LTFS

Fueling the Tape Transformation

HORISON
Information Strategies

BY: Fred Moore President
www.horison.com

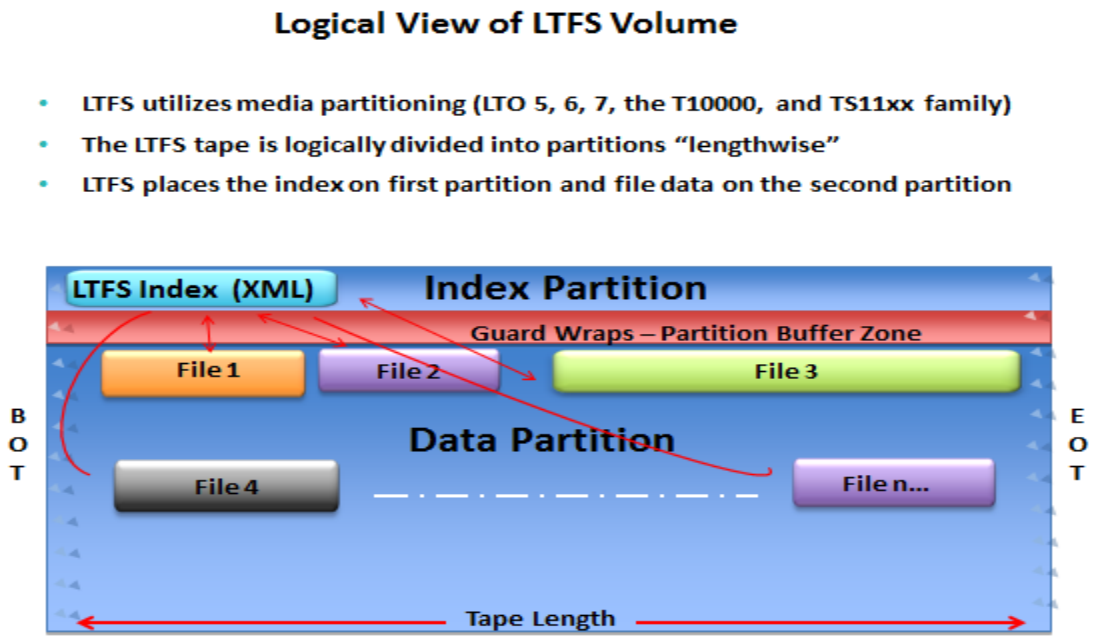
Abstract Since its inception, tape has been a sequential access storage device. Though this proved beneficial for backup and many "data streaming applications", the ability to randomly navigate a tape cartridge to locate a specific file or object was physically impossible. The LTFS™ (Linear Tape File System) Single Drive Edition (LTFS SDE) format was developed by IBM to address tape archive requirements and was first announced at the National Association of Broadcasters (NAB) annual trade show on April 12, 2010 for the LTO-5 standard tape format. The LTFS format specification was adopted by the LTO Consortium (IBM, HP, and Quantum) and is fully supported by the LTO roadmap which defines the specifications for the entire future LTO tape drive family. LTFS provides, for the first time, tape partitioning functionality for LTO drives, the latest enterprise TS11xx and T10000 drives from IBM and Oracle StorageTek, and tape libraries from several storage suppliers, further validating and extending its value.

LTFS works in conjunction with LTO and enterprise tape technologies for ease of use and portability for open systems tape storage. With LTFS, one partition holds the content and the other(s) holds the content's index, so the tape can be self-describing to improve archive management and retrieval. With the operating system's graphical file manager and directory tree, and file folders utilizing data on an LTO Ultrium tape cartridge is as easy as dragging and dropping the file. LTFS enables direct, intuitive and graphical access to data stored in the LTFS format on standard tape cartridges for reading, writing and exchanging descriptive metadata.

Key point: ***The tape transformation is well underway and LTFS is a big reason why.***

How Does LTFS Work?

IBM continues to be the principal developer of LTFS and offers LTFS as open source software in a variety of editions for single drive and robotic tape libraries. IBM officially renamed their LTFS solution as IBM Spectrum Archive™ on May 5, 2015. Other tape suppliers offer their own editions of LTFS that are compatible with IBM's version. LTFS is the first open file system that works in conjunction with tape technology and clearly sets a new standard for ease of use and portability for LTO open systems tape. With LTFS, accessing data stored on an LTO tape cartridge is as easy to use and intuitive as accessing files on a USB flash drive or any removable hard drive.



