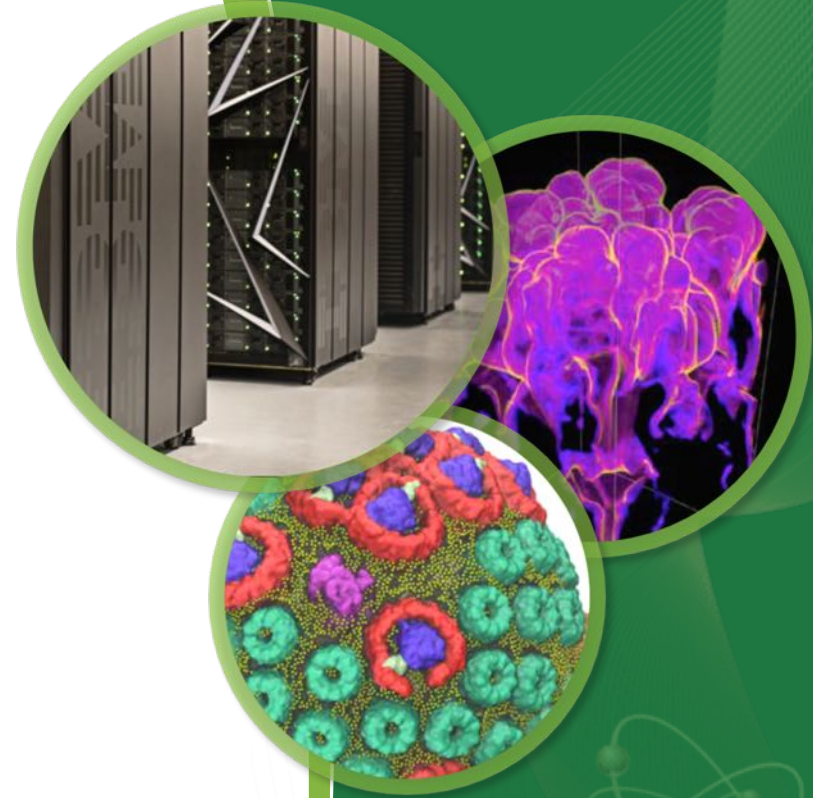


File Systems & Data Transfers Overview

Tom Papatheodore
Oak Ridge Leadership Computing Facility
Oak Ridge National Laboratory

June 26, 2018



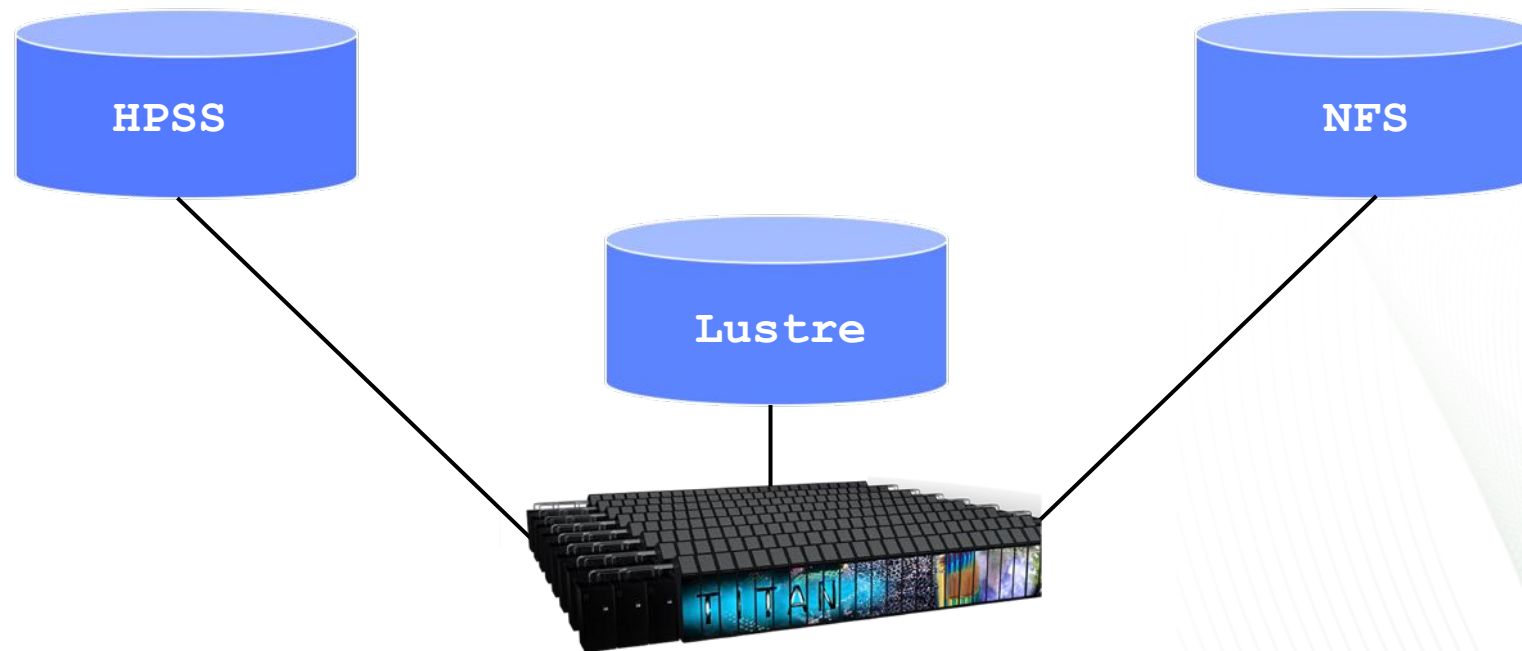
Lustre File System (Atlas)

- Temporary storage for production I/O
- 32 PB Capacity
- 1 TB/s read/write (aggregate)
- Mounted on most OLCF systems



File Systems & Data Storage

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

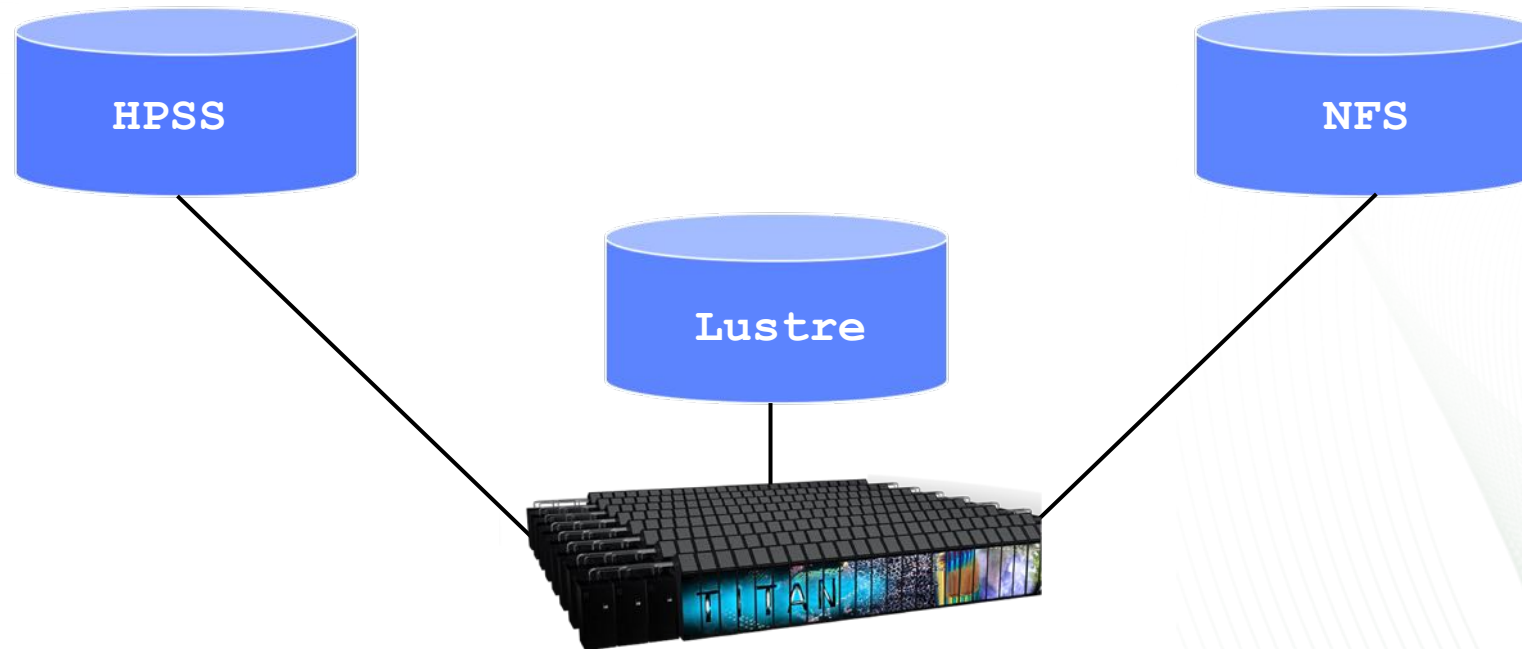


File Systems & Data Storage

This is where you're dropped
when you first log in
(/ccs/proj/[userid])

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

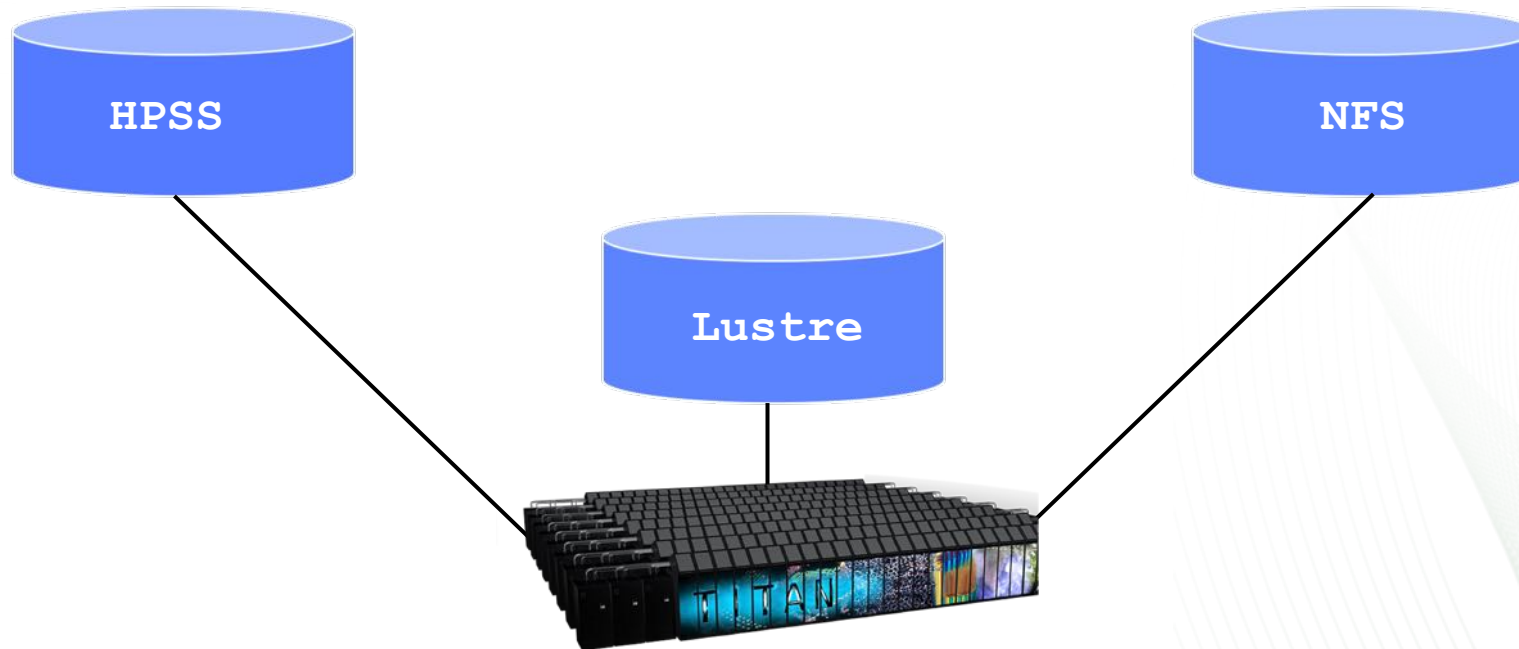
Provides long-term storage for routinely
accessed user or project data, respectively.



File Systems & Data Storage

Lustre (`/lustre/atlas/`)

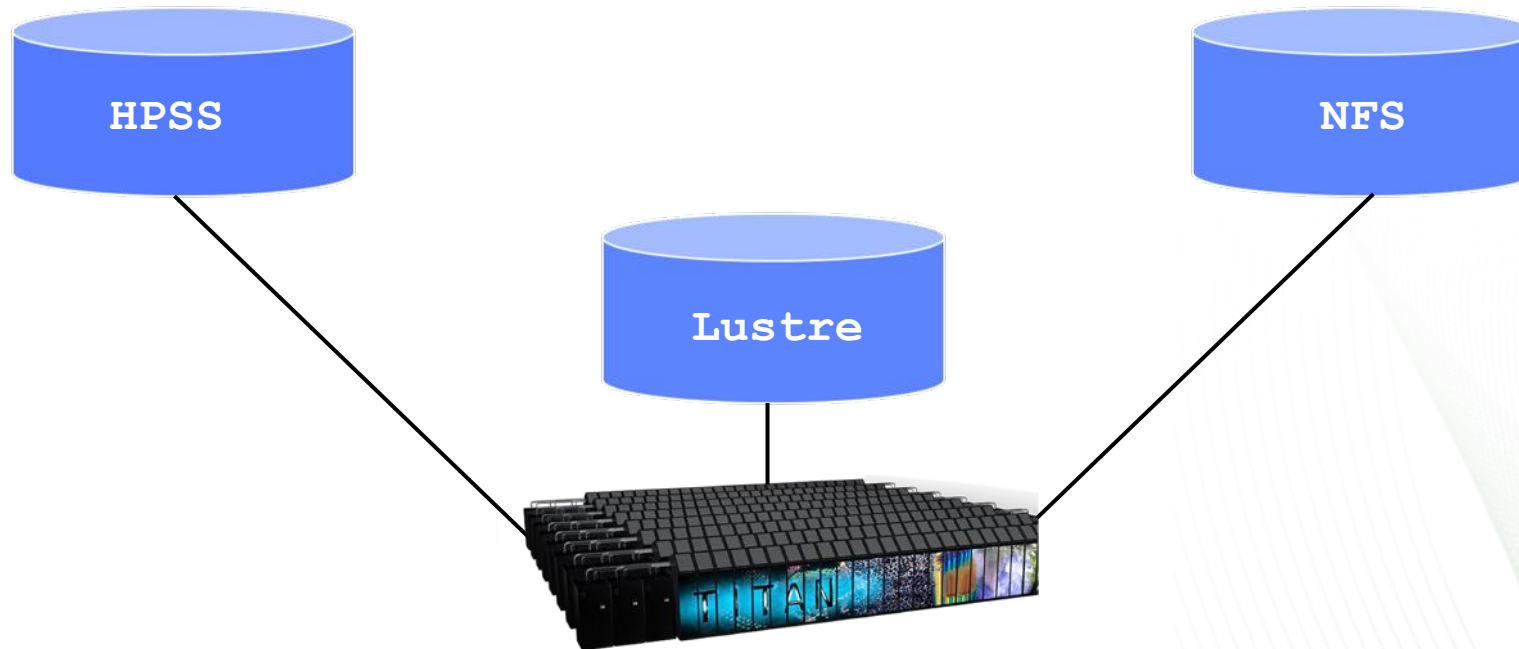
Provides temporary storage for production data created on Titan.



