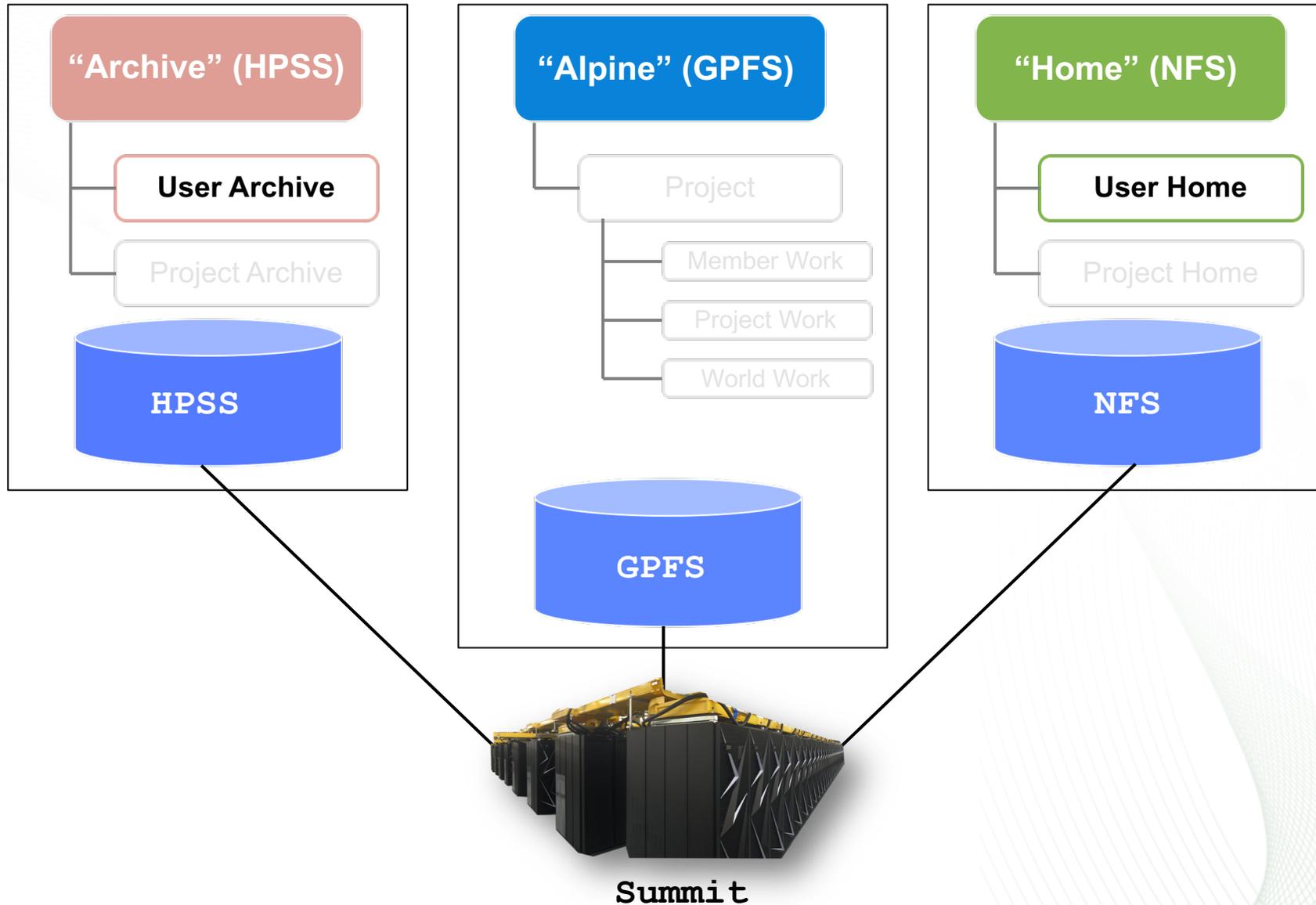
File Systems & Data Storage



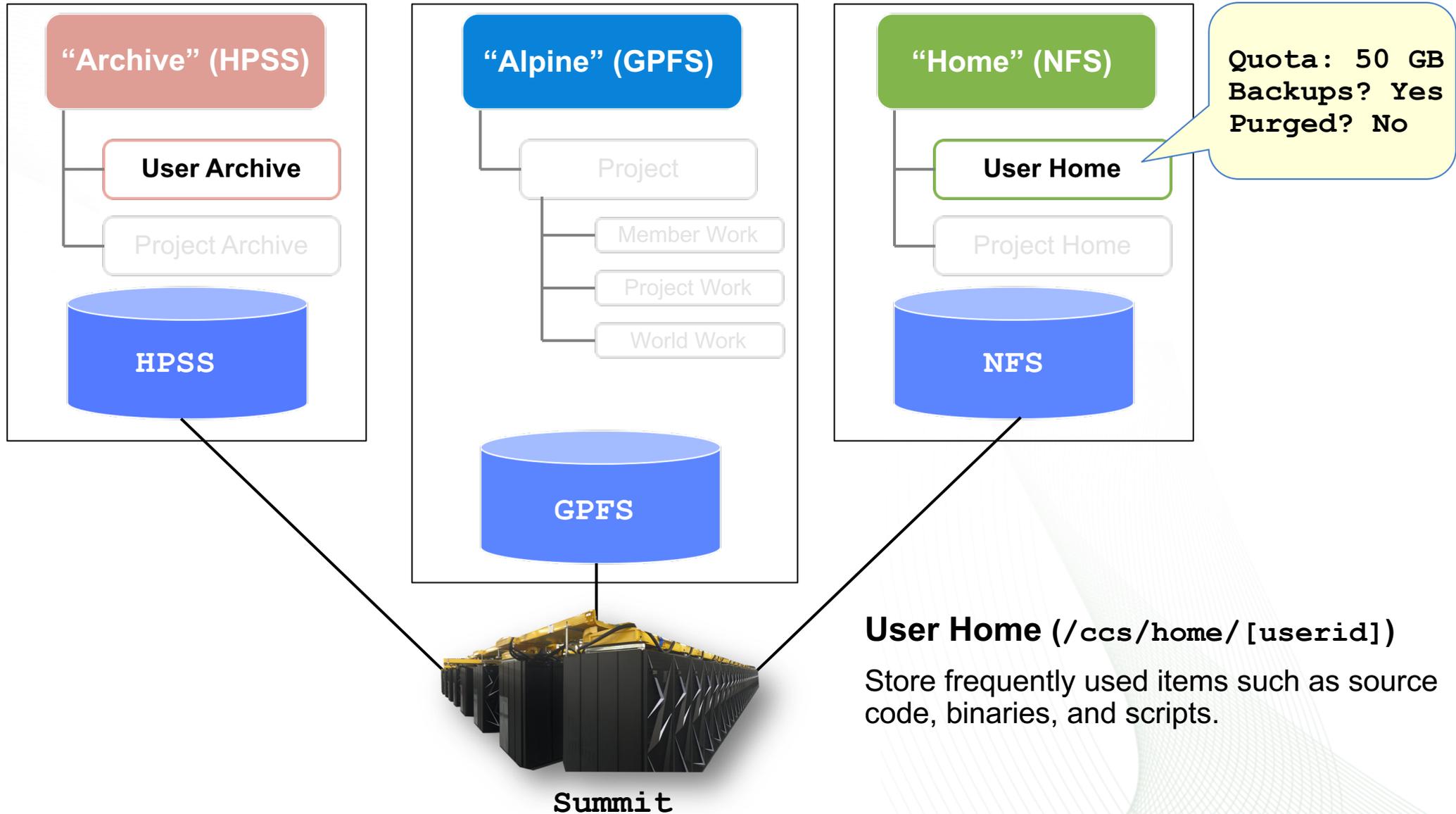
Within these storage areas

- Each user has multiple user-affiliated storage spaces
- Each project has multiple project-affiliated storage spaces

