Project 1244: IEEE Storage System Standards

John L. Cole
Computational and Information Sciences Directorate
United States Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5067
jack.cole@ieee.org
tel +1-401-278-9276
fax +1-401-278-2694

Abstract

Approaching its tenth anniversary, the IEEE Storage System Standards effort is in the process of balloting Media Management System (MMS) standards. These represent the first standards for the IEEE Storage System Standards Working Group (SSSWG), and the first storage system standards for the world. In the early years, SSSWG produced the Mass Storage System Reference Model (MSSRM), directly influencing the design of many successful commercial products and the MMS standards themselves.

The IEEE Storage System Standards Committee (SSSC), sponsor of SSSWG, will work in the coming year to complete work on the suite of MMS standards, and begin work on new projects. New projects for SSSC in 2000 include tape standards, tape recommended practice, and a project to develop a Guide for Storage System Design. Existing and new collaborations with other groups developing storage-related standards will be fostered in 2000.

The SSSC is driven by the urgent need for interoperable storage system software, and storage systems that are highly scalable and functional in distributed, heterogeneous environments.

1 Background

1.1 In the Beginning

The IEEE Storage System Standards effort began unofficially with individual discussions in the 1980s to standardize and guide development of hierarchical storage management systems. In the summer of 1990, the IEEE approved the SSSWG charter and the first (and for a long time the only) project that resulted in the un-balloted MSSRM. Development of the MSSRM progressed through several versions until the last revision, version 5, was approved by an internal SSSWG vote in September 1994.

1.2 SSSWG Charter

The IEEE Storage System Standards Committee is chartered to model generic mass storage systems, and based on such modeling, to develop widely accepted, readily implemented standards with minimal licensing requirements.

In addition to working on standards, the SSSWG may develop recommended practices and guides. The SSSWG is primarily concerned with Distributed Storage System Design, and the SSSWG, without favor, includes Storage Systems of every scale in its studies. An object-oriented approach is desired in all SSSWG efforts, and net-attached storage is intrinsic to its model.

SSSWG must also consider promising emerging technologies in its modeling, even though standards for parts of the model may not be immediately practical as a result. The model may remain partly an abstraction expressing desirable features, and the associated standards expressing practical requirements relating to current technologies.

The purpose of these standards is to promote use of best technologies resulting in interoperable, fully-scalable systems permitting ready access of information throughout distributed, secure, heterogeneous net-attached storage systems.

1.3 MSSRM

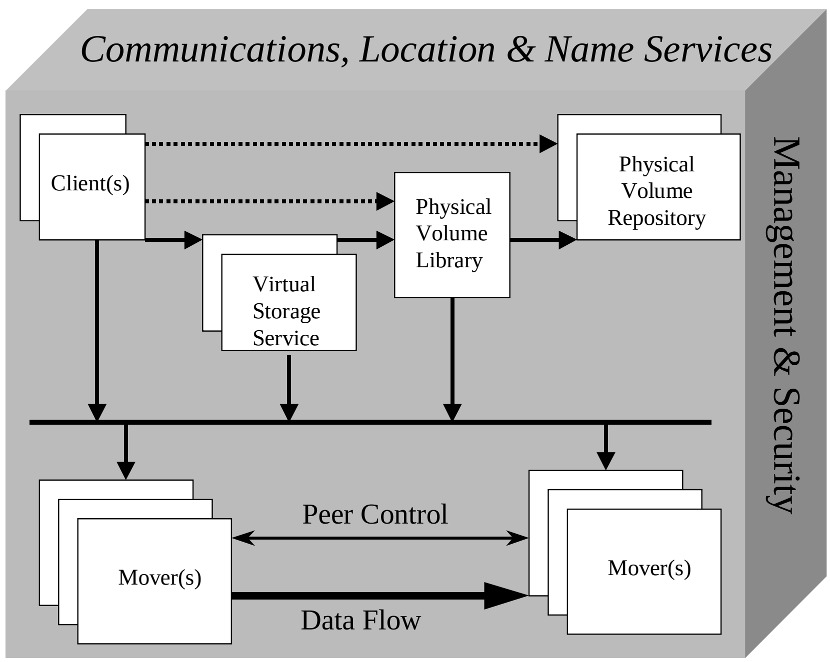

Figure 1. Components of the MSSRM

The IEEE Mass Storage System Reference Model (MSSRM), although not balloted as a standard, has been highly successful in service as a guide for development of many well-known storage systems and components of systems, commercially and otherwise, in use today.

