Summary of
IEEE Archival Storage Life Cycle Management Workshop

September 24, 2007
San Diego, California

Sponsored by:

IEEE Technology Committee Mass Storage Systems (TCMS)
a committee of the IEEE Computer Society

Bob Coyne
Vice Chair for Metadata Meetings,
IEEE TCMS Executive Committee

Merritt Jones
Chair, IEEE TCMS

Organizers:

Bob Coyne - IBM
Reagan Moore - SDSC
Robert Chadduck - NARA
Harriet Coverston - Sun Microsystems
Harry Hulen - Consultant

IEEE Archival Storage Life Cycle Management Workshop

The IEEE Archival Storage Life Cycle Management Workshop evolved from discussions among several of the leaders of the IEEE Technical Committee Mass Storage Systems (TCMS), a committee of the IEEE Computer Society. The workshop was held in conjunction with the 24th IEEE Conference on Mass Storage Systems and Technology. This conference was held in San Diego on September 24-27, 2007, with the one-day workshop taking place on September 24.

1. Background

The idea of this workshop was to bring together experienced professionals representing data resource services providers and consumers in government, academia and industry to discuss and document ideas and suggestions to improve data migration and operations effectiveness for large institutional data repositories built upon the file system metaphor.

Large file repositories consisting of many 10s of petabytes exist today, and prospects of many 100s of petabytes appear in users' plans and requests for proposals. These very large repositories tend to become locked into legacy hardware and software technologies due to various combinations of their size, lack of standards and interoperability, lack of tools, lack of information on formats, and proprietary claims of vendors.

A significant number of user organizations, including national research laboratories, government agencies, large universities, and industries such as entertainment and health care are interested in this problem. On the vendor side, there is awareness of the market potential for very large file repositories, both for repositories that are now being stood up as new ventures and for repositories that are already facing the problem of obsolescence of their hardware and/or software platforms and concepts of operation.

2. Workshop structure

The IEEE TCMS distributed a call for participation in the workshop. Attendance was intentionally limited to about 30 seats. Many of the attendees were invited because of their prominence in the archival storage field, and volunteer attendees were approved based on credentials and white papers they submitted with their application. The attendees included leading file and storage system users and engineers from ARSC, Boeing, BAE Systems, CERN, ECMWF, FNAL, Instrumental, IBM, LLNL, NARA, NASA, NCAR, NERSC, Silicon Graphics, SNIA Data Management Forum, Sun Microsystems, UCSD/SDSC, University of Illinois, and University of Stuttgart.

The workshop was structured to cover one full day, with three presentations in the morning sessions and two group discussion sessions in the afternoon. To begin the workshop, Bob Coyne presented workshop motivation and specific objectives. The second presentation was an invited talk by Reagan Moore of SDSC entitled, "Data Migration and Policy Management." The third presentation was an invited SNIA Data Management Forum overview by Michael Peterson entitled, "Solving the Coming Digital Archive Crisis." Copies of the workshop presentations, handouts, and materials are available on the IEEE Computer Society TCMS website at http://storageconference.org/archive

Robert Chadduck of the National Archive and Records Administration led the afternoon group discussion, and Bob Coyne moderated a closing session to summarize key points and to consider options for ongoing plans including a summary statement from the workshop.

All group discussions and presentations were vendor neutral, focusing on concepts. Some of the key points made during the group discussion about the workshop challenge topic are provided in discussion notes section of this document.

3. Workshop Discussion

3.1. Workshop Challenge

The key challenge for the workshop attendees was consideration of the notion that, for the file system metaphor, ownership and stewardship of data and the digital data media assets on which the data resides require a mechanism to free data in bulk from the legacy system and attach or move it to a new system. The mechanism could be in the form of a utility or other set of tools, or it could be an unrestricted disclosure that provides all of the protocols, data formats, and metadata formats necessary to enable the migration of the user's data in bulk.

The notion of copying data in bulk would include the ability to copy data in a sensible order, such as the ability to move all the files on each tape cartridge in sequential order to new media. The notion of moving data in bulk would require the ability to move the actual media, such as tape cartridges themselves, without copying files to new media. The latter notion is better suited to effect scalable and operationally efficient data migration of petabytes from one system to another.

These mechanisms inherently require access to the metadata that contain the file names, status, length, location, date, and other relevant information about the files, and this metadata must be accessed as a whole, either in its native internal form or in an intermediate non-proprietary but nevertheless complete form.

