Overview of the New Fortress Archive

Presented by: Ramon L. Williamson II
Senior Storage Engineer
• Information Technology at Purdue (ITaP) is upgrading the research computing archival storage system Fortress. Currently based upon EMC's DiskXtender (DXUL), Fortress is being upgraded to new, more powerful software: the High Performance Storage System (HPSS)

• As part of the upgrade, metadata from the current archive will be converted such that it can be imported into the new system, so all user data in the current archive will be present in the new archive transparently.
Why change archive systems?
The reason for changing to a new system are two-fold:

1) The current archive software DiskXtender Unix/Linux (DXUL) was end-of-lifed by the company that offered it as of August 2011 and they no longer support the code.

2) Archive usage at Purdue has increased 10-fold in the past 5 years, stretching the capabilities of DXUL to its limit at times. A new archive solution was already being sought when the end-of-life announcement was made.
Why HPSS?
A search was made of available archival systems to find a replacement for DXUL. Solutions were ranked by

1) Scalability
2) Secure and efficient interfaces
3) Ability to provide high performance access to archival data in excess of 1PB
4) Ability to import data from the old system to the new one with the least inconvenience to the end user.
HPSS was chosen as the archive replacement for many reasons:

1) Guided by IEEE Mass Storage Reference Model version 5
2) Written in ANSI C
3) Modular design
4) Secure interfaces
5) Active support by a team of engineers and programmers from a consortium of IBM and National Labs.
6) Conversion of DXUL metadata was offered which provided for a transparent migration path to end users.
The World of HPSS
Accessing Fortress
The information here will be available on the Fortress web site:

http://www.rcac.purdue.edu/userinfo/resources/fortress/userguide.cfm
Kerberos Authentication

• All current users of Fortress will have accounts, but will need a Kerberos keytab to provide passwordless access to the archive.
• To get a keytab, go to the web page:

https://web.rcac.purdue.edu/hpss/krbinit.cgi
• Type your username and career account password, and a download of your keytab will commence.
• Create a directory in your home directory called .private and place the keytab there, making sure it is named hpss.keytab and owner rw access only.
• Redo this operation when you change your Career account password.

Be sure to visit the Fortress user guide for the most up-to-date information!
Access Methods

**HSI**

HSI – The Hierarchical Storage Interface – provides a Unix-like environment for working with files within HPSS. It can be run interactively or in batch mode.

HSI utilities are installed on all cluster resources including front ends and nodes. On non-RCAC resources, binaries can be provided for HSI (currently only Red Hat, but others can be compiled).
ramonw@hansen-sys ~> /opt/hsi/bin/hsi

Purdue University
High Performance Storage System (HPSS)

This is the Purdue Data Archive, Fortress. For further information see http://www.rcac.purdue.edu/userinfo/resources/fortress/

If you are having problems with HPSS, please call IT/Operational Services at 49-44000 or send E-mail to dxul-help@purdue.edu.

NOTICE: This system is still in the testing phase and may experience outages and resets.

IT'S QUITE POSSIBLE THAT YOUR DATA MAY BE LOST!

Username: ramonw  UID: 168207  Acct: 168207(168207) Copies: 1  Firewall: off [hsi 3.5.8 Wed Sep 21 17:31:14 EDT 2011] [Fortress HSI]/home/ramonw->
Commands for accessing HPSS are much like UNIX commands: ls, cd, mkdir, cp, mv, touch, find

Access to the local filesystem is through commands similar (but different) to FTP Commands: ll s, lcd, lpwd, get, put, mget, mput, !<command>

Commands are entered as in a command shell; when done, type exit
HSI – Continued

In batch mode, you can use HSI by placing semicolon-delimited commands after the HSI command:

% hsi command1; command2; command3...

Also, you can put the commands in a file and redirect into HSI:

% hsi < command.list
ramonw@hansen-sys ~> /opt/hsi/bin/hsi -q "lcd data_results; mkdir data_results; cd data_results; put *.fits"
mkdir: /home/ramonw/data_results
put 'bs15621_10p.fits' : '/home/ramonw/data_results/bs15621_10p.fits' (1540800 bytes, 40655.4 KBS (cos=10))
put 'bs15621_10x.fits' : '/home/ramonw/data_results/bs15621_10x.fits' (552960 bytes, 22212.6 KBS (cos=10))
put 'bs15621_1p.fits' : '/home/ramonw/data_results/bs15621_1p.fits' (1540800 bytes, 67502.1 KBS (cos=10))
put 'bs15621_1x.fits' : '/home/ramonw/data_results/bs15621_1x.fits' (552960 bytes, 27293.1 KBS (cos=10))
ramonw@hansen-sys ~>
ramonw@hansen-sys data_results> tar -cf - *.fits | /opt/hsi/bin/hsi -q put - : results.tar
put 'stdin': '/home/ramonw/results.tar' (4198400 bytes, 57477.6 KBS (cos=10))
ramonw@hansen-sys data_results>
HTAR