Along the way, the MSSRM has been revised a number of times, and renamed the IEEE Reference Model for Open Storage Systems Interconnection (OSSI). Version 5 of the OSSI Reference Model was approved for public release in September 1994.

There were seven IEEE-approved Project Authorization Requests (PARs) relating to the MSSRM — one PAR for the model itself and six for modules of the MSSRM representing sets of services identified by the SSSWG as those essential to the composition of viable storage systems. The PAR for the MSSRM has been revised and renewed as a different project, and these six PARs, which represent the modules into which the MSSRM was partitioned, are withdrawn:

SOID (1244.1) Storage Object Identifier
PVL (1244.2) Physical Volume Library
PVR (1244.3) Physical Volume Repository
MVR (1244.4) Data Mover
MGT (1244.5) Storage System Management
VSS (1244.6) Virtual Storage Service

In addition, other associated documents were created, such as the "Virtual Storage Architecture Guide"[1].

The Reference Model was intended to provide a framework for the coordination of standards development for storage systems interconnection and a common perspective for existing standards. Through development of this structured framework, the Model would expose areas where standards were necessary or in need of improvement.

The technology and application independence of the Model would accommodate descriptions of advanced technologies and expansion in user demands. This flexibility would also support the phased transition from existing implementations to storage system standards.

It was not the intent of the Model to serve as an implementation specification, to be the basis for appraising the conformance of actual implementations, or to define precisely the standards for services and protocols of the interconnection architecture. Rather, the Model was intended to provide a conceptual and functional framework allowing teams of experts to work productively and independently on the development of standards for storage systems. These remain, roughly, the intentions of the IEEE Model.

All who participated in drafting the MSSRM must be proud that this model was one of the earliest efforts to presage present concepts of storage objects and of net-attached storage.

1.4 Change of Direction

The present direction of the SSSWG is quite different.

Although work continued after 1994 to develop the PVR and PVL modules, it became apparent to the SSSWG that the approach taken to encompass all aspects of storage systems made it impossible to compose practical standards. To that end, the SSSWG began a device-driver level standard called "The Media Changer Service Standard", or "MCS".

Work on MMS began shortly after SSSWG came to the realization that writing standards for PVL and PVR was an intractable problem. These and other components of the MSSRM were and are very robust, very all encompassing, and very difficult to reduce to standards. In order to make progress and actually publish standards in the lifetimes of the SSSWG members, a different direction was needed.

As the SSSWG met with individuals interested in MCS, it became aware of an effort by some of the MCS participants to develop a minimalist media management system known as "OpenVault" (www.openvault.org).

After hearing presentations on the OpenVault effort, it seemed to the members of SSSWG that OpenVault embraced the same one, true reality of storage system needs which SSSWG saw, and that standards based on a more minimalist approach would be possible in a reasonable time. OpenVault was still developing; some of the people developing OpenVault were early members of SSSWG; OpenVault capitalized on ideas from the MSSRM; and OpenVault, it appeared, could co-evolve with an IEEE set of standards. Today, the IEEE MMS and OpenVault are close, although not equivalent, as a seesaw development of the two has progressed over more than two years.

1.5 MMS Architecture [reference 2]

This describes the motivations for an overall architecture of the IEEE Media Management System. Although the architecture may suggest a particular design or implementation, it is not the IEEE's intent to favor a specific implementation of the MMS. Indeed, it should be possible to implement the MMS in a number of ways, ranging from a "lightweight" implementation in a scripting language such as *perl*, or a full implementation written in a traditional programming language such as C, C++, or Java.

The MMS is a software system for managing physical media. The system has the following properties:

- It is ***media-neutral***, allowing the management of computer tapes, disk media, disks, optical disks, CD-ROMs, as well as non-computer media such as videotapes or reels of film.
- It is ***scalable***, being comfortable in environments as small as a single individual's office or home and as large as a multinational corporation, educational or scientific institution, or government archive.
- It is ***platform neutral and operating system independent***, working with existing computer systems from multiple vendors with varying degrees of media-handling sophistication.