The attendees were polled for a "yes" or "no" opinion about the need for software and service providers to enable or provide a means to migrate data in the manner of the preceding paragraph. The unanimous agreement was "yes". It was clear from the lively discussion that a consensus approach on how to motivate, enable or provide a commercial or contractual means was beyond the scope of the workshop.

Significantly, attendees are in full agreement that stewardship for large scale file repositories breaks down if users must copy 10s, 100s or even 1000s of petabytes when they seek to move between incompatible file and storage system hardware and software. Several users observed that media migration for the purposes of media replacement, repack, and technology insertions has become a 24x7 process within normal daily operations. There is no time, capacity, or budget to provide for a disruptive move between incompatible storage systems that requires more data movement than is already present in their day-to-day operations.

3.2. SNIA Data Management Forum

Feedback on the Storage Networking Industry Association's Data Management Forum presentation by Michael Peterson included two primary concerns. The attendees, who are largely early adopters indicative of high performance scientific and technical computing, noted that the SNIA Data Management Forum outlook lacked requirements for bulk data migration and lacked

emphasis on performance. Concern was also expressed about the introduction of yet another application interface standard such as XAM, which has not achieved sufficient traction to influence the archival storage community represented by the attendees. (The SNIA XAM-Interface is an application-level interface coordinating information metadata between applications and storage systems.)

There was a general lack of enthusiasm about models and concepts waiting to exploit object storage at the device level. The attendees expressed skepticism about adopting solutions proposed to exploit object storage devices, about which there is much discussion on the part of the research community and little motion on the part of industry.

Michael made the point that SNIA is recommending self-healing to reduce data migration overhead. While reduction of cyclic media refresh and physical data migration overhead was a welcomed notion, the workshop attendees wanted to hear about methods to migrate very large data stores between storage services in bulk and in a timely manner without great pain and expense. Again, file and storage solution lock-in was noted as a systemic problem.

Michael mentioned "Self-Describing Self-Contained Data Format" (SD-SCDF), a concept that is part of the SNIA data migration vision. The group agreed that improving the ability to preserve and re-use data is a necessary objective, yet the reasonable SNIA message was overshadowed by the association's lack of focus on high performances and bulk storage requirements. While the workshop attendees acknowledged that SNIA and data preservation concepts are focused on improving interoperability, data migration, and data path component life cycle management, the lack of emphasis on scalability and, to a lesser degree, the lack of commodity object storage devices, are in many cases significant impediments.

Though popular in the trade press and a growing number of end users, “de-duplication” solutions are not generally adopted by this group, as their data is less prone to duplication of the kind that can be discovered and acted upon.

Archival object aggregation and packaging that is in use among workshop participants is largely more primitive than Open Archival Information System (OAIS), an ISO standard for data preservation. Members representing high fidelity archive programs at NARA, NASA, NOAA and Library of Congress were the notable and important exceptions within the workshop. These organizations seek to advance packaging concepts, library services and data preservation models akin to these efforts.

3.3. SNIA DMF and IEEE TCMS Alliance Discussion

We discussed the benefits of an alliance between SNIA and the IEEE in the archival storage area and agreed to proceed in developing it. Bob Coyne will be the conduit and in this capacity voiced a concern about asking the vendors within SNIA to adjust to the high-performance storage and IEEE input. Michael Peterson offered assurances and encouraged them to try.

4. Workshop Notes

Here we present the unedited notes of the workshop session.

4.1. Problem statement

PROBLEM: Replace bitfiles and storage systems and services managing 40 plus petabytes of bitfiles and 40 petabytes of second copies from one vendor/open file system to another in less than one year. How to do this via moving not copying?

1) Focus on migrating file and storage metadata while retaining file data on old media until the current operations concept moves old files to device and media technology provided by a new vendor/open solution.
2) Expose metadata and formats so information can be read natively allowing data to be moved in bulk. This is a real intellectual property concern for some vendors.
3) Need an information representation and mapping between file and data block and media, and all file attributes and formats provided by the old bitfile and storage system. It was agreed that metadata is a generic term and that no assumption can be made about metadata interoperability between vendor/open bitfile/storage system solutions. Common representation for file system properties is required.
4) Move information off old systems and into new and transform at that time, as necessary.

4.2. How effective is a reference model in empowering the customer?

1) "A weak forcing function" (unless customers really embrace it and enforce the use of it in the purchasing process.) These have not been successful in the past. Examples: the Mass Storage Reference-Model, MS-66 (archival open tape format) and OAIS.
2) The need for a best practices document to drive visibility within organizations and to give those practices credibility. (SNIA survey identified that achieving organizational change is an essential element. Thus, a best practices document is an instrumental tool and a lever.)

