Tape Manifesto for the 21st Century

The Era of Modern Tape is Here

Fred Moore, President
Horison, Inc.
www.horison.com

Tape Today

Have you wondered what has been happening in the tape storage industry, if anything?
Did you know?

- Tape is cheaper to acquire than disk,
- tape is less costly to own and operate than disk,
- tape is more reliable than disk,
- tape now has media partitions for faster "disk-like" access,
- the capacity of a tape cartridge is higher than disk drive capacity, and
- the media life for tape is 30 years or more for all new media.

If you didn't - you are not alone – but it's time to bring your understanding of tape up to date.

In response to this "awareness" challenge, the tape industry began to re-architect itself and the renaissance is well underway. As this century began, several important technologies were implemented yielding numerous tape improvements including unprecedented cartridge capacity increases, vastly improved bit error rates, much longer media life and faster data rates than any previous tape or disk technology. Many of these innovations have resulted from technologies borrowed from the disk or HDD (Hard Disk Drive) industry and have been used in the development of both LTO (Linear Tape Open) and enterprise tape products. Clearly disk technology has been advancing, but the progress in tape has been even greater over the past 10 years. Today's contemporary tape technology is nothing like the tape of the past. It's time to bring your views of tape up to date as the future for tape technology has never been brighter. *The era of modern tape is here.*

The Era of Modern Tape Arrives - Major Tape Enhancements Since 2000

By 2000, a new era of tape was fully underway as the tape industry was busy re-engineering itself. Key tape developments yielded much longer media life, vastly improved drive reliability, lower acquisition and TCO (Total Cost of Ownership) and much faster data rates than any previous tape technology. Troublesome tape issues of the past including stretch, edge damage, tear, loading problems, and media alignment from older tape formats such as Travan, DAT, DDS, DLT, and 8MM tape were successfully addressed. The most recent LTO-6 format has expanded the "history buffer" in the compression engine, giving it a 2.5:1 compression ratio, up from 2:1, yielding a 6.25 TB per cartridge capacity.

Reliability

PRML (Primary Response Maximum Likelihood) is the most popular and effective error detection scheme used in modern disk drives. Borrowing from disk, LTO drives switched to PRML from RLL (Run Length Limited) encoding as PRML attempts to correctly interpret even the smallest changes in the analog signal. Because PRML can correctly decode a weaker signal it enabled a much higher recording density while allowing tape to surpass disk in reliability.

The head positioning servo tracks were embedded in the tape rather than being located on the edge as was the case with older tape products. This provided increased positioning accuracy of the drive's read-write heads while keeping the format consistent throughout the life of the product line and ensuring backward compatibility across the current and two prior generations of LTO products.

Technology and Reliability	**BER (Bit Error Rate)**
Enterprise Tape (T10000x, TS11xx)	1 x $10E^{19}$ bits
Midrange Tape LTO-5, LTO-6	1 x $10E^{17}$ bits
Enterprise HDD (FC/SAS)	1 x $10E^{16}$ bits
Enterprise HDD (SATA)	1 x $10E^{15}$ bits
Desktop HDD (SATA)	1 x $10E^{14}$ bits

Source: Vendor's published product specifications.

Security and media life

Security features were added to LTO and enterprise tape offerings to address increasing compliance and legal requirements including Write-Once-Read-Many (WORM), data encryption to protect data at rest, and various write-protect capabilities. In addition, the media life range for all LTO and enterprise class tape now reaches 30 years or more, up from 4-8 years on prior tape media, making tape a very secure long-term storage and archive medium.

Improved Tape Access Time

A new tape file system specification for LTO tape called LTFS (Linear Tape File System) became available with LTO-5 in 2010 and has changed the rules for tape access. LTFS allows tape data to be accessed in a manner similar to disk or other removable media providing ability to drag, drop and share data without regard to platform. The faster tape file access capability provided by LTFS becomes more important as tape cartridge capacities and the number of files per tape steadily increases. Even with tape management software, file access was not always intuitive, as it often required users to know which cartridge volume contained the files they needed. LTFS takes a significant step in making tape storage easier to use.