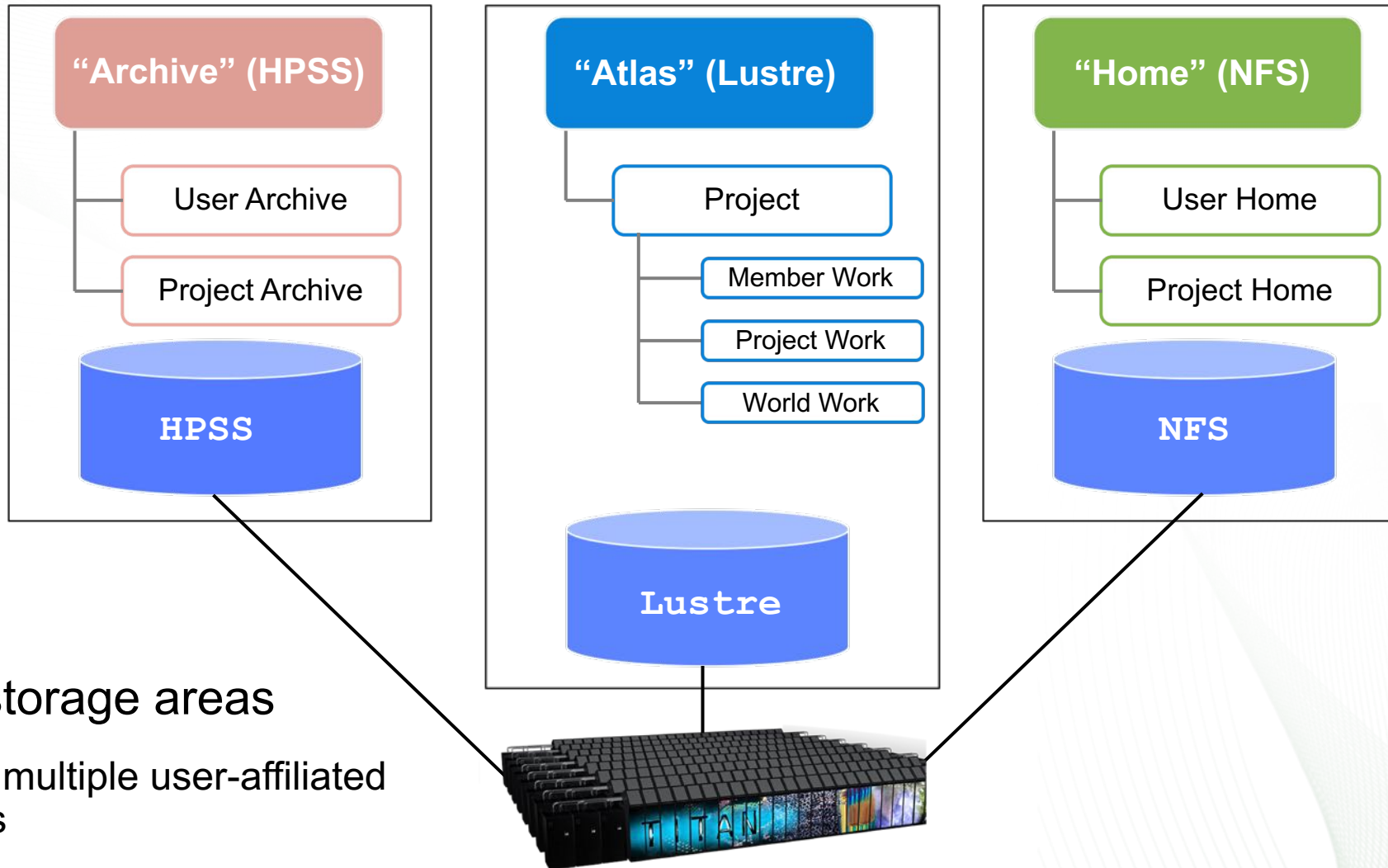
File Systems & Data Storage

HPSS (/home/ and /proj/)

