IDC
Analyze the Future

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

WHITE PAPER

LTFS Is a Game-Changing Technology: Extends Use Cases for Tape-Based Archiving

Sponsored by: Iron Mountain

Laura DuBois Robert Amatruda

May 2013

EXECUTIVE SUMMARY

In today's "immediate gratification" society, users and IT professionals alike expect fast access to data when and where they need it. At the same time, data growth combined with an industry that is reticent to dispose of data results in an abundance of information and infrastructure that must be managed intelligently. In the current era of big data analytics, firms want to leverage archived data to deliver new business value. However, storage costs continue to be a concern for IT organizations. As a result, balancing the use of lower-cost media with the need to meet expected SLAs for retrieval time is an ongoing challenge and objective.

Linear Tape File System (LTFS) is a game-changing new technology that extends the use cases for tape, enabling users to drive costs out of storing large volumes of archive data. Firms across sectors such as healthcare, telecommunications, media and entertainment, and government are using LTFS because it provides faster access to data and offers intuitive file operations, an open format, and greater ease of use. Because of the material benefits and leading use cases for LTFS, the ecosystem of technologies that are supporting this new standard continues to grow. As evidenced from interviews with users, LTFS enables firms to cost effectively retain fixed content as well as do more with their information when business needs dictate.

SITUATION OVERVIEW

Today's computer rooms and datacenters face a near doubling of data annually. Not only is more data being created by users and applications, but the increased adoption and use of technologies such as GPS, RFID, social media, and smart sensors and meters result in a growing corpus of machine-generated data. According to IDC's Digital Universe study, the amount of data created will grow by a factor of five between 2010 and 2015. This leaves IT professionals with the daunting task of managing growing storage infrastructure while ensuring that corporate information stored in that infrastructure is managed according to retention, retrieval, or analytics/reuse business rules.

As a result of both business and operational decisions, much of the data that exists in today's storage infrastructure is no longer active. Much of it might be viewed as obsolete. It's data that is no longer accessed or used. However, record retention requirements, eDiscovery mandates, or analytics or business intelligence initiatives across a firm may mean this information needs to be retained. Rather than keeping this infrequently accessed data on tier 1 or tier 2 disk storage, firms increasingly look to use tape for what

is often referred to as deep archive. Firms should not overlook the economic benefits of tape or the power, cooling, and datacenter floor space that tape affords.

However, the need to more easily and quickly access archived data in response to internal requests, regulatory audits, or eDiscovery events or to leverage the data for business purposes is what really drove the adoption of disk for active archive scenarios. Active archiving served compliance and legal use cases but did not address the need to retain large volumes of fixed content that had current or future business value. However, the cost of disk-based archive solutions combined with the growth rate of unstructured data has made this strategy untenable. The ideal solution would offer the TCO/ROI of tape and the accessibility of disk and enable firms to use their archive for mining and business value (i.e., analyze, reuse, or repurpose archive data).

INTRODUCING LTFS AS A GAME-CHANGING TECHNOLOGY

Technology Overview

So what is LTFS, and why is it a game changer? At a simplistic level, LTFS is a standard file system that makes tape easy to use in the data center and in standard IT workflows. It can make an individual tape look like a USB thumb drive and a tape library look like look like network-attached storage. Rather than accessing data on tape by mounting the media in a tape drive and accessing data sequentially by seeking the right location, firms can use LTFS to access data in a more random way using standard tools such as Windows Explorer to navigate to specific files via a directory structure that represents the contents on tape. At a more specific level, LTFS is both a specification and an implementation of the LTFS format for how data is stored on tape as well as a software mechanism that is used to present a file system interface to data stored on tape. This file system view makes accessing files stored on the LTFS formatted media similar to accessing files stored on other forms of storage media such as disk.