Total Cost of Ownership

A comprehensive TCO study by ESG (Enterprise Strategies Group) [1] comparing an LTO-5 tape library system with a low-cost SATA disk system for backup using de-duplication (best case for disk) shows that

disk deduplication has a 2-4x higher TCO than the tape system for backup in several use cases over a 5 year period. A variety of other published studies conclude that disk has a TCO ranging up to 15x higher than tape for long-term data archiving. The TCO advantage for tape is expected to become even more compelling with future technology developments.[1] A Comparative TCO Study: VTLs and Physical Tape, By Mark Peters, ESG, Feb., 2011.

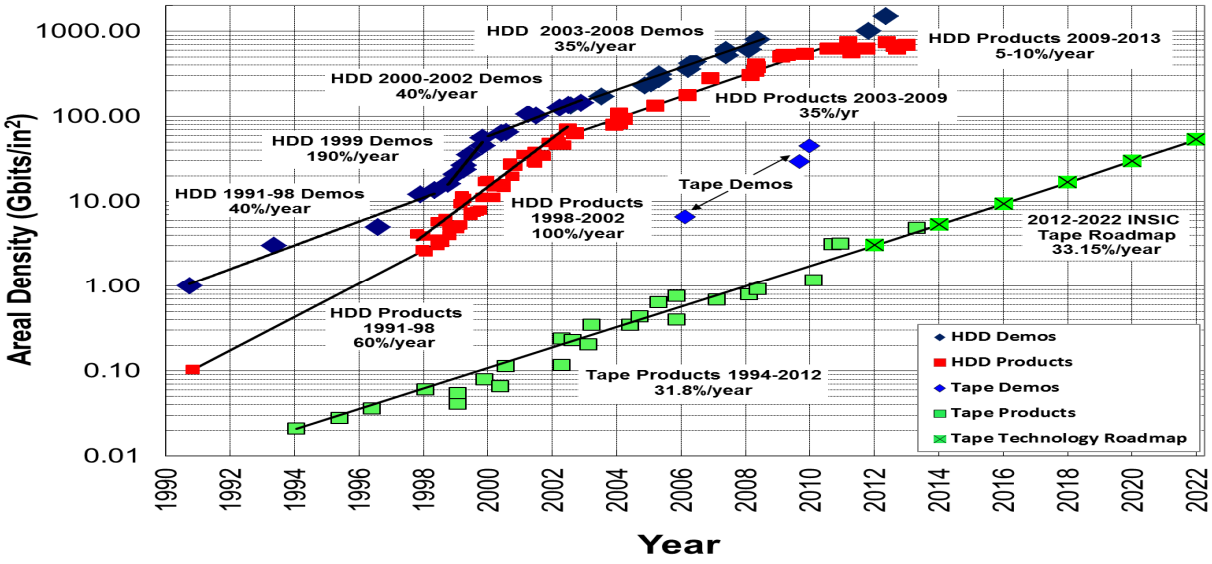
Key point: *Numerous tape enhancements have been implemented that are redefining the role of tape.*

Future Data Recording Projections

A significant step was taken regarding the long-term future of tape on Jan. 2010 by IBM Research in Zurich when IBM recorded data onto an advanced prototype barium-ferrite tape media developed by Fujifilm Corporation at a density of 29.5 billion bits per square inch that could produce a native capacity of 35 TB. Several new tape technologies were developed as a result of this R&D effort including improved precision control of read-write head positioning, more than a 25-fold increase in the number of tracks, new detection methods to improve read accuracy, and a new low friction read-write head. These enhancements represent a significant step toward achieving areal densities for tape of 100 billion bits per square inch and beyond. The tape industry has pushed tape capacity, reliability and media life to new levels. Reliability for enterprise and open systems tape drives and media are one to three orders of magnitude higher than the most reliable Fibre Channel disk drive.

Key point: *Honestly, did you realize tape has a higher capacity than disk and is more reliable than disk?*

Areal Density of Hard Disk and Tape Laboratory Demonstrations and Products



These scenarios (see chart above) suggest that the annual tape areal density growth rates will be either maintained at traditional 40% values or exceed traditional growth rates and could possibly approach 80%. These scenarios also suggest that annual HDD areal density growth rates will not maintain their

traditional 40% values using vertical recording and will likely slow to 20% values. Specifically, the tape bit cell is 300-X to 500-X larger than the HDD bit cells. Thus the area available to increase HDD capacity is getting crowded and HDD areal density growth is slowing. The net result of these areal density scenarios is a sustained volumetric and total capacity storage advantage for tape technology over HDD technologies for many years. Tape cartridge capacities of 125 TB native are within reach given these strides. This also signals a lower total cost per gigabyte for tape than other technologies for the foreseeable future.

