Progress in Defining a Standard for File-Level Metadata

Joel Williams
Systems Engineering and Security
7474 Greenway Center Drive
Greenbelt MD 20770
joel.williams@ses-inc.com
Tel: +1-301-441-3694
Fax: +1-301-441-3697

Ben Kobler
NASA/GSFC
Greenbelt MD 20771
ben.kobler@gsfc.nasa.gov
Tel: +1-301-614-5231
Fax: +1-301-614-5267

Introduction

In the following narrative, metadata required to locate a file on a tape or collection of tapes will be referred to as *file-level metadata*. This paper describes the rationale for and the history of the effort to define a standard for this metadata.

The Problem

Extremely large data systems, such as the Earth Observing System Data and Information System (EOSDIS), must rely on hierarchical File Storage Management Systems (FSMS) to stage files to disk as required for fast access, and to migrate files to tape for more economical storage when there is no requirement to keep them on disk. There is no standard format for such files when they are moved to tape, and so each FSMS uses a proprietary format. Files, particularly those which have been updated frequently, may be scattered over several tapes, and the information required to reconstruct the files is likely to be stored on disk separate from the tapes on which the files reside. Some file-level metadata information may be embedded as header information on each block on the tapes, so that any program reading the file would have to identify this header and understand that it is not really part of the file.

Changing from one FSMS to another would therefore most likely require the re-writing of all of the tape files written by the original system. For a large archive this would be extremely expensive.

Initial 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 work was funded by the Goddard Space Flight Center (GSFC) Earth Observing System Data and Information System (EOSDIS) Project. In the paper the following conclusions were made:

- 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.

- While there was no current standards activity addressing the problem, there was sufficient interest in the customer and vendor community to support such an activity.
- While the requirement to refresh media as it ages somewhat reduces the potential for cost savings from media-based data interchange, it does not eliminate it.

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 file-level metadata. This metadata needs to contain sufficient information to enable the receiving system to reconstruct the file system represented by the tapes. Transferring this metadata would enable the receiving system to incorporate the tapes with a minimum of effort.

In this context, we are not concerned with the semantics of the information contained in the files themselves. We are only concerned with the information required to identify the file (for example its name) and to associate it with one or more delimited bytestreams on one or more tapes. The bytestreams themselves would have to be ordered, as in the case of multi-reel files or striped files.

Initial Proposal for a Standard Tape Format

Encouraged by the MITRE study, the NASA GSFC EOSDIS project asked Joel Williams (then of the MITRE Corporation, currently of Systems Engineering and Security, Inc.) to develop a *Straw Man* standard and to gauge the reaction of the vendor and user community to this standards effort and to the proposed *Straw Man* standard itself.

The *Straw Man* standard was a tape format standard. The fundamental concept of the standard was to put a directory on each tape of the archive so that by reading the directory an application could determine where the files or file segments on the tape were located. The *Straw Man* was inspired by two proposed standards which include on-tape directories, the DD1 (ISO/IEC CD 14417) and the DD3 (ANSI X3.267) standards, and also by the EMASS practice of placing a directory on D2 tapes. In addition, during this same time period when the *Straw Man* was beging developed, IBM announced its Magstar product, which has a directory at the beginning of the tape. Subsequently, Sony has announced a tape which has a directory on a chip on the tape cassette.

The *Straw Man* proposal was for a logical tape format, and would have been written at the application (FSMS) level. It could therefore apply to any tape technology, although it did require partitioning of the tape in order to be able to update the directory without having to re-write the entire tape.

There are two different types of on-tape directories: those that contain information about the file (such as its name, for instance) and those that primarily contain information allowing the fast positioning of the tape.

The DD3 standard, as outlined in Figure 1, is of the second type. It allows one to position the tape quickly, but the (file name, position) mapping must be done at a higher level.

