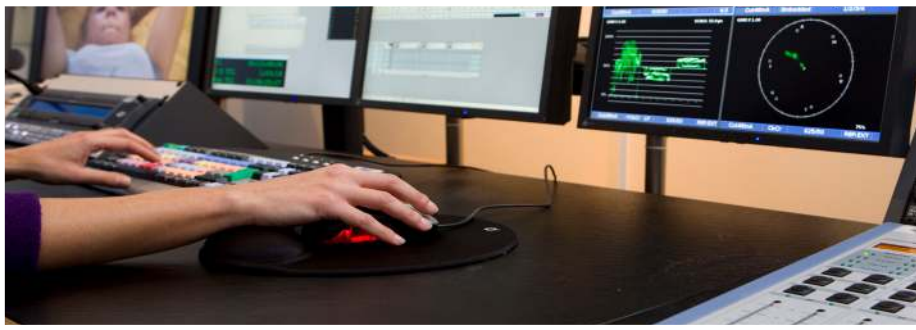
LTFS Hits the Mark in Media & Entertainment:

An In-Depth Introduction to LTFS for Digital Media
by Media Technology Market Partners LLC



LTFS
Linear Tape File System

In the media and entertainment industry, content is king. The majority of this content is now produced in digital form and virtually all of this content has digital distribution; that content is now digital data. Protecting that content, the lifeblood of this industry, with the right data storage solution is more important than ever.

MTMP Media Technology Market Partners

LTFS Hits the Mark in Media & Entertainment:

An In-Depth Introduction to LTFS for Digital Media

Executive Summary

In the media and entertainment industry, content is king. The majority of this content is now produced in digital form and virtually all of this content has digital distribution; that content is now digital data. Protecting that content, the lifeblood of this industry, with the right data storage solution is more important than ever.

This digital content is growing rapidly in size due to the demands for more data intensive formats (HD, 4K, 8K, 3D, etc.) and more digital effects. Storage options for these digital files can be expensive and cumbersome, adding to production costs that are already soaring to record levels. Disk, tape and solid state disk (SSD) are the most prevalent technologies, each with its own set of tradeoffs.

With the advent of Linear Tape File System (LTFS), a new open source file system available with LTO-5 technology, tape has broadened its already significant appeal to the media and entertainment industry. Coupled with the large capacity of an LTO-5 tape, LTFS provides a robust, self-describing open format interchangeable standard with improved ease of use via operating system integration. Archiving data to tape is now as easy as using drag-and-drop with a USB flash drive. It has no application software dependencies, offers support for large and numerous files, and can have a lower total cost than traditional managed tape storage.

In short, LTFS has come of age for Digital Media at just the right time. LTFS and associated products have already been recognized by the industry, winning several awards, culminating with a prestigious Technical Emmy Award in 2011.

This whitepaper offers deeper insight into how LTFS provides significant advantages for media and entertainment applications. It includes an overview of LTFS, evaluates the pros and cons versus other storage technologies, and takes a look at how LTFS enhances industry workflows and processes.

Table of Contents

What is LTFS?

Linear Tape File System (LTFS) is a multi-vendor open format specification for storing data files and an index of those files together on data tape. It relies on the support of the tape technology for partitioning that was introduced in the LTO technology roadmap with the LTO-5 format specification. The metadata stored in one partition provides an index or directory of the data stored in the other partition.

An LTFS driver is software that extends operating system functionality, enabling the data on an LTFS tape to be accessed by a file system similar to that found on disk drives using commands like copy, move, and delete. Similarly, applications such as file explorers can now access data on tape via user friendly drag and drop interfaces.

The combination of the tape format and operating system drivers is LTFS.

Why was LTFS developed?

Tape for data storage is an established, trusted, and proven technology. Although the tape "model" hasn't changed dramatically over the years, the speed, storage density and features of data tape has improved significantly, ultimately providing reliable and inexpensive storage as a sequential storage medium.

LTFS was developed to present a standard file system view of the data stored on the tape media. With LTFS, accessing files stored on the LTFS formatted media is very similar to accessing files stored on other forms of storage media, such as disk or removable USB flash drives.

While this functionality is useful for a wide range of applications, it particularly appeals to those with unstructured data. This is especially true for rich media related to the media and entertainment industry with its numerous and large files of content (i.e. video, digital cinema, audio) to manage and preserve. The high data transfer rate and dense cost effective storage of LTO tape were appealing to the industry, but somewhat difficult to use… until LTFS.

How does LTFS work?

LTFS makes data on tapes accessible in a manner like accessing data on disk drives, USB flash drives, etc. These devices contain blocks of data as well as an index (e.g. FAT32-File Allocation Table), which holds metadata information about a file (e.g. name, size, creation date, extended attributes, etc.) and a map of which blocks of data belong to which file.

Prior to LTFS, tapes contained only blocks of data and file marks in sequential order, without a directory to locate a specific file. External applications such as backup and archive storage

LTFS at a Glance

An LTO 5 tape cartridge is divided into 2 partitions

- One partition contains an index
- One (larger) partition holds the data

The index partition can be updated without affecting the data partition

- Analogous to disk drives, USB flash drives, etc.
- Index data is cached for faster operation

Human readable XML index file maps which blocks belong to which files

- Parseable by software or manually to minimize dependencies for robust access

Operating system drivers

- Data on tape appears as a standard file system
- Data on tape is usable by generic applications which may be unaware they are accessing data on tape
- Provides the ease of use similar to other storage devices

management software maintained the mapping of which blocks belonged to which files. If that mapping were lost, the data on the tape would be present but not easily usable.

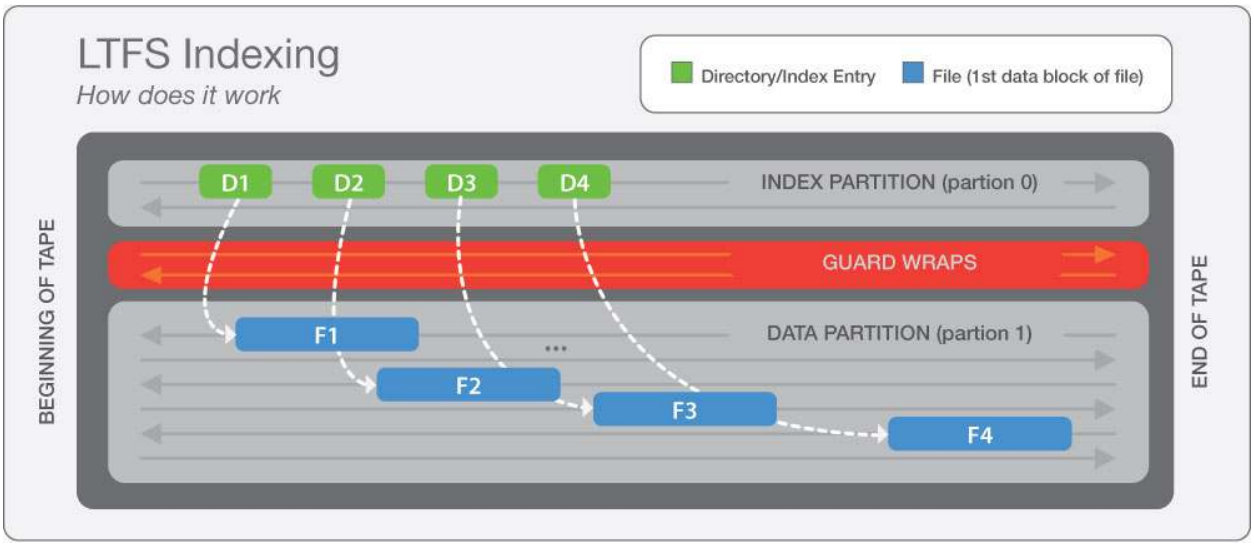
Due to the serial nature of tape, there is no inherent "in place" update capability. If a block is rewritten, all subsequent blocks must also be rewritten. Alternatively, updating an index at the current end of a tape is not practical as it would move to a new physical location every time data is added.