Key Factors That Changed the Rules for Tape

Tape Reliability Better Than Disk Reliability *Note: BER (Bit Error Rate) Hard Read Errors per Bits Read*	T10000x Enterprise 1 x $10E^{19}$ bits **Tape drive BER** TS11xx Enterprise 1 x $10E^{19}$ bits LTO-5,6 Open Systems 1 x $10E^{17}$ bits Enterprise HDD FC/SAS 1 x $10E^{16}$ bits **Disk drive BER** Enterprise HDD SATA 1 x $10E^{15}$ bits Desktop HDD SATA 1 x $10E^{14}$ bits
Capacity and data rate	The increasingly popular LTO tape drive family is now in its 6th generation with LTO-6 and two future generations are now defined in the LTO roadmap. Today, over 85% of all tape drives and 90% of all cartridges shipped are LTO. LTO-6 media has a native capacity of 2.5 terabytes (6.25 terabytes compressed @2.5:1) and a data rate of 160 MB/sec (400 MB/sec compressed), higher than disk.
Capacity growth rates	Roadmaps favor tape over disk with 35 TB capabilities now demonstrated by Fujifilm and IBM using barium ferrite materials.
Security features	LTO tape drives offer both WORM and encryption for higher levels of security.
Long life media (Shelf life)	The shelf life for all new enterprise tape media and LTO media is rated at 30 years or more making it ideal for long-term archival storage. Disk drives typically have a 3-4 year lifespan before replacement or failure.
Improve tape file search and access time performance (LTFS)	LTFS arrived with LTO-5 and partitions the tape cartridge into two distinct, individually addressable, unequal partitions with the first quick read partition containing descriptive metadata that enables the quick search capability (random-like) of the data contained in the second partition.
*Energy efficiency heavily favors tape **(green initiatives)***	This is becoming a goal for most data centers for archival data - "If data isn't being used, it shouldn't consume energy". The 5 year HDD energy cost is approximately 25-X higher than that of tape for equivalent subsystems.
Acquisition price	Tape has a lower purchase price ($/GB) than disk.
Tape TCO Better Than Disk	5-year TCO for disk is 2-4X times higher for backup and 15-X higher for archive storage than tape.

Source: Horison, INC.

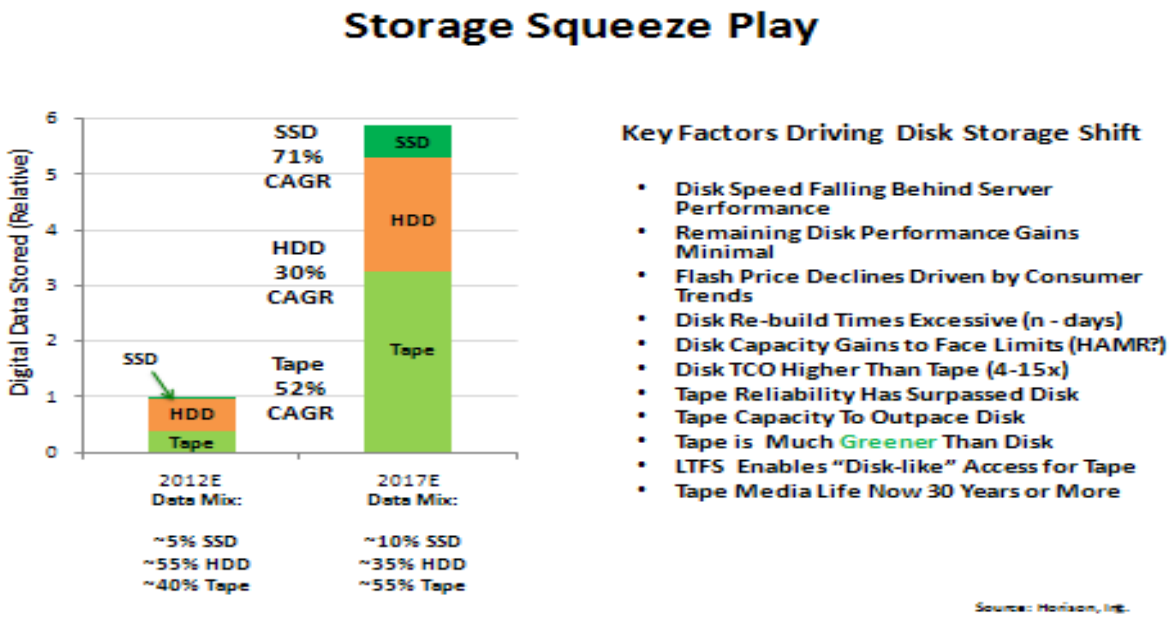
Key point: *Tape recording technology faces few limitations in the foreseeable future enabling tape to address the large-scale storage demands that lie ahead.*