File Systems & Data Storage: User-Centric

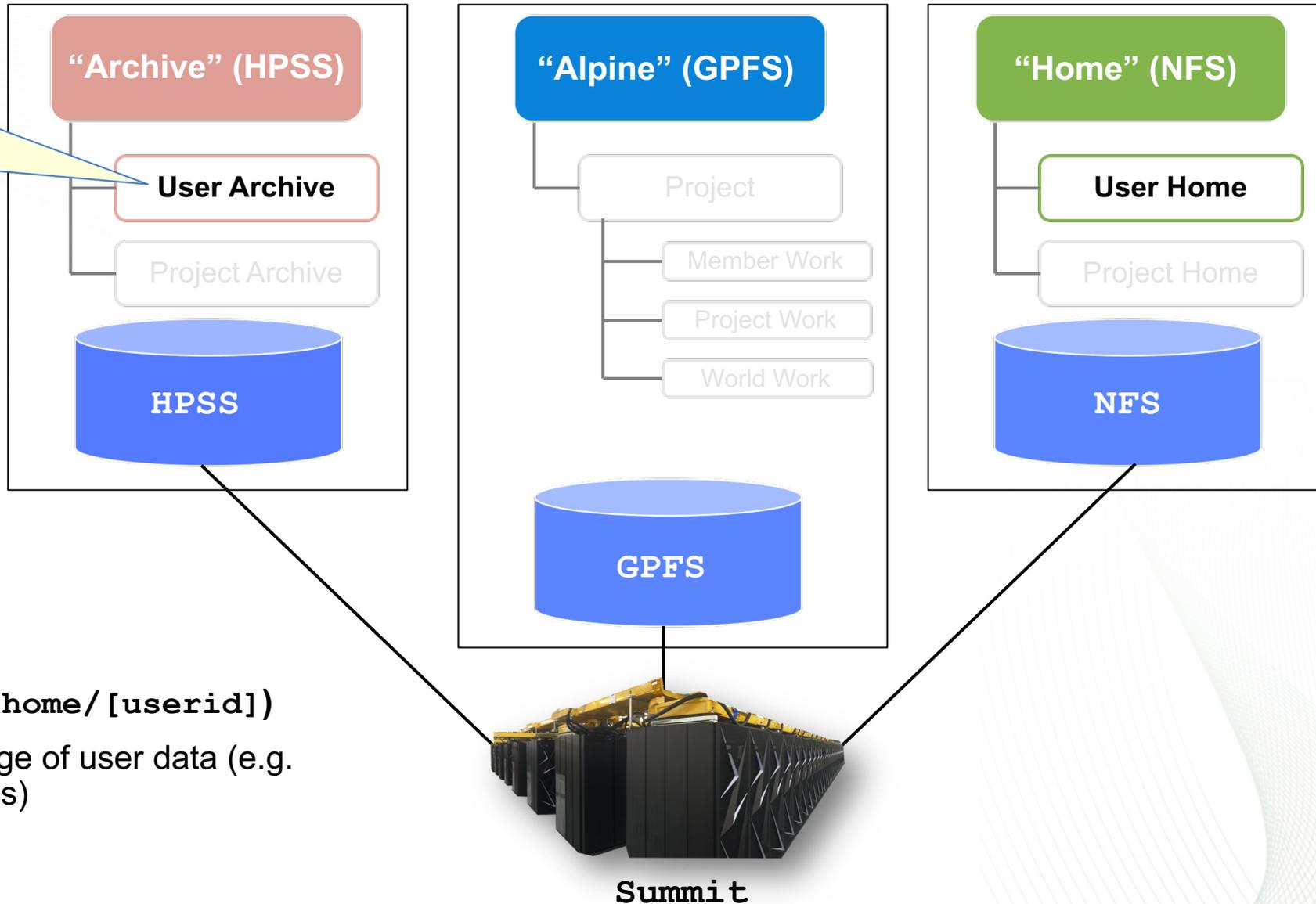


File Systems & Data Storage: User-Centric



File Systems & Data Storage: User-Centric