- It is ***distributed***, allowing access to media and the devices that store and perform data transfer operations on the media by more than one system. A single MMS may manage devices that are connected to many host computer systems, including devices that are physically connected to multiple hosts. Connectivity between elements of the MMS requires the availability of standard TCP/IP.
- It provides a reasonable degree of ***security and protection*** for access to the media by ensuring that specific media may be mounted only by those applications which have authority to access that media. All parties are authenticated, and network communication is digitally signed so that it is extremely difficult to forge.
- It is ***content-neutral***, and does not have any inherent understanding of the content of the media; indeed, with some media, such as videotape or film, the MMS many not even have access to the content of the media.
- It is ***application independent***, providing appropriate media management functions for diverse applications ranging from backup and hierarchical storage management, to broadcast television automation. Media belonging to multiple applications may be managed by a single MMS; these applications may be multiple instances of the same program, or of different applications.
- It is designed to be ***modular*** to allow independent groups to work on components of the MMS independently; the modularity is provided by strong, flexible interfaces that can evolve over time.
- It is ***language-neutral***, permitting programmers to write applications that interact with the MMS in almost any programming language, and, indeed, to allow the MMS itself to be written in almost any programming language.
- It allows ***multiple implementations*** to interoperate seamlessly.

The key to the architecture of MMS is to clearly define the basic functionality that the MMS must provide, and to declare specific points in the functionality to provide defined interfaces that allow independent components to interoperate.

1.6 MMS Described

The IEEE Media Management System (MMS) suite of ten standards arguably could have been a single standard, although the first five total 304 pages in aggregate. The intent is that these five standards will evolve separately. Also, that builders of storage systems or components could comply with each standard individually.

A tutorial of MMS was given in March 1999, but to review the components briefly to see what has been balloted and what remains, some brief information about MMS is offered.

The following drawing depicts potential system boundaries (thick lines) between the core (or cores) of a media management system (or systems) and the protocols used to communicate among the components. And it suggests the distributability of a MMS. In this drawing “DM” is the drive manager; “dmp” is the drive management protocol; “LM” is the library manager; “lmp” is the library management protocol; and “mmp” is the media management protocol.