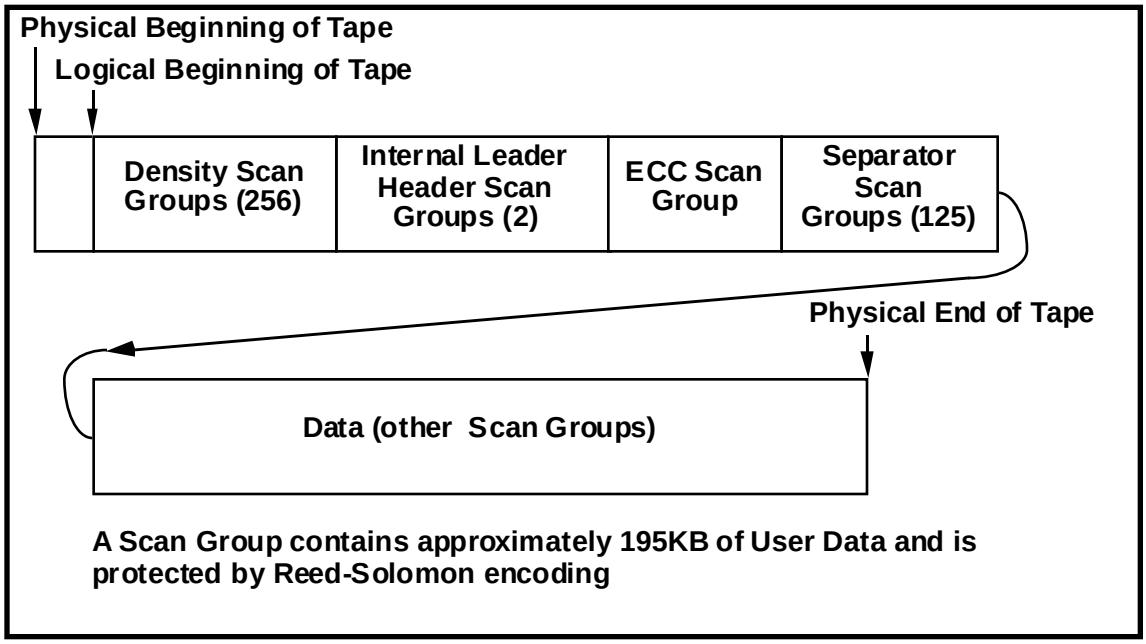

Figure 1

Each of these scan groups contains approximately 195KB of user-written data, including end-of-record markers, but exclusive of error correcting codes. The Internal Leader Header scan groups are reserved for directory information, and whenever the tape is mounted, they are read into the drive memory, then modified before the tape is dismounted. It contains information that allows for fast positioning of the tape, and additional information such as