Disk Challenges Are Mounting

Technology roadmaps and approaching technology limits indicate HDD storage is entering a squeeze play in the data center. HDDs are increasing in capacity -but not in performance - as the IOPS (I/Os per Second) for HDDs have basically leveled off. Therefore, as HDDs increase in capacity, their ability to deliver the same level of performance declines. The potential for more concurrently active data sets increases as HDD capacity grows and the increased contention for the single actuator arm causes response time delays. This concern has fueled the emergence of SSD (Solid State Disk) using flash memory in order to deliver faster service levels for storing high-performance and response time critical data. HDDs also have a higher TCO and use considerably more energy than tape or SSD. Excessive RAID rebuild times are becoming longer and now can take several days to rebuild a failed HDD. This means that the disk subsystem will run in degraded mode during the rebuild period. As HDD capacities continue to increase, the rebuild process will become unacceptable for most IT organizations.

A common objective for many CIOs today is that *"if data isn't used, it shouldn't consume energy".* In response to this directive, the movement of archival data from HDD to more reliable, much more energy efficient and more cost-effective tape storage is well underway at the other end of the storage hierarchy. Tape capacities will continue to grow and total ownership costs will decline, while HDD performance increases will be negligible.

The shift in the storage landscape is underway as high performance data finds its way onto flash SSD and lower activity and archive data migrate to modern tape. As a result, the amount of data stored on HDDs will continue to grow, but at a slower rate than in the past, facilitating a "storage squeeze play" – see chart below.



For the foreseeable HDDs will remain the home for primary storage, mission-critical data, OLTP, and databases. HDD shipments will continue to grow *but* growth rates will slow as performance sensitive data begins to move from HDDs to SSD and archival data moves to tape. Over the next five years, forecasts for growth in the total amount of data stored on HDDs is expected to fall from 40%-45% annually to 30%, below the 52% CAGR for tape and 71% CAGR for flash SSD. As a result, HDDs are caught in the middle as storage administrators strive to optimize their storage infrastructure by getting the right data in the right place – and the technologies are there to do so.

Summary - Comparing Key Factors for Tape and Disk

Factor	Tape	Disk (HDD)
Price/TCO (5 years)	TCO for HDD ranges 2-15X higher than tape for backup and archiving. Price per GB much lower for tape	HDD purchase price typically 20% of 5-year HDD TCO. Disk has more frequent conversions and upgrades
Long-life media	30 years or more on all new tape media	~4 years for most HDDs before upgrade or replacement, 7 years or more is typical for tape drives
Maximum Capacity	Current max. tape cartridge is 5 TB (>10 TB compressed @ 2:1)	Max. HDD capacity is 4 TB
Reliability	Tape BER (Bit Error Rate) has surpassed disk and is 1 to 3 orders of magnitude higher than best HDD	Disk BER not improving as fast as tape
DR - Move data to remote location for with or without electricity	Yes, can move data remotely with or without electricity. Natural disasters can require physical media movement	Difficult to move disk data to a remote location for DR purposes without requiring electricity
Inactive data does not consume energy	Yes, this is becoming a goal for most data centers. "If the data isn't being used, it shouldn't consume energy"	Rarely for disk, potentially in the case of "spin-up spin-down" disks *Note: data striping in arrays often negates the spin-down function*
Provide the highest security levels	Yes, encryption and WORM capability available on essentially all midrange and enterprise tape	WORM and encryption available on selected disk products but usage not yet widespread
Capacity growth rates	Roadmaps favor tape over disk with 35 TB native capability jointly demonstrated by Fujifilm and IBM	Continued steady capacity growth - roadmaps project disk to lag tape
RAID re-build times becoming prohibitive	NA	Can take several days to rebuild a nTB HDD in a busy RAID array
Key Applications	Large files and large streaming data applications, active archives, less frequently used data, data that must be securely stored indefinitely	Active, frequently accessed data, mission critical data, databases, OLTP, revenue generating applications
Labor content	Typical tape administrator can manage PBs ($1x10^{15}$) of tape	Typical disk administrator can manage TBs ($1x10^{12}$) of disk