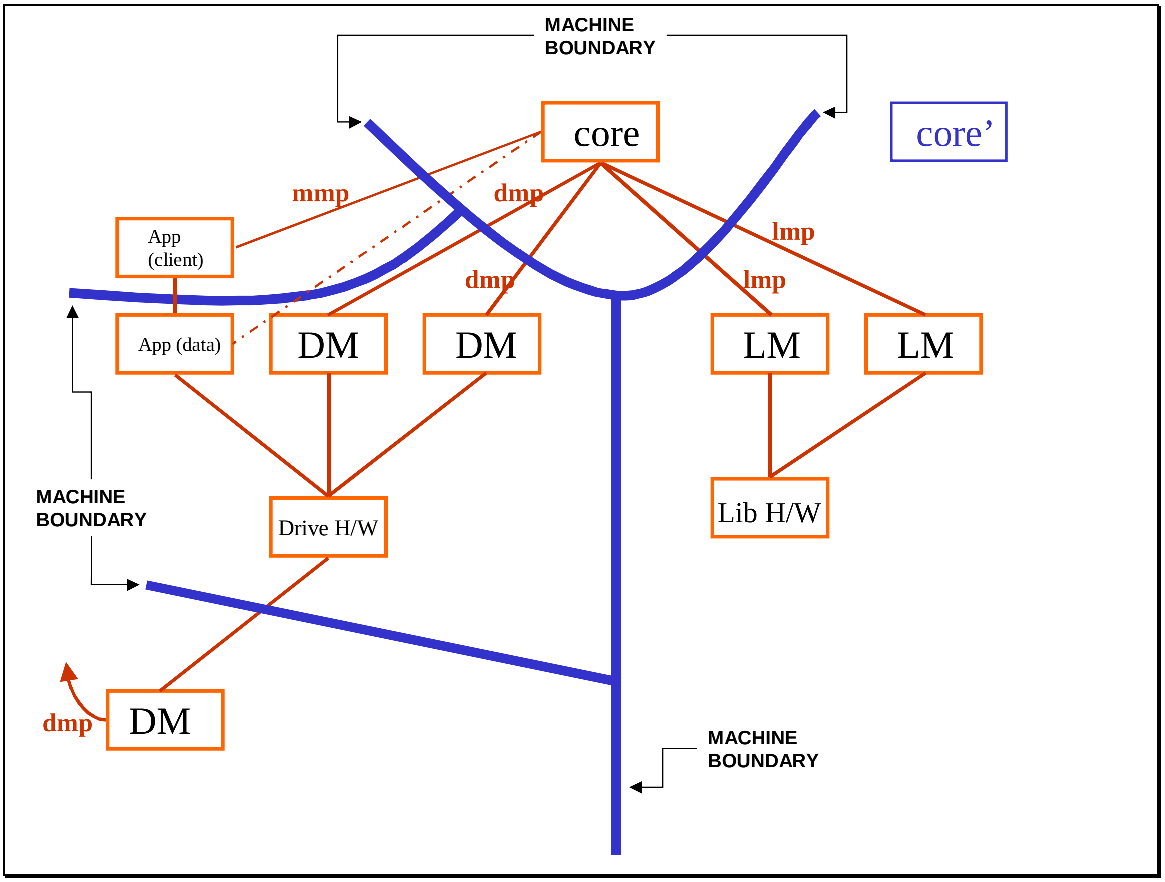

Figure 2. MMS Distribution of Components Across Machine Boundaries

There is a suite of ten MMS projects and a one project to develop an IEEE Data Mover under the 1244 series. The first five of the MMS projects are draft standards in balloting now. Work will proceed on the other five and MOVER, which is not part of the MMS, this year.

1244.1 - Media Management System (MMS) ***Architecture***.

Specifies the architecture of a distributed, platform-independent, system to manage removable media, including both disk and tape, using robotic and manual methods. The general schema for managing media, the expected components of the software system, and the data model to be supported by the software system for managing this media are described by this standard. Details of components of the Media Management System are specified by companion standards.

1244.2 - Session Security, Authentication, Initialization Protocol (***SSAIP***)

is the initial "handshake" protocol used by components of the MMS to establish identity, authority, and initial communication.

1244.3 - Media Management Protocol (***MMP***)

used by client and administrative applications to allocate, deallocate, mount, and dismount volumes, and to administer the system. The MMP includes levels of privilege so that, for example, a client application cannot perform administrative functions, or an operator console program cannot perform higher-level management functions.

1244.4 - Drive Management Protocol (***DMP***)
is used between two software components of the MMS: the central management core and a program that manages a drive which is used to access removable media.

1244.5 - Library Management Protocol (***LMP***)
is used between two software components of the MMS: the central management core and a program that manages an automated library or a vault. The minimum functionality required to implement an MMS is the SSAIP and MMP. Most practical implementations will include the DMP and LMP. Additional protocols are defined to extend the MMS to interoperate with other MMSes and with other media management systems:

1244.6 - The Media Manager Interchange Protocol (***MMIP***)
defines a protocol to allow interchange of information between autonomous Media Managers.

1244.7 - The Media Manager Control Interface Protocol (***MMCIP***)
defines a protocol which permits interfacing the data management component of the MMS with existing library management systems.

1244.8 - The C Language Procedural Interface
defines a set of standard programming interfaces which facilitate construction of components of the MMS, particularly client, administrative, and operational applications, library managers, and drive managers. The initial definition will be for the C programming language. The interface will be designed so that implementation in languages such as C++ or Java could be easily accomplished.

1244.9 - MMS User Mount Commands
defines a set of standard commands to allow a user to mount, unmount, acquire, and release media. These commands are specified as a part of a command line interface for systems that offer such interfaces, such as the UNIX shell or NT command line interface. Commands may be embedded in scripts to produce more complex or custom functions, or to allow an application program that is not written for MMS to be adapted for use with MMS.

1244.10 - MMS Standard Administrative and Operational Commands
defines a set of standard administration and operation commands of an MMS. The standard defines a command- line, minimally interactive interface for basic interaction with the MMS; these commands could be used to construct interactive interfaces using scripting-based systems such as web CGI scripting or tcl/tk.