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



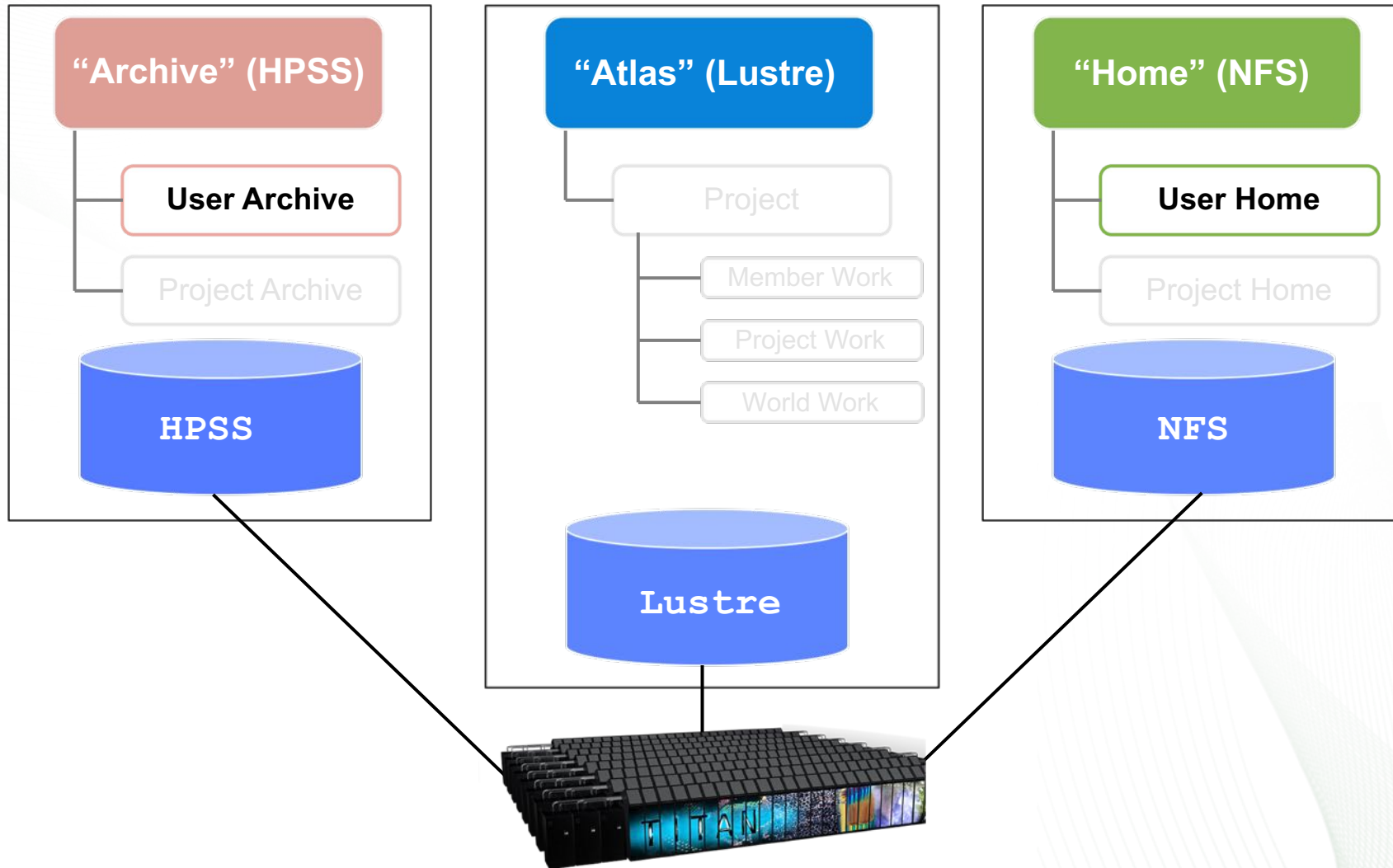
File Systems & Data Storage



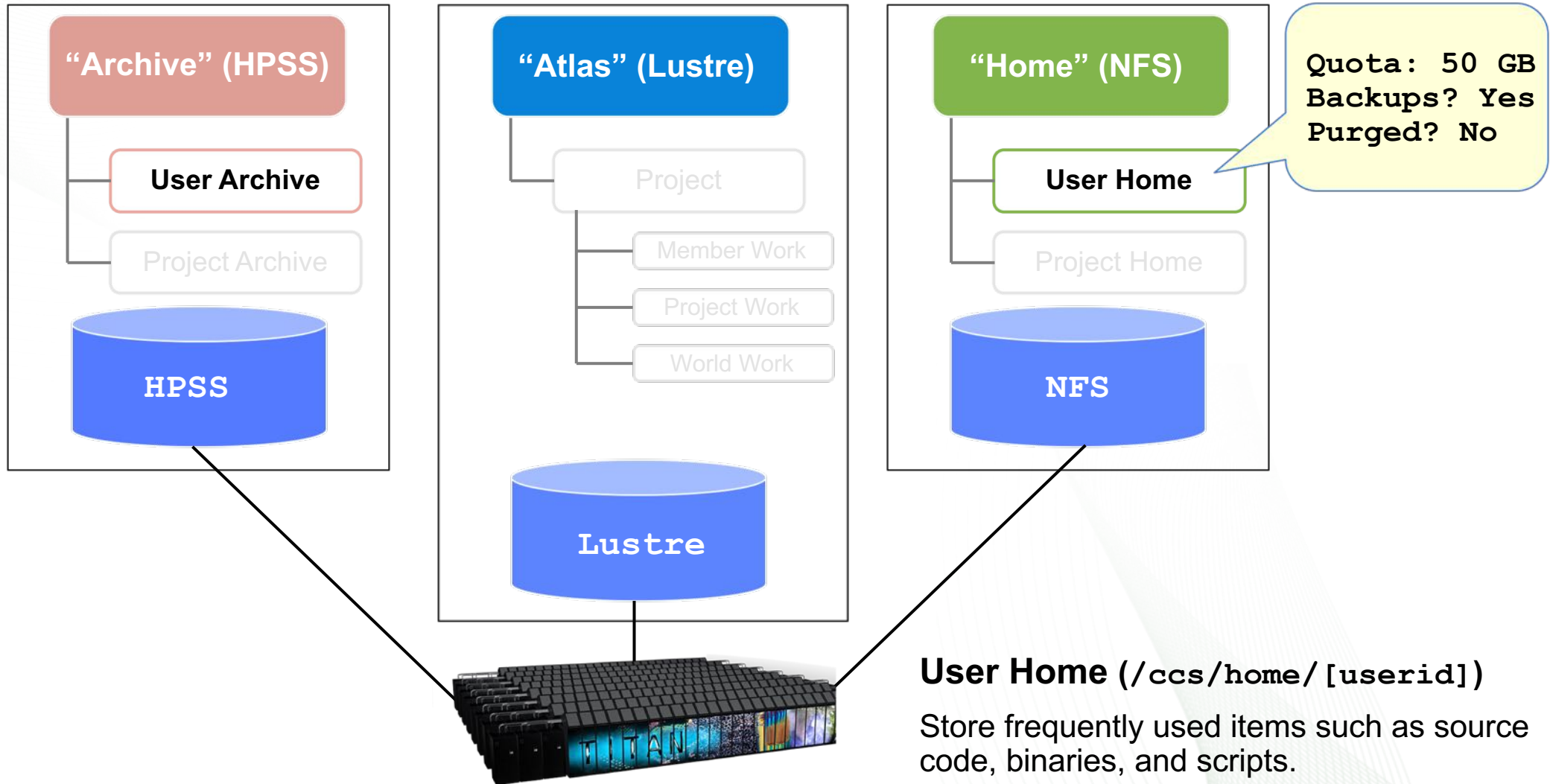
Within these storage areas

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

File Systems & Data Storage: User-Centric

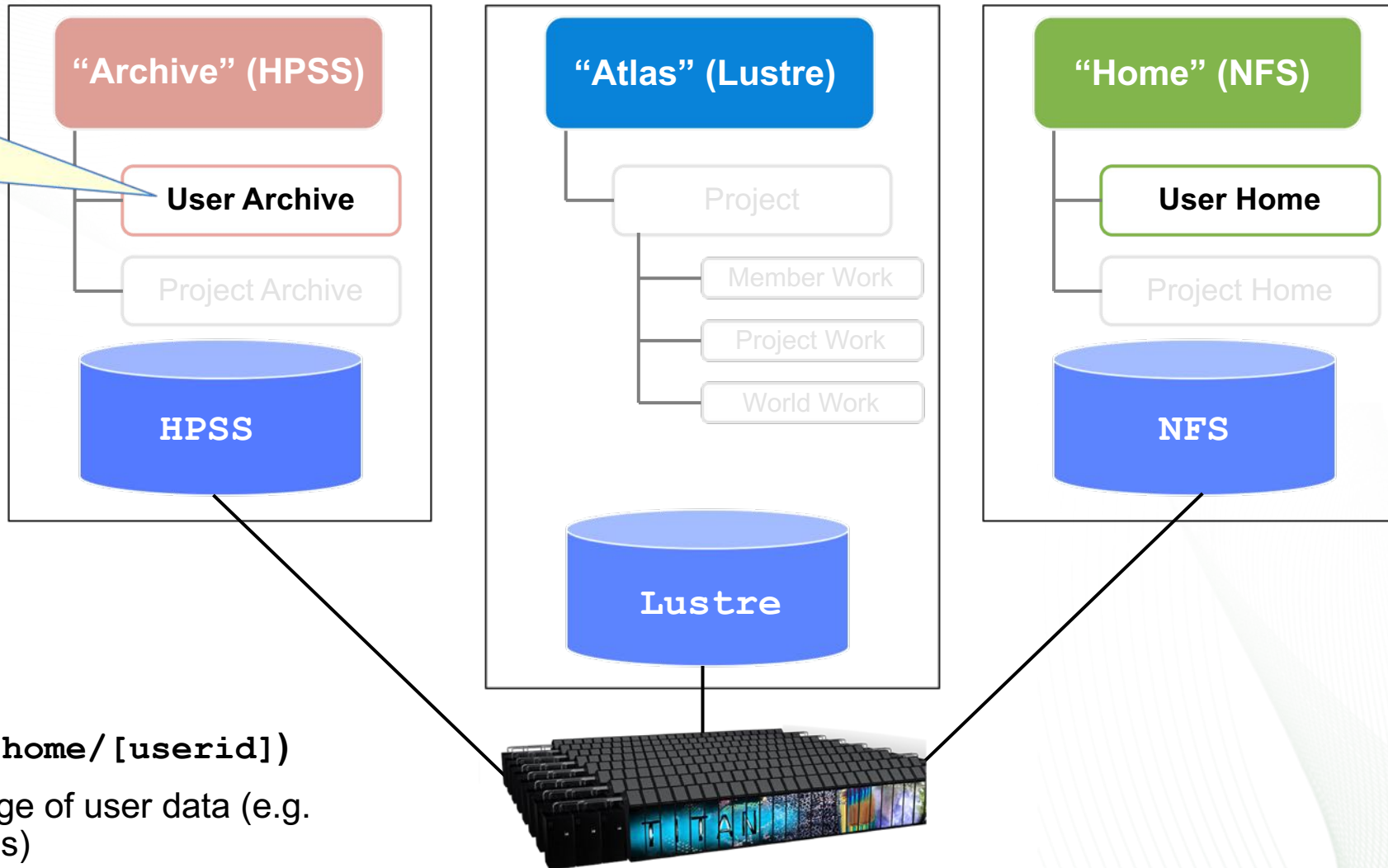


File Systems & Data Storage: User-Centric



File Systems & Data Storage: User-Centric

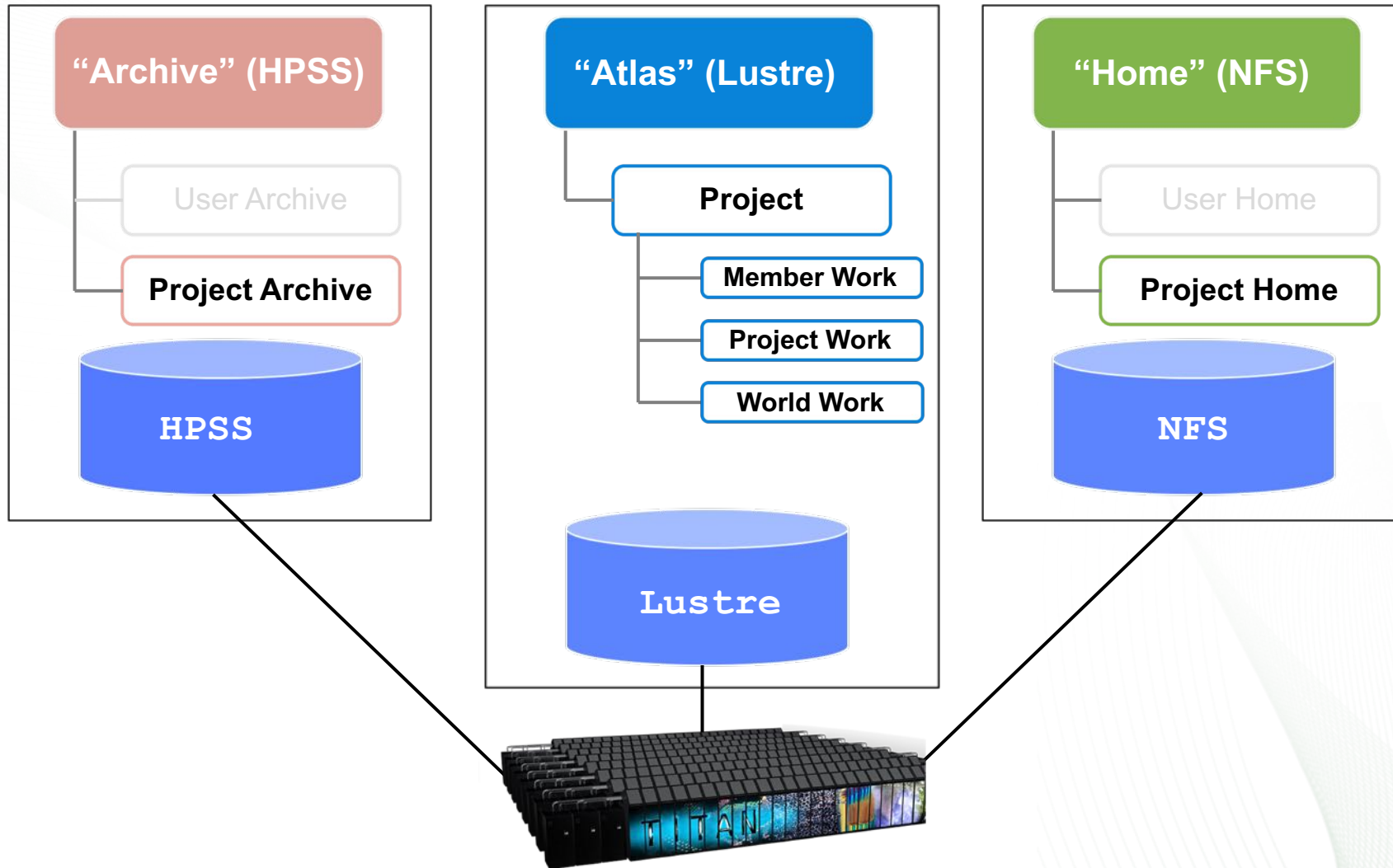
Quota: 2 TB
Backups? No
Purged? No



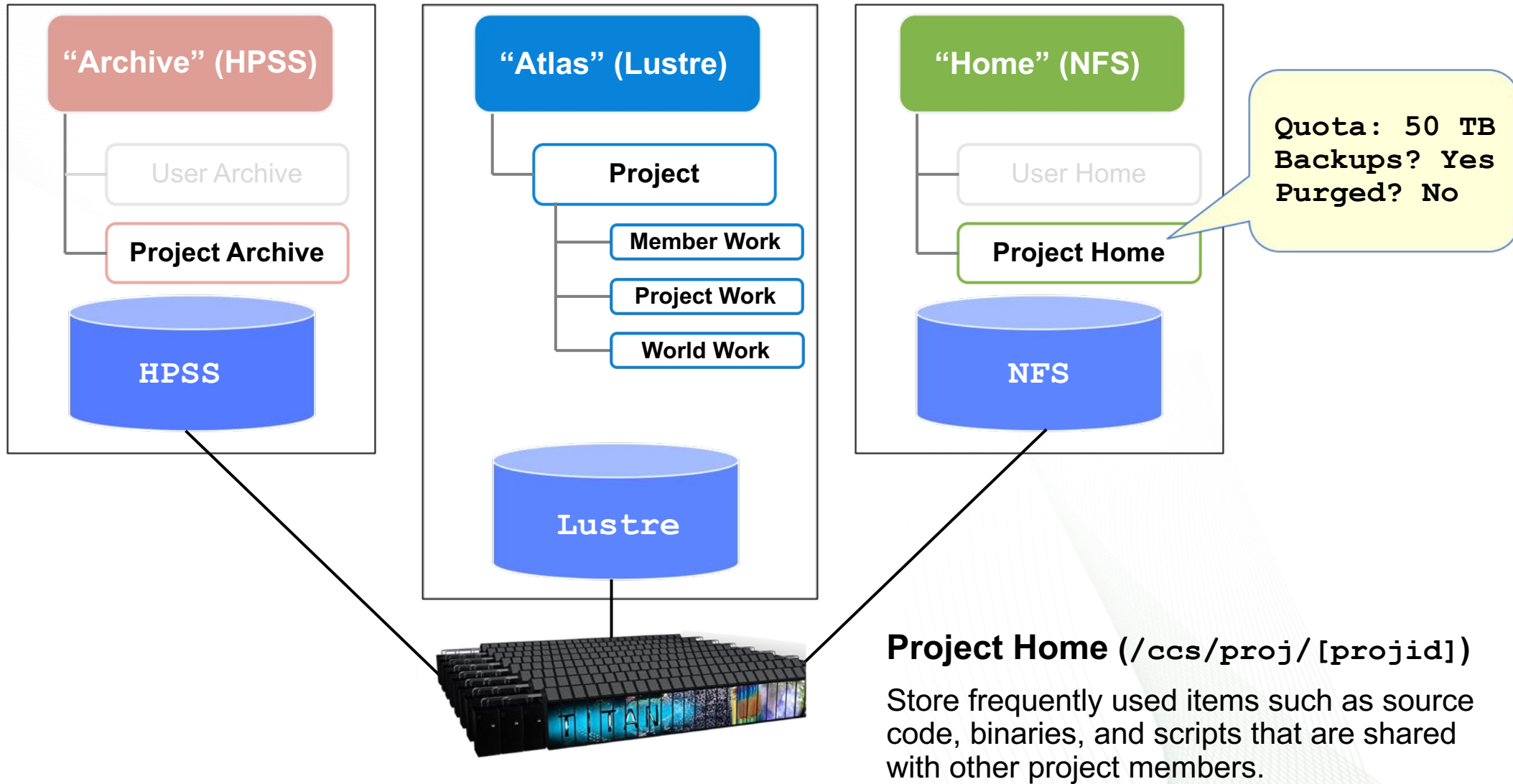
User Archive (/home/[userid])

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

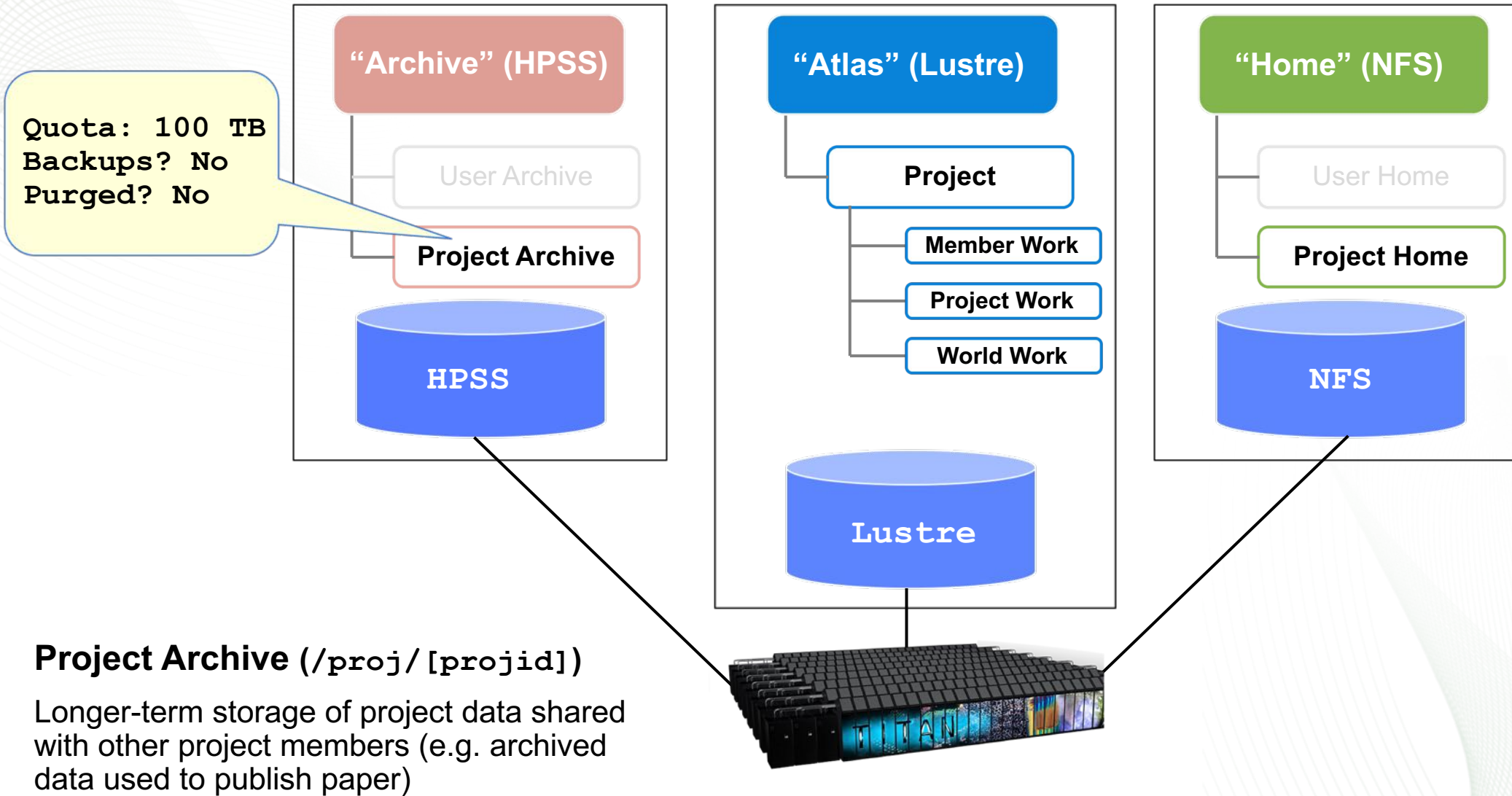
File Systems & Data Storage: Project-Centric



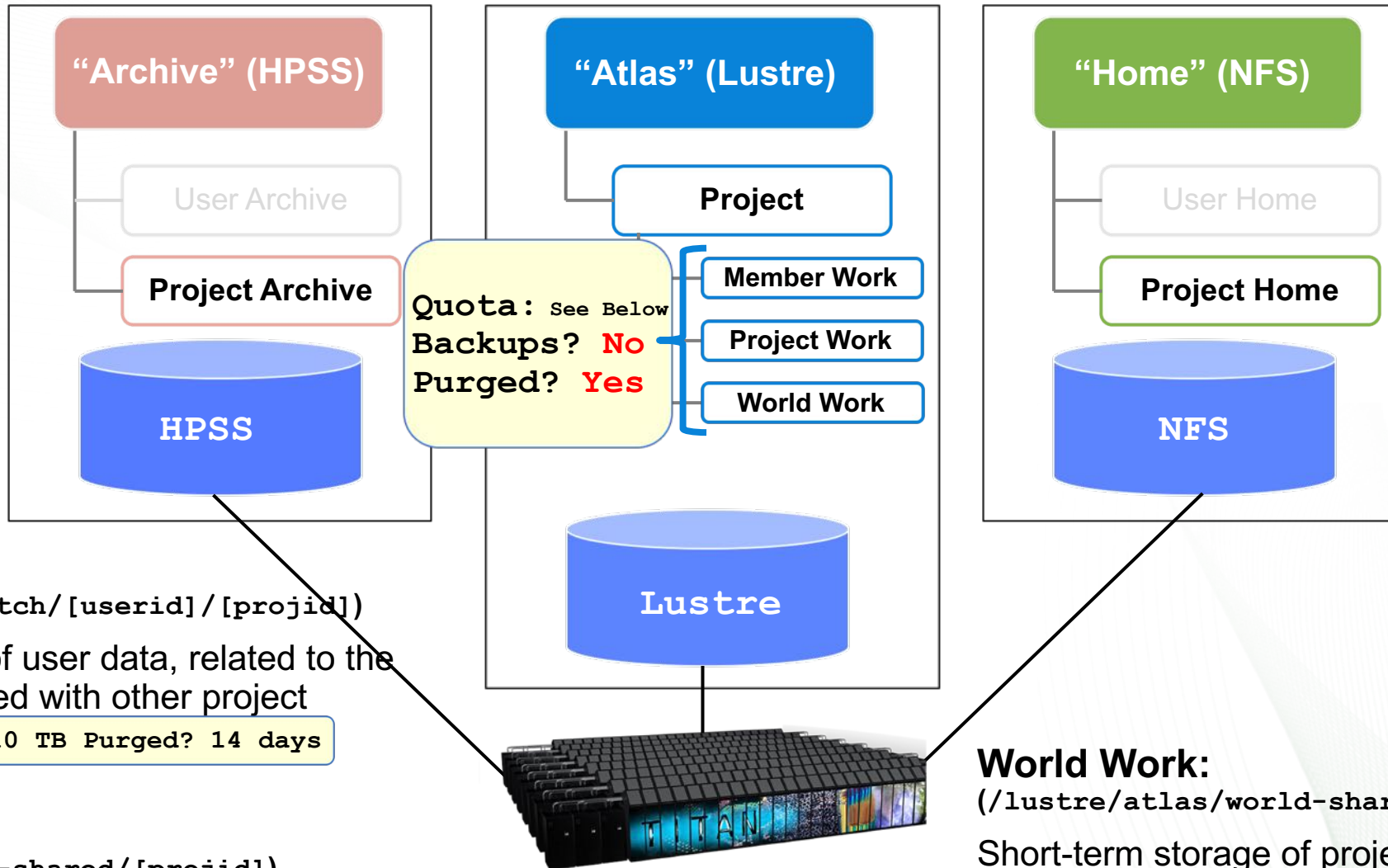
File Systems & Data Storage: Project-Centric