LTFS creates a self-describing, dual-partitioned tape cartridge, meaning that a specific application is no longer needed to determine what's on the tape because the metadata (the descriptive information of what is on the tape) in the Index Partition describes the files, data contents and pointers to their physical tape location in the Data Partition. The LTFS metadata in the Index Partition enables faster searching and accessing the files in the second partition via a GUI (Graphical User Interface) provided by the LTFS operating system. The two LTFS partitions are shown in the chart above. Note: LTO-6 extended the specification to allow 4 separate partitions.

LTO tape is physically laid out with four wide data bands sandwiched between five narrow servo bands. The tape head assembly, that reads from and writes to the tape, straddles a single data band and the two adjacent servo bands for alignment. The tape head can have 8, 16, or 32 data read/write head elements and 2 servo read elements. The set of 8, 16, or 32 tracks are read or written in a single, one-way, end-to-end pass that is called a "wrap". The tape head shifts laterally to access the different wraps within each band and to access the other bands.

Physically, the Index Partition is organized across the length of the LTO tape and requires about 5% of the total cartridge capacity. The Index Partition is created to use one whole wrap on the media. The LTFS partitioning scheme uses an additional wrap as a Guard Wrap between the Index Partition and the Data Partition. The Guard Wrap serves as a physical buffer zone to ensure that writes to one partition cannot interfere with data on the other partition.

When a tape with an LTFS index is inserted, the LTFS will load the metadata for all the active files on the tape cartridge into server memory for faster future access. LTFS enables creating metadata or "tagging" of individual files providing faster access making it possible to double-click, open, drag, drop, copy and paste files to tape in the same familiar way as performed on disk. While it is *not* a replacement for disk, LTFS represents a major breakthrough in that it enables many new applications for tape. LTFS currently supports the Linux, Apple OS X (Mac), and Windows operating systems.

LTFS Functionality – Points to Remember	
LTFS Drive Ann. LTFS Library Ann. LTFS Enterprise Ann. LTFS Open Systems	April 12, 2010 for LTO-5 drives (LTFS Single Drive Edition LTFS SDE) July 12, 2011 (LTFS Library Edition LTFS LE) June 10, 2013 (LTFS Enterprise Edition LTFS EE) Sept. 19, 2014 (LTFS Library Edition for Windows using third party libraries with LTO-5 and subsequent LTO drives
LTFS Software Cost	The single tape drive version of the LTFS software is open source and carries no charge. Check with the specific tape library vendor for LTFS pricing details.
Operating Systems	Linux, Apple OS X (Mac), Microsoft Windows
Tape Drive Support	Beginning with LTO-5 and subsequent LTO versions, IBM TS1140, TS1150, TS1155, Oracle T10000C and T10000D.
Tape Mgmt. SW	No additional tape management, backup software or utilities are required.
Where to get LTFS	Download LTFS software from the vendor's website.
LTFS Support for Prior and Future LTO Generations	LTFS tape partitioning was introduced with the LTO format beginning with LTO-5, and therefore, earlier generations such as LTO-3 and LTO-4 tape drives do not support LTFS. Note: The LTO roadmap indicates that LTFS tape partitioning will be supported in all future LTO drives.
Cartridge Contents Inquiry for Fast Access and Retrieval	After the tape library is initially inventoried in memory, tape cartridges do not have to be remounted to retrieve its content index data and basic cartridge information like volume name, date, serial number and pointers saving time.
Partition Update Capability	Each partition (Index and Data) can be accessed and updated independently. The ability to access small sections of data on tape is significantly improved.
Protection for Index Partition	With LTFS Version 2, the index partition is periodically copied to the data partition on the tape for backup in case the primary copy is unavailable.
Data Sharing	Files stored on LTFS can be shared between multiple operating systems, different LTFS editions, and between different applications.

Source: Horison, INC.

Key point: ***LTFS is not a disk replacement but it changes the rules for tape access and enables much faster tape access than any previous tape format.***

LTFS Support Extends to Tape Libraries (LTFS LE)

LTFS SDE was initially announced to improve the management of individual tape drives. The next logical step was to support tape libraries. On July 12, 2011 IBM announced LTFS Library Edition (LTFS LE) V2.1 extending the capability of the IBM LTFS to automated tape libraries including the IBM TS2900, TS3310, TS3100, TS3200 and the large-scale TS3500 and TS4500 tape library systems. LTFS LE enhances access to large capacity tape archival systems by improving search and retrieval times to entire libraries of LTO tape.

