

File-Based Media Workflows using LTFS Tapes

Arnon Amir, David Pease, Rainer Richter,
Brian Biskeborn, Michael Richmond, Lucas Villa Real
IBM Almaden Research Center
650 Harry Road, San Jose, CA 95120

[arnon, pease]@almaden.ibm.com, [rainer, bbiskebo, mrichmon, lucasvr]@us.ibm.com

ABSTRACT

While digital video cameras have existed for over two decades, *digital video cassettes* are still the primary storage medium in professional video archives. One of the major inhibitors in the transition to file-based workflows and media archives is the lack of an affordable, portable and archive compatible storage medium for the vast amounts of content produced.

We address this need by a) defining the Linear Tape File System (LTFS) tape format for storing files, file properties, hierarchical directories and extended attributes; b) building file system software that allows LTFS tapes to be used in the same way as portable storage devices and c) leveraging LTFS to create efficient file-based media workflows.

In the exhibit we present LTFS on LTO-5 (Linear Tape Open, Gen 5) tapes. We demonstrate file-based workflows with storyboards, video proxies and partial video restore of MXF (Material Exchange Format) professional video content. LTFS on LTO-5 tape can be 20 times higher in capacity, 10 times faster and 40 times cheaper than digital video cassette media. Furthermore, it combines the benefits of tape-based and file-based workflows. The new tape format streamlines file-based production, from video capture and transport to long-term archive. The tape format and file system implementation are available as open source.

Categories and Subject Descriptors

H.3.4 [Information Storage And Retrieval]: Systems and Software

Keywords

Self-describing tape, DuraBytes, LTO-5, LTFS, MXF, archive

General Terms

portable tape storage, LTO, LTFS, MXF

1. INTRODUCTION

File based archives for media have been replacing traditional video tape vaults and VTR's. Rows of video tapes on shelves with metadata on printed labels are being supplanted by automated robotic data tape libraries managed by integrated Media Asset Management systems providing metadata search, thumbnails and proxy views. Clip lists are created via proxy and the corresponding high resolution essence delivered directly to edit bays or playout servers. This major paradigm shift, known as the *Digital Workflow*

Transformation, benefits flexibility, fast access to assets, and lower storage costs. One of the major challenges in this major transformation is digital video storage [6, 1]. File-based video production requires the storage, transportation, processing, and archiving of vast amounts of data contained in media files. One hour of professional raw video in AVC-Intra 100 format (100 Mbps) requires 45 GB of storage space. A major broadcaster may have hundreds of thousands of hours of video in their archive, amounting to many petabytes.

It has been said, "It's hard to beat the bandwidth of a truck full of tapes." Traditional video tape archives, as illustrated in Figure 1(A), have the advantage that once material arrives it can be "archived" with a high degree of parallelism (the number of staff working the tape vault) and bandwidth (the speed of the staff in putting tapes on shelves). In contrast, file based archive systems (Figure 1(B)) must copy files from a source, typically a SAN or set of portable disk drives used for transport, to the data tapes being managed in the tape library. Bandwidth is constrained by the speeds of the source devices, server and network bottlenecks, the number of tape drives and their write speed.

Data tape archives, common across the IT industry, have high capacity and storage density, high longevity, high throughput, and are cost effective [8]. However, their tapes are not portable; all file metadata (e.g., file names, dates, and other attributes) and directory information is kept not on tape but in a central database on a *Storage Manager* server.

Table 1 is a comparison of various common video storage media. LTFS tapes are both portable and archive-able (long term), have fast throughput and are most economic for long-term archive of large amount of data.

In the *DuraBytes* project we have developed a new, self-describing tape format, the Linear Tape File System (LTFS), and used it to enable efficient media workflows.

2. LTO TAPES AND THE LINEAR TAPE FILE SYSTEM (LTFS)

Linear Tape Open (LTO) [2] is an open tape standard for which several companies produce and market compatible LTO drives and media. The current generation of LTO is Generation 5 (LTO-5). LTO-5 cartridges have an uncompressed capacity of 1.5TB. The drives have a streaming data rate of 140MB/second, and do both compression and encryption on the fly. LTO-5 is a *dual partition* tape. It supports up to two partitions on tape; each partition can be written without affecting the other.

Our new Linear Tape File System (LTFS) [7] implements a POSIX-compliant [5] file system using LTO-5 dual-partition

Table 1: Archival properties of data storage devices and media (at time of writing). Cost of storage media is per hour of HDCAM SR video at 440Mbit/Sec. LTO tape compression is not counted, and would improve capacity and data rates.