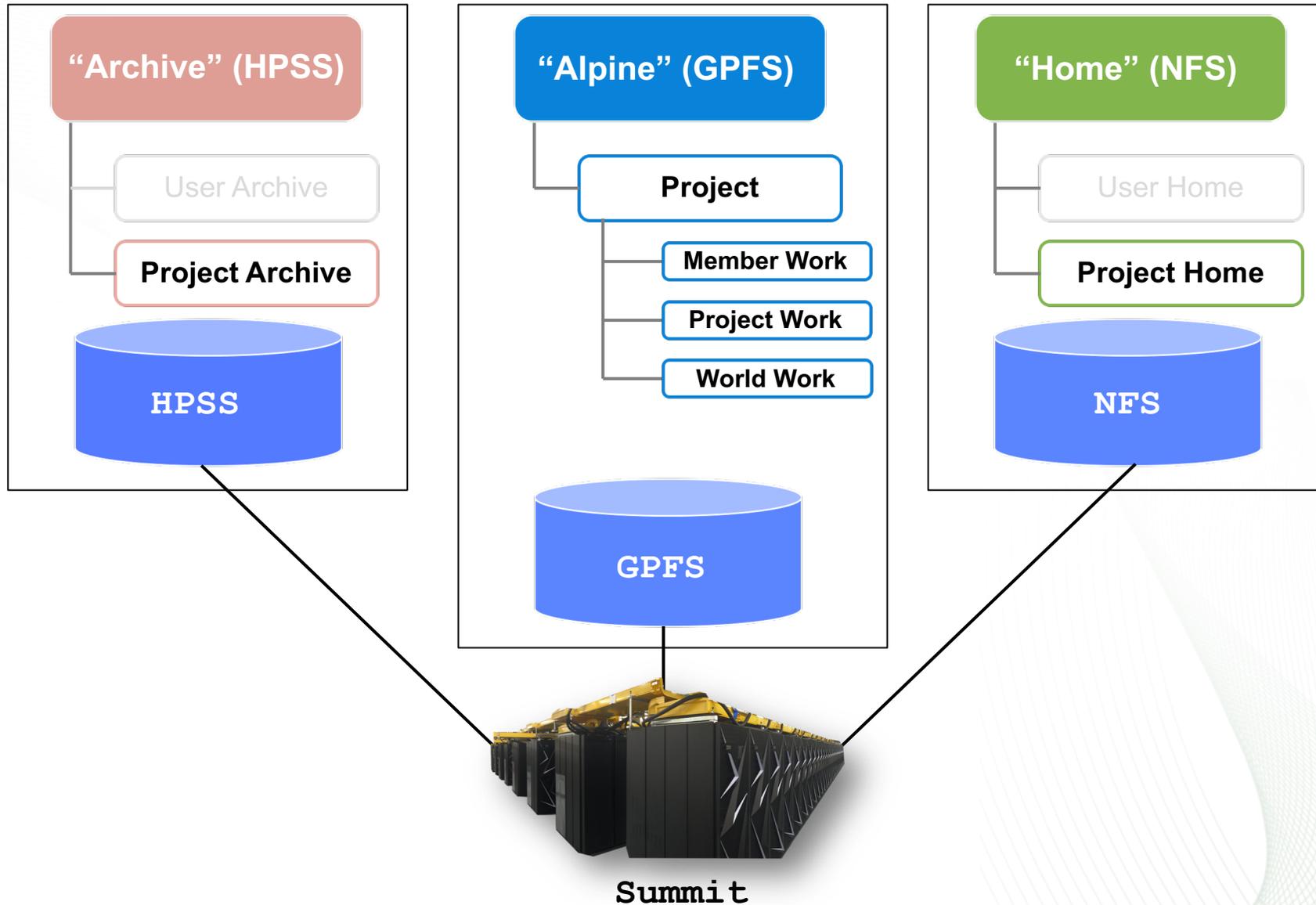
Quota: 2 TB
Backups? No
Purged? No



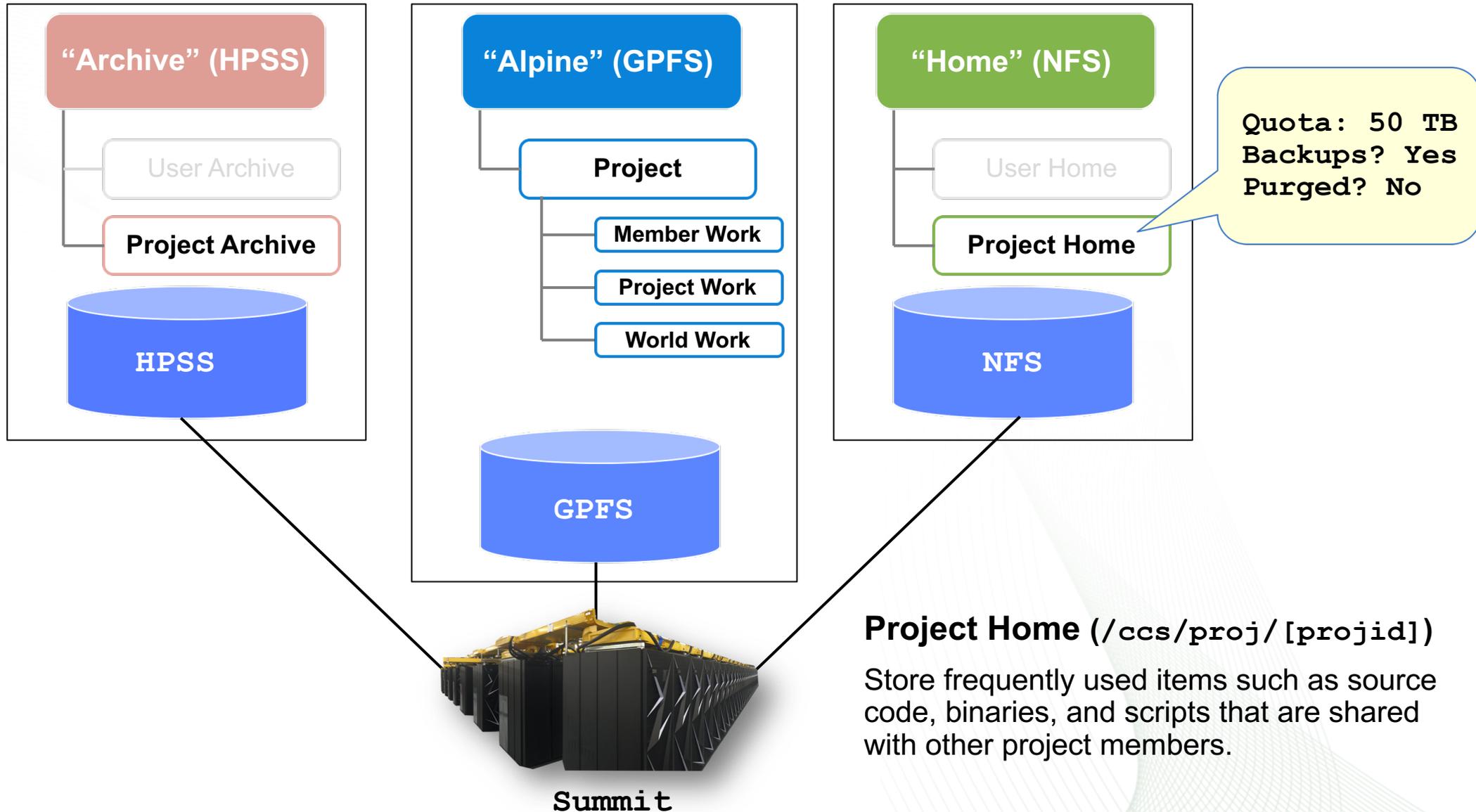
User Archive (/home/[userid])

Longer-term storage of user data (e.g. backup of user files)