But with LTFS, tape doesn't have this limitation. The standard LTO-5 tape cartridge is segmented into two partitions, one for the data and one for the index, so the index partition can be modified as needed **without** affecting the data partition.

The combination of the partitioned LTO-5 cartridge, format of the LTFS index, and conventions for the index placement is the magic behind LTFS. This creates a self describing tape where a user can see the tape cartridge and its contents in the operating system directory tree browser and can copy, paste, or drag and drop files/folders to and from the tape. LTFS also supports extended attributes which enable custom file metadata.

Within the partitions, the tape contents are still stored via the trusted decades old format of blocks of data and file marks. It is the interpretation of the contents of the blocks that is LTFS, **NOT** some radical alteration of established tape technology. LTFS is as remarkable for what it is not as for what it is.

Note that the LTFS specification defines the format and layout of the tape, but not how it got there. Tape system vendors supply LTFS drivers for various operating systems which read and write according to the LTFS specification. However, anyone is free to create their own software to utilize the LTFS specification instead and some already have. An LTFS format verifier application is available to verify conformance with the specification.



What are the Benefits of LTFS?

LTFS combines the traditional advantages of tape (reliability, longevity, portability and low cost) with a modern familiar interface providing users faster easier direct access to files on data tape.

Open Specification

The LTFS format specification and implementation scheme are in the public domain and have been adopted by the LTO Program group, which provides open licensing of the LTO Ultrium technology to multiple vendors. Keeping user data stored in an open format has the advantages of allowing for multiple providers, increased options for users and interchange between both competitive and complementary offerings.

Portable

LTFS tapes contain an index which describes the contents. There is no required connection with the software system that wrote the tape. This differs markedly from proprietary backup and archive storage management applications which usually require special tape formats or metadata transfers from the application to provide access to data on the tape. In this regard, LTFS tapes, portable disk drives and generic tar formatted tapes (tar is the tape archive program for UNIX and now widely used in other operating systems) are equivalent in portability.

Easy to Use

LTFS is an extension of the operating system, so now tape is easy to access in a manner like accessing disk drives, USB drives, data DVD's, etc. A user only capable of using a mouse to drag and drop files can now be a tape user, heralding a new era of accessibility.

Note that while an LTFS tape *appears* to the system just like a disk, it still *acts* like traditional data tape with its latency and serial access. Numerous applications will work well without changes.

Multi-Platform Support

To complement the open specification of LTFS and the portable, self-describing, characteristics of an LTO tape written using LTFS, most LTO tape system vendors provide LTFS versions for all three dominant operating systems used by the media industry: Linux, Windows and Mac OS. This further enhances the benefits of interoperability, portability and ease-of-use.

Reliable and Robust

The LTFS specification contains a number of features designed specifically to enhance data recoverability under various circumstances. Multiple copies of the index are stored in both the index and data partitions. Older copies of the index are retained so the tape can be "rolled back" to a previous state with use of LTFS functions. The index format is in a human readable XML format.

LTFS adds to LTO-5 tape reliability features, which include tape drive encryption to secure sensitive information and read-after-write verification to help ensure data integrity.

Economical

While portable disk drives may have an initial acquisition cost advantage over the initial cost of a tape drive and cartridge, the expected life, long term costs, robustness and reliability of a disk system may not be satisfactory in comparison. For moderate and larger storage applications, tape drive costs are amortized over a larger amount of data storage. Thus, less expensive LTO media ($35/TB, or for example, $0.83/hr of 50 Mbps video, as of February, 2012) tip the scales in favor of tape. This is especially true when the total costs of ownership are fully calculated. For energy costs, this is significant because tapes require no power or cooling when idle. Additionally, LTO-5 tape cartridges will be readable by LTO tape drives two generations in the future (e.g. an LTO-7 drive will be able to read an LTO-5 tape), helping to preserve investments and easing migration.

Technology Comparison

Comparing LTFS formatted tape versus tape written with Tar or by proprietary HSM/backup/archive applications shows LTFS is equal or superior in all categories.

LOWEST ◄► BEST	LTFS Tape	Tar Tape	Managed Tape
Open Format	■	■	□
Portable	■	◪	□[1]
Ease of Use	■	◪	□
Reliability[2]	■	◪	◪
Speed of Access	■	□	■
Economical	■	■	□

NOTES:

1. *Managed tapes have to refer back to the applications that manage them*
2. *LTFS is the only technology with redundant easily readable indices.*

Compared to disk alternatives, either disk arrays or removable disks, LTFS formatted tape may be the better choice depending primarily on speed of access requirements.

LOWEST ◄► BEST	LTFS Tape	Disk Array	Removable Drives
Open Format	■	■	■
Portable	◪	□	■
Ease of Use	◪	◪	■
Reliability[1]	■	◪	□
Speed of Access	□	■	◪
Economical[2]	■	□	◪