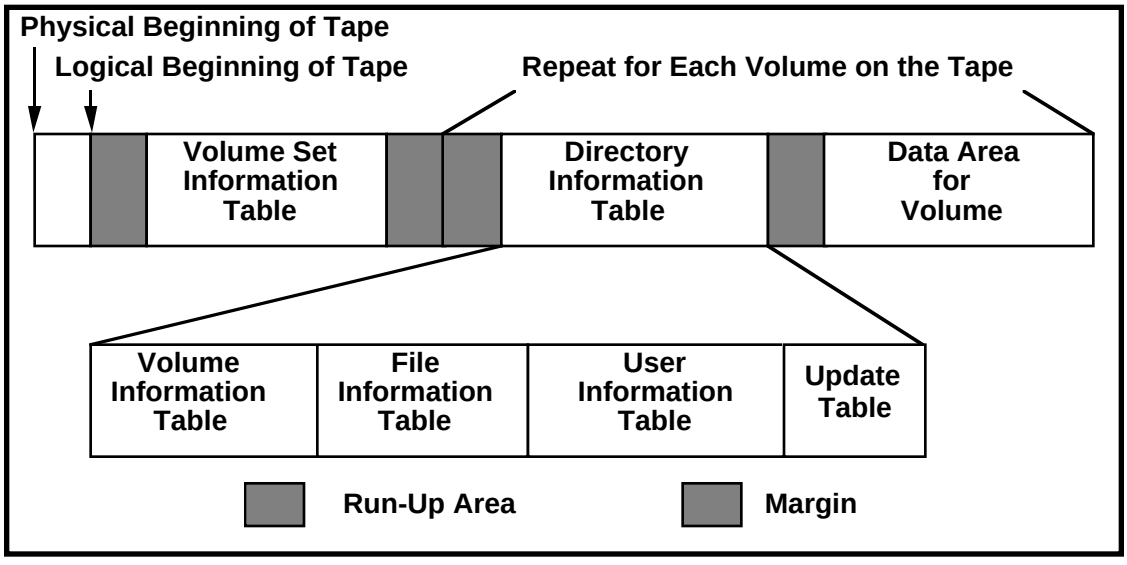

Figure 2

the volume id, the number of mounts, the time and date of the last five mounts, and the tape manufacturer.

The Magstar directory provides similar functionality, and also includes extensive information concerning any errors which may have happened when the tape has been accessed.

The DD1 proposed standard is outlined in Figure 2. It contains all of the file-level metadata which would be required to locate a file on the tape, given the file name.

The following blocks are defined:

- The Volume Set Information Table. This is at the beginning of the tape, and contains information on the number of volumes contained on the tape, and identifies the cassettes in a volume set. This information supports files which are striped across different tapes. There is only one Volume Set Information Table on each tape.

- The Directory Information Table. There is one of these for each logical volume on the tape, and it consists of the following four blocks:

    - The Volume Information Table, which describes the volume
    - The File Information Table, which contains information used for positioning the tape to files in the Data Area
    - The User Information Tables, which contain information on each file in the following Data Area, such as the file name, version number, creation date, etc.
    - The Update Table, which is used to ensure that the directory has been updated properly.

- The Data Area, which contains the files in the volume described in the Directory Information Table.

The *Straw Man* proposal looks very similar to the DD1 proposal, and is outlined in Figure 3.

In addition to information such as the file name and its location, the directory contains the following information:

- Pointer to the next file segment if the file is continued
- Pointer to the first file segment if the file does not begin in this location
- Pointers to the other stripes if the file is striped

In this way multi-reel files and striped files are supported.