HTAR is another add-on utility that speaks natively to the HPSS System. Command syntax for most usage looks like the UNIX tar command, the difference being that the tar file being acted upon (creation or extraction) resides in the archive. Direct system calls and buffering algorithms optimize the performance.

Upon creation, a tar file is created in the archive and an index file to provide faster access to files in the tar upon

-Hcrc flag allows for checksumming the tar file on creation or For use with the –t flag.
ramonw@hansen-adm data_results> /opt/hsi/bin/htar -Hcrc -cvf data_results.tar *

HTAR: a   bs15621_10p.fits
HTAR: a   bs15621_10x.fits
HTAR: a   bs15621_1p.fits
HTAR: a   bs15621_1x.fits
HTAR: a   /tmp/HTAR_CF_CHK_19528_1317698780
HTAR Create complete for data_results.tar. 4,191,744 bytes written for 4 member files, max threads: 6 Transfer time: 0.071 seconds (59.247 MB/s)
HTAR: HTAR SUCCESSFUL

ramonw@hansen-adm data_results> /opt/hsi/bin/htar -Hcrc -tvf data_results.tar

HTAR: -rw-r--- [0x088cb91e] ramonw/pucc 1540800 2011-10-03 17:15 bs15621_10p.fits
HTAR: -rw-r--- [0x209266d8] ramonw/pucc 552960 2011-10-03 17:15 bs15621_10x.fits
HTAR: -rw-r--- [0xb65a3103] ramonw/pucc 1540800 2011-10-03 17:15 bs15621_1p.fits
HTAR: -rw-r--- [0x1d9b87b7] ramonw/pucc 552960 2011-10-03 17:15 bs15621_1x.fits
HTAR: -rw------- [0x6042c521] ramonw/pucc 256 2011-10-03 23:26 /tmp/HTAR_CF_CHK_19528_1317698780
HTAR Listing complete for data_results.tar, 5 files 5 total objects
HTAR: HTAR SUCCESSFUL

ramonw@hansen-adm data_results>
FTP Interfaces

• Only kerberized ftp clients can connect to the archive:
  • Kerberized ftp
    • Usually in /usr/kerberos/bin
    • Not optimized for HPSS use; may be missing some commands needed for HPSS
  • /opt/hpss/bin/pftp_client
    • Compiled as part of HPSS as a client
    • Installed on all RCAC resources
    • Available for non-RCAC resources
    • Optimized for use with HPSS and can multithread transfers to optimize throughput
FTP Interfaces Continued

Note that the port to connect to is 4021, not the usual 21.

To use either ftp, you need to create a kerberos ticket to use it. This is done with the keytab file you downloaded earlier:

% kinit -k -t ~ramonw/.private/hpss.keytab ramonw
Connected to havoc.rcac.purdue.edu.
220-# HPSS 7.3 Parallel FTP Daemon on havoc.rcac.purdue.edu
# coming from hansen-sys.rcac.purdue.edu
#

220 havoc.rcac.purdue.edu FTP server (HPSS 7.3 PFTPD V1.1.1 Mon Sep 19 10:46:46 EDT 2011) ready.
Parallel stripe width set to (1).
334 Using authentication type GSSAPI; ADAT must follow
GSSAPI accepted as authentication type
GSSAPI authentication succeeded
Preauthenticated FTP to havoc as ramonw:
232 GSSAPI user ramonw@CORE.RCAC-HPSS.PURDUE.EDU authenticated
230 User ramonw@CORE.RCAC-HPSS.PURDUE.EDU logged in as ramonw@CORE.RCAC-HPSS.PURDUE.EDU
Remote system type is UNIX.
Using binary mode to transfer files.
***** NOTE: FTP Daemon supports feature discovery *****
***** NOTE: Protocol set to PDATA_AND_MOVER *****
***** NOTE: Daemon supports Transfer Agent
***** NOTE: Server supports Parallel Features *****
 ***** Auto-Parallel Substitution Enabled. *****
***** NOTE: PFTP Server supports 64-bit protocol
ftp>
ftp> lcd data_results
Local directory now /autohome/u100/ramonw/data_results
ftp> cd data_results
250 CWD command successful.
ftp> mput *.*fits
mput bs15621_10x.fits? y
local: bs15621_10x.fits remote: bs15621_10x.fits
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for bs15621_10x.fits.
226 Transfer complete.
552960 bytes sent in 0.0445 seconds (11.853 MBytes/sec)