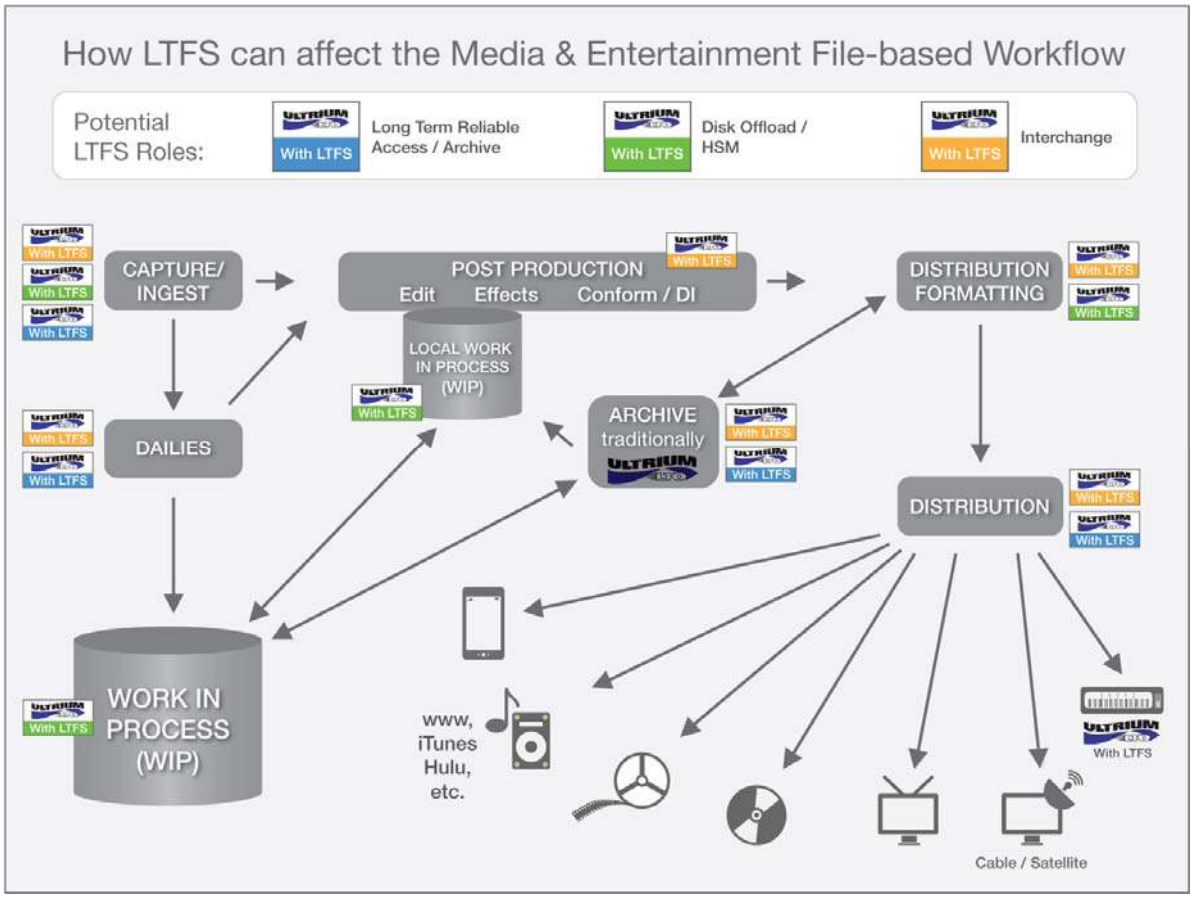
NOTES:

1. *LTFS is the only technology with redundant indices. Also disks are mechanically fragile.*
2. *In medium to large storage use cases, storage controller, disk media and power costs exceed tape library and media costs*

Where does LTFS fit into media workflows?

LTFS opens up new opportunities for media and entertainment storage and distribution, providing lower-cost options to many of the systems currently in place.

The diagram below presents a high-level view of a media workflow from capture through distribution. Before LTFS, LTO tape had been used for archive and backup, including the backup of content created by Capture/Ingest. Now with LTFS, LTO tape can be a primary interchange medium and provide a less expensive and easy-to-integrate working storage option for any of the functions not requiring low latency access to the content. The diagram shows which functions are affected and how.



LTFS directly affects two major trends having significant impact on media and entertainment companies and digital media producers: the shift to file-based workflows and ever increasing storage demands.

With the advent of digital technologies, moving image (video and film) content producers and distributors are transitioning from analog/linear workflows based on film or videotape technology to digital/non-linear workflows based on the manipulation of data files. Historical workflows relied on physical video tapes and film elements as the medium for moving content from one process to the next. In the digital workflow the file is the unit of work being moved, whether by network file transfer, removable disk drive or LTO tape.

Many organizations have already completed this transition. This has been commonly referred to as "moving to a tapeless workflow," though it is more accurate to call it moving to a "*video*tapeless"

workflow. With LTFS, users can go “Back to the Future” and enjoy the low cost and robustness of tape while supporting the most modern of file based media workflows.

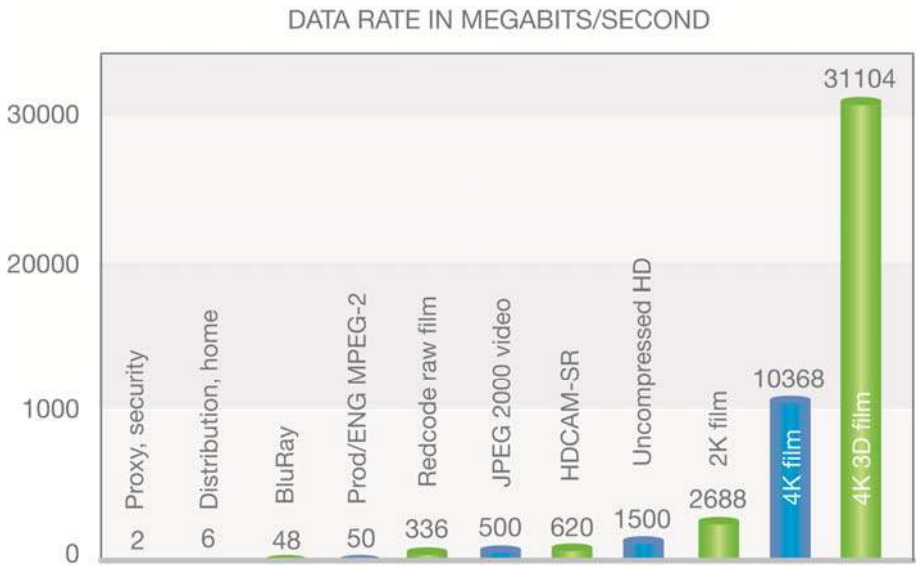
COMPARISON OF LTFS / LTO VS. VIDEO MEDIA

Format	GB	Cost per Unit[1]	$/GB	LTFS Cost Advantage
LTFS/LT05 tape	1,425[2]	$ 52.50	$ 0.04	—
G-Tech G-Drive	3,072	$297.00	$ 0.10	3X
HDCAM SR tape	698	$217.00	$ 0.31	8X
XDCAM disc	50	$ 67.50	$ 1.35	37X
P2 SSD card	64	$828.00	$12.94	350X

1. *Cost per unit based on market prices as of February, 2012*
2. *Capacity of an LTFS formatted LTO5 tape data partition*

The benefits of LTFS make it an excellent fit in a number of common file-based media workflow use cases.

The other fact of life in media & entertainment is the demand of ever-increasing visual resolution and complexity (HD, 4K, 3D, etc.) creating more and larger files that must be managed. Keeping hours of such media online on disk quickly becomes cost prohibitive. This is especially true with uncompressed production-quality content as the following chart dramatically illustrates. LTO tape, especially with LTFS, can address the challenge.



** The selected bit-rate for a given format is meant as a representative example for that format. Each format’s bit-rate can vary due to multiple factors.*

LTFS can work with LTO hardware-based lossless compression which can provide bandwidth and capacity benefits depending on the data content.

LTFS in Production

LTO tape is already firmly established in the media production environment. This is as a result of a mandate to many feature motion picture productions by insurance companies that content (data) captured on set or on location be archived to LTO tape on a daily basis. LTFS helps facilitate that process by being easier-to-use and more robust. The open and self-contained LTFS format is useful if the tapes are to be sent offsite, archived or shared with a variety of recipients.

Whether the production is for motion pictures, scripted television or live events (news & sports) the content captured by the cameras needs to be securely stored and readily available for the next processes in the workflow. LTFS is particularly well suited to productions shooting with digital cameras as there is no required pre-processing or packaging of the files produced by the cameras. The existing digital files can simply be copied to the LTFS tape volume. A number of vendors provide on-set solutions and several of these have already made the transition to open LTFS format tapes.