Source: Horison, Inc.

Key point: *HDDs are caught in the middle as storage administrators strive to optimize their storage infrastructure to better address high performance applications with SSD and archival demands with tape.*

Archive, Compliance, Fixed Content Applications Drive Tape Demand

Tape has been expanding its historical role as a backup solution to a much broader set of requirements including data archives and disaster recovery services. Digital archives, unstructured data, digital images, multi-media, video, social networks, compliance data and are the fastest growing data category experiencing a CAGR (Compounded Annual Growth Rate) of ~65% annually. [Note: this category of storage is also referred to as Tier 3 storage]. Just ten years ago, large businesses generated roughly 90% of the world's digital data. Today it is estimated that 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 a large business data enter or service provider.

Businesses face growing challenges in managing archives, controlling costs, and meeting regulation requirements. Difficulties include data capture (ingest), storage, retrieval, protection, and providing timely retrieval. The value of an archive is increasing as the benefits of working with larger and larger datasets (Big Data) enable analysts to project business trends, prevent diseases, and improve security and national defense. Presenting an ever-moving target, the limits of archives are now on the order of petabytes ($1x10^{15}$), exabytes ($1x10^{18}$) and will approach zettabytes ($1x10^{21}$) of data in the foreseeable future. Meeting these storage requirements with disk will be financially prohibitive for most businesses.

Key point: *With a TCO of 1/15th of disk for archival storage, the pendulum has shifted to tape to address much of the exploding Tier 3 demand.*

What about Tape in the Cloud?

The inherent consolidation of unstructured data into large-scale archive repositories in the cloud suggests that another category of storage is emerging –*Tape in the Cloud* – which will further expand the economic model for tape as the preferred archiving technology. Cloud storage users currently rank security compliance, and service levels as their top cloud concerns. By definition, cloud users must be willing to have their data reside side by side with that of a possible competitor, which is why encryption mechanisms can become more important. When disks are used for cloud storage, disk drive arrays use RAID to break up and spread data from several users at the block level across multiple disks for data resiliency. Additional software is required to ensure that any given customer's data is securely isolated from every other user on a given disk or array.

In tape environments, however, each tape cartridge is a separate object. The customer or cloud provider has control over what's on each one. Encryption and WORM capabilities provide security for data at rest and are fundamental to delivering a secure archive strategy. If just one bit is in error, encrypted data cannot be de-crypted and the data is rendered useless. Tape is removable media enabling physical cartridges to be easily transported to another location in the case of power outages which have become

all too common in recent years. Without electricity, media portability becomes the last line of defense for data protection.

Key point: *Cloud providers should look to tape for the lowest-cost, long-term archival storage offering. Storing archival data in the cloud represents a future growth opportunity for tape and a lower cost solution for cloud providers.*

Summary

Today's reality is that the magnetic tape industry has made considerable progress in the past 10 years. The continued role for disk is in place but tape has expanded its position as a viable complement to disk for the foreseeable future due to its lower total cost of ownership, improved reliability, and significantly lower energy costs than other alternatives. As a result of this progress, the tape industry is aggressively re-positioning itself to address the many new large capacity storage opportunities which now represents over 65% of the world's stored digital data and that number is growing. The latest technology improvements in the tape industry suggest tape will continue to be more cost-effectively suited for the enormous archival opportunities that lie ahead, whether on-site or in the cloud. For tape it's a new game with new rules.

End of report.