HORISON
Information Strategies

BY: Fred Moore President

LTFS

Fueling the Tape Transformation

**2015
Tape Technology
Update Series**

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. With arrival of LTFS (Linear Tape File System), the long standing rules of tape access have finally changed. The LTFS Single Drive Edition (LTFS SDE) format was developed by IBM to address tape archive requirements and 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 supported by the LTO roadmap which defines the specifications for the LTO tape drive family. LTFS provides, for the first time, tape partitioning functionality for LTO-5, LTO-6, the latest enterprise TS11xx and T10000 tape drives from IBM and Oracle StorageTek, and tape libraries from several storage suppliers, further validating and extending its value. With this long overdue tape access capability now available, LTFS enables businesses to capitalize on the sizeable economic advantages of modern tape in many ways that weren't practical before.

*Key point: Unless you've somehow missed it, tape is officially "trending" and LTFS is a big reason why.*

How Does LTFS Work?

IBM is the principal developer of LTFS and offers LTFS as open source software in a variety of editions for single drive and robotic tape libraries. Other tape suppliers offer their own editions of LTFS that are compatible with IBM. LTFS is the first 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 this file system, accessing data stored on an LTO tape cartridge is as easy to use and intuitive as accessing files on a USB flash drive. LTFS essentially transforms LTO technology making it as open and easy to use as 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) in the Index Partition describes the files, data contents and pointers to their location in the Data Partition. The LTFS metadata 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 described in the chart below.

Logical View of LTFS Volume

- LTFS utilizes media partitioning (LTO5, LTO6, T10000, TS1140 and TS1150)
- The LTFS tape is logically divided into partitions "lengthwise"
- LTFS places the index on first partition and file data on the second partition
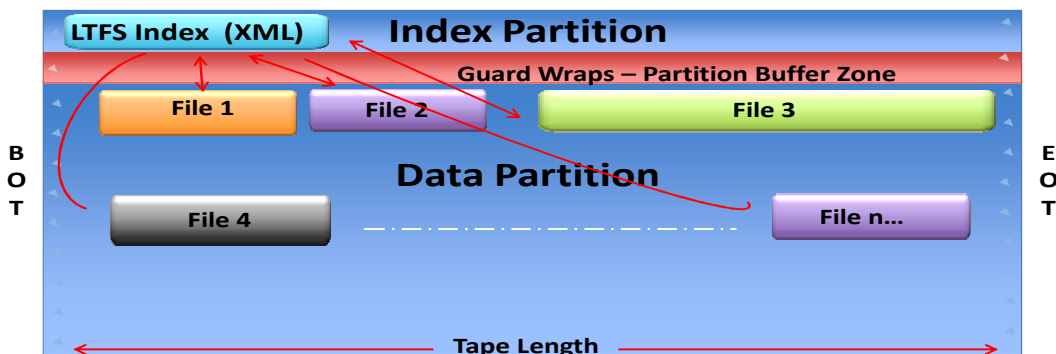
Physically, the Index Partition is organized across the length of the LTO tape and requires about 5% of the total cartridge capacity. The Index Partition is created to use one whole wrap (the set of tracks written at the same time is referred to as a "wrap") on the media. The LTFS directories are designed to be simple, multi-platform accessible, with tags and values for faster retrieval in an easy to read format. 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 complete replacement for disk, LTFS represents a major breakthrough in that it will enable many new applications for tape. LTFS currently supports the Linux, Mac OS X and Windows operating systems.

LTFS Functionality – Points to Remember	
LTFS Drive Ann.	April 12, 2010 for LTO-5 drives (LTFS Single Drive Edition LTFS SDE)
LTFS Library Ann.	July 12, 2011 (LTFS Library Edition LTFS LE)
LTFS Enterprise Ann.	June 10, 2013 (LTFS Enterprise Edition LTFS EE)
LTFS Open Systems	Sept. 19, 2014 (LTFS Library Edition for Windows using third part libraries with LTO-5 and LTO-6 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	Apple McIntosh OS X, Windows 7.0, Windows Server 2008 R2, and Linux.
Tape Drive Support	LTO-5, LTO-6, IBM TS1140, Oracle T10000C and T10000D.
Tape Management Software	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 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 content index data and basic cartridge information like volume name, date, serial number, pointers etc. 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 systems, different LTFS editions, and different applications.