For production LTFS efficiently supports several important requirements:

Camera media reuse

Digital cameras encode motion images directly to solid state devices (SSD) or removable disks in the camera. These media are quite expensive (3x to over 350x the equivalent media cost on LTO). Fast transfer of their contents to tape with LTFS enables reuse of this expensive media, reducing the number of SSDs or disks that must be purchased or rented.

Backup

Backup of daily footage to LTO tape is a common requirement as the loss of a day's worth of production is very costly. LTFS facilitates backup by enabling small portable independent systems to easily write daily content to tape.

Transport

The density and cost of LTO-5 tapes with the self describing capabilities of LTFS combine to create a very effective transport medium. Large amounts of data can be sent more quickly and economically than network based transmission methods. This is especially compelling for digital productions which can produce terabytes of data for every day of shooting. Encryption features of LTO tape help secure the data in transit.

Economic direct access to data

For any file-based production workflow an LTFS enabled tape drive can feed workstations or networks with content directly and relatively fast, similar to a disk and unlike most traditional tape systems. An application via the operating system always has a direct and persistent view of a mounted LTFS tape and the files it contains. Consequently, in a workflow where access to a file is expected to be fast but not instantaneous, such as a stock footage collection or archive footage of an ongoing news story, an LTFS tape is an effective and economical choice for storage.

Archive

LTFS formatted tapes can be easily imported into an LTFS compatible archive by simply reading the index and adding the file metadata to an archive manager's catalog. Conversely, traditional solutions that utilize separate media for transport and archive require all the data be recopied. With LTFS, there is no need to read the much larger data partition or transfer the data to other storage media. The transport media and the archive storage media are one and the same under this scenario. The "import bandwidth" of tapes being added directly to a library en masse *far* exceeds any solution which requires movement of the actual data.

The LTFS index partition can also deliver benefits not available with other tape alternatives. Some digital cameras simultaneously create a low resolution proxy video corresponding to the high resolution capture. The low res proxy files are frequently used as a convenient way to review video content since playback of the high resolution video is computationally and bandwidth prohibitive. On an LTO-5 tape, the LTFS index partition is allocated with about 35 GB of space. While it holds the index data, it can also hold small user data files. Data written to the index partition will be readily accessible near the physical beginning of the tape. LTFS enabled applications can quickly load the proxy files and allow users to review the contents of the tape without the latency associated with reading the much larger data partition.

LTFS in Post Production

Post production is where all the content elements of a project come together: camera-captured content, digital effects, graphics and sound. It is in the post production process where all these elements are repeatedly edited, enhanced, assembled and reviewed. With the advent of digital intermediates the entire project becomes a collection of data files and post production has become a very technologically challenging part of the media workflow. This aspect of media production involves medium to high bit rate content (very large files), multiple processes and often multiple companies and departments working on the files. Making the content (the files) readily available is increasingly both an economic and logistic challenge.

In an ideal post production storage scenario all the content (files) for a production (film, TV show, commercial, etc.) would be stored in a high performance shared storage system with access control and integrated backup. The storage environment would ideally be accessible by all the post production functions and vendors. The reality is many of the functions and vendors will not have direct access to such a shared storage system. Files will have to be moved either by network or by physical media (removable drives or tape). Given the size of the files and the costs of high speed networks, a self-describing, open format, LTFS LTO tape containing the needed files can meet this critical requirement. LTFS, available for the key platforms in media production – Linux, Mac and Windows – provides an excellent and economic portable interchange method.

Another reality in post production is the ongoing need to manage the shared storage systems. The data files for a film project can easily need 10's of terabytes of storage during post production. Multiple projects are active at any one time and it is not unusual for project priorities to change – delays, postponements and rush jobs. Again economics typically drive a tiered storage solution utilizing a combination of high performance disk for online data, moderate performing lower cost disk storage for nearline and tape for offline storage, portability and interchange.

Bandwidth "By the Box"

Tape functions remarkably well as a data transmission and exchange technology. As the saying goes, "It's hard to beat the bandwidth of a truck full of tapes." Though coined in relation to video tapes, the concept has been reborn with a vengeance with LTFS.

The availability of worldwide expedited shipping and flat rate boxes provide both high speed and very low cost options. The US postal service will deliver a flat rate box which can hold 28 LTO tapes (about 42 TB) anywhere in the US for $15 in about 3 days. Hypothetically, this represents a data transfer rate of 1.3 Gigabits per second or roughly the speed of a one Gigabit leased line at a fraction of the cost (incremental cost, not considering setup costs for either channel).

Expedited, the box can go from New York to Los Angeles in 15 hours for $200 (770 MB/s, 6.2 Gbps) or Paris to LA in 47 hours for $300 (250 MB/s, 2 Gbps). With sophisticated technology like larger boxes, bandwidth can be scaled even higher.

To secure data in transit, LTO provides a standardized hardware-based data encryption scheme. Custom source encryption can also be easily implemented as the LTFS formatted data are in separate files.

While tape has always been a natural choice for the offline tier, LTFS enables tape to be a strong candidate for the nearline and even, in select cases, the online tier. LTFS enables applications developed for the traditional nearline disk tier to access data on LTFS tape without being aware the data is actually on tape. Latency (but not bandwidth) is the key performance metric which determines suitability for a particular use case.

In one popular scenario files are stored inexpensively on LTFS tape and brought online "just in time" for sharing, editing and further enhancement on disk. Once this work is completed, the new or modified files are written back to tape, leaving expensive online storage available for newer active files. Each technology is efficiently utilized where its cost and benefits are most effective.

Tape remains the medium of choice, or even the only choice, for large storage applications. While cost per gigabyte for tape and disk continue to shrink, tape maintains the advantage over disk when the overall total cost of ownership is taken into account. LTO tape provides the high bandwidth and economical storage density required, and LTFS makes it easier to locate and access the files stored on tape.

The point at which tape is more cost effective than disk is not based just on hardware cost. Complexity and the associated cost of managing it also figures prominently in the decision-making process. LTFS reduces tape technology's "barrier to entry" and lowers the cost of tape-based solutions. Tape becomes more attractive when you also consider factors such as redundant disks and total-cost-of-computing (e.g. power and cooling).

LTFS for Distribution

LTFS formatted LTO tape provides the industry a common dependable exchange medium for film scans, video and other digital assets. Its open specification and cross platform interchangeability distinguish it from other options. Add to that the reliability and robust nature of LTO tape and LTFS becomes increasingly appropriate for media distribution.

While network-based solutions are used for many distribution needs, they require significant network bandwidth, technology and expense. By contrast a single LTFS formatted LTO-5 tape can hold over 60 hours of broadcast distribution quality content in a single self-describing tape cartridge. That cartridge can be any place in the world in less than 48 hours. Equally important, the receiver of that tape cartridge only needs a stand alone LTO-5 tape drive attached to a PC or Mac with LTFS to read and use the content.

With an LTFS tape cartridge now being so easy to use, encryption of tape contents may be important for some applications. Encryption can occur before files are written to tape or via standard LTO tape drive hardware-based encryption which is compatible with LTFS

LTFS for Digital Media Archiving

The need for highly predictable and reliable storage technology to support digital archiving is even more important today as growing amounts of valuable content is being captured, produced and distributed in digital formats.