File Systems & Data Storage: Project-Centric



File Systems & Data Storage: Project-Centric



Member Work:

(/lustre/atlas/scratch/[userid]/[projid])

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

Quota: 10 TB Purged? 14 days

Project Work:

(/lustre/atlas/proj-shared/[projid])

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

Quota: 100 TB Purged? 90 days

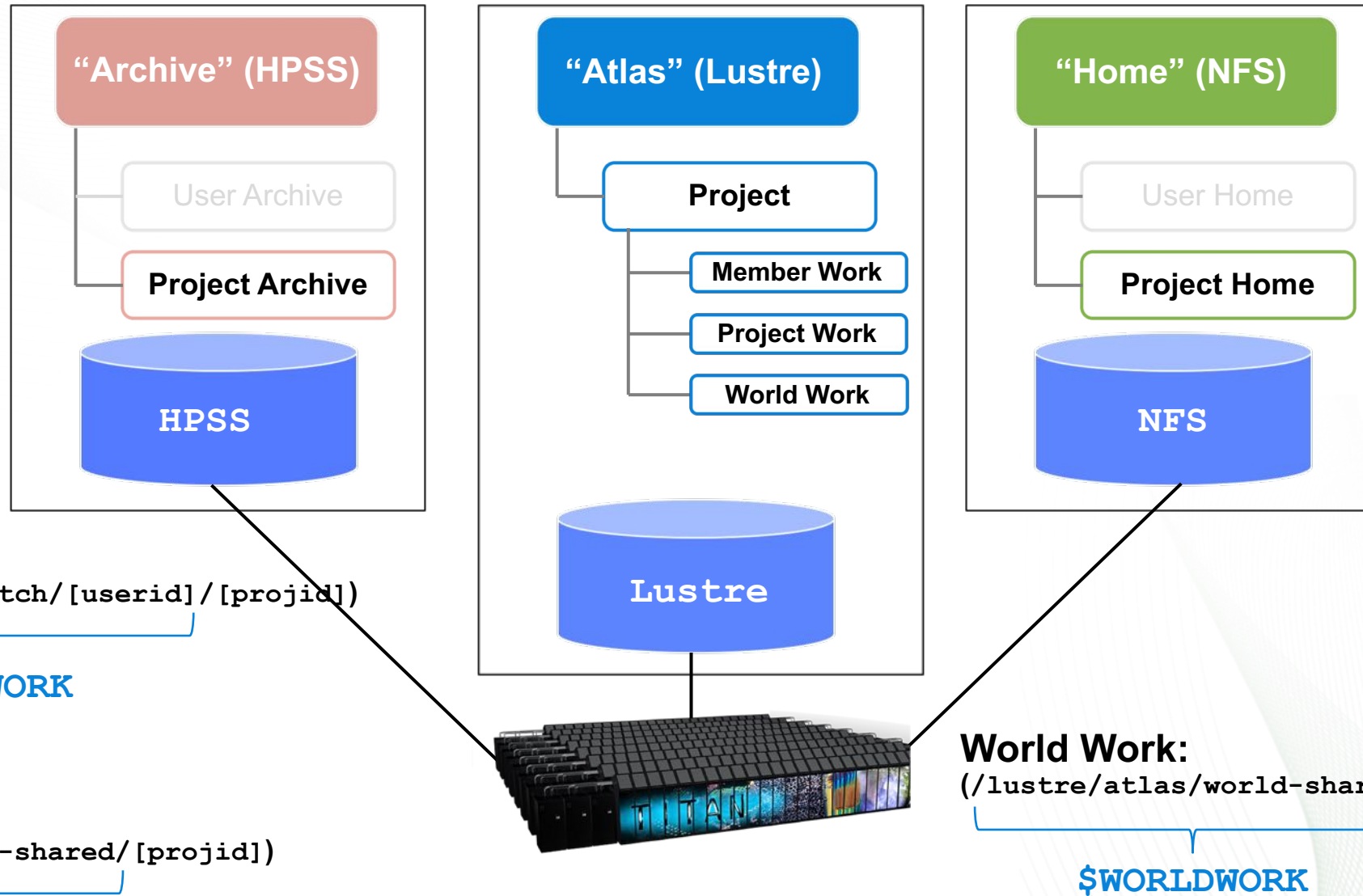
World Work:

(/lustre/atlas/world-shared/[projid])

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

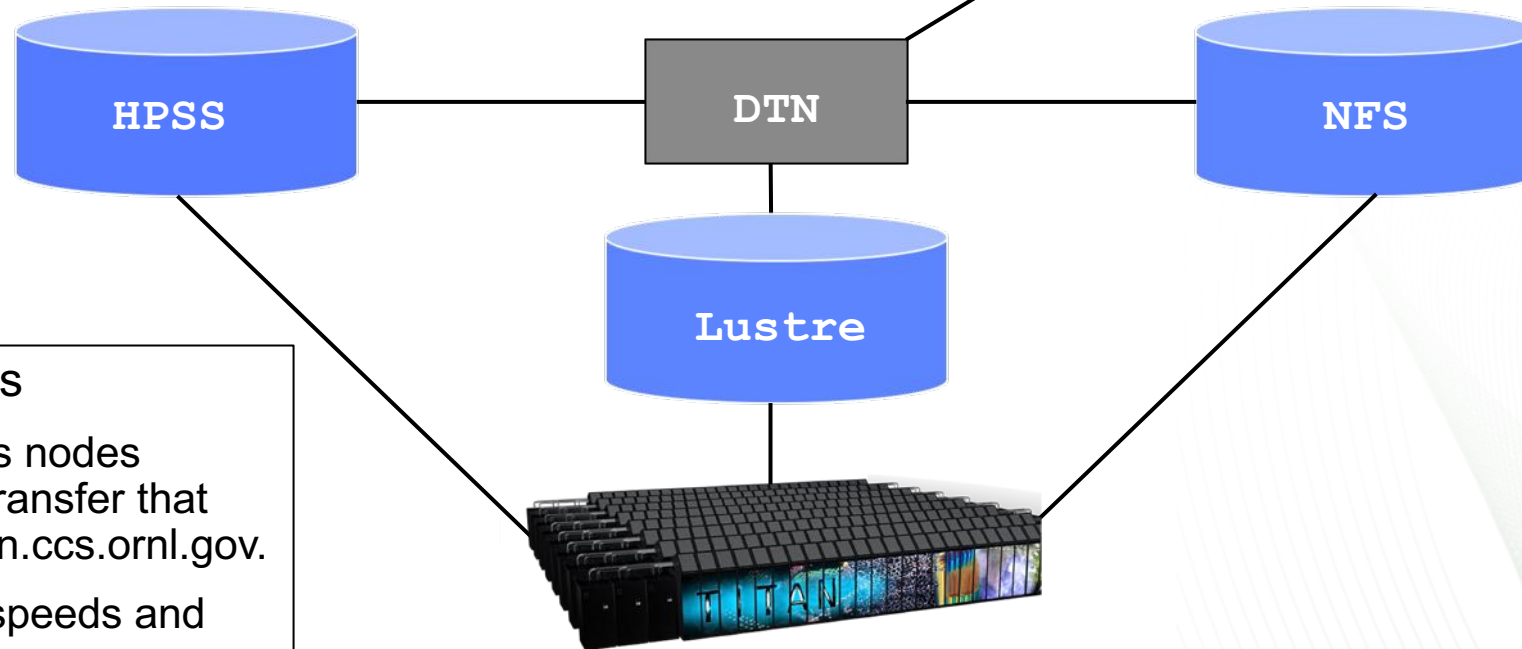
Quota: 10 TB Purged? 90 days

File Systems & Data Storage: Project-Centric



Data Transfers

- How do you transfer data within OLCF systems?
- How do you transfer data to/from external systems?



Data Transfer Nodes

- The OLCF provides nodes dedicated to data transfer that are available via dtm.ccs.ornl.gov.
- Improves transfer speeds and reduces the load on computational systems' login and service nodes

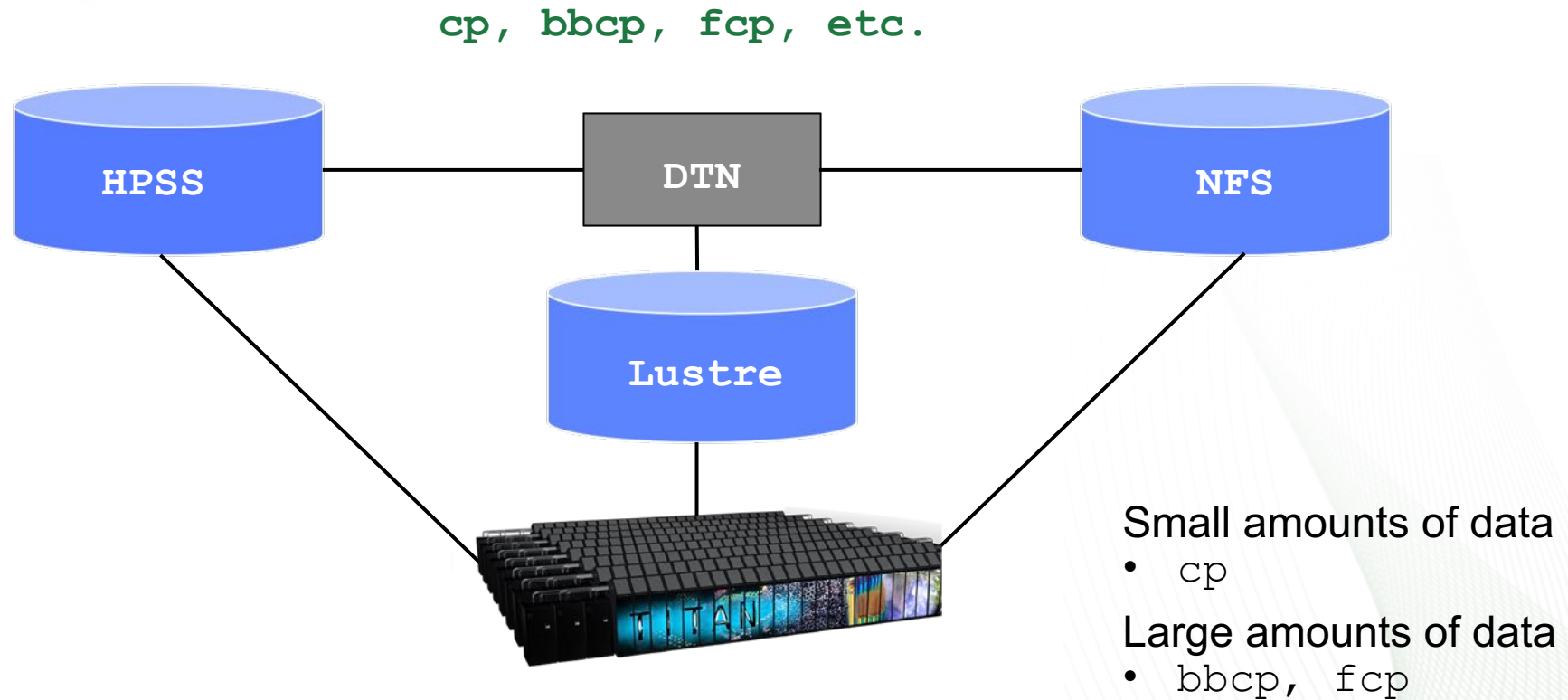
External System



...

Data Transfers: local data transfers

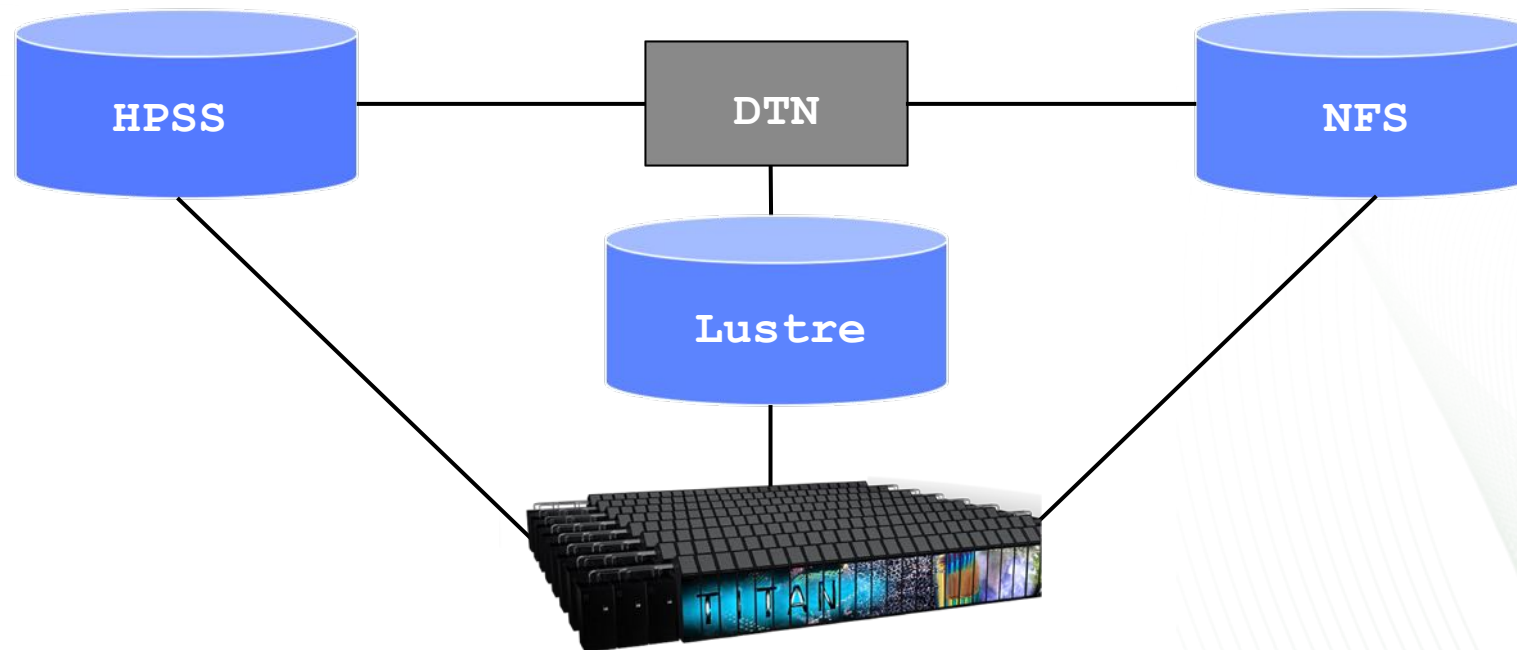
- How do you transfer data within OLCF systems?



