High Performance Storage System

An overview of HPSS

July 2021

Jim Gerry       Senior Architect
Today’s talk…

• What is HPSS?
• Who is using HPSS?
• HPSS value propositions.
• Interfaces, platforms and scaling HPSS to meet the difficult storage requirements.
What is HPSS?

• HPSS is a joint Department of Energy and IBM development collaboration.
  o Utilizing all partner’s strengths and experience.

• HPSS is software for long-lived data repositories, for customers that understand the value and benefits of tiered storage.
  o Simply stated, not all data needs to be on expensive media, but
  o Access to data on all tiers, and movement of data between tiers MUST be automated and performed in a hardware efficient manner.

• HPSS is an IBM service offering.
  o Services are best for the stewardship of long-lived data.
Publicly disclosed sites
Top 10 publicly disclosed sites – 1Q2021

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Site Name</th>
<th>PB</th>
<th>M Files</th>
<th>Since</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECMWF</td>
<td>European Centre for Medium-Range Weather Forecasts</td>
<td>556.2</td>
<td>405.1</td>
<td>2002</td>
</tr>
<tr>
<td>UKMO</td>
<td>United Kingdom Met Office</td>
<td>444.2</td>
<td>826.6</td>
<td>2009</td>
</tr>
<tr>
<td>SSC</td>
<td>Shared Services Canada</td>
<td>279.4</td>
<td>29.7</td>
<td>2017</td>
</tr>
<tr>
<td>Meteo-France</td>
<td>Meteo France - French Weather and Climate</td>
<td>241.8</td>
<td>687.6</td>
<td>2015</td>
</tr>
<tr>
<td>NOAA-RD</td>
<td>National Oceanic and Atmospheric Administration Research &amp; Development</td>
<td>239.3</td>
<td>109.1</td>
<td>2002</td>
</tr>
<tr>
<td>LBNL-User</td>
<td>Lawrence Berkeley National Laboratory - User</td>
<td>218.8</td>
<td>235.2</td>
<td>1998</td>
</tr>
<tr>
<td>BNL</td>
<td>Brookhaven National Laboratory</td>
<td>203.9</td>
<td>216.3</td>
<td>1998</td>
</tr>
<tr>
<td>MPCDF</td>
<td>Max Planck Computing and Data Facility</td>
<td>187.8</td>
<td>334.4</td>
<td>2011</td>
</tr>
<tr>
<td>ORNL</td>
<td>Oak Ridge National Laboratory</td>
<td>140.8</td>
<td>417.8</td>
<td>1997</td>
</tr>
<tr>
<td>CEA TERA</td>
<td>Commissariat a l’Energie Atomique - Tera Project</td>
<td>136.1</td>
<td>30.3</td>
<td>1999</td>
</tr>
</tbody>
</table>
The value of the HPSS service offering

- Annual support fee for 50 PB and 5 EB is the same and remains relatively flat from year to year.

- Delivery services.
  - Architect the vendor neutral storage solution.
  - Verify hardware is installed and meets expectations.
  - Install and configure HPSS according to delivery milestones.
  - Training is accomplished during the delivery process.
  - Production readiness review is the final delivery milestone, where the keys to a production-ready HPSS system are handed-over.

- Personalized support and relationships are cornerstone to the HPSS service offering.
Solutions with best of breed features for tape.

- HPSS enables faster tape reads and writes.
  - Extreme scale single file tape transfers beyond that of a single tape drive using tape striping.
  - Near native tape transfer speed can be achieved even when reading and writing small files using aggregation, buffered tape marks, and full aggregate recalls.
  - Automatic file-grouping by directory on migration for efficient recalls for improved recall performance.
  - Recommended Access Order (RAO), time-based access ordered (TAOS), and logical offset ordered tape recalls (LPOS) for improved tape recall performance.

- HPSS cuts redundant tape costs with RAIT.

- End-to-end data integrity finds data corruption and redundant tape minimizes loss.

- HPSS maximizes automated tape library mount rates.

http://www.hpss-collaboration.org

HPSS Slides
### HPSS Scales

- **Add HPSS Metadata Storage** to scale file count capacity and Db2 performance.
- **Add Disk Cache Storage Units** to scale disk cache bandwidth and capacity.
- **Add Tape Libraries** to scale tape capacity, tape drive count and tape mount rate.

---

#### Sample HPSS Hardware Architecture

- **Add HPSS Protocol Nodes** to scale HPSS client connections and client performance.
- **Add HPSS Db2 Off-Host Nodes** to scale HPSS Db2 performance.
- **Add HPSS Core Server Computers** to scale HPSS file transaction performance.
- **Add HPSS Disk Movers** to scale tape drive count and tape bandwidth.
- **Add HPSS Tape Movers** to scale tape drive count and tape bandwidth.

---

**HPSS Slides**

[http://www.hpss-collaboration.org](http://www.hpss-collaboration.org)
HPSS interfaces

HPSS Storage Broker  SwiftOnHPSS  HPSS for IBM Spectrum Scale  HPSS VFS FUSE  Parallel FTP  HSI HTAR  Client API and PIO for 3rd Party Applications

S3 Cloud Storage  File System Storage

Massively scalable global HPSS namespace enabled by Db2
RHEL Core Server & Mover computing platform
Extreme-scale high-performance automated HSM

Traditional HPSS Block Storage

Block or Filesystem Disk Tiers
Hardware Vendor Neutral
IBM + Oracle + Quantum + Spectra Logic + Enterprise + LTO Tape

Flash
Intel
Power
Disk
Tape

http://www.hpss-collaboration.org
New Storage Broker Interface for HPSS.

- Provides a unified interface to multiple storage systems including clouds, file systems, and HPSS.
- Delivers durability and portability of archived data using preservation containers.
- Increases data visibility and exploitation by mirroring preservation container metadata to a high-performance database.
- Provides project data management and sharing.
- Simplifies data access and movement.

http://www.hpss-collaboration.org
Thank you.
HPSS is best of breed for tape!
End-to-end data integrity (E2E-DI) protects against silent data corruption

- Only data we intended to write to tape will make it to the tape media.
- If you have a pre-calculated checksum (Adler32 is pictured, but many others are supported) save it with your file.
  - HPSS E2E-DI read-after-write validate data on disk and tape.
  - HPSS E2E-DI finds silent data corruption.
- If you don’t have a file checksum, HPSS will calculate one for you.

1. User sends file to HPSS disk and file checksum is stored in Db2. When the file is migrated to tape, the HPSS Mover will:
   - A. Continuously read the file data from HPSS disk and calculates file checksum.
   - B. Interlace the T10-LBP CRC for each tape block being written to tape.
   - C. The tape drive will:
     - Receives the data blocks with CRC
     - Writes the blocks to media
     - Reads and verifies each CRC

http://www.hpss-collaboration.org
HPSS is best of breed for tape!
Recommended Access Ordered (RAO) tape recalls

- The most efficient recall ordering technology tested by HPSS.
- Cuts the seek times between files.
- Without RAO tapes spend more time in the drive and require 2x to 5x more drives to meet recall requirements.
**HPSS is best of breed for tape!**

**HPSS RAIT**

- RAIT = Redundant Array of Independent Tapes.
- An HPSS software technology for striping data on tape with the added benefit of one or more rotating parity to protect against data corruption or damaged cartridges.
- HPSS RAIT is field proven and used in production for over seven years.