Data tape has long been the technology of choice for archiving. The overriding metric for a successful archive is whether the data can be retrieved when needed. With the addition of LTFS, tape can now provide ease of locating and restoring files. LTFS minimizes dependencies, maximizes recoverability and facilitates the use of LTO tape for long term archives of high value data.

The archives of a media company contain the company's most valuable assets (the MGM, Universal, and Miramax film libraries being recent examples in the news). The archive masters of films, TV shows

and music have real economic value and losses if they are lost. Historically archives were analog film and recorded magnetic tape but increasingly the archive of a project will only exist as digital files. LTFS was designed to have, and achieves, superior recoverability versus other IT storage technologies.

What keeps Digital Archivists Awake at Night?

As the content ages and access becomes infrequent, that content is moved to more economical media, removable disk drives or more commonly tape managed by backup and archive storage management software.

The content migrates to this offline storage and eventually, after more years of inactivity, personnel changes and multiple system upgrades, to a state that could be termed “way offline”. The data allegedly exist, but Archivists are not really sure if it can be accessed. One worst case scenario is finding a storage medium of unknown origin and being asked to recover its contents; LTFS to the rescue.

If the medium was a tape, the Archivist would need to locate a compatible tape drive. For LTO this would be the same generation drive that wrote the tape or up to 2 generations later, a fairly broad span.

Once the data are accessible, the problem is one of data format. Was the tape written via a tar command? Is the tar header still intact? Or was the tape written by a backup and archive storage management application? If so, which one and where is the critical database which maps file metadata to blocks of data on tape? Is the entire file on the tape or only a portion? The content may be on the tape but not usable without this mapping. The tape has a dependency on systems and databases which may no longer exist and thus hinder, perhaps fatally, recovery of the data.

In contrast, LTFS excels in this scenario. Once the data on tape are accessible, a scan of the tape would reveal 2 partitions and the distinctive pattern of block sizes and filemarks. The Archivist could install a current version of LTFS software and read the tape contents regardless of the original source system, operating system or version. If for some reason no LTFS driver was available, the Archivist could copy the contents of the index to his computer, read the man readable XML index, and identify which blocks of data on the tape belonged to which file and the associated file metadata, e.g. name, create time, size, etc.

Disks are problematic as archive media. Kept spinning they require power and cooling. Spun down and kept in place, they use valuable rack space. Removed and put on shelves, they may be difficult to reconnect. One of the major limitations versus tape is that the IO (Input/Output) channel is combined with the storage. A 50 pin SCSI drive will always require a 50 pin SCSI HBA and appropriate drivers in the operating system. A better solution is tape storage which is independent of the IO channel. A Gen 3 LTO tape written by a SCSI attached drive can be read in a Gen 5 fiber attached drive. A Gen 5 tape written today via SAS will be readable by a Gen 7 drive attached via a protocol that has not been invented yet!

LTFS and LTO are the solution to a Digital Archivists restful night’s sleep.

Is LTFS the answer by itself?

Yes, it can be for small or straightforward uses. For many situations however, driven by the archive size and use cases, additional software will be needed. Different users will want different levels of management for their solution, from automated and policy-based managed HSM (tiered storage) to simple manual drag and drop.

In a single drive LTFS implementation the tape already contains an index viewable through Linux, Windows or Mac utilities and the user has drag-n-drop access. But that index only covers the contents of that particular tape; there is no centralized or library-level index that tracks contents across multiple tapes. For smaller scale implementations it is practical for a user to individually track the tape contents and receive the benefits of LTFS.

For larger storage environments, implementations of LTFS are now available from several LTO library manufacturers. These applications map the contents of a tape library or library partition as a set of directory trees, one for each LTFS formatted tape. **The LTFS tape format is unchanged and identical to the single tape implementation.** Index data are cached for rapid access and tapes are loaded as needed to access data. The effect is similar to a NAS.

While the various tape library implementations of LTFS may provide a viable option for some users that have larger storage requirements, current LTFS library support offerings are not complete solutions for users requiring more media aware management capabilities. Media aware backup and archive storage management software has historically been addressed by solutions such as FlashNet from SGL, Diva from Front Porch Digital, AREMA (formerly ADMIRA) from IBM, SAM-FS from Oracle, StorNext Storage Manager from Quantum and solutions from Masstech. These have been, and in many cases still are, the keepers of the data in digital media archives. They receive data to archive, store it on data tape (typically in proprietary formats) and retain an index of which blocks of data on which tape(s) belong to which file. To retrieve data, the application, index information and data tape all need to be accessible.

Backup and archive storage management and HSM software that implement LTFS can add important functionality beyond LTFS single drive or library solutions. They can make LTO tape more efficient and easier to use by enabling users to locate files across thousands of tapes and via queries on other metadata fields (e.g. timecode, name of the director, etc.). They also provide efficient tape drive and library management, storage tier data migration, storage policy management and interfaces to 3rd party systems. As of early June, 2012, several media aware backup and archive storage management software providers have delivered or announced LTFS support in their solutions including: Atempo, Cache-A, Masstech and SGL among others.

Media companies will gain content accessibility and operational flexibility with LTFS-enabled backup and archive storage management solutions. The storage management applications will **no longer be essential for unlocking data in the archive.** Any one of the LTFS formatted tapes could be pulled and its contents read at a workstation using LTFS. Additionally, should there be a catastrophic failure of an LTFS-based storage management application database; the contents can be rebuilt by reading the index partition of each of the managed tapes. Similarly, data stored on LTFS tapes can be moved between storage management products from different vendors – providing more options than would proprietary formats.

One of the technical considerations in converting a storage management solution to LTFS is that capacity must be managed by individual tape cartridge. LTFS was purposely designed to hold data stored as files so that each LTFS tape remained self describing and stand alone. In the current LTFS versions, there is no support for spanning a file across multiple tapes. Doing so would break the “standalone” paradigm. Spanning across removable media is not simply an LTFS tape consideration. The same consideration would apply to other removable media such as hard drives, DVDs, or flash memory sticks.

There are two primary reasons for spanning a file across multiple tapes: to handle extremely large files (currently larger than 1.425 TBs for LTFS-formatted LTO-5 tapes), or to fully utilize all the storage capacity of a cartridge by splitting a file across two tapes in order to completely fill a tape.

Traditional backup and archive storage management software stores data on tape as blocks so spanning is usually supported and software solutions supporting LTFS formatted tapes may provide spanning capabilities. For LTFS-formatted LTO-5 tapes with a capacity of 1.425 TBs the need to span extremely large files is rare. Even atypically large files of 100GB would still result in tapes 96% full. The instance of spanning extremely large files will decrease further with the availability of future generation, higher capacity LTO products.

How else is LTFS suited to media workflows?