ftp> mput bs15621_1x.fits? y
local: bs15621_1x.fits remote: bs15621_1x.fits
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for bs15621_1x.fits.
226 Transfer complete.
552960 bytes sent in 0.0297 seconds (17.746 MBytes/sec)

ftp> mput bs15621_10p.fits? y
local: bs15621_10p.fits remote: bs15621_10p.fits
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for bs15621_10p.fits.
226 Transfer complete.
552960 bytes sent in 0.0445 seconds (11.853 MBytes/sec)
mpput bs15621_1x.fits? y
local: bs15621_1x.fits remote: bs15621_1x.fits
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for bs15621_1x.fits.
226 Transfer complete.
552960 bytes sent in 0.0297 seconds (17.746 MBytes/sec)
mpput bs15621_10p.fits? y
local: bs15621_10p.fits remote: bs15621_10p.fits
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for bs15621_10p.fits.
226 Transfer complete.
1540800 bytes sent in 0.0668 seconds (21.999 MBytes/sec)
mpput bs15621_1p.fits? y
local: bs15621_1p.fits remote: bs15621_1p.fits
200 Command Complete.
200 PORT command successful.
150 Opening BINARY mode data connection for bs15621_1p.fits.
226 Transfer complete.
1540800 bytes sent in 0.0563 seconds (26.082 MBytes/sec)
ftp>
CIFS

- As with the current archive you can access HPSS via CIFS, or SAMBA your Career account password is used. The only difference here is that a server of a different hostname is used to serve CIFS.
- To connect to your HPSS home via CIFS use the following address:

  smb://entropy.rcac.purdue.edu/homes

- This access method is not optimized for HPSS, and some HPSS functions may not work with this method.
- Be aware accessing or inputting many small files is not suggested for this access method and that some characters in filenames may not make the transition to HPSS.
NFS

Due to performance issues both on client and server sides, access via NFS will be restricted to RCAC cluster front ends and data.rcac.

SFTP/SCP

Since there are no direct logins into the HPSS server, access via these methods will be limited to the cluster front ends and data.rcac via the NFS mountpoint.

It is suggested that you use HSI and HTAR for access to the archive.
Caveats
Trashcans

DXUL included the trashcan utility, which placed recently deleted files in a .trash directory. These files remained for several hours so that users could retrieve files deleted by mistake.

HPSS does not have such a facility. Files, once deleted, cannot be restored. We suggest you use the “-i” flag to the rm command to avoid deleting needed files.
Server Names

Fortress has been the name of the archival service at Purdue for many years.

HPSS uses a distributed approach with single-purpose servers acting together to provide all of the archive services.

As such, we have 5 different servers to provide the different Services:
Havoc – Core Server and FTP access
Bedlam and Turmoil – Data Movers (cache to tape, tape to cache)
Tumult – NFS Services
Entropy – CIFS/Samba Services

FTP services will still be accessible by the Fortress address as well (although you may see havoc mentioned at login).
Paths

The directory structure for the archive was retained, so that a user home directory is still:

/archive/fortress/home/username

However, in HPSS, the default is just

/home/username

In some access methods, you may see the latter designation for your current directory. Soft links are in place so that the old path will work.
Small files

HPSS is much better at handling small files than DXUL. But...

- Small files burden the metadata database
- So do directory listings containing thousands of files

Try and use the HTAR facility if you need to ingest lots of small files.
Class of Service

HPSS uses Classes of Service (COS) for handling different use cases for files, such as number of tape copies, cache retention, tape handling, resource management.

We currently only have implemented 2 COSes – 1 for small files and One for everything else. These are handled automatically.

We may add new COSes in the future as usage patterns changes, e.g. special archiving projects, so handling COSes will be an added skill you might need in the future.
Getting Help
Site-Specific Help

Fortress Web page:
http://www.rcac.purdue.edu/userinfo/resources/fortress/userguide.cfm

RCAC User Support:
Email to rcac-help@purdue.edu or Phone: 49-44000

HPSS

http://www.hpss-collaboration.org
HSI/HTAR

http://www.mgleicher.us