File Systems & Data Storage: Project-Centric



File Systems & Data Storage: Project-Centric

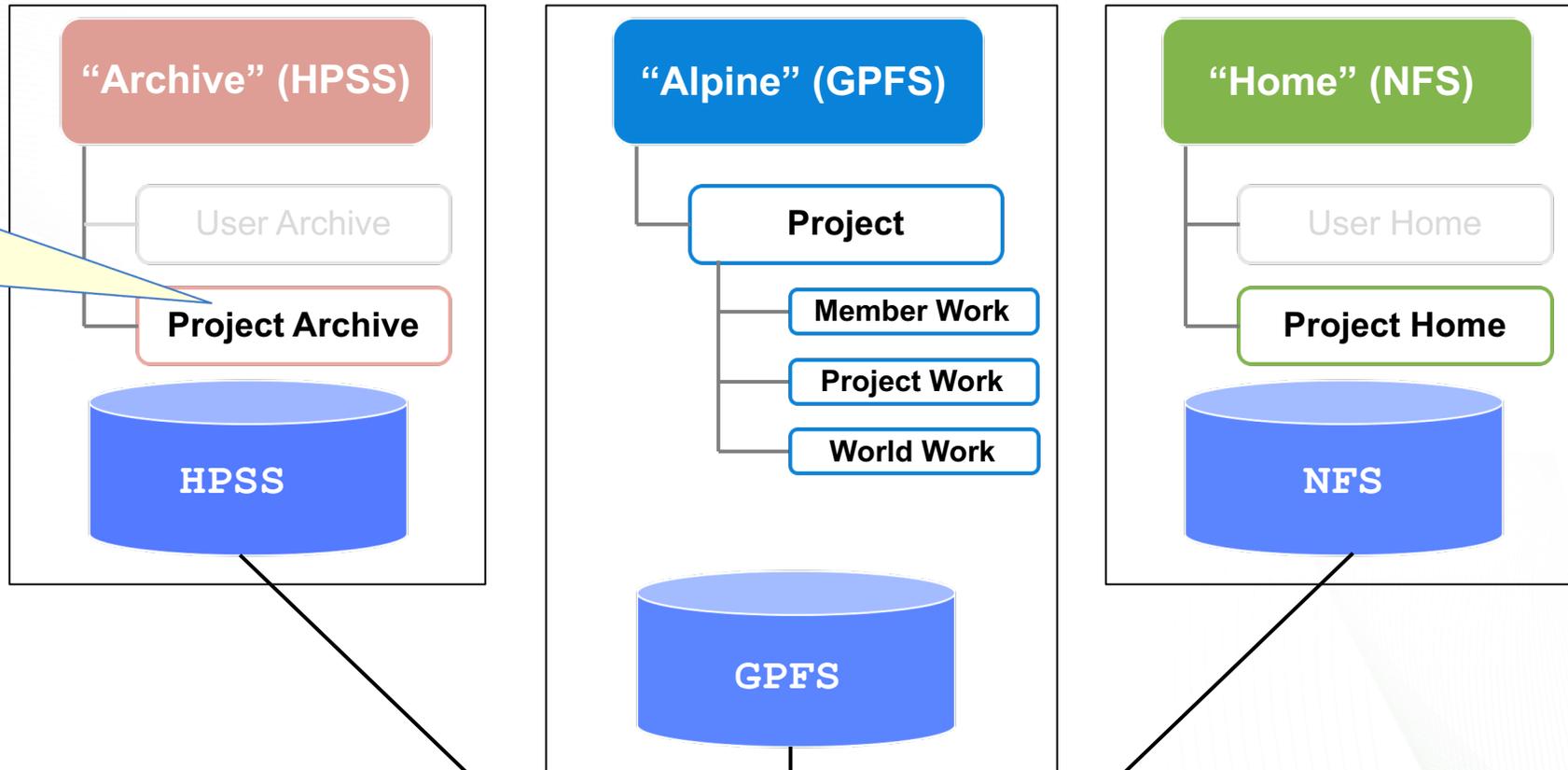


Project Home (/ccs/proj/[projid])

Store frequently used items such as source code, binaries, and scripts that are shared with other project members.

File Systems & Data Storage: Project-Centric

Quota: 100 TB
Backups? No
Purged? No

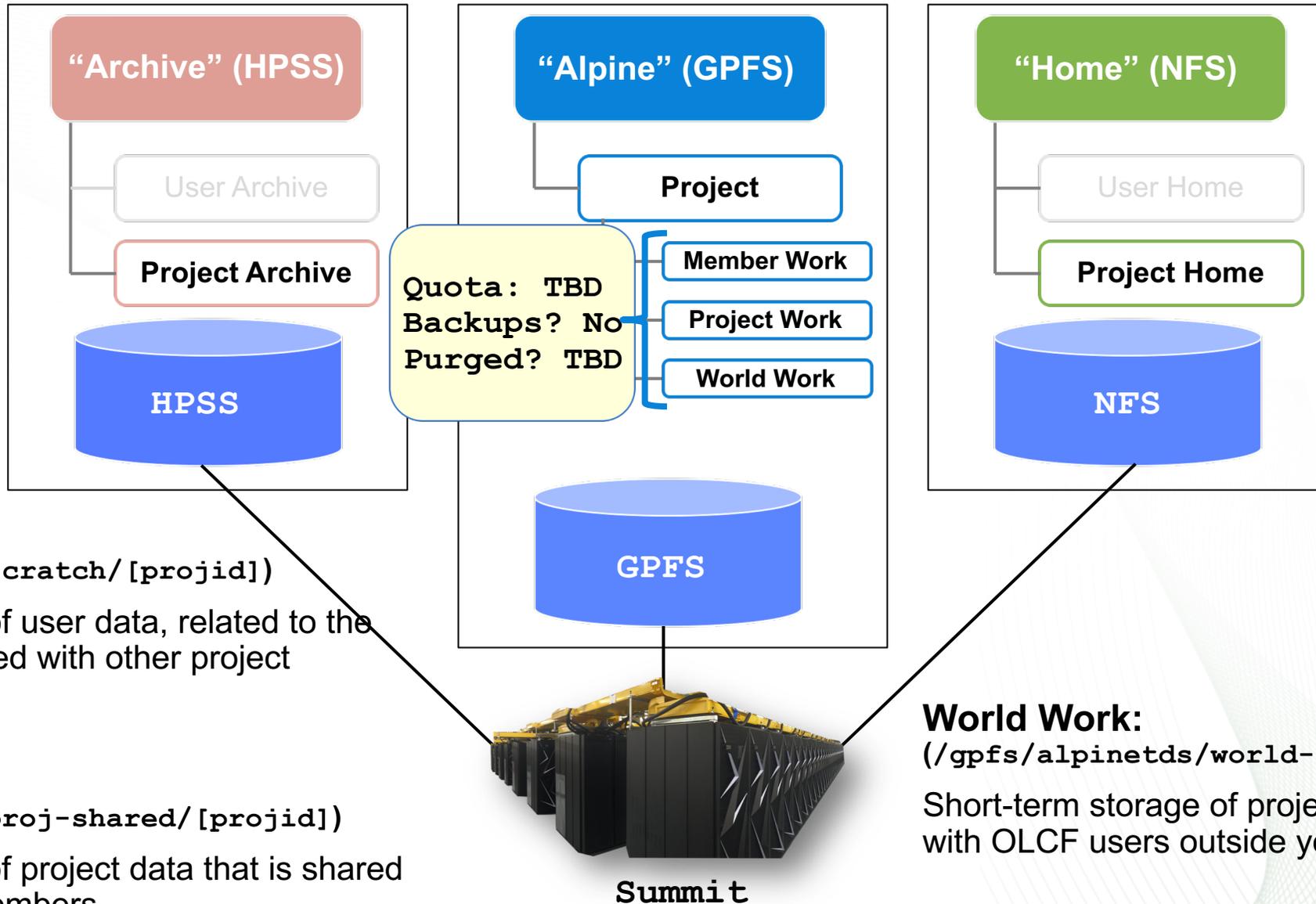


Project Archive (/proj/[projid])

Longer-term storage of project data shared with other project members (e.g. archived data used to publish paper)

Summit

File Systems & Data Storage: Project-Centric



Member Work:

`(/gpfs/alpinetds/scratch/[projid])`

Short-term storage of user data, related to the project, but not shared with other project members.

Project Work:

`(/gpfs/alpinetds/proj-shared/[projid])`

Short-term storage of project data that is shared with other project members.