Benefits of LTFS

What are the user benefits of implementing LTFS? The benefits of LTFS address many if not all of the shortcomings users have experienced with previous tape formats. Using LTFS offers:

- An **open standard format** – increasing the likelihood that data will be readable years from now without needing specialized, proprietary tools or software.
- A familiar directory-based navigation and direct access to files on tape. This results in **faster access to data stored on tape** over traditional methods because directories can appear on the desktop in the same way as a disk directory listing.
- A standard interface to the datacenter, using protocols such as NFS and CIFS that **makes tape look like a NAS storage infrastructure**.

 ©2013 IDC

- Intuitive file operations where content can be stored on tape by simply dragging and dropping files to and from file directory allows tape to be used by an extended group of business users. Thus, LTFS offers **greater ease of use** and minimizes the use of specially trained tape operators.

- Application and operating system independence through the use of an open, nonproprietary data format and a file system independent of specific software applications. This translates to **lower capital and operating costs.**

LEADING USE CASES FOR LTFS

IDC conducted in-depth interviews with individuals in many organizations that are currently using LTFS and discussed the use cases for LTFS and the benefits that have been realized by using this innovative and game-changing technology. Our research found that the usage of LTFS spans a broad range of industries, including telecommunications, media and entertainment, financial services, service provider, government, professional sports, and healthcare. We discuss four of the most interesting use cases in this section.

DigitalFilm Tree Benefits from Intuitive LTFS-Enabled File Archive, Faster Recovery, and Lower Cost

The tsunami in Japan in 2011 had a material impact on the availability of HDCAM SR tapes, which were the de facto container used for the delivery of video masters in the media and entertainment industry. Availability and pricing of HDCAM SR tapes triggered the more widespread use of LTO for the delivery of video masters and as a backup medium. The more widespread use of LTO as an alternative to HDCAM SR tape also increased awareness of LTFS. As a result, LTFS has quickly proliferated into video production and post-production organizations as a means to archive both raw and gold master content while enabling easy, open file sharing and integration into file-based video editing and production workflows.

DigitalFilm Tree (DFT) is a post-production, creative, consulting, and software development company. Since 1998, DFT has played a definitive role in designing post-production and IT workflows for the media and entertainment industry. Clients include a wide variety of media professionals across the television, motion picture, independent production, and technology industries.

DFT fosters an academic approach to post-production and IT design. Prior to the tsunami, DFT used LTO tape for internal backup and recovery purposes. However, in the media and entertainment industry at the time, some television shows or feature films were using LTO, but the technology was not ubiquitous. Given the shortage of SR tapes, DFT began to suggest that studio clients use LTO-4 and then soon thereafter LTO-5, and, very quickly, the LTO tapes essentially replaced many of the roles that HDCAM SR played historically.

The challenge DFT and the media and entertainment industry faced with the adoption of LTO was ease of use for non-IT staff. According to DFT executives, "In media,

most of the production and post-production houses are not really IT-type environments, where you have an IT guy hanging out. For a lot of people, it was like learning a new language." As a supplier, HP played a critical role. HP provided LTO hardware, LTFS-enabled tape drives, and the commensurate software and technical staff onsite. According to DFT, training people — especially users without an IT background — on TAR LTO was a laborious process. LTFS as a folder on a user's desktop allowed DFT to train someone in minutes to format, mount, and then archive to LTFS. According to DFT, "It has been a 90% reduction in training personnel on archiving."

The use of LTFS has offered DFT some other material benefits. According to DFT executives, "We started off with doing TAR LTO, and one of the major drawbacks of TAR LTO was just this long recovery process. A normal daily production shoot of about a terabyte a day is backed up in 40 minutes onto LTFS. Previously, an LTO TAR backup was a four- to five-hour process to back up an episode. It's been a 75% reduction in actual archiving time." From a cost perspective, LTO tape is more cost effective than SR. With LTFS enabling more widespread adoption of LTO, DFT was able to benefit from the economics of LTO tape. DFT used to spend $500/day shooting to SR stock, but that expense has been replaced by a $40 LTO/day cost or $80/day for two tapes for redundancy. According to DFT executives, "Studios quickly realized the savings … a lot of shows instantly saved about $150,000 to $200,000 in tape stock cost."