Storage Media	Capacity [GB]	Portability	File-based	Density TB/ft^3	Longevity	Throughput [MByte/Sec]	Energy	Cost \$/Hr
HDCAM SR	410	high	no	very low	30 yr	55 (video)	very low	120.00
DVD	4.7	high	yes	low	5-10 yr	21 (at 16x)	low	10.00
Blu-Ray	50	high	yes	low	5-10 yr	54 (at 12x)	low	4.50
External 2.5" HDD	500	high	yes	medium	3-5 yr	20-60 (USB-2)	medium	9.00
Internal 3.5" HDD	2000	low	yes	medium	5-7 yr	40-160	high	7.80
Internal SSD	500	low	yes	medium	2-7 yr	12-100	low	135.00
P2, SxS flash card	64	high	yes	low	2-7 yr	100	low	750.00
LTO-5 (Managed)	1500	low	no	high	30 yr	140	very low	3.00
LTO-5/LTFS	1500	high	yes	high	30 yr	140	very low	3.00

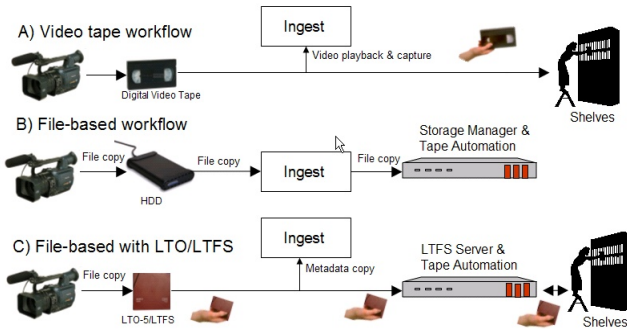


Figure 1: Ingest and archive workflows.

tapes. LTFS makes a tape as easy to use as a removable storage device. After loading and mounting an LTFS tape cartridge, a user can see the directories and files stored on the tape through the standard tools available on their system (e.g. a graphical file explorer, application open menus, command-line tools, etc.). They can traverse the directories (or folders), read and modify files from within their applications, and open applications by double-clicking file icons. Files can be updated, and in general can be used as if they were on typical storage devices (disk, flash, DVD etc.).

The LTFS technology consists of two major components: an on-tape index format, and a file system implementation. The tape index is written to the smaller partition dedicated to that purpose, while the data is written to the larger partition that occupies the remainder of the tape. The index is recorded in an easy-to-understand XML format, and is in the process of becoming an open industry standard.

The file system has been implemented on Linux and Mac OS X; development is currently underway for a native Windows version. Tapes written on any of the platforms are completely interchangeable with the others. The Linux and Mac OS versions are freely available as open source ¹.

3. MXF WORKFLOWS WITH LTFS

Video tape workflows benefit from the direct archival of tapes, while traditional file-based workflows provide high flexibility, manageability, automation, and global access. A workflow based on LTFS tape combines the best of both paradigms. The file system mapping of the data on tape provides improved and flexible access to metadata, clip extraction, and the media files, while the use of LTFS enables the “bandwidth” of physical tape movement. Tapes be directly “archived” into a tape library or stored on shelves with no replication of the essence files.

¹See <http://www.ibm.com/systems/storage/tape/ltfs/>

The Material Exchange Format (MXF [4]) is an industry standard for packaging together multiple media files and associated metadata. It is used for data interchange, transportation, post-production, and archive. MXF is supported by almost all of the major systems vendors (Panasonic, Sony, Avid, Adobe, Apple, IBM etc.) and adopted by broadcasters (BBC, NPR, Turner etc.).

The example MXF digital workflow starts from a camera, travels via LTFS tape and ends in the archive. Today’s broadcast cameras produce MXF clips directly on a flash card (Panasonic P2 or Sony SxS). The media files are then copied from the flash card to LTFS on LTO-5 tape. We analyze the video, detect shot boundaries, extract key-frames, and create storyboards [3]. The essence MXF files are stored on the LTFS data partition and the metadata is stored on the index partition for faster access. This metadata is used for content ingest into Media Asset Management. The high resolution essence can be efficiently accessed using the MXF index, restoring only the needed portions.

LTFS tape combines the benefits of tape long enjoyed by the IT industry (density, longevity, power consumption, and economy) with the portability, familiarity and ease of use of file system storage devices. In doing so, it overcomes the limitations which have hindered wider adoption of tape based archives, particularly in the media industry. The established video cassette paradigm but with file based access is now possible and can be utilized to create efficient workflows and long term archives.

4. REFERENCES

- [1] The Digital Dilemma, Strategic issues in archiving and accessing digital motion picture material, 2007. The Science and Technology council of the academy of motion picture arts and sciences.
- [2] Aberdeen Group. LTO Tape Has It All, Dec 2001.
- [3] A. Amir, G. Ashour, and S. Srinivasan. Automatic generation of conference video proceedings. *JVCI*, 15(3):467 – 488, 2004.
- [4] B. Devlin, J. Wilkinson, M. Beard, and P. Tudor. *The MXF Book: An Introduction to the Material eXchange Format*. Focal Press, Apr 2006.
- [5] IEEE. *IEEE Standard for Information Technology: Portable Operating System Interface (POSIX). Part 1: System Interface*. IEEE Standards Association, USA, 2001.
- [6] A. Kovalick. *Video Systems in an IT Environment: The Essentials of Professional Networked Media*. Focal Press, Jun 2006.
- [7] D. Pease, A. Amir, L. V. Real, B. Biskeborn, M. Richmond, and A. Abe. “LTFS: The Linear Tape File System. In *Proc. 26th IEEE conf. on MSST*. IEEE, May 2010.
- [8] D. Reine and M. Kahn. Disk and Tape Square Off Again – Tape Remains King of the Hill with LTO-4, Feb 2008. Clipper Notes.