1.7 MOVER

Very little thought or work has gone into developing a data mover standard since MSSRM version 5. If there are volunteers to develop this standard, and it is not overcome by trends in net-attached storage, the SSSC will ballot a draft standard before 2002. The present PAR for a data mover standard is:

1244.11 - MOVER

provides a standard storage system data mover architecture and interfaces for use by the IEEE Media Management System and other storage system software. MOVER transfers data between two endpoints in a distributed storage system." This deceptively simple statement belies the difficulty in describing MOVER as a standard.

The MSSRM concept of MOVER is synopsized in the Mover Architecture drawing and description that follows:

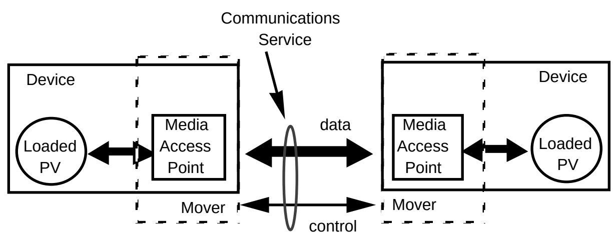

Figure 3. MSSRM Concept of Mover Architecture, "PV" is "Physical Volume"

"The Mover performs operations on media access points and affects data transfer. Media access points are the means of accessing physical volumes and sections of memory accessible to Mover clients."

"A Mover performs two distinct functions: 1) it changes or monitors the read/write state of a device (e.g., positioning within the physical volume, reporting status and errors, and loading and unloading physical volumes as necessary). 2) it transfers data and source/sink information (to effect the transfer of data) between devices, devices and memory, or from memory to memory."

2 Present and Future

2.1 Balloting Status, Note About Standards

The draft standards in ballot now are:

1244.1 - Media Management System (MMS) Architecture
1244.2 - Session Security, Authentication, Initialization Protocol (SSAIP)
1244.3 - Media Management Protocol (MMP)
1244.4 - Drive Management Protocol (DMP)
1244.5 - Library Management Protocol (LMP)

The SSSC could have held a ballot without involving the IEEE Standards Department Balloting Service, but chose to use that service. At this time balloting is carried out partly online and partly by old-fashioned methods. Registration of interested parties and balloting itself are carried out online, but the invitations are sent through the U.S. Postal system. Certain time periods are allowed for sending messages and receiving responses, so that balloting consumes months of time. In the case of the first five MMS draft standards, balloting is being accomplished online, and began December 15, 1999. This was an unfortunate timing with preparations for the holidays and concerns over Y2K effects. Roughly fifty individuals registered interest in balloting the MMS draft standards, and 31 responded to the official ballot invitations from IEEE. The first ballot of these standards ends on January 14, 2000. Depending on the response, revisions and re-balloting (re-circulation it is called) may be required. The intent is to complete all of the process in time for submission to the March 2000 IEEE meeting in which approval is sought. If this succeeds, these standards will be published by summer of 2000.

A note about the nature of standards: they are not fixed in stone. Standards are living documents, receiving modifications during their lives, being re-validated periodically, occasionally being withdrawn, and being re-balloted when the aggregate of modifications becomes too great. The work of groups like SSSWG to develop draft standards and initial balloting are just the beginning of the journey for standards. The anxiety by some that standards somehow bestow a lock-in on anything is founded on a misunderstanding of standards.

2.2 Remaining MMS and MOVER Standards

The remaining MMS standards to be developed are:

- 1244.6 - The Media Manager Interchange Protocol (MMIP)
- 1244.7 - The Media Manager Control Interface Protocol (MMCIP)
- 1244.8 - The C Language Procedural Interface
- 1244.9 - MMS User Mount Commands
- 1244.10 - MMS Standard Administrative and Operational Commands

The group believes that 1244.6 can be completed in 2-3 weeks, 30 pages of writing, and an additional 20 pages of XML DTD. One SSSWG member volunteered to serve as editor of 1244.8 starting with commercial work of another member in this area. The projects 1244.7 and 1244.11 were deemed too complex to grapple immediately. Project 1244.9 (user mount commands) will be accomplished for UNIX only, and will permit scripting. Project 1244.10 will be presented as a trial use standard only. In fact, discussion entertained the idea that all three of the Programming and Command Line Interfaces should be trial use standards (1244.8-1244.10).