Timecode and Keycode Usage

There are ways other than traditional IT storage that LTFS can play a significant and unique role in supporting media production and distribution workflows. LTFS provides a level of granularity and control consistent with timecode and keycode on video tape and film elements, respectively. Timecode and keycode have been the mechanism to track, search and organize video and film elements for transforming (editing) disparate elements into finished productions. They have also been key in the synchronization and distribution of such productions via play-out systems for broadcasting. Direct subfile access to video or film content at the frame or field level is also important for accurately and efficiently retrieving source material for editing. Until now, this has been the domain of expensive and complex disk systems or video-tape and film systems. The ability of LTFS to support this level of granular access greatly increases data tape's usefulness for applications in media production and distribution.

In general, traditional data tape storage systems did not allow for easy access to a single digital element (for instance, a specific frame of video or audio clip within a file). While there was support for "partial file recall," it was difficult to integrate into media applications and often required separate recall and processing steps.

LTFS Applications for other industry verticals

While early LTFS-based solutions have focused on media and entertainment applications, there are many other applications for which it is well-suited. Here are a few examples:

Medical Care – Medical applications such as imaging technology, X-ray, and ultrasound now feature increased resolution and in some instances include video. Regulations and best-practices call for a secure long-term storage solution for diagnostic, research and legal purposes.

Security - LTFS opens the door for new video surveillance applications. While individual surveillance video feeds are generally not high bandwidth, the sheer number of parallel channels (often numbering in the thousands for a single facility such as a casino or airport) produces a high aggregate bandwidth. LTFS facilitates access as a workstation can be used to review footage independent of the system that recorded it.

Biological Data – As partial and complete human genome mapping becomes less expensive and therefore practical, the enormous amount of data is extraordinary (3 billion base pairs per human!). Researchers leverage these massive data sets to examine correlations between genetic mappings and the specific diseases each is studying. The portability and density of LTFS are well suited as are its seamless support of LTO technology's lossless data compression.

Cloud Data – Data clouds represent the accumulation of potentially vast amounts of data. LTFS formatted tapes provide a very economical and easy to use method for transferring large amounts of data (even between disparate operating systems) to create, replicate, move or backup clouds. The same tapes used for a transfer can be retained to provide a local restore capability.

With LTFS, systems for editing, distributing or playing out digital content can use data tape to selectively and granularly store, search and retrieve elements down to the frame or field level. Applications that formerly had to implement custom storage manager API's can now access files via standard file system methods of opening files and seeking to offsets in the same manner used for disks.

During the writing process, applications can control the relative placement of data on tapes to facilitate efficient streaming recall. Backup and archive storage management applications generally provide a few placement options but not with the same level of granularity. LTFS allows files to be arranged and written to tape so that the recall of a timecode contiguous series of files (a series of contiguous DPX film frame files or timecode continuous video clip files) generates a high speed streaming recall from the tape. While this may occur using a non-LTFS storage manager, it's with far less certainty.

This granular access to digital elements that LTFS provides allows the 'near-line' (and conceivably some on-line) activities of creating and distributing media to take advantage of the cost and reliability benefits of data tape.

How does LTFS compare to the alternatives?

The combined benefits of LTO technology and LTFS: openness, robustness, capacity, bandwidth, and self describing distinguish it from other technologies. Following is a comparison to some alternatives:

Raw Tape

Writing blocks of data to tape via custom applications or OS commands leverages the features of LTO tape. However, raw is not self describing, open or documented. A recipient would require specific knowledge about how data were written and an indexing scheme to define which blocks belong to which file.

Tar

The tar command in Linux (short for Tape ARchive) is a common and well established method of archiving data to tape. It is open and portable like LTFS, which was intended in part to be "a better tar than tar." The serial and append-only nature of tape limited the options available to tar developers

LTFS circumvents these limitations by exploiting the partitioned tape architecture option introduced with the fifth generation of LTO tape. The index partition of an LTFS formatted tape enables LTFS metadata to be updated, stored redundantly, and accessed quickly without having to scan the entire tape. This is superior to tar files because each tar file has a single copy of the index of the data -- and only that tar file contains that particular index. There is no central tree structure for the tar tape, so tar indices are scattered throughout the entire tape. A tar file extract requires two tape access steps per relevant tar file, one for the index and then a second for the data. LTFS can, after a one time per tape index load, seek directly to any of the data.

Tar also has no "rollback" capability. LTFS tapes retain older copies of the index as well as the data (tape is append only). An LTFS tape can be returned to a previous state in the event of an inadvertent overwrite or deletion of a file.

Backup, Archive, and Storage Management Software Proprietary Tape Formats

Using a backup and archive storage management application to manage tapes in a tape library is by far the most common architecture for tape deployments. Their centralized databases store file metadata which can be searched and scale to track millions of files stored on thousands of tapes.
The main concern is their use of proprietary formats or methods to store the data on tape. This locks in users to a specific vendor and systems, introducing a dependency on the application and its database

for recovering data. This dependence can preclude backup and archive storage management applications from some very large archive solutions where long term retention is paramount and risk tolerance is extremely low.

Several backup and archive storage management application vendors have already adopted the LTFS format for storing the data on tape. As users indicate preferences for open storage formats more will follow suit. As one vendor representative put it, "when customers learn about this, we will have to support it."

Portable Disks

Portable disk drives typically connected via USB or FireWire are, like LTFS tapes, portable and self-describing and have the advantage of low latency. However they tend to be slower, more fragile and have a shorter expected shelf life.

RAID Disks

By writing data in parallel to multiple drives, RAID arrays have the high speed and low latency ideally suited to online data. However, they are difficult to move between locations and generally more expensive especially after power and cooling costs are considered.

LTFS Considerations

The limitations of LTFS are associated with the physical tape it resides on and are primarily related to latency. To access data on a tape, the tape must first be loaded into a drive, unspooled to the read heads and then moved to the correct location for reading to begin. If all drives are busy, or a previous tape needs to be rewound and ejected, access times can stretch to minutes until a drive becomes available. Once a tape is mounted average access times to beginning of file can be a reasonable 30 to 90 seconds.

Tape data transfer rates of up to 140 MB/s are quite high and superior to the practical bandwidth of many disk drives and gigabit Ethernet. Note that a high transfer rate is only achieved while the tape is able to stream. If read requests are not sequential, the tape drive has to stop streaming and seek to a new location on the tape and net bandwidth can drop precipitously. Intelligent writing and reading of LTFS tapes by applications will be critical for performance oriented solutions. Tape drive compression can increase the net data rate.

The LTFS software provides information about the location of files on tape, enabling storage managers and other applications to properly sequence read requests. During file writes, it is important that files

What to look for in an Application Supporting LTFS

When evaluating an application supporting LTFS, consider the following:

How do users access the files on the LTFS tapes?

- Are they shared out via a NAS-like appliance or are custom interfaces required? How easily can your existing software and workflow integrate to LTFS?

How well does it scale?

- What is the prudent upper limit on the number of files it can manage and still meet performance metrics?
- Does it mediate and manage multiple concurrent user requests?

How do users locate their files?

- Can they search on metadata fields or just by filename?
- Can they select individual files? Contiguous groups? Arbitrary sets?
- Can they specify sets of files via patterns or lists?

How efficient are the data transfers?

