

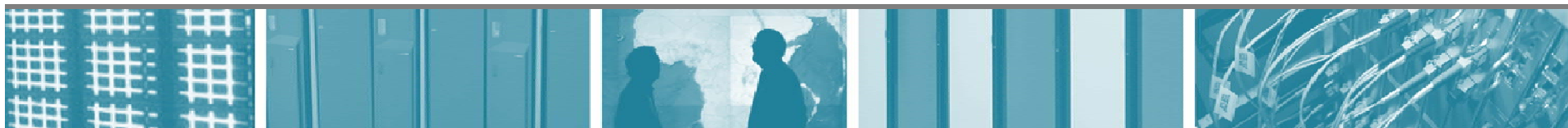
# Storage at OLCF



*Presented by:*

**Mitchell Griffith**

Oak Ridge Leadership Computing Facility (OLCF)



U.S. DEPARTMENT OF  
**ENERGY**



**OAK RIDGE NATIONAL LABORATORY**

MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

# OLCF Storage Policy

- <http://www.olcf.ornl.gov/support/user-guides/olcf-policy-guide/#396>
- Manages Storage Resources in the OLCF
- Give fair share of storage resources to projects and users
- **Project or user can request exceptions to the policy**
  - <http://www.olcf.ornl.gov/support/getting-started/special-request-form/>
- There are 3 types of storage resources at OLCF
  - NetApp
  - Lustre
  - HPSS

# NetApp (NFS shared areas)

- Shared home (/ccs/home/\$USER)
- Shared project (/ccs/proj/[projectid])
- Has a hard quota
- Limits
  - /ccs/home/\$USER 5GB
    - quota/quota -s
  - /ccs/proj/[projectid] 50GB
    - df -h /ccs/proj/[projectid]

# User Work Directories (lustre)

- Work directories
  - /lustre/widow0/scratch/\$USER
  - /lustre/widow1/scratch/\$USER
  - /lustre/widow2/scratch/\$USER
  - /lustre/widow3/scratch/\$USER
  - Only one is the 'primary' work directory
    - /tmp/work/\$USER
    - \$WORKDIR
  - No default quota, but all files not accessed in 14 days are eligible to purge
- Each lustre filesystem is 2.5PB in size.

# Project Lustre Space

- /tmp/proj/[projectid]
- 5TB soft quota
- Quotas enforced via email
- No backup, catastrophic failure means data is lost
- Project space is on one of the 4 available lustre filesystems
  - /tmp/proj/[projectid] is a symlink to /lustre/widow[0-3]/proj/[projectid]

# Lustredu

- Lustredu
  - Provide lustre usage
  - Does not tax the lustre metadata server
  - Module loaded by default
  - Available only on machines that mount /lustre/widow[0-3]

```
[user@dtn01.ccs.ornl.gov:/ccs/home/user] lustredu /tmp/work/$USER
```

```
Last Collected Date          Size      File Count  Directory
2013-01-22 11:30:02      27.84 TB          120  /lustre/widow2/scratch/user
```

```
[user@dtn01.ccs.ornl.gov:/ccs/home/user]
```

# HPSS (archival storage)

- What is HPSS?
  - HPSS is software that manages petabytes of data on disk and robotic tape libraries. HPSS provides highly flexible and scalable hierarchical storage management that keeps recently used data on disk and less recently used data on tape.
  - Hierarchical storage system
  - Access time can be slow if accessing from tape
    - HSI get command: `get a.pdf`
    - HSI output: Scheduler: retrieving file(s)
    - Waiting on a tape drive (resource contention)

# HPSS (archival storage)

- Home areas (/home/\$USER)
  - 2TB size/2,000 files limit
- Project Areas (/proj/[projectid])
  - 100TB size/100,000 file limit
- Showusage
  - Show usage for user and project on hpss

```
[user@home2.ccs.ornl.gov:/ccs/home/user] showusage -s hpss
```

HPSS Storage in GB:

Project	Project Totals Storage	user Storage
user	2090.03	2090.03
abc123	6.52	0.01
abc123pri	0.10	0.10
def456	561.83	0.00



# HPSS Layout

- (5) SL8500's
  - 10,000 tape slots per silo
- 38,944 tapes in use
  - Tape capacity (500GB, 1TB, 5TB)
- 112 tape drives
  - 16 T10K-A
  - 60 T10K-B
  - 36 T10K-C

# HPSS Layout (2)

- Disk Cache
  - Disk cache is striped(multiple disks, faster device)
  - Upgrade disk cache this year (660TB in disk cache)
  - (12) disk movers in production
- Tape
  - (17) production movers

# HPSS Layout

- Class of Service (COS)
  - Single copy COS Iscos

COS	Name	Min	Max
5081	Xsmall	0	131,071 (128K)
6001	Small	131,072 (128K)	16,777,215 (16M)
6002	Medium	16,777,216 (16M)	536,870,911 (512M)
6054	Large_T	536,870,912 (512M)	8,589,934,591 (8G)
6056	X-Large_T	8,589,934,591 (8G)	281,474,976,710,656 (256T)

- Single copy class of service is default.
  - [http://www.olcf.ornl.gov/kb\\_articles/hsihtar/#Setting\\_the\\_Number\\_of\\_Copies](http://www.olcf.ornl.gov/kb_articles/hsihtar/#Setting_the_Number_of_Copies)

# HPSS Interfaces

- Archival Space/HPSS
  - 2 interface options
    - HSI
    - HTAR
  - Preferred only 1 stream at a time
  - Batch options

# HSI Batch

```
in.hpss
```

```
get <<in  
a  
B  
c  
in
```

```
O:[hpss-nccs]/home/user/a-> in in.hpss
```

```
get <<in
```

```
get 'a' : '/home/user/a/a' (2012/10/08 07:03:41 5 bytes, 12.0 KBS )
```

```
get 'b' : '/home/user/a/b' (2012/10/08 07:03:44 5 bytes, 12.0 KBS )
```

```
get ,c' : '/home/user/a/c' (2012/10/08 07:03:44 5 bytes, 12.0 KBS )
```

```
O:[hpss-nccs]/home/user/a->
```

# HPSS Tips

- HPSS is archival storage
  - Many very small files are bad for HPSS
  - Too large of files are bad as well (disk cache fills up)
  - Optimal file size for HPSS is between 2GB and 256GB.
  - Use HTAR to archive several smaller files into one
    - A member file of an HTAR file cannot be > 64GB
    - Lustredu can be used to estimate the size of a source directory in lustre

# Authentication Issue

> hsi

result = -11000, errno = 0g]

Unable to authenticate user with HPSS.

result = -11000, errno = 9

Unable to setup communication to HPSS...

\*\*\* HSI: error opening logging

Error - authentication/initialization failed

## Workaround

➤ hsi -A combo

➤ Send email to [help@olcf.ornl.gov](mailto:help@olcf.ornl.gov)

# Data Removal

- Please remove any unwanted data from all areas (NFS, Lustre, and HPSS)
- When projects end, data is subject to deletion after 3 months after the project ends
  - OLCF only reserves the right to delete
  - Projects should remove data from OLCF resources they do not want potentially shared.
- When a user account ends, data is subject to deletion after 3 months
  - OLCF only reserves the right to delete data
- We are working to automate account and project cleanup



# Data Movement (offsite)

- [http://www.olcf.ornl.gov/kb\\_articles/employing-data-transfer-nodes/](http://www.olcf.ornl.gov/kb_articles/employing-data-transfer-nodes/)
- SCP/SFTP
  - Can be slow, but this method is supported on most clients
- BBCP
  - Faster file copy than scp
  - Requires the bbcp binary at source and destination
- GridFTP
  - Uses DOE grid certificates or SSH for authentication
  - Recommended way to transfer from DOE labs (NERSC, ANL, etc).

# Questions?