Work will proceed on development of a 1244.11 MOVER draft standard subject to the normal limitations of interest, volunteerism, and time.

2.3 SSSC

Over the last several years the SSSC focus has been only on SSSWG and MMS standards. The new sponsor chairs for storage system standards will work to broaden the focus of SSSC to include other projects, working groups, study groups, and collaborations. The SSSC itself will become an actual committee of several people instead of just the sponsor chair.

2.4 Collaborations

During 1999, the chair and other members of SSSWG exchanged mail, held teleconferences and meetings with members of the Distributed Management Task Force (DMTF, www.dmtf.org) and members of the Storage Networking Industry Association (SNIA, www.snia.org) to provide the IEEE definitions of storage objects to the Common Information Model (CIM) being developed by DMTF. Major industrial entities are basing their Web-based Enterprise Management (WBEM, pronounced "web-um") products on CIM, and so this collaboration is very important. The SSSC will continue to pursue collaborations such as this in 2000 and beyond.

2.5 Tape Standards

Three tape standards were suggested by a member of SSSWG, and project authorizations requested. Approval for these projects is assumed, and SSSC intends to purse these in 2000:

Portable Tape Driver Architecture (1563.1, Recommended Practice) provides a reference model for tape driver architectures that is portable across multipleoperating system environments, fully featured, and high performance.

A fully realized architecture that industry can base their implementations on that will reduce the effort required to support a new tape device on a given platform and thereby increase the available choice of drives on any given platform. This will benefit the application vendor and the end customer.

Common Tape Driver Semantics (1563.2) defines a common set of operations and semantics for access to tape drives across multiple operating systems platforms.

Eases the task of porting and supporting applications that use tape storage across multiple operating system environments. This will enable application vendors to port to more platforms and thereby increase the end customer's available choices.

Common Format For Data On Tape (1563.3) defines a self-identifying format and record structure for the storage of data and meta-data on tapes, a structure that contains the key to understanding the format of the data stream as well the data itself. An analogue from the networking world would be the Document Type Definition (DTD) structure used to describe documents in XML (eXtended Markup Language).

Enables data written by one application to be accessible by other applications without those applications having to know how each other encodes data written to tape.

2.6 Guide for Storage System Design (P1600)

This project, a revision of the original P1244), will produce a clear, abstract, model exposing the design features required for storage systems to provide transparent, secure information access in highly distributed, heterogeneous computing environments.

The model produced will describe design alternatives and rationales applicable within the spectrum of valid storage system architectures.

Emphasis in the model will be placed on net-attached storage, object-oriented design, open source software, minimal licensing alternatives, and maximum scalability.

The work under this project will serve to revise the popular IEEE Mass Storage System Reference Model version 5 of 1994, and use it as a basis for related IEEE Recommended Practices and Standards.

Purpose. This Model will guide implementers toward interoperability in meeting such demands by suggesting best use of current and emerging technologies.

This Model will inspire commercial designs of storage systems and system components from a broad spectrum of implementers, resulting in a high level of interoperability throughout the world.

3 Observations, Motivations, Problems

The entire area dubbed "storage" suffers from a set of common problems.

First, the focus on storage is grossly misleading. Storing things, including data, can be very simple. It is the act of accessing data that makes it information. So the thrust of all efforts in "mass storage" is only for the sake of "information access". And, as a practical aspect, this is what is observed.

Second, storage system technology urgently needs to advance much more rapidly than it is advancing now to meet the challenge of information access. Perhaps the test of when storage system technology sufficiently advances is that "Any sufficiently advanced technology is indistinguishable from magic"[3]. That is, when information appears magically on request, success will have been achieved.

There are many symptoms and documented aspects of this urgent need, more than this paper can accommodate and stay on topic. As example, one aspect cited by Jim Gray is that storage capacities are increasing at the rate of 100x/decade while storage throughput is improving at only 10x/decade [4].