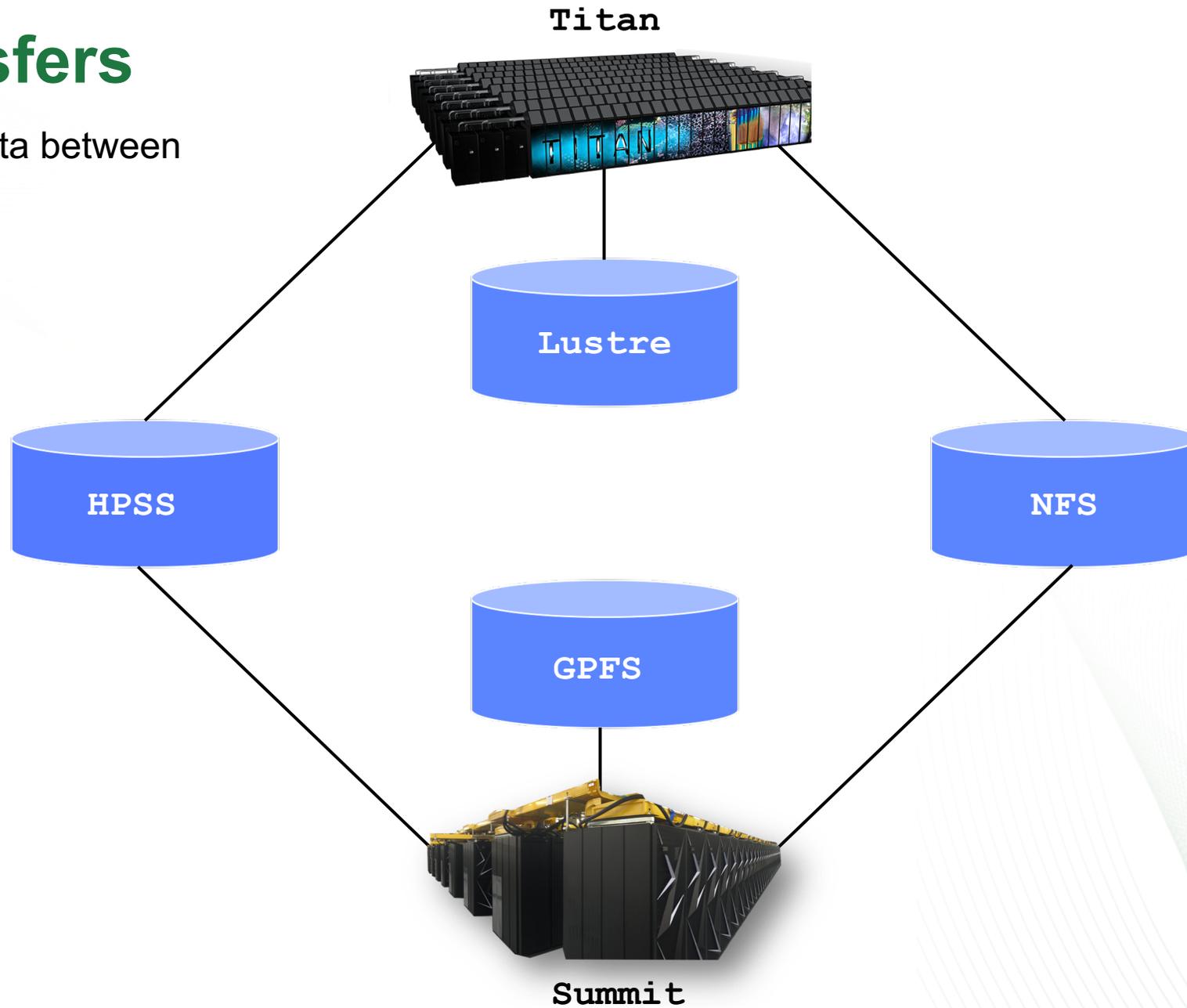
World Work:

`(/gpfs/alpinetds/world-shared/[projid])`

Short-term storage of project data that is shared with OLCF users outside your project.