4.3. Move context not physical data

1) Data management is an active process that requires planning for obsolescence, migration to new technology.
2) Can a move of context replace a physical data copy? Implies the need to move the context independently of the physical data move. The context constitutes representation information for the preservation environment.
3) Bulk operations are essential. The movement of entire collections will become more prevalent.
4) Context includes the management policies that govern authenticity and integrity of the stored data. Examples include the ownership of the files, criteria for validating data integrity, criteria for minimizing risk of data loss through replication.

4.4. Parallel file system uniqueness adds complexity

1) Parallel file systems tend to be unique across vendors differentiating themselves from other vendor solutions. New paradigms introduce complexity such as object storage (data content and structuring, access control) and importing of archival features through storage classes (include tape in file system).
2) Not clear how the higher level of sophistication can be characterized. With respect to data bulk data migration, it is necessary to document what is provided by a feature rich archive.

4.5. What can users do to encourage progress towards high performance, bulk data migration?

1) Several attendees requested the development of best practices guidelines for the HPC environment. Content should include: lessons learned, why important, dollar impact in decreasing risk, endorsements, requirements, and real world errors.
2) Henry Newman encouraged the group to consider a second session with expanded participation from other groups. He suggested that we could do more formally through NSF or NARA or DARPA. A tangible set of deliverables would be the objective for the meeting.
3) Require better end-of-life planning in procurements even within vendor solutions.
4) Require bulk move capabilities as a procurement requirement.
5) Require self-describing media.
6) Require scalable archival storage systems management infrastructure.
7) Require access to user data without a vendor being available.
8) National data systems may be dispersed over an entire continent. Require that archival systems be integrated as an indispensable part of infrastructure similar to communication infrastructure. This implies collaboration with ISO and networking infrastructure.

4.6. Observations on why sites migrate data

1) Recover floor space seemed the number one response. Density can be a major requirement.
2) Gain capabilities such as higher bandwidth needed for ingesting larger collections.
3) Reduce power and cooling expense
4) Mission requirements can force a data migration or vendor switch
5) Media refresh migrates file/block data every three to five years even though manufactures increase useful life of media.

4.7. Observations on why sites use data grids to virtualize data collections

1) Storage vendors have focused on virtualization of storage systems. This means the development of a common interface standard that can be used by all applications independently of the choice of vendor product. Data grids focus on the virtualization of data, the management of properties about the data that is being stored, independently of the choice of vendor product. This means that operations can be performed upon the data grid independently of the protocols supported by the vendor storage system. In practice, this

requires writing a specific storage system driver for each different type of storage system protocol. The data grid maps standard actions performed by client applications, to sets of standard procedures / micro-services. The standard procedures are in turn mapped to standard operations that can be performed at the storage system. A storage-specific driver maps the standard operations to the required protocol.

2) Current research on data grids is investigating the concept of “structured information resource” access. Each storage system manages structured information such as the mapping of a file to a set of blocks, or the mapping of access controls to files. A structured information resource driver queries the storage system for the information needed to manipulate a file. Operations on the file are then performed based upon the information.

3) The concept of bulk data movement inverts the concept of “structured information resource”. The new storage system is given the context information extracted from the original storage location. This information is then mapped to operations that the new system is capable of performing and that execute the desired procedure. In a preferred implementation, the mapping should be possible by either the data grid or the storage system. This would allow the movement of media between systems, under the control of the data grid.

5. Workshop Attendees

1. Balac, Natasha
2. Brantly, Dale
3. Butler, Michelle
4. Cavena, Dave
5. Chadduck, Robert
6. Coverston, Harriet
7. Coyne, Robert
8. Crouse, Don
9. Dequenne, Francis
10. Feiner, Phil
11. Grayzeck, Edwin
12. Haas, Peter
13. Hick, Jason
14. Hulen, Harry
15. Jordan, Chris
16. Kerr, Ann
17. Kuehn, Dennis
18. McDonald, Robert
19. McGill, Gene
20. Merrill, John
21. Mirvahabi, Nadi
22. Moore, Reagan
23. Newman, Henry
24. Perelmutav, Timur
25. Peters, Marc
26. Peterson, Michael
27. Petravick, Donald
28. Pipes, Randy
29. Ponti, Carnolo
30. Rehm, Kevan
31. Seiffert, Kurt
32. Shoopman, Jerry
33. Yacenda, Bob