

Indiana University

Bloomington, Indiana

HPSS 2004 User Forum
May 4-6, 2004
Sante Fe, New Mexico

Our Group

- Manager:
Andrew Arenson, aarenson@iupui.edu
- User Support (DFS, pftp, hsi, etc):
Haichuan Yang, haiyang@indiana.edu
Kristy Kallback-Rose, kallbac@indiana.edu
- HPSS Administrators:
Chris Garrison, ecgarris@iupui.edu
Jeff Russ, russ@indiana.edu
- DCE Administrator:
Haiyan Li (haiyli@indiana.edu)
- Our Web Site: <http://storage.indiana.edu>

Current Software

- HPSS 4.5
- AIX 5.1 ML04
- Encina 5.0 + PTF 3
- DCE 3.2 + PTF 4
- DFS 3.1 + PTF 4
- Sammi
- Gresham STK Driver (fast-locate feature)

Distributed HPSS

- HPSS is distributed between Bloomington and Indianapolis. Connected via high speed fiber network.
- SFS and core servers in Bloomington.
- Movers and DFS/HDM run at both sites.
- Dual copy classes of service store the second copy at the remote site for disaster recovery.

Current Hardware

- **Bloomington:**
 - SP2 with Silver nodes:
 - 1 wide node for SFS and the core servers
 - 3 wide nodes for disk/tape movers
 - 4 thin nodes for DFS/HDM
- **Indianapolis:**
 - 1 H70 for disk mover and DFS/HDM
 - 5 B80s for disk and tape movers

Tape Hardware

- **Bloomington:**
 - STK 9310 Tape Silo
 - 6 SCSI 9840 tape drives, 4000 tapes
 - 6 FC 9940B tape drives, 1000 tapes
 - IBM 3494 Tape Library
 - 1 3590 tape drive for metadata backups
- **Indianapolis:**
 - STK 9310 Tape Silo
 - 12 SCSI 9940A tape drives, 2000 tapes
 - 6 FC 9940B tape drives, 1000 tapes
 - IBM 3494 Tape Library
 - 1 3590 tape drive for metadata backups

The Users

- 1100 users throughout Indiana:
 - Bloomington Campus: 975
 - Indianapolis Campus: 170
 - Regional Campuses: 25
- 180 TB of data stored
- 7 million files
- Users get a quota of 500 GB by default

What HPSS Is Used For

- Research: Astronomy, Biology, Chemistry, Geology and Physics.
- Other: Anthropology, Business, Fine Arts, Folklore, History, Library Science and Theater.
- IU School of Medicine: genomics, proteomics and radiology.
- Digital Libraries: audio, images, video and scanned materials.
- Administration: student records, scanned materials and HR data.
- Workstation and server backups.

DFS Problems

- If it wasn't for DFS my life would be easy.
- DFS aggregates filling up before data is migrated to HPSS.
- Users will sometimes attempt to move too much data through an aggregate (an 18-GB file won't fit on a 16-GB aggregate but the software tries anyway).
- DFS access makes it too easy for a user to copy the 250,000 files on his PC's hard disk to HPSS.
- The good news: all problems fixed in HPSS 5.1.

HDM Problems

- The HDM's TCP process dies occasionally.
- Long migrate and purge times caused by having to read every A-node to determine which files are candidates for migration and purging.
- The HDM's connection to the DMG gets lost. Must recycle the DMG to get things going again.
- The ACL log gets out of wack (grows to 2 GB via a negative lseek).

What We Learned From DFS

- Too many users per aggregate.
- The aggregates are way too small:
 - Aggregate sizes we used are 8 and 16 GB.
 - A single user can fill an aggregate with 1 file.
 - An aggregate size of 128 or 256 GB would have been better.

Phasing Out DFS

- Existing Filesets:
 - 79 HPSS-only Filesets
 - 15 Archive Filesets
 - 1134 Mirrored Filesets
- The archive filesets contain about 5000 files holding about 200 GB of data. The data must be copied out and then back into HPSS.

Hardware Problems

- Tape:
 - 9840 - few problems (almost a perfect tape drive).
 - 9940A/B - tapes get stuck, leader blocks come off tapes.
- Disk:
 - SSA: no problems the past year but the hardware is old and aging (5 years old, 140 9-GB drives).
 - SCSI, 1 enclosure appears to be permanently jinxed.

Metadata Backups

- Using the HPSS SFS backup utility:
 - Two copies are made to 3590 tape locally.
 - One copy is made to 3590 tape at Indy.
- A raw dd backup is done to local 3590 tape every Sunday morning during the system maintenance window.

SFS Problems

- Currently running Encina TX 5.0 + PTF 3
- SFS will dump core every 2-3 weeks.
- Working with HPSS and Encina Support.
- Encina support sees evidence of data corruption in the traces.
- Additional tracing has been turned in an attempt to locate the source of the corruption. Waiting for another dump to occur.
- Increased the SFS idle timeout to 360.

Almost Daily SFS Problems

- Almost daily we are seeing:
 - BFSR0038 (add segment to metadata error)
 - BFSR0131 (failure in updating a disk map)
 - BFSR0183 (failure in processing storage seg unlink record)
- The errors cause the purge of the small file disk cache to fail. Eventually the purge will succeed but sometimes the cache will fill up to 100% and SFS/HPSS has to be restarted.
- Temporary work around is to keep the small file disk cache purged down to 0%. This gives us a longer time period to run before the cache fills or the purge succeeds.

Summer 2004

- The Indianapolis equipment will be moving to a new machine room in July.
- Rather than move the 9310 silo STK will install a new one and then just move the drives and tapes.
- HPSS will probably be down during the move since the Bloomington disk caches fill up if the second copies can't be written.

Things We Will be Testing

- HPSS 5.1 on AIX 5.1
- LINUX Disk mover on DELL 2650 server
- EMC Clarion ATA disk system on AIX 5.1 and LINUX
- XFS on DELL 2650 server.

Odd Problems

- Sammi Login Box comes up empty (no place to enter username and password). Once it took 6 hours to properly display.
- Doing a mksysb on the core server without specifying the no-compression option results in a damaged backup (Unpack: file out of phase).

Wish List

- SFTP/SCP access to HPSS.
- GRIDFTP access to HPSS.
- Undelete for users.
- Online device add/delete.

Conclusions

- HPSS has been very successful at IU and new applications are being found almost every week.
- File system access to HPSS is essential to our users and we need to find a replacement for DFS.