Data Transfers: local data transfers

```
Titan
$ ls /home/ccs/[userid]
test_dir my_file test.exe
$ cp /home/ccs/[userid]/test.exe /lustre/atlas/scratch/[userid]/[projid]/
```

Copy file from NFS (User Home) to Lustre (Member Work)

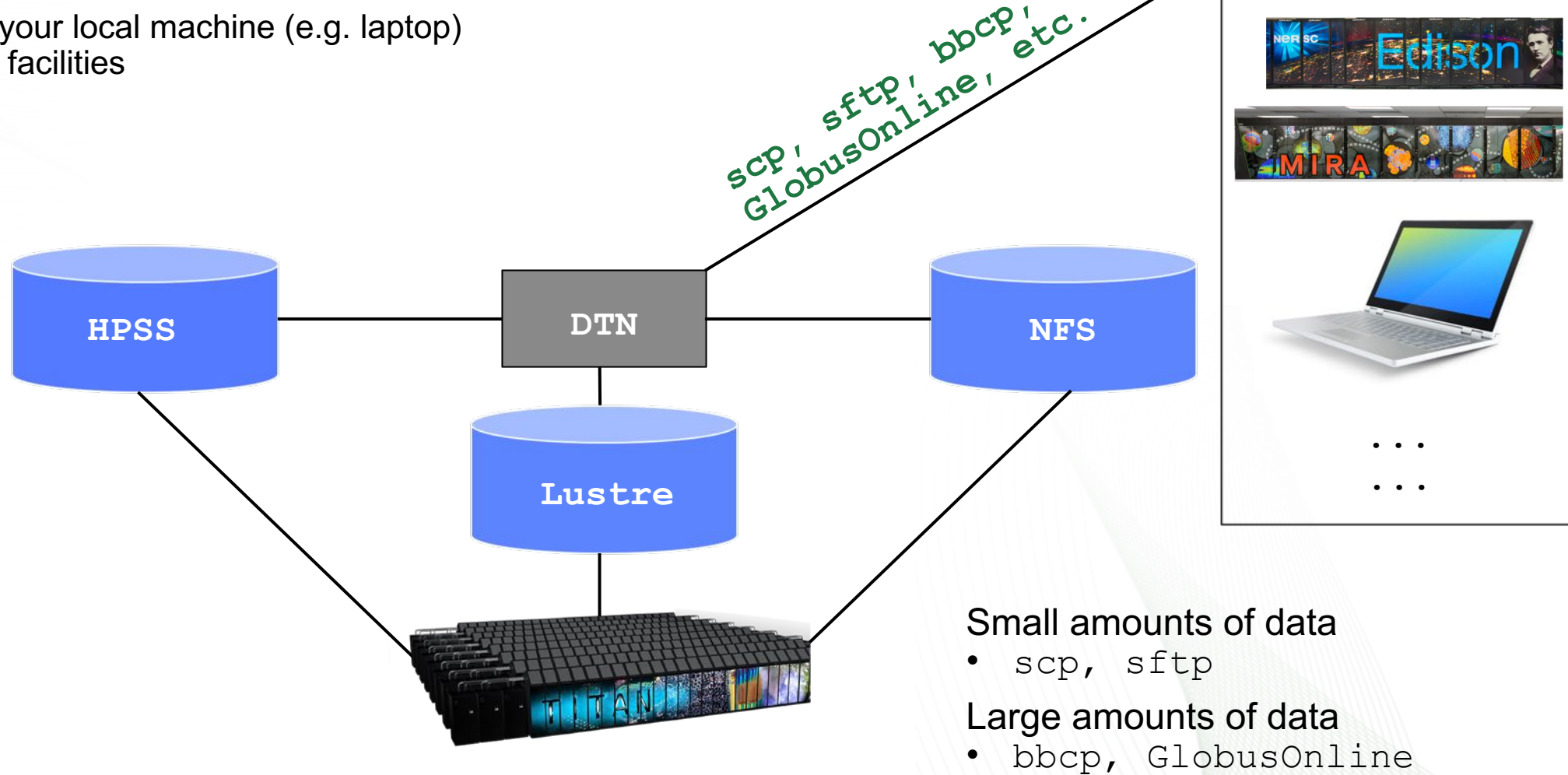


<https://www.olcf.ornl.gov/for-users/system-user-guides/titan/file-systems/#local-transfers>

Data Transfers: remote data transfers

How do you transfer data to/from the OLCF systems?

- Send data to/from your local machine (e.g. laptop)
- Send data to other facilities

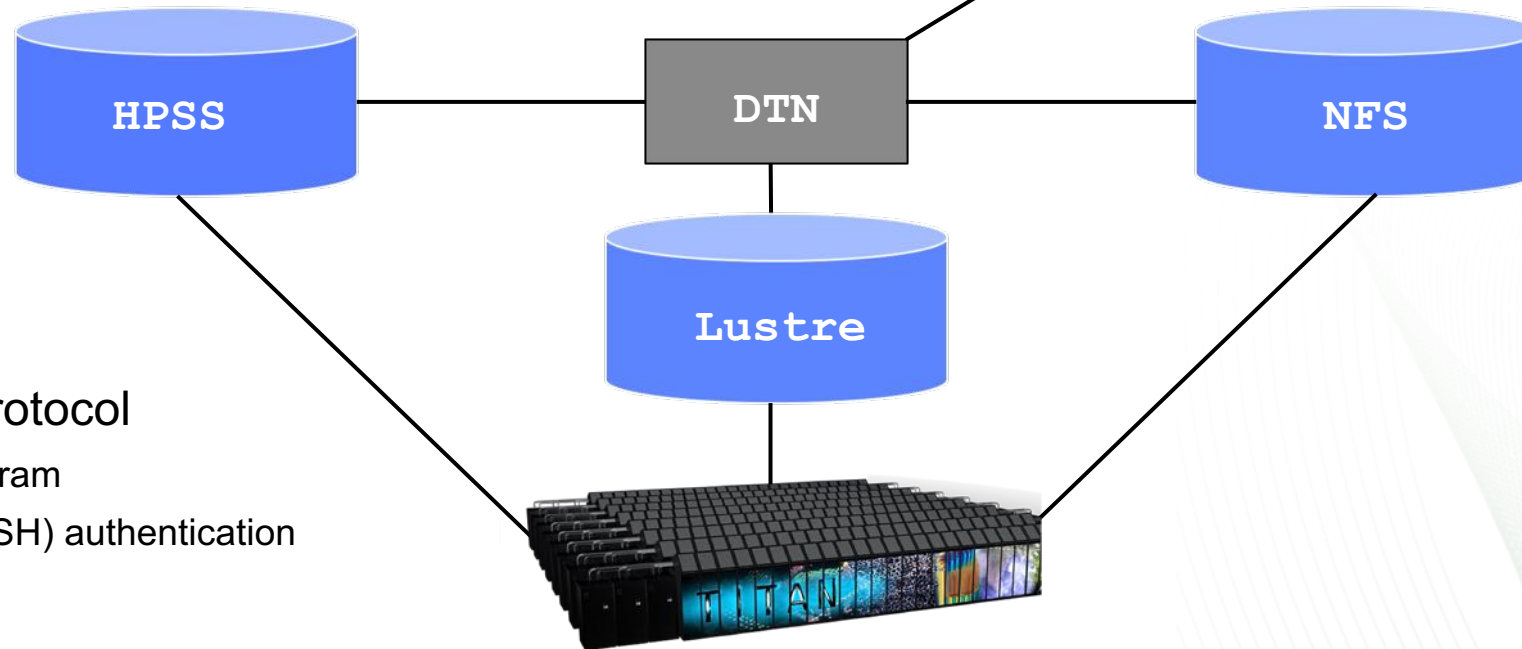


Data Transfers: remote data transfers

```
Local
$ ls
test_dir my_file

$ scp myfile userid@dtm.ccs.ornl.gov:/lustre/atlas/scratch/[userid]/[projid]/
```

Copy file from laptop to Lustre at OLCF



External
System



scp: secure copy protocol

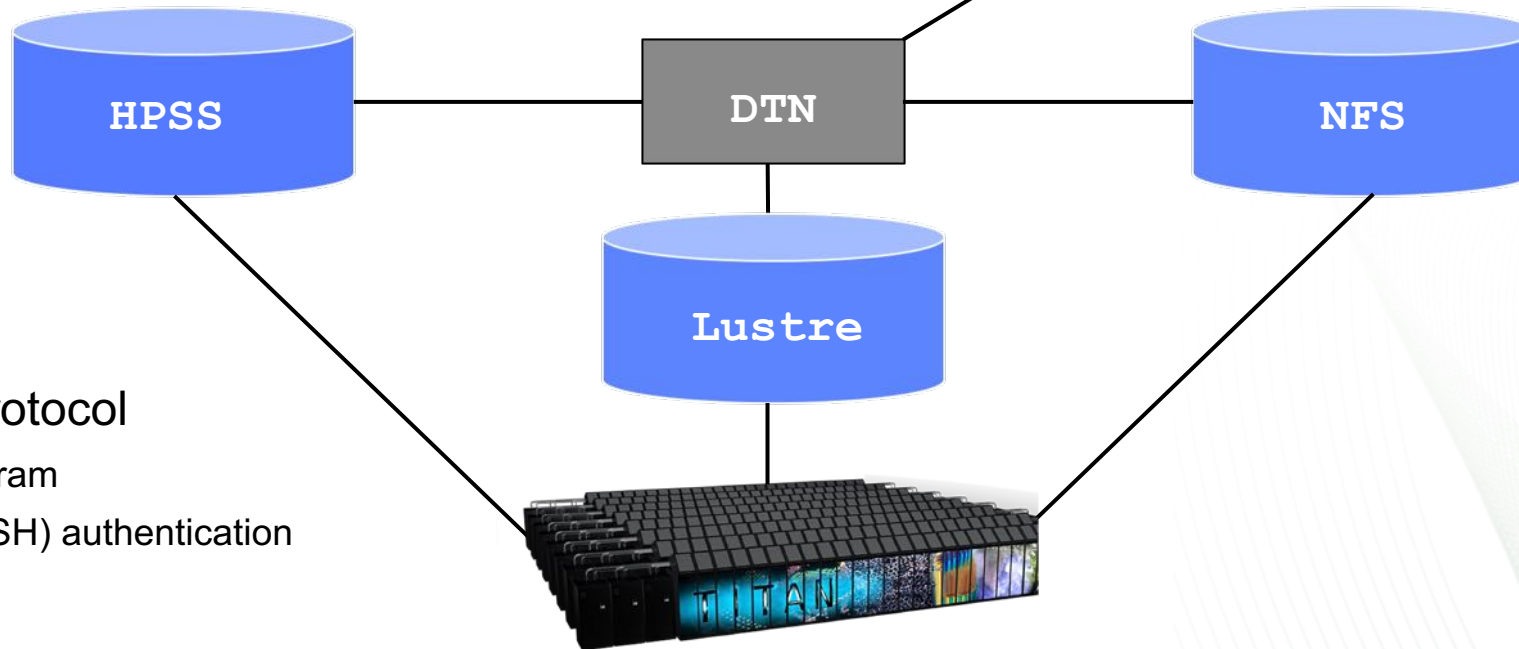
- Remote file copy program
- Uses Secure Shell (SSH) authentication for data transfers.

Data Transfers: remote data transfers

```
Local
$ ls
test_dir my_file

$ scp -r test_dir userid@dtm.ccs.ornl.gov:/lustre/atlas/scratch/[userid]/[projid]/
```

Copy directory from laptop to Lustre at OLCF



scp: secure copy protocol

- Remote file copy program
- Uses Secure Shell (SSH) authentication for data transfers.

Data Transfers: remote data transfers (GlobusOnline)

- Navigate to www.globus.org and click “Log In”
- Choose “Oak Ridge National Laboratory” from the drop-down menu and click “Continue”



Log in to use Globus Web App

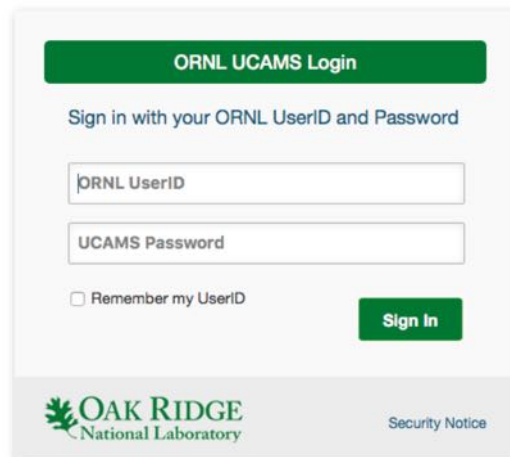
Use your existing organizational login
e.g., university, national lab, facility, project

Oak Ridge National Laboratory

Didn't find your organization? Then use [Globus ID](#) to sign in. (What's this?)

Continue

- You will be redirected to enter your ORNL UCAMS credentials



ORNL UCAMS Login

Sign in with your ORNL UserID and Password

ORNL UserID

UCAMS Password

☐ Remember my UserID

Sign In

 OAK RIDGE
National Laboratory

[Security Notice](#)

Data Transfers: remote data transfers (GlobusOnline)

Search for
“OLCF ATLAS”

The screenshot shows the 'Transfer Files' window in GlobusOnline. At the top, there's a header with 'Transfer Files' on the left, 'Get Globus Connect Personal Turn your computer into an endpoint.' in the center, and 'RECENT ACTIVITY' with three circular icons on the right. Below the header, there are two identical panels. Each panel has an 'Endpoint' field with a placeholder 'Start here...', a 'Path' field, and a 'Go' button. A blue arrow points from the text 'Search for “OLCF ATLAS”' to the 'Endpoint' field of the left panel. Another blue arrow points from the text 'Search for endpoint at other center' to the 'Endpoint' field of the right panel. The main area of each panel contains the text 'Start by selecting an endpoint.' Below the panels, there is a 'Label This Transfer' section with a text input field and a note 'This will be displayed in your transfer activity.' At the bottom, there are 'Transfer Settings' with five checkboxes: 'sync - only transfer new or changed files', 'delete files on destination that do not exist on source', 'preserve source file modification times', 'verify file integrity after transfer' (which is checked), and 'encrypt transfer'.

Search for endpoint
at other center

Data Transfers: remote data transfers (GlobusOnline)

- Click “Continue” to authenticate to each endpoint (PIN + TokenCode)

An app is requesting access to your account

Globus (<https://www.globus.org>) is requesting temporary delegate access to your account. To approve, please sign in with your OLCF credentials.

OLCF Username

Username

OLCF Passcode

Passcode

(PIN+TOKENCODE)

Transfer Files

Get Globus Connect Personal
Turn your computer into an endpoint.

RECENT ACTIVITY ☐ 0 ☐ 0 ☐ 0

Endpoint

Path

Please authenticate to access this endpoint

When you click continue you will be redirected to the endpoint's login page (you will be returned here once you have authenticated).

Endpoint

Path

Please authenticate to access this endpoint

When you click continue you will be redirected to the endpoint's login page (you will be returned here once you have authenticated).

