

High Performance Computing Modernization Program Mass Storage

September 21, 2012

Designing Storage Architectures for Digital Collections

Library of Congress

Gene McGill, mcgill@arsc.edu

Arctic Region Supercomputing Center/HTL

What is the HPCMP?

- Five large DoD Supercomputing Resource Centers (DSRCs) and various smaller facilities
 - Air Force Research Lab (AFRL DSRC), Dayton, OH
 - Army Research Lab (ARL DSRC), Aberdeen Proving Ground, MD
 - US Army Engineer Research and Development Center (ERDC DSRC), Vicksburg, MS
 - Navy DoD Supercomputing Resource Center (Navy DSRC), Stennis Space Center, MS
 - Maui High Performance Computing Center (MHPCC DSRC), Kihei, Maui, HI
- ARSC HEUE Test Lab (ARSC HTL), Fairbanks, AK
- Other smaller, more focused centers

What is the HPCMP? (cont)

- ~1,500 users
- Diverse scientific & engineering applications, systems are general-purpose, no single usage profile
- Ten large, unclassified, HPC clusters
 - Total TFLOPs: 1421; 6.1 PBs local, high speed disk;
- Mass Storage servers
 - SAM-QFS, SL8500, SL3000, T10000[A|B|C] + Remote DR.
- 25 PB 1st copy, nearly all files get DR copy
 - Less than 2% of the users own 50% of the PBs
 - Growth is accelerating

Legacy Architecture

- Batch jobs run on HPC clusters
 - Jobs use fast, expensive, local disk
 - Insufficient space on HPCs to store restart files between jobs
 - Users store files on mass storage servers that shouldn't live long-term
 - Restart file, intermediate results, ...
 - ... and users rarely go back and remove what they don't need
- Users tell us:
 - Tools to manage the millions of files on mass storage servers are lacking
 - HPC clusters not good for pre- & post-processing, this discourages post-processing and saving only high-value PP'd files
 - Users don't think all files need a DR copy
- Program management wants to live within current tape slots

The HPCMP Enhanced User Environment

- We don't want to be an Archive like this group thinks of them!
 - We are implementing tools that could be used for an Archive...
 - ... for the purpose of reducing rate of growth of data.
- The solution:
 - Add Utility Server (US)
 - Interactive use for pre-, post-processing
 - Install PB-scale center-wide file system (CWFS)
 - 30 day residency commitment vs. ~10 days on HPC disk
 - Add Storage Lifecycle Management tool (SLM) to manage the archive
 - Enhance metadata for users
 - Implement storage management policy

The HPCMP Enhanced User Environment

- Nirvana Storage Resource Broker (SRB):
 - Commercial offering, w/ vendor support
 - Application runs on top of a database
 - Automated Storage Lifecycle Management using metadata-driven policy
 - SRB presents a global namespace of potentially many storage resources, some of which may be geographically distributed.
 - HPCMP is layering SRB on top of existing SAM-QFS file systems

The HPCMP Enhanced User Environment

- SRB (cont):
 - Policies we're implementing in SRB
 - ARCHIVE: Local tape copy is automatic
 - DR: Remote copy by user request per object, default is no DR copy
 - EXPIRE: Automatically removes objects when they reach the end of their retention period
 - Enhanced metadata
 - Retention period
 - Default is 30 days,
 - Users can specify decades but it cannot be infinite
 - Users must confirm retention every 3 years

The HPCMP Enhanced User Environment

- SRB (cont):
 - Enhanced User metadata:
 - Dublin core
 - Name Value scheme
 - Multi-level security features added in
 - Up to 16 site-defined flags per object
 - Opportunity for doing science in the metadata
 - Automated processing during ingest can mine data to populate metadata needed for scientific visualization later
 - Avoids accessing files in favor of using metadata in SRB/DB
 - Metadata is query-able.

The HPCMP Enhanced User Environment

- Currently in Pilot phases at two DSRCs

Acknowledgements

- Many people working on this:
 - Tom Kendall, Executive Agent, Army Research Lab DSRC
 - Stephen Scherr, HPCMP
 - Robert Hunter, HPCMP
 - Constantin Scheder, Nirvana SRB Chief Architect
 - Reid Bingham, Nirvana
 - And many others....