- If there are dozens or thousands of files being recalled, can it maximize tape streaming? Does it use the location of the files on tape to sequence reads?
- What is the net bandwidth for various recall scenarios?
- How easily can inadvertently fiendish users create a "perfect data storm" scenario?
- Can it write data to optimize read patterns?

are written contiguously. If multiple users or applications are writing to the same LTFS tape at the same time, the blocks of their files will interleave on the tape precluding the capability to do streaming reads. Here again LTFS-savvy backup and archive storage management applications can manage "naive" user requests to optimize performance.

Deleted or "overwritten" files remain on the tape since new data is always appended. On the positive side, this enables older versions of files to be recovered in case of inadvertent deletion; however it also enables older files to be recovered despite intentional deletion. Deleted files can be truly deleted and their space can be reclaimed if remaining valid tape contents are copied to another tape and the entire tape is reformatted. These are general tape processes and not specific to LTFS.

What is the status of LTFS in the Media Marketplace?

Awards and Recognition

Since its introduction in 2010 LTFS and related products have been recognized with various awards.

- **2012 Storage Magazine "Hot Storage Technology for 2012"**
- **2011 Primetime Technical Emmy Award**
- **2011 Hollywood Post Alliance Engineering Award**
- **2011 National Association of Broadcasters Broadcast Engineering Pick Hit**
- **2010 Broadcast Engineering Pick Hit – For-A LTR100HS Archive Recorder**

Basic Applications

Users have an ever-increasing number of options for applying the benefits of LTFS to their workflows. For straightforward workflows, users can construct basic applications using the LTFS specification and file systems implementations provided by the LTO program partner companies.

LTFS DIY (do it yourself)

It is possible to write LTFS-formatted tapes via custom software instead of an LTFS driver. However it is complex and error prone. The use of an independent LTFS format verifier is highly recommended.

LTFS Stand Alone Drive Software

This is the starting point for most users and, for some, all they will require. Four LTO tape drive provider companies (HP, IBM, Quantum and Tandberg Data) offer LTFS software drivers. Vendors may provide free of charge downloads of the software drivers that may support Linux, Mac and Windows environments. Additional tape technology and system vendors are now providing LTFS support as well.

LTFS software drivers enable single LTO-5 drives to be accessed via the operating system. LTFS interfaces through the standard file system interfaces on the platform on which it runs (POSIX in Linux and Mac, Windows file system interfaces in Windows). As such, it can be used with any command-line tools and scripting environments that work with the native file systems. Multiple drives connected to a server can be accessed in parallel. While use of a stand alone version does not support tape libraries or their automation, individual drives in a tape library can be accessed via LTFS software drivers. There

needs to be some mechanism to move tapes to and from the drives which could be manual, via tape library web interfaces, scripts or open source library controllers.

LTFS Library System Support

Several LTO tape library providers have LTFS support for their LTO-5 library systems, including HP's StoreOpen Automation application and IBM's LTFS Library Edition. These offerings combine the tape content access of the single drive LTFS versions with control of the tape library. All LTFS formatted cartridges in a tape library are visible and accessible. As a tape cartridge is accessed, it is loaded into a tape drive and then unloaded after access has completed if the drive is needed for a different tape. Users must coordinate access to optimize tape load and unload cycles.

LTFS Products in Media and Entertainment

A growing number of application providers have adopted LTFS as a supported technology. Traditionally data tape would only be appropriate for archive use cases, but LTFS gives LTO tape advantages for broader use. This section provides a list of announced or shipping products and services using LTFS. The list is as of early June, 2012, and due to the dynamic marketplace is necessarily incomplete. The applications and services are listed alphabetically within categories.

Company	Product
On-Set Applications and Appliances	
1 Beyond, Inc. www.1beyond.com	Wrangler LTO-5 NetDrive
	Wrangler FlexVTR™ Archive Recorder
	Wrangler LTO-5 Offload Station
FOR-A www.for-a.com	LTR-120HS (LTO-5 Video Archiving Recorder, AVC-Intra/DVCPRO)
	LTR-100HS (LTO-5 Video Archiving Recorder, MPEG-2)
Passive Archive, NAS-like Solutions	
Crossroads Systems www.crossroads.com	StrongBox (Network-attached storage (NAS) solutions providing a file-based data archive combining LTFS tape with disk.)
Active Archive – Backup and Archive Storage Management Applications/Appliances	
Atempo www.atempo.com	Atempo Digital Archive
Cache-A www.cache-a.com	Pro-Cache5 (Multi-user, multi-function archive appliance)
	Power-Cache (Multi-user, multi-function rack-mount archive server)
	Prime-Cache5 (Multi-user, multi-function quiet desktop archive appliance)
DigitalMedialogix www.digitalmedialogix.com	LTFS for Media Consolidation
	LTFS for Hybrid Media Backup
IBM www.ibm.com	AREMA (formerly known as ADMIRA) (Enterprise archive and workflow management for LTFS-based archives)
	LTFS Storage Manager (Provides storage lifecycle management of multimedia files using IBM LTFS)
Masstech www.masstech.com	Topaz (Media asset management solution)

Panasonic www.panasonic-broadcast.eu/en/	AJ-SF100G LTO Archive Software (P2 media archive software)
SGL www.sglbroadcast.com	Flashnet (Content archive and storage management solution for the broadcast industry.)
StorageDNA www.storagedna.com	DNA Evolution (Data management solution for the file-based media workflow.)
Tolis www.tolisgroup.com	BRU Producer's Edition (Manages creative artist session archives.)
Workflow Use Cases	
YoYotta www.yoyotta.com	YoYo Non-linear Data Management
Service Providers implementing with LTFS	
Arkivum www.arkivum.com	Arkivum Assured Archiving (Providing a media-oriented archive cloud service.)
T3Media (Thought Equity) www.t3media.com	Cloud-based storage, access and licensing services for master-quality video.

What Might the Future Hold for LTFS?

- The positive impact provided by the 2011 Technical Emmy Award will accelerate the wider adoption of LTFS as users appreciate the concepts, want the benefits and deploy the technology.
- LTFS will continue to migrate to other tape technologies as they support partitioning.
- New LTFS enabled applications will emerge and existing backup and archive storage management applications will be enhanced with LTFS support and LTFS unique features.
- It is expected that the LTFS open standard will be submitted for adoption by a vendor independent standards body(s).
- Studios, networks and other content owners will begin to specify the use of LTFS. Post production, effect houses and archives will comply, creating a tipping point for broad LTFS acceptance.
- LTFS will become the major format for media data interchange and long term digital media archiving.