Source: Horison, INC.

Key point: *LTFS is <u>not a disk replacement</u> but it changes the rules for tape storage and enables much faster 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 keeping the index and data in-synch. 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 cost equivalent of 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 Support Extends to Enterprise Class Tape Drives (LTFS EE)

Following support for single drives and libraries, the IBM LTFS Enterprise Edition (LTFS EE) support was announced Oct. 6, 2014 for the IBM TS1150 (10 TB native capacity) tape drive and extended support to IBMs high capacity TS3500 and TS4500 enterprise tape libraries. The enterprise drive announcement signals further expansion of tape partitioning functionality by extending LTFS to the enterprise drive market in addition to the open systems market using LTO drives. The only other enterprise tape supplier, Oracle StorageTek, announced their version of 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.

General Parallel File System (GPFS) Support Becomes Available

LTFS EE also integrates with IBM's General Parallel File System (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 are capable of managing 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.
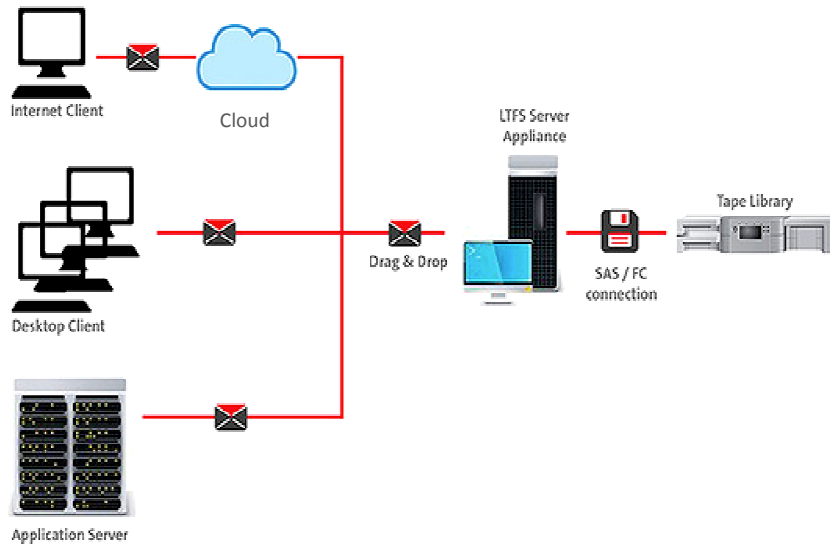
Key point: Rather than storing static, unchanging files on costly disk storage, LTFS EE is designed to enable policy management of tape as a critical storage tier in the enterprise, HPC, Supercomputer and GPFS environments.

LTFS Highlight Summary

- A dual-partition file system first implemented on LTO-5 tape drives
- LTO tape drive support (LTO-5, LTO-6 and follow-on versions)
- Enterprise tape drive support (IBM TS1140,TS1150, 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 similar to 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 HP, IBM, Oracle StorageTek and Quantum and can be downloaded from the following vendor URLs
  For HP   http://www.hp.com/go/LTFS
  For IBM   http://www-03.ibm.com/systems/storage/tape/ltfs/browse.html
  For Oracle   https://oss.oracle.com/projects/ltfs/
  For Quantum   http://www.quantum.com/ServiceandSupport/OpenSource/LTFS/Index.aspx

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 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, expect the LTFS format to become the standard for linear tape.