National Library Considers LTFS to Enable Preservation of Cultural Audio and Video Content

As individuals, governments, and corporations continue to move from a physical paradigm to a digital paradigm, the role of research libraries continues to evolve. Research libraries once concerned only with storing physical assets (books, photographs, other physical artifacts) now need to retain digital content. A single research library may store millions of digital objects representing tens if not hundreds of petabytes of archive data. These institutions keep data for more than 100 years, and LTFS offers an open format for more simplified media migration. Given that this archive data has broad societal value, the self-describing index stored on tape that is possible with the LTFS standard is also valuable to ensure content and tape reliability over time. IDC recently spoke with a digital archive library of sound and video recordings about its outlook on LTFS and the benefits it anticipates from onboarding the technology.

The national library is an archive environment, not just a backup environment, and most of the content needs to be preserved for at least 100 years. The center receives physical media content from different entities (TV networks, radio stations, state/local government, etc.) and is responsible for making it available and storing it in perpetuity. Today, the center accepts content from third parties in LTO TAR format and in its media intake facility. There, the center analyzes the content, runs AV scans on the content, and then transfers the content across the network into the final archive, which is a SAN-attached enterprise tape library. This national library is evaluating the use of LTFS to help with ingestion of tape content; specifically, using LTFS verification processes to help verify tape content is sound.

 ©2013 IDC

It is critical for the library to be able to recognize bit rot or degradation of media or media that might have some problems before any content is lost. Although the library's policy is to keep two copies of all content, archive administrators need to know when a copy has degraded enough that it's not safe and migrate content off the degraded media to safe media. The ability to validate and verify all the content on a tape without having to stage it back to disk reduces time and cost, especially with the scale of the environment. The archive library has more than 3 petabytes archived, with a design point to scale to hundreds of petabytes. The archive library wants to peruse quickly the full table of contents (i.e., the self-describing LTFS metadata) without necessarily having to mount all the tape, which is one of the attractive features of LTFS.

Another challenge the library faces is that when it receives standard LTO TAR formatted tapes from third parties, the library needs to know and retain the right backup software and underlying operating systems used to write data onto the tape to restore the data. The library hopes that LTO LTFS formatted tapes as an open standard can eliminate those dependencies. The library is satisfied with the enterprise tape drive and automation product it currently uses, so it will not be migrating to LTO for its preservation archive. However, it wants to accept LTO LTFS formatted tapes for introducing content into its environment.

MLB Network Integrates LTFS Tape Workflows into Its Media Asset Management System, Gains Economic Advantages

Many organizations collect digital media assets (e.g., digital photos, animations, videos, music, and audio). These assets can be brand related and useful for sales and/or marketing organizations or relevant to actions or events pertinent to businesses that monetize rich media content. Typically, media asset management systems enable access to audiovisual content so that content owners can easily and efficiently reuse, repurpose, and resell content. LTFS integration into media asset management systems enables more streamlined access to rich media content for use and reuse.

Major League Baseball (MLB) oversees the game of professional baseball in North America, including 30 franchises in 28 cities across the United States and in Canada. MLB Network, the league's 24 x 7 cable television network, captures video content assets from venues in its two SAN storage environments, each with over 70 terabytes usable space, providing over 4,000 hours of video content storage. These storage pools serve as primary editing SANs for the editors working with the content on the SANs natively. Once an edited file or game is finished, the MLB Network personnel responsible for media management will decide what content to keep and what content to purge and will migrate a copy off to the LTO-4 environment. On a daily basis, MLB Network and MLB Productions go through 25–30 terabytes of tape.