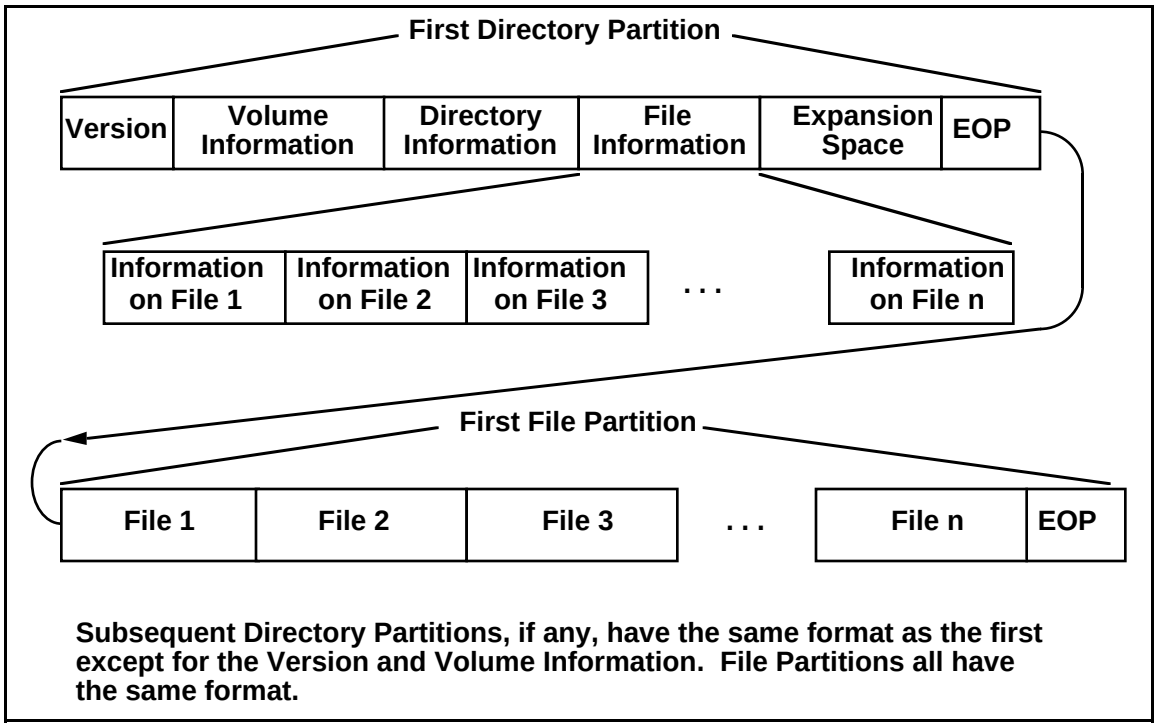

Figure 3

Presentation of the *Straw Man* and Reactions to it

This *Straw Man* proposal was first circulated at the Fourteenth IEEE Symposium on Mass Storage Systems at Monterey, California in September, 1995. Subsequently, it was briefed to THIC, the ISO/CCSDS Archiving Workshop at GSFC, the ANSI X3B5 Committee, the AIIM Optical Tape Study Group, and to individuals at the National Security Agency. Several changes were made to the original proposal to lower the overhead of having to re-write the directory whenever the tape was updated.

The decision was made to form a Study Group under the auspices of the Association for Image and Information Management (AIIM). The group's name is the *File-Level Metadata for Portability of Sequential Storage (FMP) Study Group*, and the first meeting of the group was on April 1 at the AIIM International Convention in Chicago Illinois. The group is chaired by Fernando Podio of the National Institute of Standards and Technology (NIST).[1]

The First FMP Meeting

The first FMP meeting was held April 1-2 in Chicago. The following organizations were represented:

Ampex
Applicon
Datatape
EMASS

[1] For further information about the FMP study group, contact Fernando Podio at fernando.podio@nist.gov (Fernando Podio)

Fermilab
NASA GSFC
LDS Church
HPSS
Kofax Image Products
Lawrence Livermore National Lab
Legacy Data Systems
Library of Congress
Los Alamos National lab
Lots Technology
LSC Inc.
Micro Design International
MITRE
National Media Lab
NIST
Research Libraries Group
Systems Engineering and Security
Storage Technology
Terabank Systems

The *Straw Man* proposal was presented at this meeting, and various other presentations were made. The consensus of the group was that an on-tape directory containing file-level metadata was impractical for performance reasons. There was general agreement, however, that there needed to be a standard for the export of file-level metadata, and that it was advantageous to work toward that standard under the auspices of AIIM.

In this regard, the group agreed to a statement of work as follows:

> *The AIIM FMP SG will document an interchange format for file-level metadata for data stored on sequential storage media. This approach does not concern the data format on the physical media or drive.*

Figure 4 graphically depicts how this interchange standard would work.

The original system would of course maintain its own metadata in some form, which could remain proprietary. This metadata would enable it to manage the tapes under its domain. When it came time to migrate these tapes to a new system, the original system's metadata would be exported to the public standard. The new, receiving system would read this standard metadata and convert it to its own representation, which might also be proprietary. In this way, the new system would be able to take over the management of the tapes without re-writing them.

The major challenge in developing this standard for file-level metadata export is to develop something that is broad enough to cover current and anticipated practice. Cooperation from the vendor community will be important in meeting this goal.