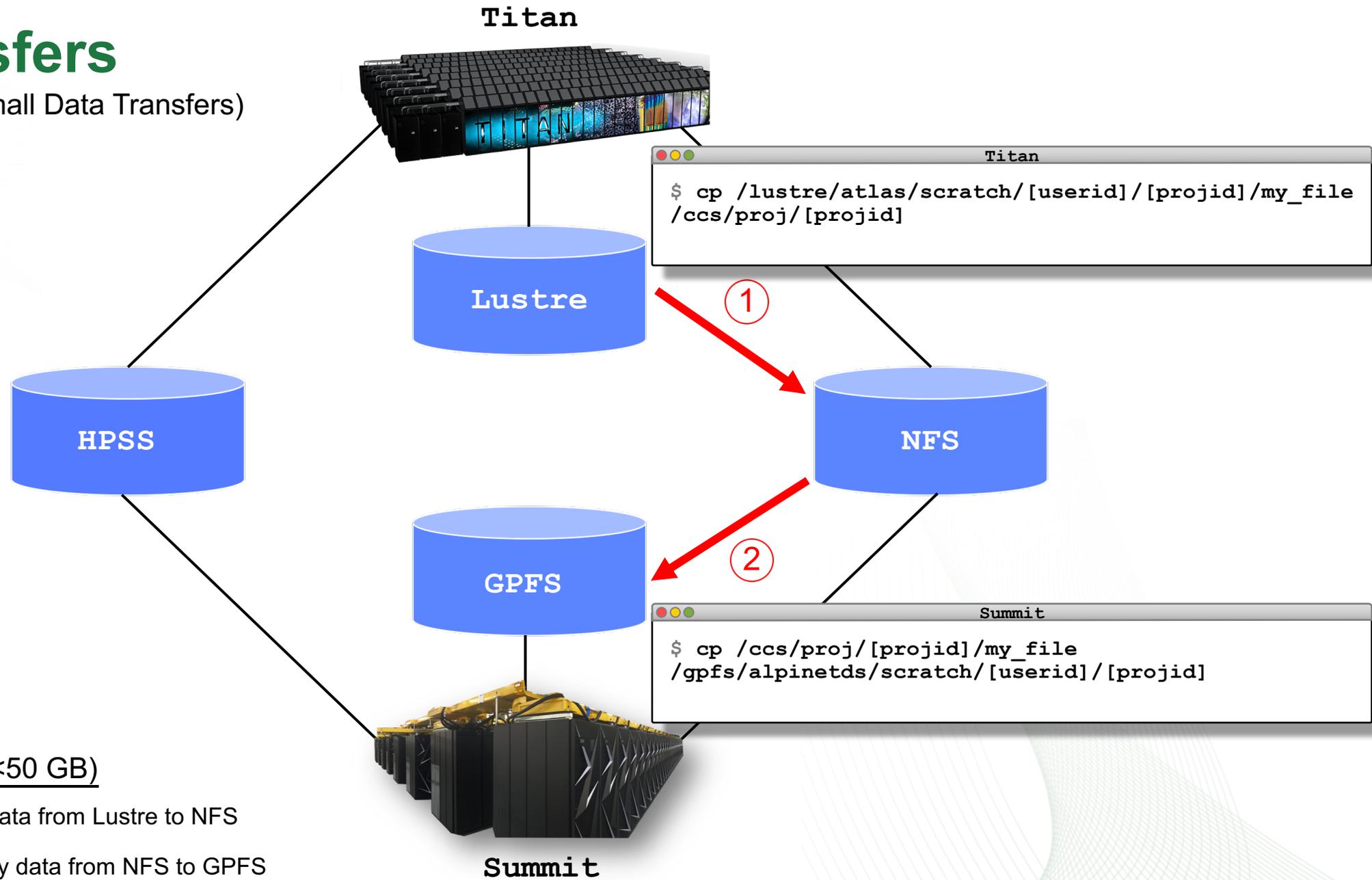
Data Transfers

How to transfer data between
Lustre and GPFS



Data Transfers

Lustre to GPFS (Small Data Transfers)

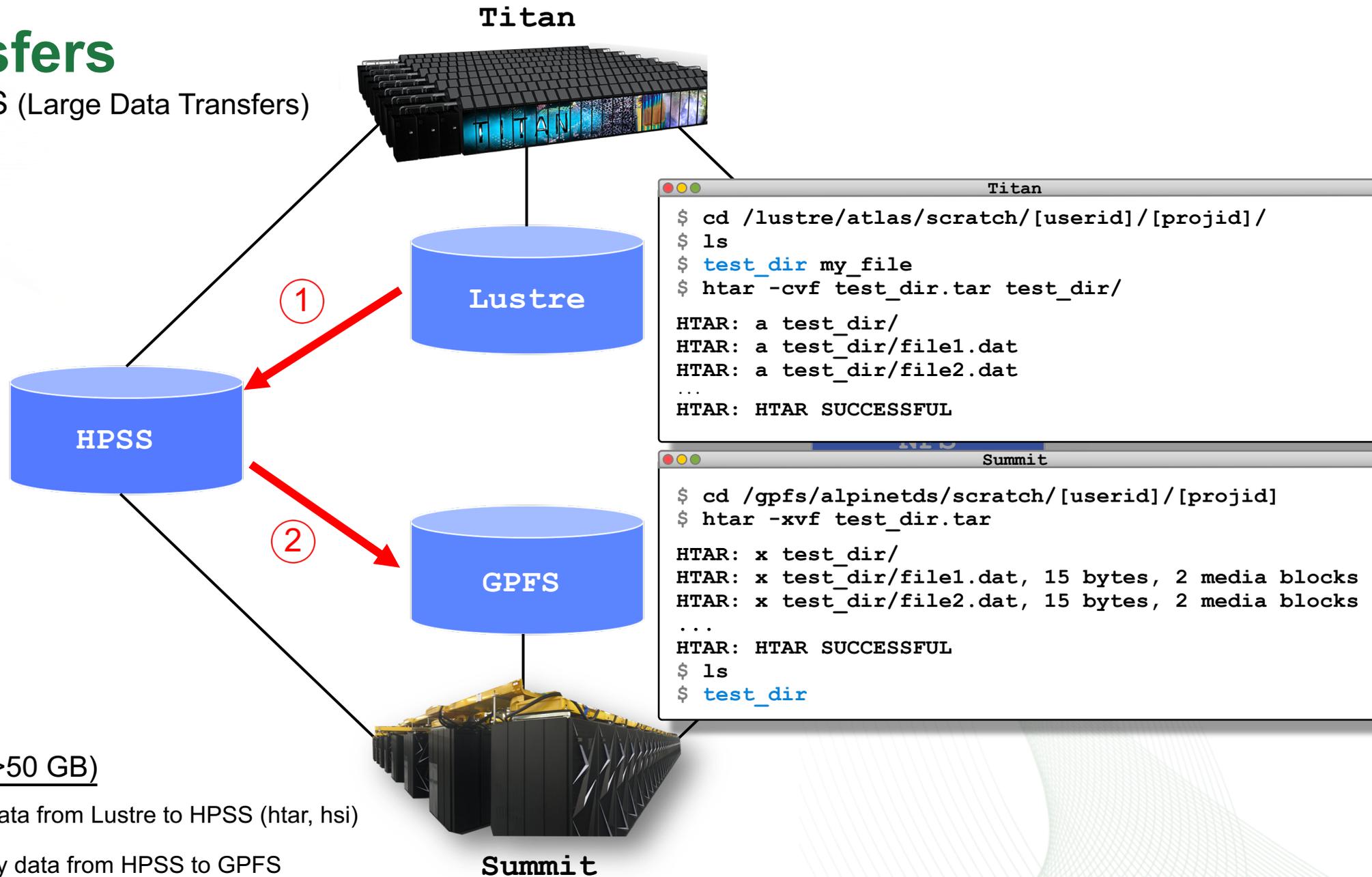


Small Data Transfers (<50 GB)