For a LTFS capable tape library, a virtual index of the data on *all* cartridges that are presently in or have been in the library is maintained in server memory allowing users to search for files without physically moving or mounting tapes thereby improving access time. The server running the LTFS software graphically displays each cartridge as a subfolder of the library's contents using a GUI format. The tape contents represented by the Index Partition are available when a tape is loaded into the drive, and can be viewed by a browser or any application that has the tape attached to it. When a new tape cartridge is mounted, or the data on an existing physical tape is updated or changed, the LTFS index in memory is updated with that cartridge's information first, then asynchronously that index is backed up and written to the tape cartridge for higher availability. LTFS will recognize when tape leaves and reenters the library and performs consistency checks to determine if tape index has changed.

Since LTFS Library Edition is based on the LTFS format specification, it allows library cartridges to be 100% interchangeable with cartridges written by single drive versions of LTFS. By using LTFS to combine tier 2 (disk) and tier 3 (tape), organizations can achieve significant operational efficiency and cost effectiveness compared to the higher costs of equivalent all disk-based environments.

Key point: ***By extending LTFS support to libraries, LTFS LE is designed to increase efficiency and reduce costs by making tape a fully supported level for automated tiered storage.***

LTFS Enterprise Edition (LTFS EE) Support Extended to Enterprise Class Tape Drives

LTFS EE enables the use of LTFS for the policy management of tape as a storage tier in IBM's General Parallel File System (IBM GPFS™ was renamed as IBM Spectrum Scale on February 17, 2015) based environment and enables the use of tape as a critical tier in the storage environment hierarchy. LTFS EE can run any application that is designed for disk files on tape. LTFS EE supports LTO 5, 6, and 7 tape drives in IBM TS3310, TS3500, and TS4500 tape libraries and the IBM TS1140, TS1150 and TS1155 tape drives are supported in TS3500 and TS4500 tape libraries. LTFS EE extends LTFS to the enterprise drive market in addition to the open systems market using LTO drives. The only other enterprise tape drive manufacturer, Oracle StorageTek, offers LTFS LE support for their T10000C (5 TB native capacity) and the T10000D (8.5 TB native capacity) enterprise tape drives along with support for their entry level SL150 library, and enterprise level SL3000 and SL8500 libraries.

Spectrum Scale (GPFS) Support Key for HPC Systems

LTFS EE integrates with IBM Spectrum Scale (GPFS) delivering a scalable, file-based storage solution for managing very large amounts of data for the HPC (High Performance Computing) market. GPFS was developed by IBM as a high-performance clustered file system providing concurrent high-speed file access to applications executing on multiple nodes of clusters and is used by many of the world's largest supercomputer data centers. GPFS data management and integrated information lifecycle tools can manage petabytes of data and billions of files on disk and tape. LTFS is designed to enable policy management of data throughout its lifecycle and views tape as a full storage tier when used in a GPFS computing environment.

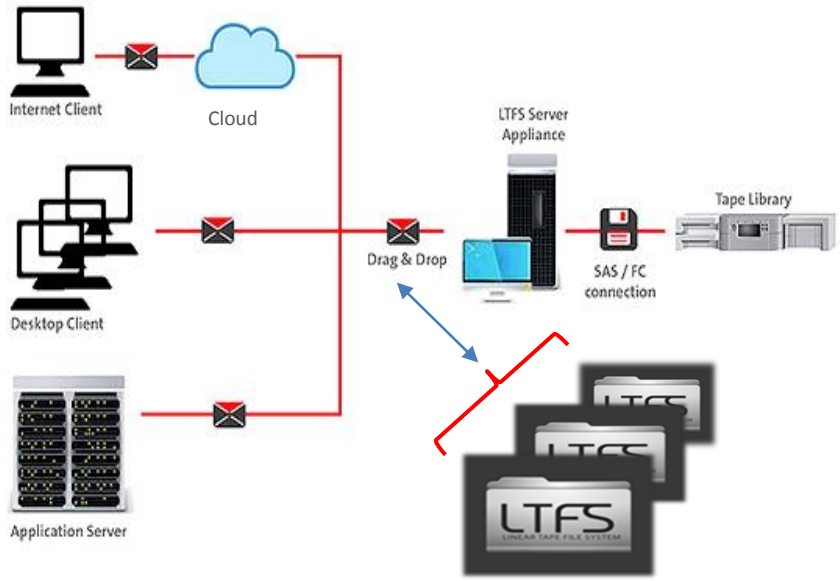
LTFS implementations can be as simple as a single tape drive, small to large tape libraries, or as part of an integrated tiered storage system with flash and hard disk. In this environment, data can move seamlessly between tiered layers based on data groups and policies.