Using LTFS to Address Archive Data, New Applications and the Cloud

The wave of Big Data applications is fully underway. How much is data growing? According to IDC's latest Digital Universe Study of April 2014, the world's digital data is more than doubling every two years with approximately 2.8 zettabytes created and copied or replicated in 2012 and is projected to be up to 44 zettabytes (44×10^{21}) by 2020. This equates to nearly 6,000 GB of data for every person on earth in 2020. Digital archives consisting of unstructured data, digital images, multi-media, video, social networks, compliance data have become the fastest growing data category experiencing a CAGR (Compounded Annual Growth Rate) of >50% annually. Many businesses that are evaluating whether to build their archive and long-term retention strategy on disk or tape, have an additional option called the Active Archive. These are relatively new solutions that use a combination of NAS (disk) and tape, and are most often combined with LTFS. Gaining marketplace momentum, Active Archive solutions are often referred to as "Tape NAS" are becoming available from several 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. 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. The recent collaboration between two SNIA (Storage Network Association - SNIA.org) technical working groups - LTFS and Cloud - is working to use LTO Technology with LTFS to address cloud data transfer issues. These groups are developing 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 large files directly to tape using the LTFS format making "tape in the cloud" a standard cloud offering.

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

• Archive, compliance and Big Data analytics applications

• Improved access capabilities for surveillance, scientific 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

Key point: Though gradually gaining momentum, expect a steady increase in the number of ISVs and cloud providers supporting LTFS in the future.

Tape Technology Leaps Forward

In addition to LTFS, tape technology innovation has been aggressive and the results are impressive. Enterprise tape drive data rates now surpass disk drive data rates and the media life for modern tape media now exceeds 30 years. Fujifilm and IBM have jointly demonstrated the capability to develop a 154 TB native cartridge using Fujifilm's advanced Barium Ferrite (BaFe) media signaling a long and promising future for tape technologies. This represents significant progress toward achieving tape areal densities of 100 billion bits per square inch and beyond. Native tape cartridge capacity now exceeds the capacity of the largest disk drive. In the process, the BER (Bit Error Rate) of tape surpassed that of Fibre Channel disk making both enterprise and LTO tape more reliable than the most reliable disk drive. These enhancements 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.

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.

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 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. An LTFS based active archive implementation also provides a higher performance archival solution where data can be easily accessed and stored on the most cost-effective media for the lifecycle requirements of that data. To help address the growing 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. See http://activearchive.com/

Selected LTFS Based Archive Appliances and Products	
Fujifilm Dternity	Dternity archiving system utilizes network-attached storage (NAS), cloud and LTFS tape storage and can serve as primary storage and as an active archive for long-term data retention. Dternity's cloud offers service-level agreement options allowing customers to choose a plan that meets their requirements.
Crossroads Systems	StrongBox appears to the network as a file share to hosts and servers, and has a disk-based cache that moves data to the tape library. StrongBox is viewed as a NAS device that can serve as primary storage and as an active archive.
Quantum	Scalar Linear Tape File System (LTFS) presents a tape library as a NAS share, enabling users to leverage familiar file system tools.
Cache-A	Cache-A offers several network-attached small form-factor LTO-5 and 6 archive appliances optimized for the media and entertainment industry using LTFS and the older TAR (Tape Archive) format.
QStar	QStar Archive Manager software creates an Active Archive by virtualizing an archive technology behind a file system and provides LTFS volume-spanning for small, medium and large tape libraries.
Spectra Logic	BlackPearl Deep Storage Appliance combines SSD and LTFS tape, offering a tape-based archival object storage system that uses the RESTful S3 interface.

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

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. With LTFS, tape's evolutionary model has been disrupted by providing significant access time improvement while opening the door for a new, more effective storage solution than ever before. Today an estimated 75-80% of all digital data is generated by individuals - not by large businesses – however the majority of this data will eventually wind up back in large data centers and will be most cost-effectively stored on tape. Therefore, as future tape cartridge capacities significantly increase, the need for improved access and data retrieval capabilities using LTFS and advanced partitioning capabilities will also increase. As a result of this significant progress on many fronts in less than 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.

Though LTFS has barely scratched the surface of its potential, it is helping fuel the tape transformation and potentially can be one of the most significant developments in the tape industry.

End of report