MLB Network, MLB Productions, and MLB Office of the Commissioner created a custom application called DIAMOND that allows data about the play action to be logged in real time while content is being recorded into the environment. The aim was to create a catalog of information for a given time code of the video. Some of that information is automatically prepopulated from other data sources (player rosters,

statistics, etc.). As the games are recorded, MLB Productions loggers add metadata tags from over 700 categories and may enhance that information by entering metadata describing "it's a high fly ball back at the warning track with a diving catch," so producers mining the video content can drill in and extract the video content that helps them tell their story. The DIAMOND application enables locating points in time of video content based on the metadata.

MLB Network uses the Grass Valley video production system and Front Porch Digital tape library management software for the over 13,000 LTO tapes stored in the Oracle SL8500 library. When full game video files are sent from the Grass Valley system to the Front Porch system, pulling the entire game from tape takes as much as one hour. The environment needs to support desktop editing as well as searching within the DIAMOND application for viewing the content and for building collections of content. The ability to perform partial file restores of video content enables high-speed access to the specific content that is required for editing. However, according to MLB Network Media Management executives, "the current Grass Valley application requires low-resolution proxy content to be served off spinning disks or at least needs to believe that the proxy content is online and on spinning disk, and that's where LTFS comes in. The LTFS environment looks like a disk mount to the application. That is an environment that is extremely important to MLB Network because it gives the availability from an application level of appearing to be a disk environment but yet is based on tape in the back end. LTFS solves this environmental area where the application needs high-performance disks or needs something along those lines or looks like it has that kind of access, but at the same point, you're really building tape behind that mount."

Another important factor was the economics. According to an MLB Network executive, "Disk is great for that instantaneous experience, but with large archive proxy libraries, you end up with content that never gets viewed, but yet you're paying the premium for high-performance instant access, and that's where the LTFS comes into play and is what we're looking at as a possible solution." MLB Network is currently testing the Crossroads StrongBox LTFS solution in its environment and identifying workflows that it can support. The first step was to interface StrongBox to the Apple Final Cut Pro edit SANs for archiving an entire video project from high-definition files to edited, finished products. The executive added, "We developed that workflow last year and got that to work very elegantly, very nicely." LTFS allowed MLB Network to leverage a lower-cost LTO LTFS tape environment to replace the high costs of spinning disks to store the equivalent amount of data.

U.S. Media and Entertainment Company Enables Faster Access to Call Center Recordings Using LTFS

Call centers across many industries must retain call center records for prescribed periods of time. For example, telecommunications firms must keep customer voice recordings for seven years. Retrieving the records pursuant to a customer dispute can take weeks because the process involves locating data on backup tapes. Using LTFS enables firms to store the audio files on disk for several years and then leverage a standard file system directory structure to more easily locate and retrieve relevant audio files that have moved to a tape medium.

A large United States–based media and entertainment company has used tape extensively for operational backups and disaster recovery. The company uses tape to meet its regulatory obligations to retain customer voice recordings for seven years. However, according to company executives, the use of tape has been problematic in terms of retrieval "because the way our voice recording works, it basically creates all of these little files in complex direct infrastructure, so if you back it up and use your conventional backup to actually find stuff and then get what you need for compliance purposes, that process can take weeks and weeks." Retrieval requests, which arise as a result of eDiscovery events, customer disputes, or the need for management to review how customer communication was handled, were taking three weeks to satisfy.

The company built in a tiered archiving process so that the first stage of the process would be to back up to disk and then to LTO LTFS tape. The company set up a staged environment to keep a year and a half of data on backup disk archive and then push the rest to tape, which is stored offsite. According to the company, the best way to move its archives while keeping the file structure relatively intact and still be able to find things was to use LTFS. To meet one-week SLAs, the operators leveraged LTFS as a persistent file system to know what volume the records were on, and they were able to reduce retrieval times by as much as 66%.

There were also economic advantages of the new LTFS-enabled archive. With 40–50 requests per month, the company estimates it saved approximately 200–250 man-hours per month in support personnel labor. The workflow for retrieval is simple. The call recording software has catalog information that allows an operator to identify a particular file that satisfies the request. The cataloged file resides in a particular folder, and based on that, the volume is identified using the recording software. Once the volume is identified, the operator loads the volume and is able to basically pull an actual file off a file system directly without having to look through the tape backup software catalog, find the file, identify the volume, and then restore the data to a particular storage space. That speeds up the process dramatically.