Label This Transfer

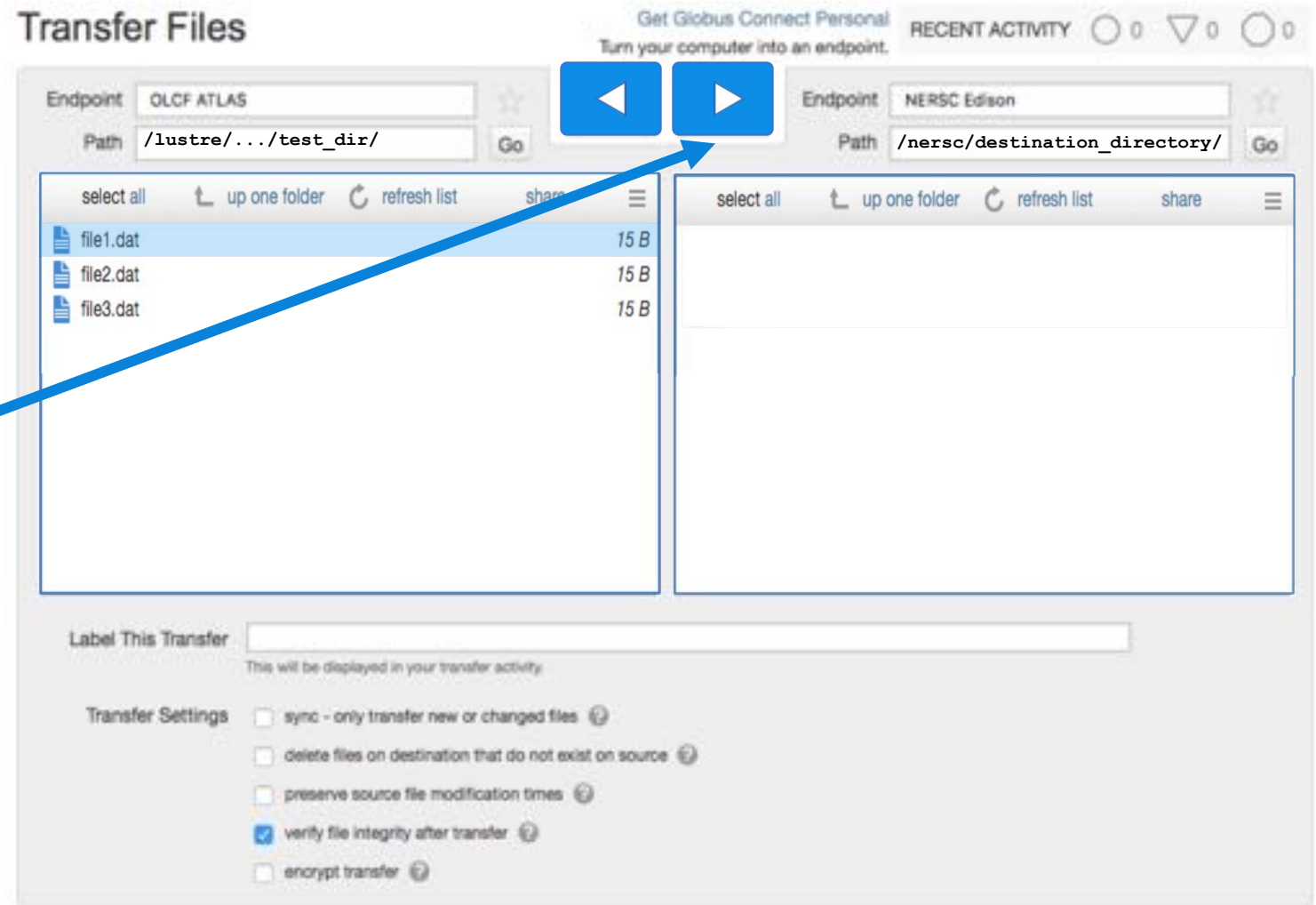
This will be displayed in your transfer activity

Transfer Settings

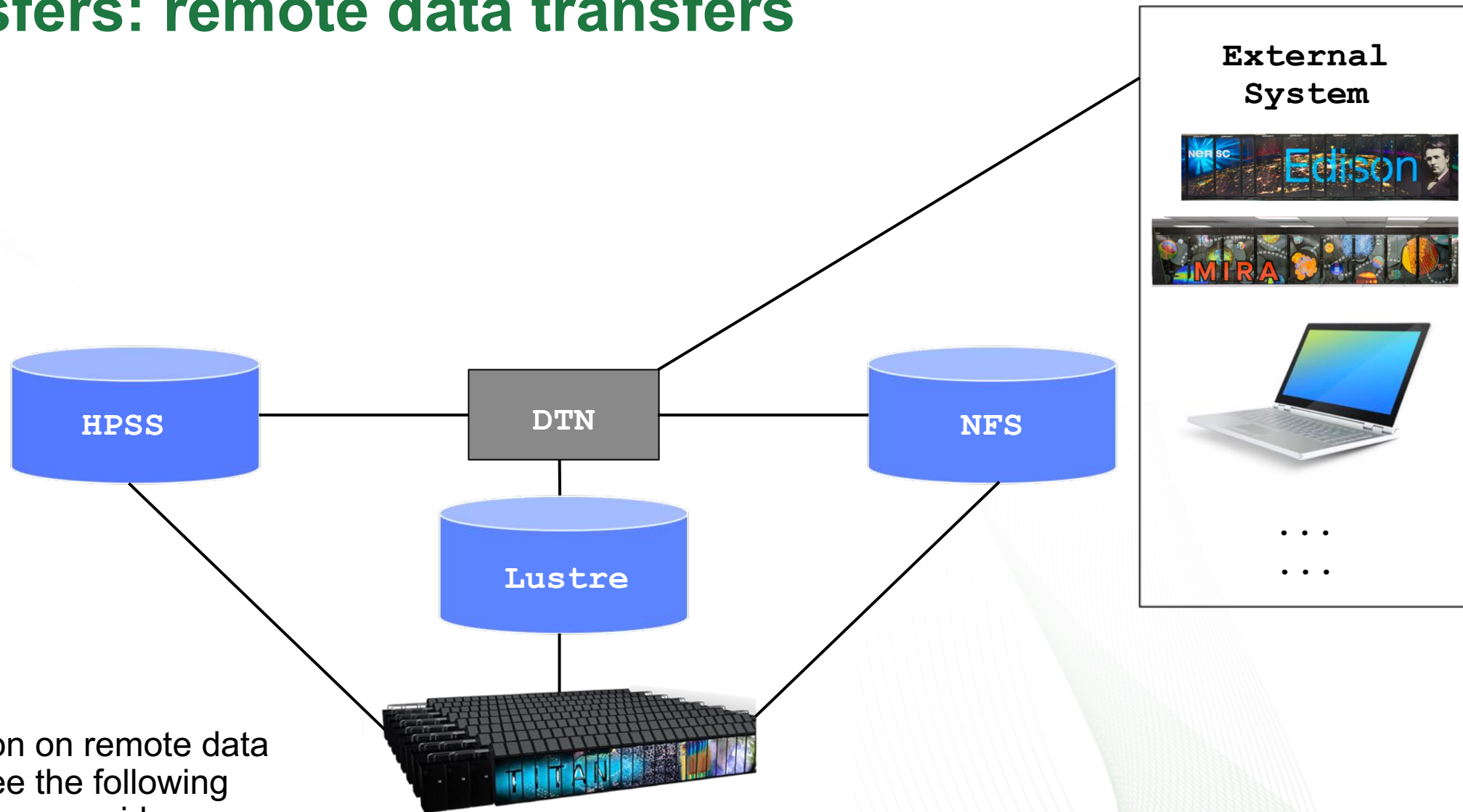
- ☐ sync - only transfer new or changed files ?
- ☐ delete files on destination that do not exist on source ?
- ☐ preserve source file modification times ?
- ☒ verify file integrity after transfer ?
- ☐ encrypt transfer ?

Data Transfers: remote data transfers (GlobusOnline)

- Navigate to the file or directory you want to transfer.
- Navigate to the destination directory you want to transfer to.
- Then click the blue arrow in the direction you want the transfer to occur.
- You will be emailed a notification when it completes.



Data Transfers: remote data transfers



For more information on remote data transfers, please see the following section of the Titan user guide:

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

High Performance Storage System (HPSS)

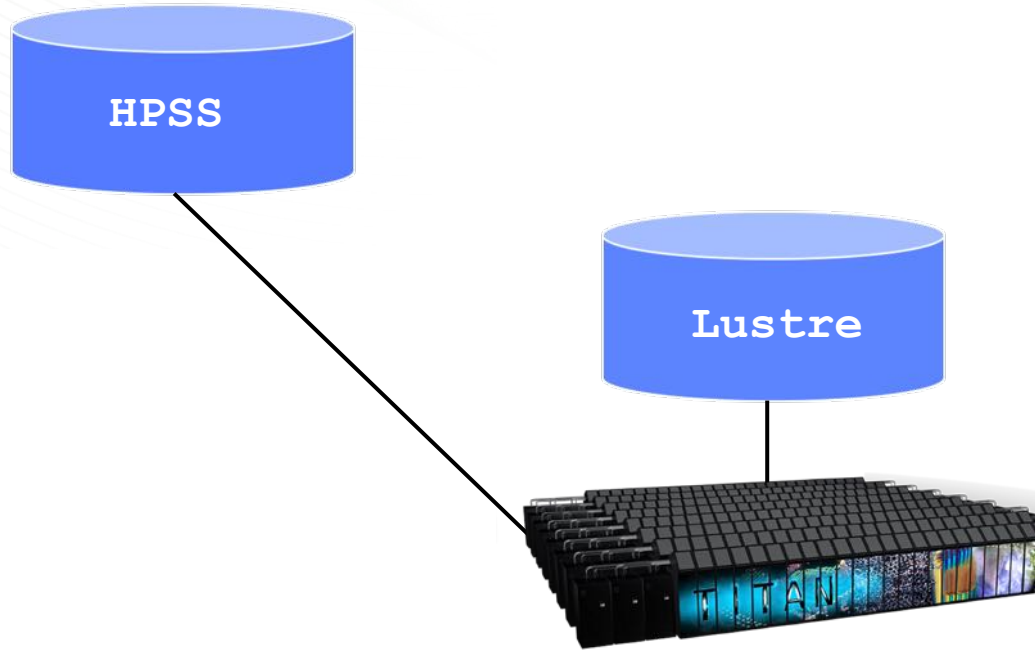
- Long-term archival storage resource
- Robotic tape and disk storage
- `/home/[userid]` & `/proj/[projid]`
- Access using `hsi`, `htar`

NOTE: The training project does not have HPSS access, so you will not be able to try this with training tokens.



Basic HPSS Usage: hsi

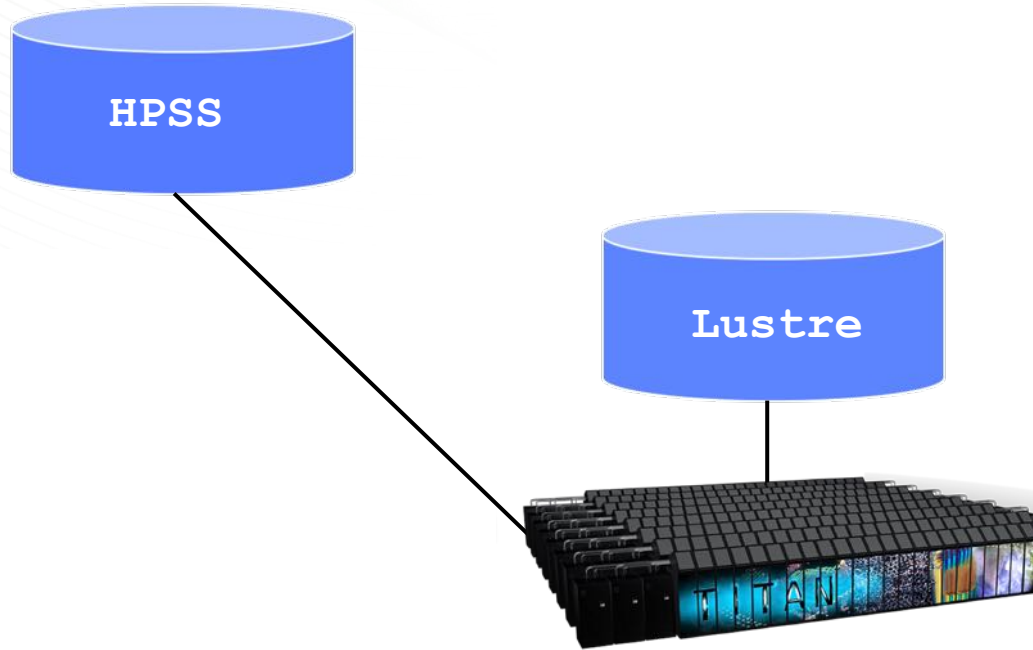
`hsi [options] command(s)`



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt
```

Basic HPSS Usage: hsi

`hsi [options] command(s)`

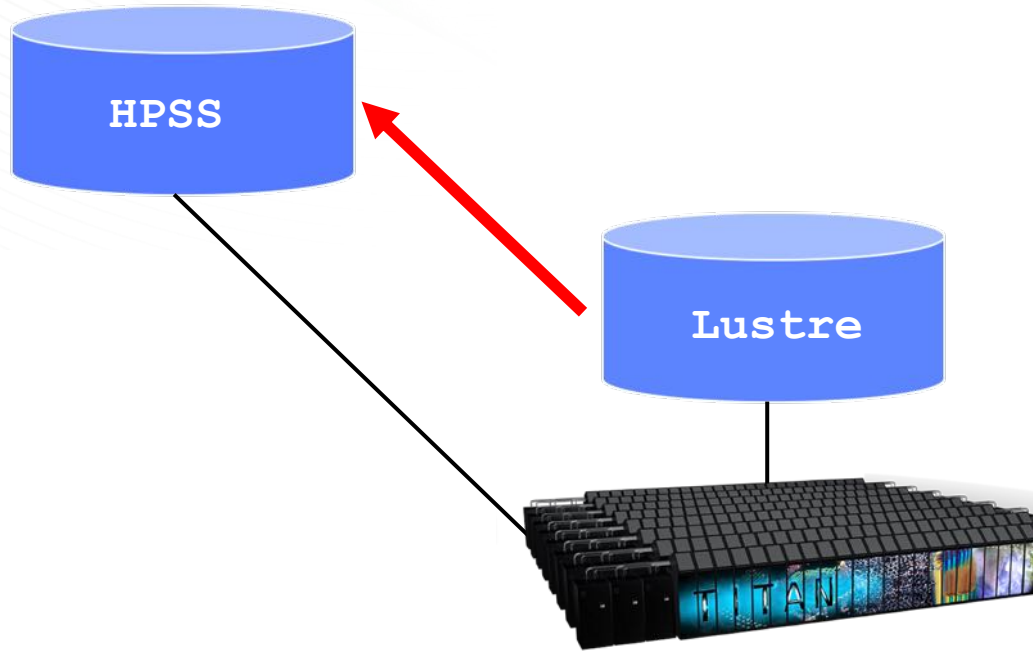


```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar
```

Basic HPSS Usage: hsi

`hsi [options] command(s)`



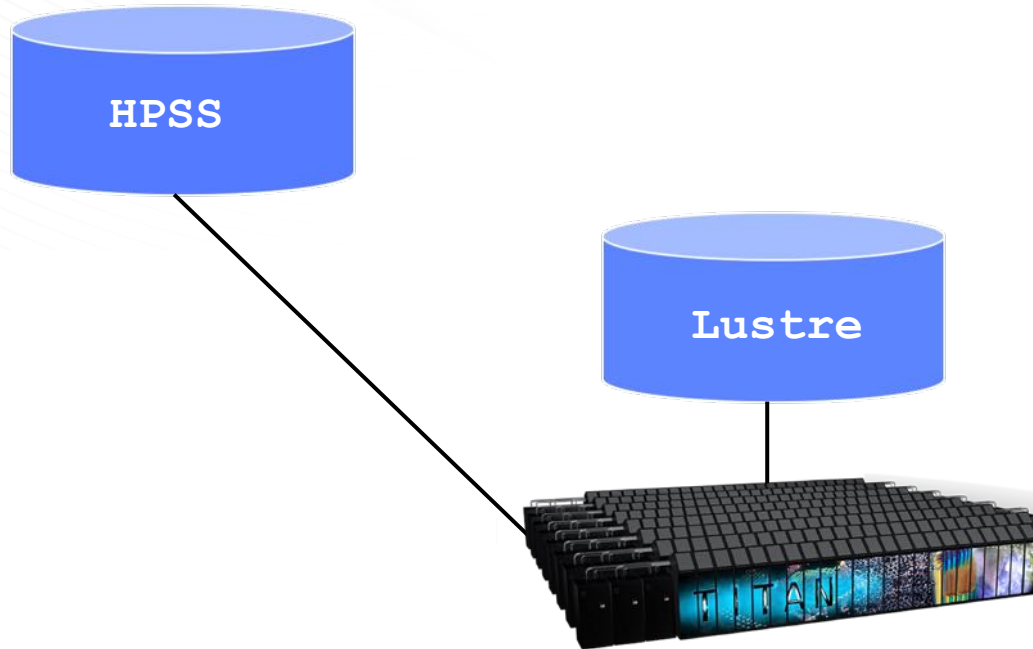
```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar

$ hsi put my_file1.txt
put 'my_file1.txt' : '/home/[userid]/my_file1.txt'
```

Basic HPSS Usage: hsi

`hsi [options] command(s)`



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

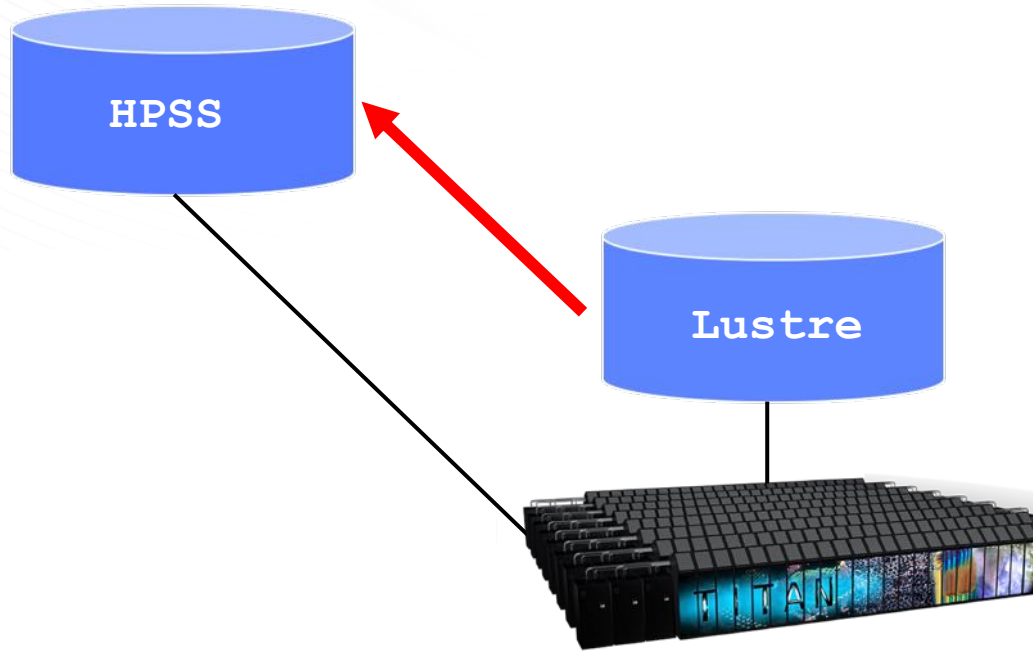
$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar

$ hsi put my_file1.txt
put 'my_file1.txt' : '/home/[userid]/my_file1.txt'

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar
```


Basic HPSS Usage: hsi

`hsi [options] command(s)`



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar

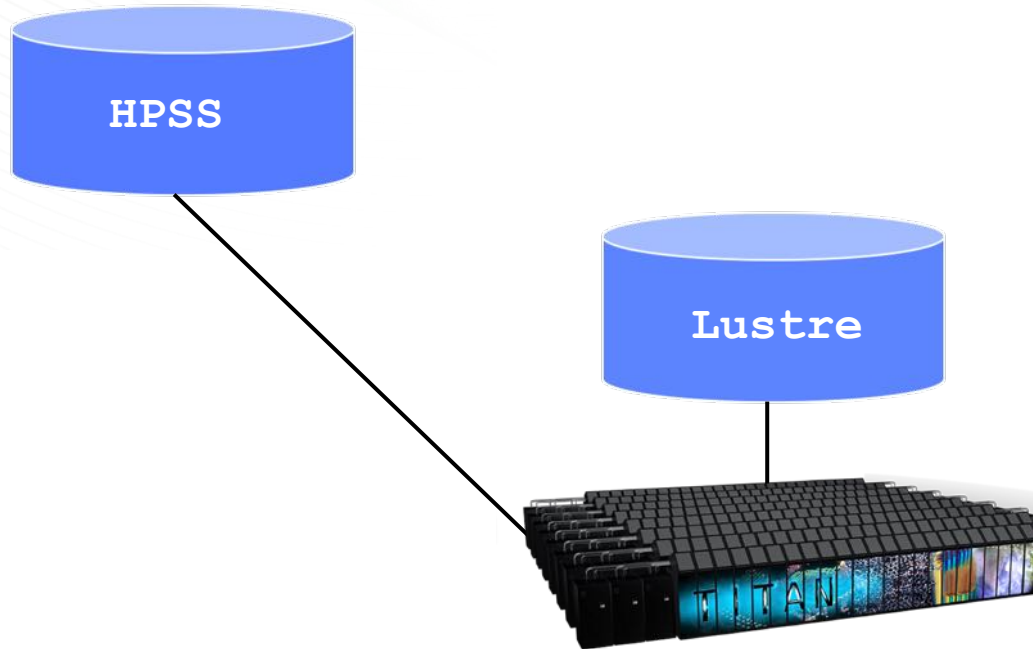
$ hsi put my_file1.txt
put 'my_file1.txt' : '/home/[userid]/my_file1.txt'

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

$ hsi "cd pre-existing_dir; put my_file2.txt"
put 'my_file2.txt' : '/home/[userid]/pre-existing_dir/my_file2.txt'
```

Basic HPSS Usage: hsi

`hsi [options] command(s)`



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar

$ hsi put my_file1.txt
put 'my_file1.txt' : '/home/[userid]/my_file1.txt'

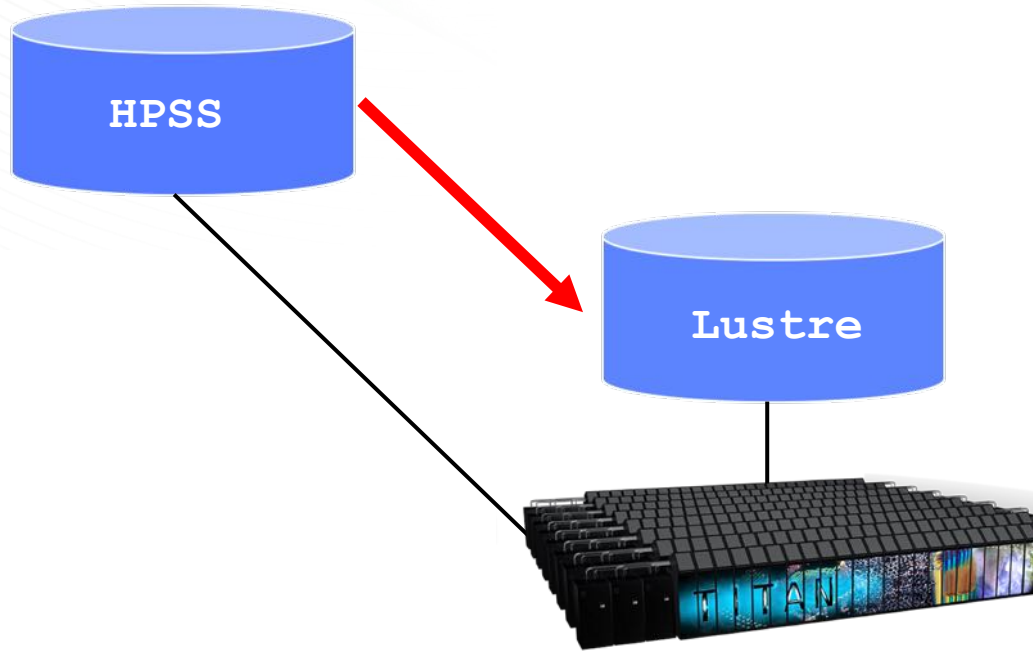
$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

$ hsi "cd pre-existing_dir; put my_file2.txt"
put 'my_file2.txt' : '/home/[userid]/pre-existing_dir/my_file2.txt'

$ hsi ls /home/[userid]/pre-existing_dir
/home/[userid]/pre-existing_dir:
my_file2.txt
```

Basic HPSS Usage: hsi

`hsi [options] command(s)`



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar

$ hsi put my_file1.txt
put 'my_file1.txt' : '/home/[userid]/my_file1.txt'

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

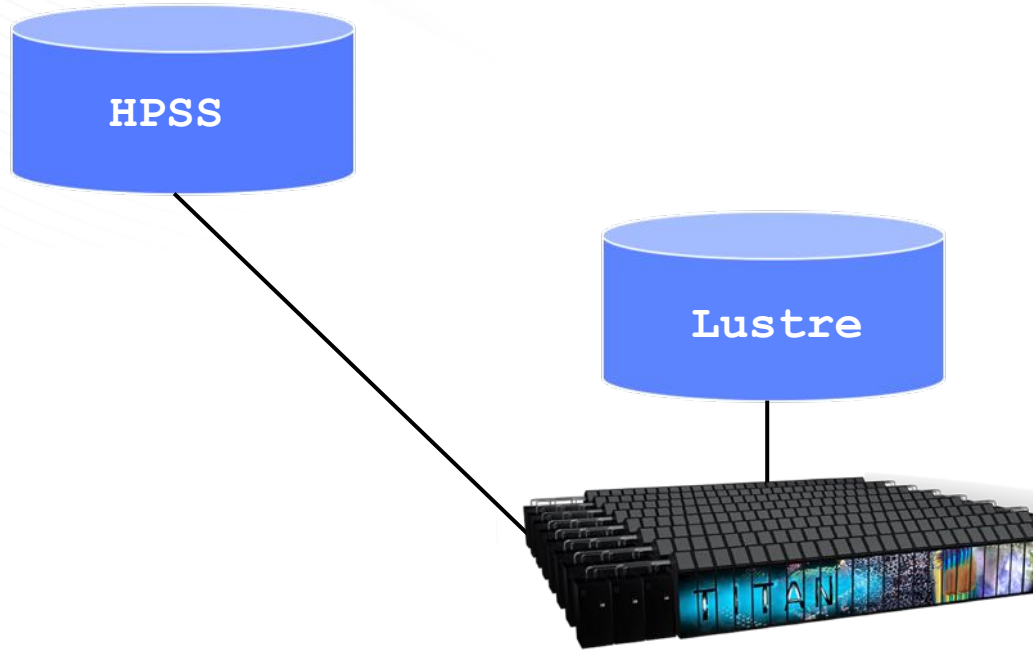
$ hsi "cd pre-existing_dir; put my_file2.txt"
put 'my_file2.txt' : '/home/[userid]/pre-existing_dir/my_file2.txt'

$ hsi ls /home/[userid]/pre-existing_dir
/home/[userid]/pre-existing_dir:
my_file2.txt

$ hsi get sample.py
get '/ccs/home/[userid]/sample.py' : '/home/[userid]/sample.py'
```

Basic HPSS Usage: hsi

`hsi [options] command(s)`



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt

$ hsi ls
/home/[userid]:
pre-existing_dir/  sample.py  some_files.tar

$ hsi put my_file1.txt
put 'my_file1.txt' : '/home/[userid]/my_file1.txt'

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

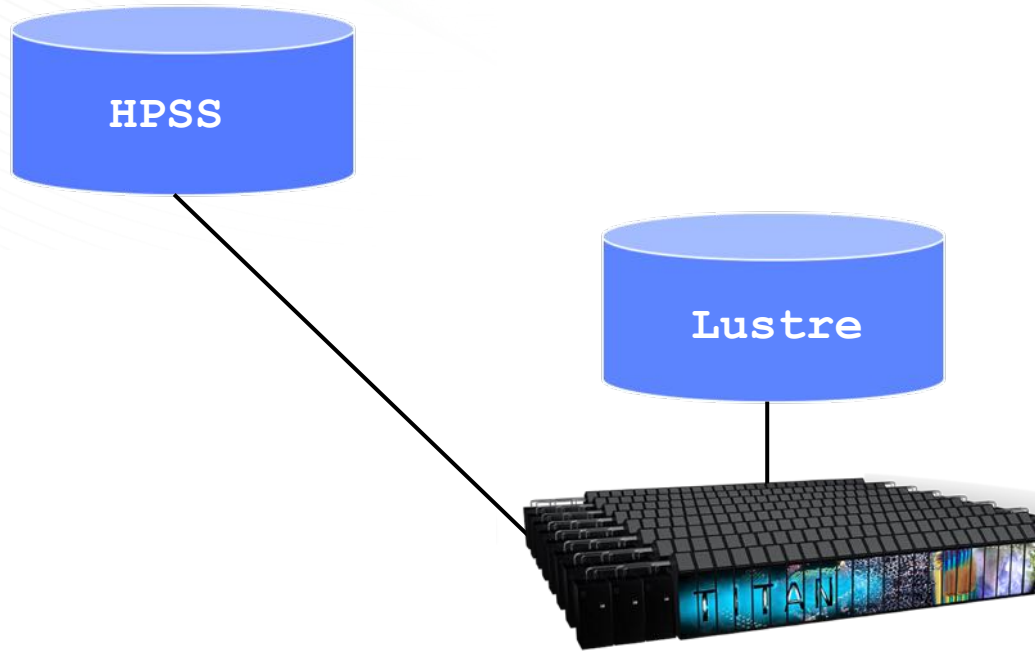
$ hsi "cd pre-existing_dir; put my_file2.txt"
put 'my_file2.txt' : '/home/[userid]/pre-existing_dir/my_file2.txt'

$ hsi ls /home/[userid]/pre-existing_dir
/home/[userid]/pre-existing_dir:
my_file2.txt

$ hsi get sample.py
get '/ccs/home/[userid]/sample.py' : '/home/[userid]/sample.py'

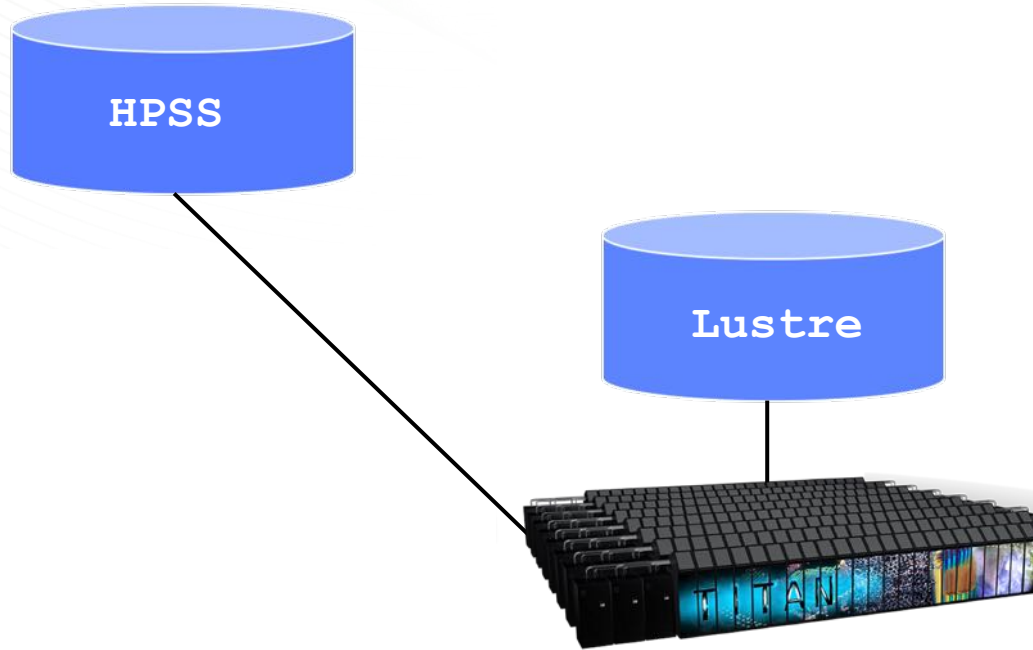
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py
```