- 1 Log in to Titan. Copy data from Lustre to NFS
- 2 Log in to Summit. Copy data from NFS to GPFS

Data Transfers

Lustre to AlpineTDS (Large Data Transfers)

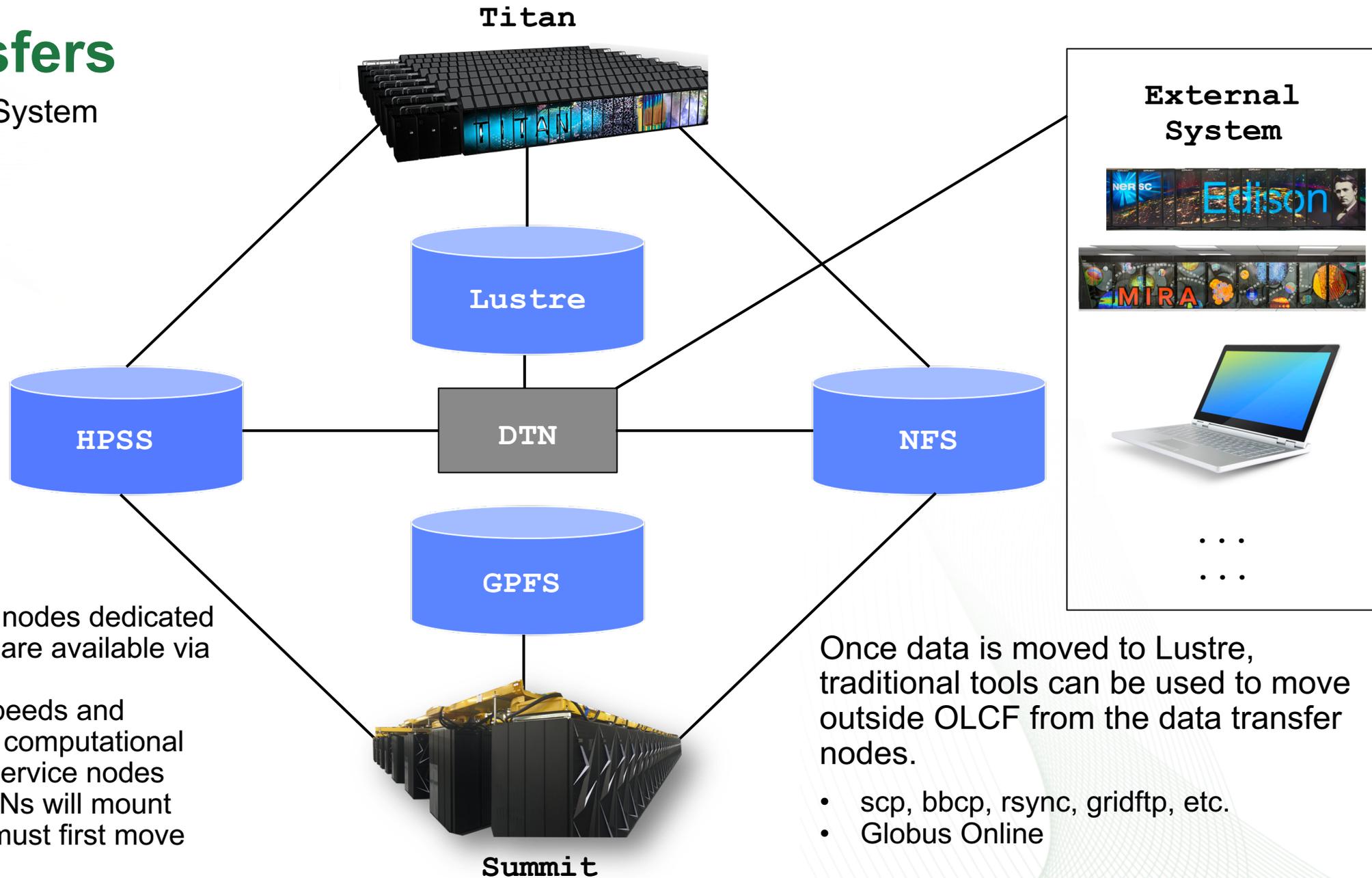


Large Data Transfers (>50 GB)

- 1 Log in to Titan. Copy data from Lustre to HPSS (htar, hsi)
- 2 Log in to Summit. Copy data from HPSS to GPFS

Data Transfers

GPFS to External System



Data Transfer Nodes

- The OLCF provides nodes dedicated to data transfer that are available via dtn.ccs.ornl.gov.
- Improves transfer speeds and reduces the load on computational systems' login and service nodes
- In the future, the DTNs will mount GPFS, but for now must first move data to Lustre.

Once data is moved to Lustre, traditional tools can be used to move outside OLCF from the data transfer nodes.

- scp, bbcp, rsync, gridftp, etc.
- Globus Online

Additional Information

For more information, please see the **File Systems: Data Storage & Transfers** page of the Summit User Guide:

<https://www.olcf.ornl.gov/for-users/system-user-guides/summit/file-systems/>

- Storage Overview
- Data Transfers and Summit
- HPSS Best Practices

For details on using Globus Online, please see the “Remote Transfers” section on the **File Systems: Data Storage & Transfers** page of the Titan User Guide:

<https://www.olcf.ornl.gov/for-users/system-user-guides/titan/file-systems/#remote-transfers>



Thank You.