THE LTFS ECOSYSTEM

LTFS is an open tape format specification and file system that has enabled an entire ecosystem of technology providers to offer tape support in a nonproprietary manner. This growing ecosystem includes independent software vendors (ISVs), storage manufacturers, and industry-specific solution providers. Support for LTFS by this ecosystem enables technology providers to offer their customers the ease of use, increased utility, and improved economics that LTFS affords.

Among the software community, LTFS has been embraced by vendors in the backup application software market such as QStar, Veeam, and Quest Software. In the media and entertainment space, LTFS has been backed by companies such as StorageDNA, Cache-A, Avid, and Grass Valley, which support the LTFS standard as a nonproprietary media format allowing for improved media, post-production, and archiving.

Within the storage industry itself, support for LTFS has spanned the tape drive and tape automation community, including HP, IBM, Oracle, and SpectraLogic, all of which have a long-standing history of tape drive, automation, and media integration.

All the participating tape hardware vendors provide full support for LTFS integration across their portfolio of products. Vendors with file-based storage offerings such as Quantum are beginning to support LTFS through their distributed file system offerings or via NAS appliances.

Providing a NAS-like interface to tape using standard network protocols such as NFS and CIFS presents a real opportunity to expand the use cases for tape. One company that recognized early on the strategic opportunity that LTFS offered was Crossroads Systems. The company launched its StrongBox solution as a turnkey data archive storage solution that provides customers with a standard NFS or CIFS interface to their tape libraries.

The use cases in this document show how LTFS is a game changer for on-premises storage. However, IDC research suggests that more companies are looking to purchase IT services (IaaS) through the cloud. The challenge that customers and service providers alike face is offering a low $/GB while also satisfying customer SLAs. Depending on the workload or use case, the role of tape in cloud services can fundamentally change the economics. For example, the Amazon Glacier service now offers a price mode of $0.01/GB per month for its colder storage services. IDC believes that the use of tape in concert with the introduction of LTFS will drive greater adoption of tape by service providers.

CHALLENGES

LTFS offers significant advantages to the use of tape and enables organizations to leverage the economic benefits of tape while also providing a much better SLA to organizations. Historically, tape as a storage medium has presented users with a cumbersome management model and longer restoration times. However, LTFS has largely mitigated many user concerns and provides a familiar file-based interface that is extremely well suited to data mobility, access, and ease of use. The LTFS standard leverages tape's favorable economics for capacity expansion and provides users with nonproprietary options for rich content and unstructured content archiving. We expect more innovations and improvements with LTFS that will bring the technology into a broader set of use cases in the near future.

The fundamental success of LTFS will hinge on the ability of technology providers to broaden the LTFS ecosystem beyond a finite set of rich media and broadcast verticals. IDC research shows that LTFS is beginning to be used in a variety of vertical segments such as financial services, media and entertainment, video surveillance, telecommunications, government, and healthcare. However, given the ecosystem of ISVs and applications that exist across industries, the support for LTFS needs to continue. More commercial software packages must support LTFS for its momentum to continue. Increased adoption of LTFS for a greater span of use cases will translate into more investment and growth in the LTO tape sector in the near future.

 ©2013 IDC

CONCLUSION

Increasingly, businesses are looking for new ways to derive additional business value from their information, including both active and archive data. But growing storage costs require an intelligent approach managing data across different storage technologies, such as using lower-cost tape media with spinning disk. LTFS is a game-changing technology that enables organizations to leverage the cost and reliability of tape with the accessibility and ease of use they are accustomed to with disk-based solutions. Companies should evaluate how LTFS can help them better meet their expected SLAs while scaling their storage and managing their costs.

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2013 IDC. Reproduction without written permission is completely forbidden.