Basic HPSS Usage: htar



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py
```

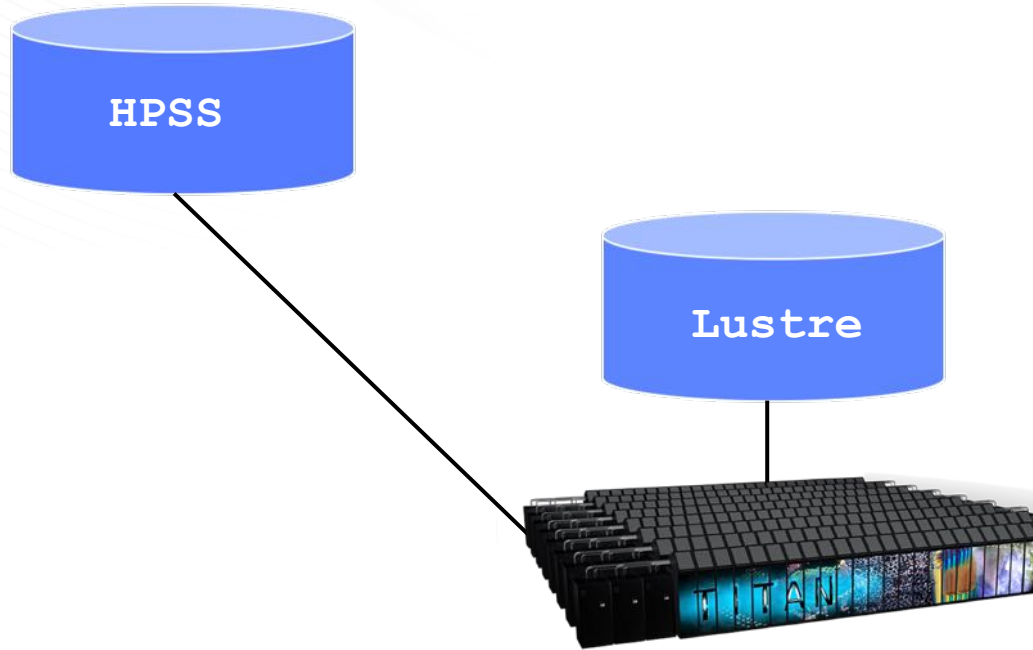
Basic HPSS Usage: htar



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py

$ ls my_dir1
file1.dat  file2.dat
```

Basic HPSS Usage: htar

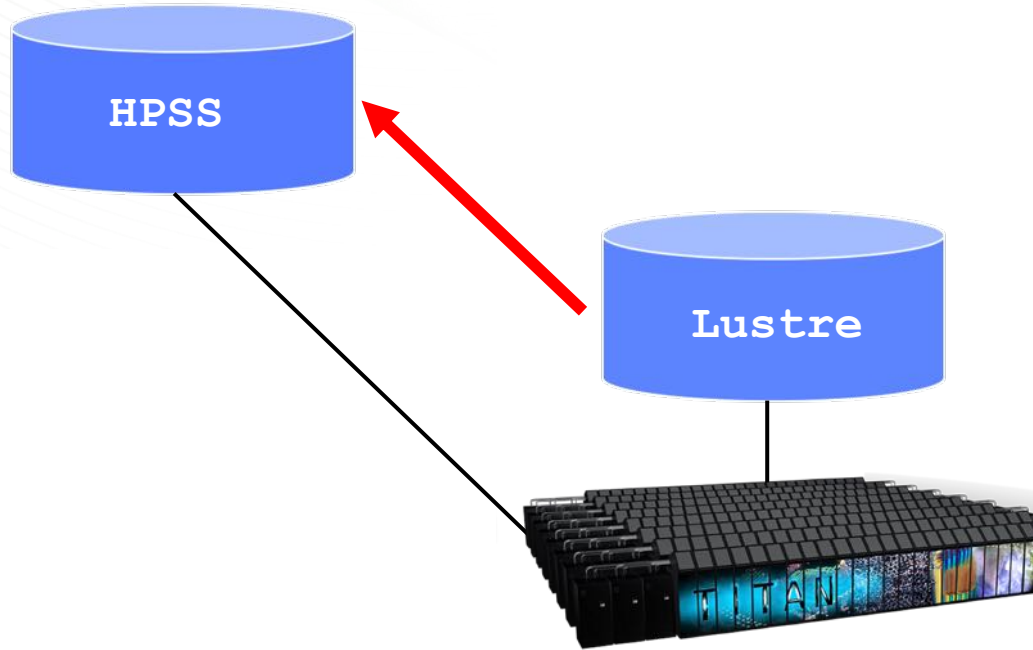


```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py

$ ls my_dir1
file1.dat  file2.dat

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar
```

Basic HPSS Usage: htar



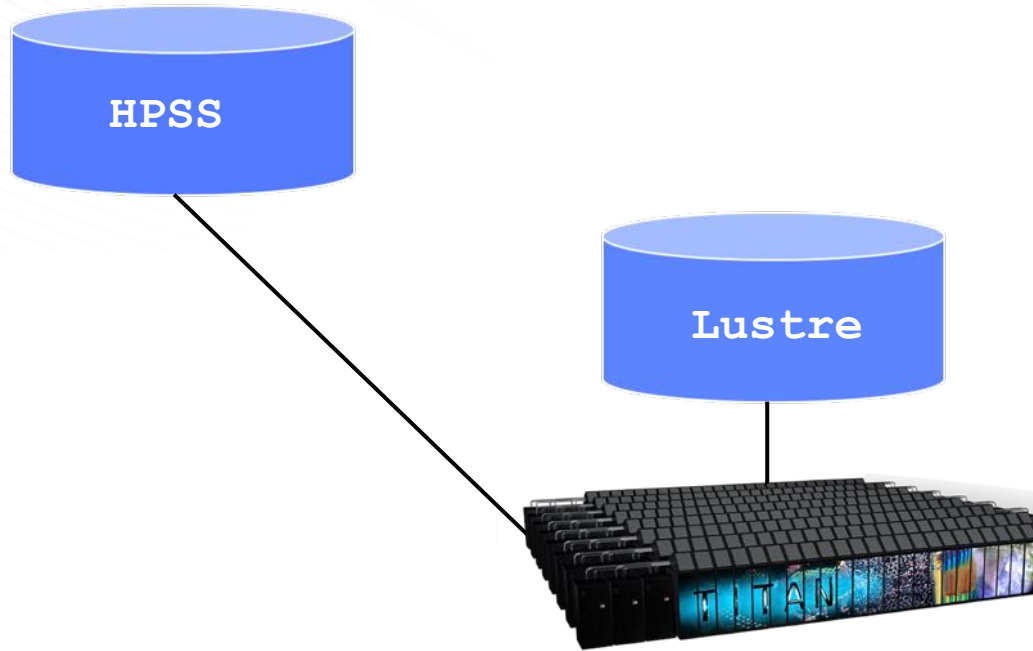
```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py

$ ls my_dir1
file1.dat  file2.dat

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

$ htar -cvf my_dir1.tar my_dir1/
HTAR: a my_dir1/
HTAR: a my_dir1/file1.dat
HTAR: a my_dir1/file2.dat
...
HTAR: HTAR SUCCESSFUL
```

Basic HPSS Usage: htar



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py

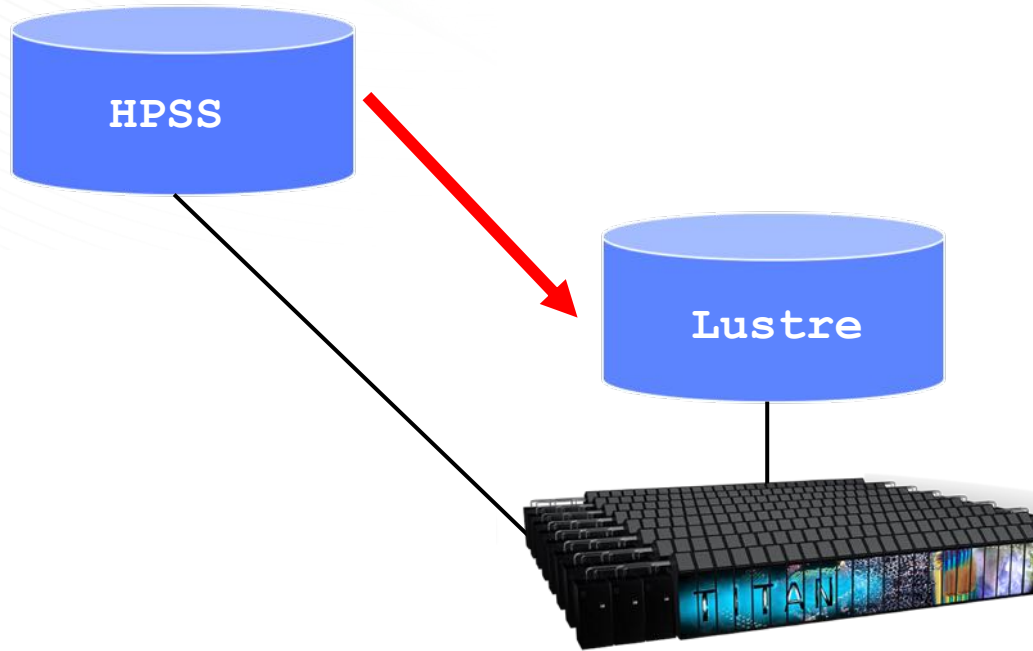
$ ls my_dir1
file1.dat  file2.dat

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

$ htar -cvf my_dir1.tar my_dir1/
HTAR: a my_dir1/
HTAR: a my_dir1/file1.dat
HTAR: a my_dir1/file2.dat
...
HTAR: HTAR SUCCESSFUL

$ hsi ls
/home/[userid]:
my_dir1.tar  my_file1.txt  pre-existing_dir/  sample.py
some_files.tar
```


Basic HPSS Usage: htar



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py

$ ls my_dir1
file1.dat  file2.dat

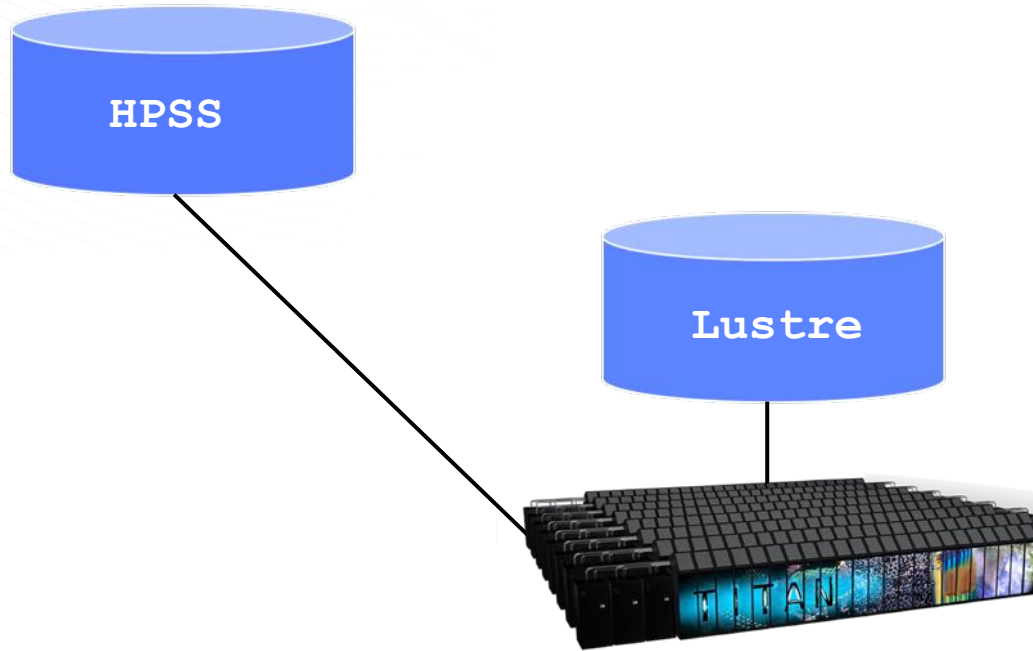
$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

$ htar -cvf my_dir1.tar my_dir1/
HTAR: a my_dir1/
HTAR: a my_dir1/file1.dat
HTAR: a my_dir1/file2.dat
...
HTAR: HTAR SUCCESSFUL

$ hsi ls
/home/[userid]:
my_dir1.tar  my_file1.txt  pre-existing_dir/  sample.py
some_files.tar

$ htar -xvf some_files.tar
HTAR: x some_files/
HTAR: x some_files/some_file1.dat, 15 bytes, 2 media blocks
HTAR: x some_files/some_file2.dat, 15 bytes, 2 media blocks
...
HTAR: HTAR SUCCESSFUL
```

Basic HPSS Usage: htar



```
Titan
$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py

$ ls my_dir1
file1.dat  file2.dat

$ hsi ls
/home/[userid]:
my_file1.txt  pre-existing_dir/  sample.py  some_files.tar

$ htar -cvf my_dir1.tar my_dir1/
HTAR: a my_dir1/
HTAR: a my_dir1/file1.dat
HTAR: a my_dir1/file2.dat
...
HTAR: HTAR SUCCESSFUL

$ hsi ls
/home/[userid]:
my_dir1.tar  my_file1.txt  pre-existing_dir/  sample.py
some_files.tar

$ htar -xvf some_files.tar
HTAR: x some_files/
HTAR: x some_files/some_file1.dat, 15 bytes, 2 media blocks
HTAR: x some_files/some_file2.dat, 15 bytes, 2 media blocks
...
HTAR: HTAR SUCCESSFUL

$ ls
my_dir1/  my_dir2/  my_file1.txt  my_file2.txt  sample.py
some_files/
```

High Performance Storage System (HPSS)



For more information on HPSS, please see the following section of the Titan user guide:

<https://www.olcf.ornl.gov/for-users/system-user-guides/titan/file-systems/#hpss-high-performance-storage-system>

Additional Information

- Titan User Guide – File Systems: Data Storage & Transfers
<https://www.olcf.ornl.gov/for-users/system-user-guides/titan/file-systems/>
- File Systems: Data Storage & Transfers – Lustre File System
<https://www.olcf.ornl.gov/for-users/system-user-guides/titan/file-systems/#spider-the-centerwide-lustre-file-system>
 - Lustre Basics, File Striping, Best Practices



Thank You.