Key point: ***Rather than storing static, unchanging files on costly disk storage, LTFS EE is designed to enable policy management of tape as a critical storage tier in the enterprise, Open Systems, HPC, Supercomputer and GPFS environments.***

LTFS Highlight Summary

- A dual-partition file system first implemented on LTO-5 tape drives in 2010
- LTO tape drive support (LTO-5 and subsequent LTO versions)
- Enterprise support (IBM TS1140, TS1150, TS1155, Oracle StorageTek T10000C, T10000D)
- Operating system support (Drivers are available for Apple McIntosh OS X, Windows 7.0, Windows Server 2008 R2, and Linux). *Note: The Mainframe operating system (System Z) does not currently support LTFS*
- LTFS makes tape file access much like a hard disk, CD/DVD-R disc or USB memory:
  - Files and directories show up on desktop as a directory tree listing
  - Enables tagging of files with any text, allowing for more intuitive searches of cartridge and library content
  - Uses familiar open, drag, drop, copy and paste to move files to and from tape
  - Can now run many applications on tape that were written to using disk files
  - Supports faster data interchange between users and systems
  - Functions such as File Open, Write, Read, Append, Delete and Close from an application are supported on LTFS enabled tape
- LTFS versions are available as open source from HPE, IBM, Oracle StorageTek and Quantum and can be downloaded from the following vendor URLs:
  For HPE StoreOpen and Linear Tape File System (LTFS) Software
  For IBM Spectrum Archive
  For Oracle/STK LTFS
  For Quantum LTFS

Logical View of LTFS Implementation



The LTFS server appliance provides LTFS client services for Apple McIntosh OS X, Windows 7.0, Windows Server 2008 R2, and Linux systems.

Internet, desktop and server-attached clients can use familiar file folder, double-click, open, drag, drop, copy and paste functions to more quickly access tape in the familiar way that files are dragged and dropped to disk.

Source: https://www.google.com/search?q=storage+newsletter+ltfs+image

***Key point:* With drive(s) and library support now available, the LTFS format has become the standard file system for linear tape.**

LTFS Addresses Archive Data, New Storage Intensive Applications, and the Cloud

How much is data growing? The enormous wave of Big Data applications and the IoT is fully underway. According to IDC's Digital Universe Study, digital data should reach 44 zettabytes ($44x10^{21}$) by 2020 and forecasts that by 2025, the global data-sphere will grow to 163 zettabytes (one trillion gigabytes). The IoT is expected to account for nearly 45 zettabytes by 2025. Digital archives consisting of unstructured data, 3-D images, multi-media, video, social network content, surveillance and compliance data have become the fastest growing data category experiencing a CAGR (Compounded Annual Growth Rate) of >50% annually. The video surveillance market is growing by leaps and bounds and it's estimated that 127 million video surveillance cameras will be sold in 2017 generating steady streams of data. Keep in mind that much of this data will never be stored or will be retained for only a brief time. Data stored for longer retention, furthermore, is frequently compressed.

Many businesses are choosing to build their archive and long-term retention strategy by combining disk and tape creating the Active Archive. The Active Archive is a newer archive solution using a combination of NAS, or just disk arrays, and tape and are most often combined with LTFS. Gaining marketplace momentum, Active Archive solutions are sometimes referred to as "Tape NAS" and are available from several storage suppliers adding a higher level of performance for more active files stored in the archive.

Cloud applications, where cloud data transfer is handled using tape, has significant potential with LTFS. Tape cartridges are commonly used for long-term archival data storage, but tapes utilizing today's LTFS can replace more expensive disk in the cloud for archival storage in many cases as storing archival

data is becoming a major cloud service. The recent collaboration between two SNIA (Storage Network Association) technical working groups - LTFS and Cloud - is working to use LTO Technology with LTFS to address cloud data transfer issues and build a standard for LTFS Bulk Data Transfer. This work intends to provide a set of instructions to describe the files to be transferred, policies for insertion or replacements, and a transfer report. Cloud providers will be able make use of this standard to develop interoperable services for the transfer of very large files directly to tape using the LTFS format making "tape in the cloud" a standard cloud offering. LTFS and tape provide a huge assist to cloud storage and archive providers.