References

2011 Technical Emmy Award

http://www.emmys.com/articles/video-highlights-engineering-emmys-host-mayim-bialik

http://www.prnewswire.com/news-releases/ibm-and-fox-networks-group-win-engineering-emmy-award-for-broadcast-technology-innovation-132332528.html

Additional Information

"Linear Tape File System (LTFS) Format Specification"
http://www.trustlto.com/LTFS_Format_To%20Print.pdf

IBM LTFS Single Drive Edition Documentation
http://publib.boulder.ibm.com/infocenter/ltfs/cust/index.jsp

"The Linear Tape File System", David Pease et al.
http://storageconference.org/2010/Papers/MSST/Pease.pdf

"Long Term Data Retention – Key Considerations", Quantum Technology Brief
https://www.mercurymagazines.com/pdf/QUANTUM1.pdf

"In Search of the Long-Term Archiving Solution – Tape Delivers Significant TCO Advantage over Disk", Clipper Group
http://www.clipper.com/research/TCG2010054.pdf

"File based media workflows using LTFS tapes", Arnon Amir et al.
Association for Computing Machinery (ACM)
Proceedings of the International Conference on MultiMedia 2010
http://dl.acm.org/citation.cfm?doid=1873951.1874269

"The IT HD VTR: PC Server Based SD/HD Recording and Playback", R. Richter
IBM Whitepaper
ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/mew00212usen/MEW00212USEN.PDF

Wikipedia entry
http://en.wikipedia.org/wiki/Linear_Tape_File_System

About the Author

Rainer Richter is a principal of Media Technology Market Partners LLC leading the company's software development and integration services. Previously he was a member of the IBM team that successfully deployed a first of a kind digital media archive on LTO-1 technology way back in 2001. Since then he has had leadership roles in additional LTO archives at major networks and studios, including a project at Fox which led to a Technical Emmy Award in 2011. Rainer received a BSEE in Electrical Engineering, summa cum laude, from Rutgers University and a MS in Computer Engineering from Santa Clara University.

Acknowledgements

The author gratefully acknowledges the contributions of various members of the LTFS community and especially George Anderson and David L. Trumbo, both of Media Technology Market Partners LLC.

About Media Technology Market Partners LLC

Media Technology Market Partners LLC (MTMP) provides system design, integrated solutions, consulting services, hardware and software to the media and entertainment industry. A core focus of the firm is LTFS (Linear Tape File System) enabled storage solutions. MTMP was an early advocate and is a continuing evangelist for the use of LTFS in storage solutions not only for media and entertainment but all digital media content. MTMP was founded in 2009 by George Anderson and David L. Trumbo. The firm is headquartered in Los Angeles, California.

Appendix

Storage Tiers

Tier	Latency	Bandwidth	Cost / GB	Typical Technology
Online	Very low (milliseconds)	Very High	High	SAN, RAID
Nearline	Low (seconds)	High	Medium	JBOD, NAS
Offline	High (minutes)	High	Low	Tape, Portable Disk

While tape has always been a natural choice for the offline tier, LTFS enables tape to be a strong candidate for nearline and even, in select cases, online tiers. LTFS enables applications developed for the traditional nearline disk tier to access data on LTFS tape without being aware the data is actually on tape. Latency (but not bandwidth) is the key performance metric which determines suitability for a particular use case.

LTFS Success Stories:

DigitalFilm Tree

DigitalFilm Tree is a post production, consulting, creative and software development company specializing in the design of post-production workflows for the media and entertainment industry. Based in Hollywood, California, Digital Film Tree has built an industry-wide reputation for quality, body of work, innovative thinking and the creative application of cutting edge technologies. Clients include television and motion picture studios, independent film makers and IT organizations.

Business Need

Implement an LTO-5 with LTFS based backup and archiving solution to improve file based production, post production, distribution and digital media archiving workflows

Results

IT improvements

- Provides a reliable, simple-to-use backup and archiving capability, creating a viable additional option for file based workflows and improving the usability of standard LTO technology
- Offers high-capacity storage media with fast transfer rates, accelerating digital workflows
- Delivers a self-describing universal format with no application dependencies, simplifying the backup and archiving process
- Works seamlessly on multiple platforms within the existing storage infrastructure, supporting post-production workflows

Business benefits

- Lowers tape costs by $120,000 per season and cuts hardware costs from $100,000 to $4,000
- Reduces daily overnight processing times by over 60 per cent, enhancing productivity
- Supports both immediate on-set and near-set processing, satisfying the industry's demanding deadlines

- Supports the transition to an alternative universal format for media data interchange and long term digital media archiving, lowering costs and boosting revenues further
- Minimizes software dependencies, maximizes recoverability and facilitates the use of LTO tape for preserving studios' most valuable digital content assets

FotoKem Industries

FotoKem is a full-service post production facility headquartered in Burbank, California. FotoKem offers a broad spectrum of services, including file-based workflows, 3D digital intermediates, digital cinema packages, mobile dailies, global data delivery, film and video finishing, audio mixing and mastering, visual effects, restoration, and production rentals. The engineering staff at FotoKem continually looks for new technologies to help them solve their customers' problems with innovative and cost effective solutions

Business Need

One of FotoKem's customers wanted to offload the content and then reuse their XDCAM discs. XDCAM discs in cameras are used extensively to capture the action in reality TV and in-the-field news gathering. The content on these discs are the source materials for the shows and have potential long term value so the content needed to be safely archived.

Requirements

FotoKem was searching for a solution that provided a reliable long term archive for XDCAM content, with better usability and robustness than tar files. They also desired a solution that was easy to use, built on reliable technology, and low cost.

Solution

FotoKem elected to use an LTO-5 tape library and LTFS managed by shell scripts. They saw the benefits of open standard storage with cross O/S platform compatibility using self describing LTO / LTFS tapes. This allowed them to easily determine tape contents with a simple browser application and provided a long-term reliable archive that was easy to use and cost effective.

Benefits

The FotoKem solution was able to store up to 100 XDCAM disc images on each LTO-5 tape, delivering a 25X+ reduction in storage media costs. It also saved considerable space and allowed for the reuse of the XDCAM discs, reducing the need and cost of purchasing new ones. Additionally, it allowed FotoKem to make two LTO tape copies inexpensively (one of which was kept offsite). The production solution developed from scratch in less than one month.

Fuji TV

Fuji Television is Japan's leading TV broadcaster. Formed in 1959, it is based in Tokyo but operates through 28 regional affiliates. It also broadcasts Japanese entertainment, news and sports content in Europe and North America.

Business Need

Needed to improve efficiency, ease of use and performance of video content archiving process

Approach

Continue with tape archiving upgrading from LTO-2 to LTO-5 tape drives with LTFS

Results

IT improvements

- Computer Graphics (CG) designers can archive files to tape through simple drag and drop
- CG files on tape appear as a disk directory listing, providing faster access to data
- Improved tape capacity by 650% from 200GB to 1.5 TB and performance by 247% from 144 GB/hour to 500 GB/hour
- CG Designers can directly access archived work

Business Benefits

- More efficient work flow with designers now able to perform archive ore operations themselves, freeing up system engineers for other projects
- Significantly reduced time taken to complete archiving, improving business productivity
- High capacity tapes enable the storage of multiple projects, reducing cost by using fewer tapes

About Media Technology Market Partners LLC

Media Technology Market Partners LLC (MTMP) provides system design, integrated solutions, consulting services, hardware and software to the media and entertainment industry. A core focus of the firm is LTFS (Linear Tape File System) enabled storage solutions. MTMP was an early advocate and is a continuing evangelist for the use of LTFS in storage solutions not only for media and entertainment but all digital media content. MTMP was founded in 2009 by George Anderson and David L. Trumbo. The firm is headquartered in Los Angeles, California. For more information, visit www.mediatechmarketpartners.com.

Company Headquarters
12314 Moorpark Street, Suite 208
Studio City, California 91604 USA
T: +818.763.6478 (U.S. & Canada)
+1.818.763.6478
E: contactmtmp@mediatechmarketpartners.com

MTMP Media Technology Market Partners