A *Straw Man* Proposal for a Standard Tape Format

Joel Williams

October 18, 1995

MITRE

A *Straw Man* Proposal for a Standard Tape Format

The Problem

Extremely large data systems, such as EOSDIS, must rely on Hierarchical File Storage Management Systems (FSMS) to stage files to disc as required for fast access, and to migrate files to tape for more economical storage when there is no requirement to keep them on disc. There is no standard format for such files when they are moved to tape, and so each FSMS uses a proprietary format. Changing from one FSMS to another would therefore most likely require the re-writing of all of the tape files written by the first system.

In addition large archives will maintain most of their data on tapes accessed by robotic servers, and for such files there is no standard which puts sufficient information on the tapes to represent multi-tape files and striped tape files. Such information would most likely be kept separately, and without a standard it would be in some proprietary format. Transferring a tape archive from one system to another would be extremely difficult.

Analysis

This situation has been analyzed in the paper dated March 15, 1995, *An Assessment of Requirements, Standards, and Technology for Media-Based Data Interchange* by David Isaac and Dana Dismukes of the MITRE Corporation. The following conclusions are taken from the Executive Summary:

- *Standards for media-based data interchange could save EOSDIS approximately $2M per storage system migration by reducing the need for additional computing capacity to support copy operations. Increased competition in procurements may result in additional savings.*
- *There is sufficient interest within the vendor community for a new standards effort and such an effort would likely succeed. Existing standards (including the System Independent Data Format or SIDF) do not adequately address the needs of hierarchical storage management systems (HSM). Nascent standards efforts in this area do not have sufficient momentum to survive without NASA promoting and supporting their efforts.*
- *Vendors are moving toward more generic and open formats and metadata. However, the formats and metadata used by emerging products are still not sufficiently general to support media-based data interchange. Vendors are not currently motivated to solve this problem by themselves, but could be motivated by requirements from NASA and other large users.*
- *Several other organizations have systems with projected data capacities that will cause them to also have requirements for media-based data interchange, but no one appears to be actively working to solve this problem.*
- *Consumer audio/video applications are not likely to incorporate the special needs of FSMSs without significant input from NASA and others with a vested interest in this capability. This has been done successfully in the past (e.g., ISO 9660 addressing standards for optical drives).*
- *The requirement to rewrite media as it ages reduces the potential for cost savings from media-based data interchange, but does not eliminate it. Media rewrites*

may be accomplished by an off-line system while format/metadata changes generally must be performed by the on-line system.

A Proposed Solution

In order to avoid the copy operation of re-writing an extensive tape archive when transferring tapes from one FSMS to another, there must be a standard way of transferring the tape-and file-level metadata. This metadata needs to contain sufficient information to enable the receiving system to reconstruct the file system represented on the tapes. Transferring this metadata would enable the receiving system to incorporate the tapes with a minimum of effort.

It is possible to transfer the metadata from one system to another in one of the following ways:

- Storing the metadata on the tapes which are to be transferred. Then the metadata automatically goes with the tapes. This automatically ensures that the metadata and the information on the tape are consistent, and allows for a partial or incremental transfer.

- Storing the metadata separately, on disk or on tape, and then transferring the metadata separately from the tapes, perhaps on tape, or perhaps over a network. Since an FSMS is likely to store the metadata on disc anyway for performance reasons, this only amounts to developing a standard format for this information.

The *Straw Man* standard for this activity represents the metadata as being on directory files on the tape. This allows the tape to be a self-describing unit, at least to the file level. The concept of having a standard format for the directory data for a collection of tapes in order to facilitate the transfer of that collection of tapes also needs to be further examined.

Characteristics of the *Straw Man* Standard Proposal

This proposal is for a high-level format which could encompass many different physical implementations. Another way of saying this is to say that the format would exist at the level of the UNIX *read()* and *write()* calls, and would not involve putting information directly on areas accessed by lower-level routines, such as the linear tracks of a helical scan tape.

The proposal would enable a receiving system to take in a number of tapes in random order and, if necessary, reconstruct the file system represented on those tapes. It would provide the data structures to reconstruct a file which spanned tapes, either because it is large, or because it has been striped, or both.

One inherent problem for which there is essentially no solution is changes in the name space. The receiving system may support an entirely different name space, and naming conflicts may require the receiving system to rename files when it incorporates a collection of tapes.

Technology Assumptions

Certain assumptions have been made about the tape technology. These assumptions appear to be true for most modern tape technology, at least for large capacity tapes.

The first assumption is that tapes may be partitioned. A partition is, for these purposes, simply defined as an area on the tape that may be updated in place, either by appending to the end of the partition, or by re-writing the partition. Such updates would not affect any other partition on the tape. This proposal requires at least two partitions, one for a directory and one for data files, and allows for, but does not require, additional pairs of partitions

It is assumed that there is some way to position a tape, given some sort of logical representation of tape position (relative byte offset, file mark number, track set id., etc.), but that this representation may vary from one technology to another.

Additional Information

The motivation for the development of this standard is to facilitate the transfer of tapes. It is likely, however, that the metadata specified in this or any other standard for tape format will include some information not directly required for tape transfer. Certainly this is more desirable than having two formats for different but perhaps overlapping sets of metadata. One example of metadata only marginally related to the transfer of information on the tape is the number of mounts of the tape. Nevertheless this is important information, and should go with the tape. Security information might also be required.

Similarly, there is additional information that would be required for tapes used for backup purposes, or for one level of a hierarchical file management system. This information is not discussed in this proposal, but would have to be addressed in any standard, perhaps simply by having an open user definable area that could be used for these purposes. Issues having to do with write-once media and two-sided platters will also have to be addressed.

Basic Framework of the *Straw Man* Standard Proposal

This framework is inspired by the EMASS tape format, as described in *The EMASS File Serv Technical Summary*, October 29, 1993, and by the proposed ISO 14417 standard. Both of these standards have directory partitions followed by file partitions, with the directory partition describing the files in the file partition. This directory partition in this proposal contains all of the information found in the directory partition described in these documents, but in addition includes explicit support for striped tapes and multi-reel files.

So in this proposal there are two types of partitions: directory partitions and file partitions. These partitions are laid out on the tape in consecutive pairs, and a directory partition describes the files or parts of files that are in the file partition that follows. Obviously, there must be at least one pair of these partitions. There are no length restrictions on either type of partition, other than that dictated by the length of the tape. Of course, it is expected that directory partitions will be significantly smaller than file partitions.

All information in the directory partitions will be stored in ASCII text. This avoids the need to specify the format of a binary representation.

The directory partition contains all of the metadata concerning all of the files in the following file partition. There are no headers or trailers on a file in the file partition. This avoids unnecessary file marks in the file partition.

Outline of the *Straw Man* Standard Proposal

As mentioned above, the tape is partitioned into a sequence of pairs of directory and file partitions, with the directory partition coming first and describing the following file partition. The first directory partition also contains information specifically about the tape, while the following directory partitions contain information only about the files in the following partition.

The first directory partition contains, therefore, the Volume Header, which has the following information:

- The Version of the Standard
- The Number of Partitions
- The Tape Identifier
- Times and Statistics Relating to the Tape

This is followed by the following Directory Information

- Sizes of the Directory and Following File Partition
- Number of Files
- File-level Information About each File
  - File Name
  - File Statistics
  - Security Information
  - File Location
  - Striping and Continuation Information

Subsequent directory partitions, if any, contain only the Directory Information

Further details of the format follow.