Open Source Means New Use Cases

The decision to make LTFS available as open source at no cost, instead of waiting for software companies to add their support, has been quite effective. Historically, tape usage has centered on backup, restore, disaster recovery, archiving and long-term data preservation applications. Since LTFS is open source, new use cases are arriving including:

- The media and entertainment industry for storing, editing and accessing video, movie, TV, broadcast
- Improved access capabilities for surveillance, scientific applications, all archives, compliance
- The Wave of Big Data analytics and IoT applications
- Active archiving integrating disk (NAS) and tape using LTFS as Tape NAS
- Cloud providers using "Tape in the Cloud" lowering costs for long-term archival storage
- The ability to take an LTFS tape and move it from various OS platforms guarantees data interchange when sending data to and from clients.

Key point: ***LTFS support is gaining momentum, and today there are 35 companies which are LTFS implementers. Expect the number of ISVs and cloud providers supporting LTFS in the future to increase.***

New LTFS-enabled Active Archive Appliances and Products Gain Traction

LTFS is expanding its functionality into software-defined appliances coupled with data management software modules. An Active Archive normally integrates network attached storage (NAS) appliances with LTFS and LTO tape cartridges (Tape as NAS) to deliver the Active Archive functionality. An Active Archive makes disk and tape storage systems complimentary and can be viewed as tiered storage enabling access to data across a virtualized file system migrating data between SSDs, disk drives, tape, and cloud. A tape library as NAS enables users to leverage familiar file system tools, and even drag and drop files directly to and from a tape cartridge, just like a disk-based NAS. Examples include:

- StrongBox Data Solutions
- Fujifilm Dternity
- HPE StoreEver Archive Manager Solutions
- IBM Spectrum Archive
- Oracle HSM
- Overland NEO Agility LTFS Archive Appliance
- Qstar Archive Manager
- Quantum Scalar LTFS and Artico
- Spectra Logic BlackPearl

To further address the enormous archive challenges, the Active Archive Alliance was launched on April 27, 2010 as a collaborative industry association formed to educate end user organizations on the new technologies that enable the most effective access to their archived data.

Key point: ***Solutions combining LTFS with disk, tape and SSDs are emerging providing faster online access, search capability and easier retrieval of long-term archival data.***

The Tape Technology Renaissance is Happening

In addition to LTFS, the pace of tape technology innovation continues to be aggressive and the results are impressive. In April 2015, Fujifilm and IBM scientists jointly demonstrated an areal recording density of 123 billion bits of uncompressed data per square inch on low cost magnetic tape. This breakthrough represents the equivalent of a 220-terabyte BaFe tape cartridge that could fit in the palm of your hand marking significant progress toward achieving tape areal densities of 100 billion bits per square inch and beyond and ensuring a long and promising future for tape technologies. Native tape cartridge capacity now exceeds the capacity of the largest disk drive. Remember that tape scales capacity by adding media, disk drives scale capacity by adding more drives. Enterprise tape drive data rates now surpass disk drive data rates and the media life for modern tape media now exceeds 30 years. The storage industry de-facto standard reliability measure of tape, BER (Bit Error Rate), has surpassed that of Fibre Channel disk by three orders of magnitude making both enterprise and LTO tape more reliable than the most reliable disk drive. These enhancements and several others have given tape significant TCO advantages over disk and recent studies indicate a 15-1 or greater advantage for tape over disk for storing archive data is common.

Key point: ***For the first time ever, tape is cheaper, more reliable, faster, has a higher capacity and much longer media life than any of its disk counterparts.***

Conclusion

Throughout its history, tape has steadily evolved providing faster data rates and higher capacities, but the enhancements were always built on a sequential access architecture limiting access time and constraining applications. LTFS is a significant step in moving tape storage away from its reputation as complex and difficult to use. With LTFS, tape's historical model has been disrupted by providing significant access time improvement while opening the door for a new, universal open file system for tape. Therefore, as future tape cartridge capacities significantly increase, the need for improved access and data retrieval capabilities using LTFS and new advanced partitioning capabilities will also increase. Because of this and significant progress on many additional fronts in the past 10 years, the tape industry is effectively re-positioning itself to address the many new high capacity, long-term storage and archive opportunities and promises to provide the most cost-effective storage solution for the foreseeable future.

Key point: ***Although LTFS has barely scratched the surface of its potential, it can become one of the most significant developments in the tape industry.***

End of report