Third, raw storage capacity is seldom sized properly with computing capabilities, even though there are direct relationships between processing power, memory, and storage. The Dept. of Energy's Accelerated Strategic Computing Initiative suggests that you need 300 bytes of archival data for each sustained megaFLOPS. This points to one of the major factors in growth of storage (information access) demands: the growth in processing power. Processing today goes on with processors of 70 to 100 million transistors, and yet single processors of more than a billion transistors is forecasted for the next decade [5]. And "QuBit" processors are contemplated with speeds "millions" of times greater [6].

And while information is being generated at rates in step with the rapidly improving technologies of processors, and storage media and hardware are improving rapidly as well, storage system software and architecture lag dangerously behind. This is in part due to the overall crisis in software [7,8]

At the same time the specialized demands for secure access to information in highly distributed and heterogeneous environments are growing.

Traditional approaches are not sufficient to meet these needs, and revolutionary or rapid evolutionary changes are needed in storage system design and software. The information which will need to be stored and accessed in the next year or two will equal twice all the data ever stored before now.

Far from leading the target, in this case the "disaster recovery" with which we should be concerned is the loss, perhaps permanent, to access of information purchased with the expense of computing and human resources. To quote John Carlin, the U.S. National Archivist, will the country lose its memory [9]?

4 Promising Trends

The IEEE is re-inventing itself in several respects, including embracing techniques that speed up every step of the standards process. In addition, the IEEE Industry Standards and Technology Organization (ISTO), affiliated with the IEEE and the IEEE Standards Association (IEEE-SA) and just formed in 1999, is moving away from the sales of standards reprints to the intelligent position of just making them publicly reviewable. The ISTO embraces corporate entities as participants, and attempts to act much like consortia in industry.

The cooperation among groups developing "standards" in storage is very promising. This is truly an area in which we must all hang together or… The greatest competitor for products and solutions against storage efforts is not within the arena of storage at all. It is the preponderance of interest in faster processors, wider and faster networks.

The trend away from "SAD" storage (server-attached disks [10]) toward Storage Area Networks (SANs) and onward to Net Attached Storage (NAS) carries great hope with it. Using idle embedded cycles on computationally rich storage devices for the remote execution of some applications serves to reduce the motion of data and fits better in highly distributed environments.

The extended abstraction of the present storage object called "file" to even more abstract storage objects will both serve the user's need for magical transparency and the global need for distributed computing [11,12].

Finally, the movement toward opensource and away from restrictive licensing will greatly aid the need for interoperability, and foster an economic bonanza for the storage industry.

5 Summary

The SSSC is balloting the first storage system standards in the world, and these are the first five of ten Media Management System (MMS) standards. These standards and the future work of the IEEE will help bring a consistent approach to building interoperable storage systems, and to address the emergency need for improvements in the systems we all use to access information.

References

[1] Many documents, such as versions 4 and 5 of the MSSRM are available for public reading at http://www.ssswg.org, and serve only as working materials since the SSSC does not intend to standardize these.
[2] G. Peck, editor, IEEE Draft P1244.1/1.15 Standard for Media Management System (MMS) Architecture, December 1999
[3] Arthur C. Clarke's Third Law.
[4] J. Gray, P. Shenoy, "Rules of Thumb in Data Engineering", IEEE Conference on Data Engineering, San Diego, April 2000.
[5] N. Ranganathan, "TCVLSI Activities: A Message from the TC Chair", IEEE Computer Society TCVLSI Technical Bulletin, Fall 1998.
[6] L. Grover, "Quantum Computing". *The Sciences*, July/August 1999, pp. 24-30.
[7] T. Lewis, "Software Architectures: Divine plan or digital Darwinism?", IEEE Computer, August 1996.
[8] D. Taylor, "Beating the Software Crisis", *Object Oriented Technology: A Manager's Guide*. Addison-Wesley, Reading, Massachussets, 1994.
[9] "Will the Country Lose Its Memory?", USA Today, September 25, 1998.
[10] G. Gibson, et al., http://www.pdl.cs.cmu.edu
[11] G. Peck, Storage Object Presentation at NASA/Goddard Mass Storage Conference, March 1998. http://www.peck.com/~geoff/storage-peck-980325.pdf
[12] J. Nielsen, "The Impending Demise of the File System," IEEE Software, March 1996.