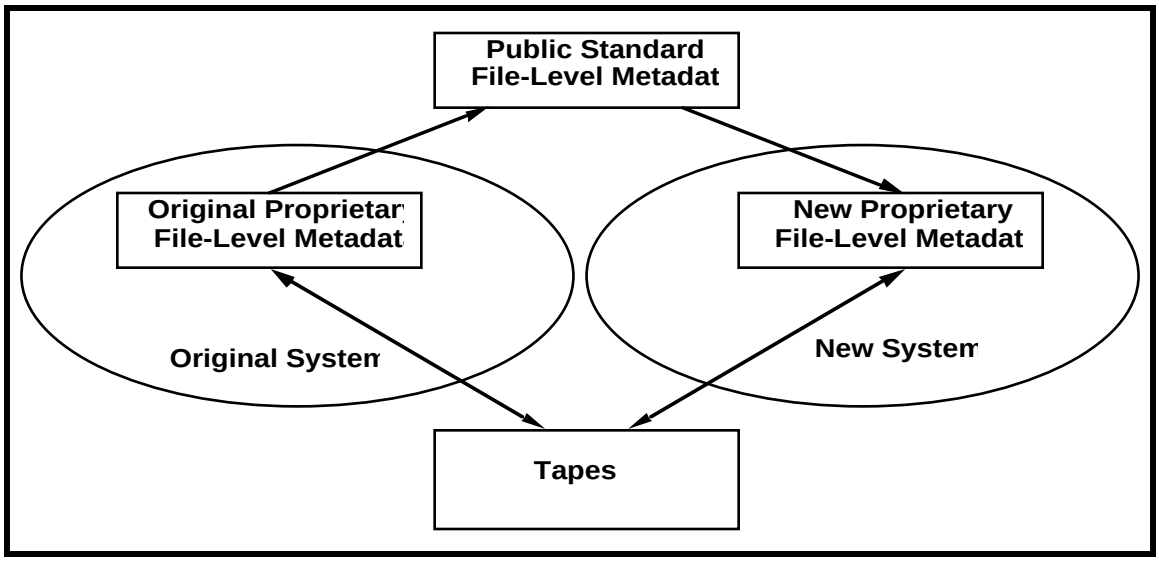
Public Standard
File-Level Metadat
Original Proprietar
File-Level Metadat
New Proprietary
File-Level Metadat
Original Systen
New Systen
Tapes

Figure 4

The Second FMP Meeting

The second meeting of the FMP Study Group occurred June 17-18 at the AIIM headquarters in Silver Spring, Maryland. The following organizations were represented:

BDM
Datatape
Department of Defense
EMASS
NASA GSFC
Hewlett Packard
HPSS
IBM
Lawrence Livermore National Lab
Lots Technology
LSC Inc.
IIT Research
National Media Lab
NSA
NASA Langley
NIST
Norsam
Systems Engineering and Security
Storage Technology
Terabank Systems

Discussions at this meeting centered on determining the data elements which constitute the file-level metadata required to be exported. These elements, it was determined, fall into three categories. The lists below characterize and contain examples from each category which were discussed at the meeting.

- Data elements having to do with the tapes

  - Tape ID
  - Tape Universally Unique Identifier (UUID)
  - Tape model or type
  - Statistics about errors on the tape
  - Compression information
  - Exporting FSMS and vendor name, operating system version, hardware identification

- Data elements having to do with the files on the tapes

  - File Name (including version information, if any)
  - File Universally Unique Identifier (UUID)
  - Method of tape addressing
  - Location of file segments, including *magic cookie* information if it exists
  - Striping information
  - Information identifying multiple copies of the file
  - Account IDs for billing purposes
  - File Family
  - Tape set
  - Volume group

- Data elements having to do with the file system represented by the tapes

    - Directory and file structure
    - Hard and soft links
    - Principal names and groupings for security purposes

The next meeting of the FMP Study Group is October 1-2 at the AIIM Headquarters, 1100 Wayne Ave. in downtown Silver Spring, Maryland.