Tape Storage Future Directions and the Data Explosion

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

Abstract - Have you wondered what has been happening in the tape industry, if anything? If so, you are not alone. As the twenty-first century began, there was a growing perception that disk had become cheaper and more reliable than tape and in response to this challenge, the tape industry began to re-architect itself. Several important technologies were borrowed from the disk industry yielding numerous tape improvements including unprecedented cartridge capacity increases, improved drive reliability, vastly improved bit error rates, much longer media life and faster data rates than any previous tape or disk technology. Tape drives offer WORM (Write-Once-read-Many) and encryption for better security, have the highest capacity of any storage device at up to 5 TB (native) and the fastest data rate of any storage device at up to 250 MB/sec (native). In addition, a new tape file system specification called Linear Tape File System (LTFS), has changed the rules of access for tape, allowing tape data to be accessed in a manner similar to disk or other removable media. It's time to bring your views of tape up to date as the future for tape technology has never been brighter. Today's tape technologies combined with disk technologies in a strategic blended implementation can help users address a myriad of objectives including performance, compliance, security, cost, energy consumption, archive and data protection.

History of Tape and Disk

The tape Industry began in 1952 and the disk Industry in 1956. In 1952, the world's first successful commercial tape drive was delivered, the IBM 726 with 12,500 bytes of capacity per reel. In 1956 the world's first disk drive was delivered by IBM, the Ramac 350 with 5 megabytes of capacity. Though no one knew it at the time, two key and lasting events linking disk and tape for the foreseeable future had just occurred; 1) the first storage hierarchy was created with online disk and offline tape storage and 2) the first storage management applications were born, namely backup and recovery, defining the primary role for tape. Backup and recovery would become the primary storage management applications for over 50 years as protecting business critical data became increasingly important.

Bottom line: *Since their inception, disk and tape have provided the storage platform for the vast majority of the world's digital data. It is estimated that 70% of the world's data is on tape.*

State of the Data Protection Environment

How much is your data worth? You need to protect it against accidental loss and theft from both insiders and outsiders. Plus, more and more people are working away from the office. Without data employees can't work and business will suffer. With few exceptions, users expect 100% availability of their applications and their information. Tape is still used by many for backup and recovery applications. However, disk based backup (native disk systems, Flash Copies/Snap, virtual tape libraries) is also being used for backup and recovery due to disk's faster recovery times, while tape as the second tier normally provides a last line of defense. It is expected that flash and snap technology backups will see increased usage and then secondary backups of those data copies can take place with tape for an offline geographically distributed tier of protection.

Disk with deduplication has gained momentum for backup. Deduplication is a disk feature for reducing the disk resources needed to store the backup copy of data. Deduplication boasts potentially large reductions in the amount of backup data that is actually stored on disk, particularly in environments where either incremental or differential backup methods are not deployed. However, deduplication can have a performance tradeoff. Where high performance streaming backup and restore is needed (e.g. large databases, offload backup from flash/snap copies, etc) tape can best fulfill this high performance streaming workload. Deduplication is best suited for applications that have multiple instances or copies of the same data. Backup is the most obvious beneficiary of deduplication, but deduplication is not necessarily effective for other applications such as archives (which do not normally contain much duplicate data), video, photos, and the growing list of fixed content tier 3 applications. As the name suggests, a high duplicity factor is needed to make deduplication most effective.

Software, hardware, and human errors, along with the increasing number of data destructive network intrusions have turned data protection into a critical and complex data management process. Natural disasters, once an infrequent occurrence, often seem like the norm as effects of fires, floods, hurricanes, earthquakes, tornados, and volcanoes often mean extended power outages. These have made the requirement for geographically distant locations and logical offline protection for backup, recovery and data archives a "must do" task for many data centers.

Long media life, high reliability, and a significantly low total cost of ownership have positioned tape as both the optimal remote and local offline storage solution for data protection and archive. Tape has gained momentum as the last line of defense if first tier backups (flash / snaps) were to fail or be corrupted. As an email service provider found, a software error corrupted primary and replicated disk copies of critical files. Fortunately, backups on off-line tape were not able to be corrupted and were used to restore. In addition, recent studies have shown tape to have significantly less TCO than comparable disk systems (more on cost in the section below) and with automated tape libraries very few labor hours are needed to manage tape. Tape storage, being offline and portable, protects data from the accidental and intentional threats that can occur with on-line storage. Data stored on tape media can be

physically removed from the system and ported to a more secure and remote location to protect the data against a system or primary site disaster.

> ***Bottom line:*** *Tape has historically been the backup technology of choice and is still used heavily in that role as well as for long term retention. In recent years disk is also being used for backup for its fast recall of high access data. Tape vendors continue to innovate and deliver compelling new features with lower economics which are positioning tape as the optimal choice for long-term archiving as well as continuing to play a key role for backup.*

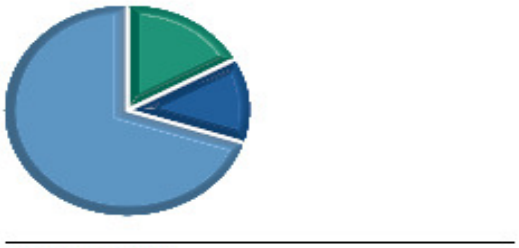
Re-architecting the Data Center with Disk and Tape

Storage managers have several capabilities available to enable them to re-architect their storage infrastructure. By implementing new disk and tape capabilities that improve storage efficiency by reducing the number of devices, lowering energy consumption, improving I/O performance, increasing allocation levels, adding data strong security measures while lowering costs. Data classification also helps answer the question "what data goes where" and encompasses aligning data with the optimal storage architectures and services based on the changing value of data. Optimal data placement is usually associated with a tiered storage implementation. Tiered storage is a data storage environment consisting of two or more kinds of storage, usually a combination of disk(s) and tape, and is delineated by differences in four primary attributes: price, performance, capacity and functionality.

Industry surveys indicate that disk and tape are often used in combination. A recent Storage Media Group study has brought to light that 77 percent of all companies are backing up either all or part of their data on tape. Another survey published by Infostor (see charts below) indicated that tape is used in backups by 83% and for archive by 82% of the respondents. In addition, disk and tape are frequently used together to address data recovery and archive issues with 70% of the respondents using disk and tape for recovery and interim storage, while 58% use both technologies for long-term data archiving (see charts below).

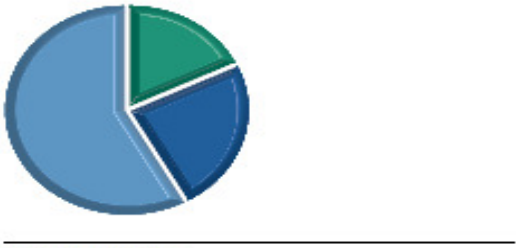
Current Systems Used

Backup and Recovery



Source: *Infostor* May 28, 2008 (pg.16)

Long-term data Archiving



Source: *Infostor* May 28, 2008 (pg.16)

Today, most all tape drives offer WORM, **data encryption with key management, and various write-protect capabilities to address this growing list of data preservation and security requirements.** Storage managers have been required to implement more stringent security measures for digital data in order to address the growing list of legal, government mandated compliance laws, and long-term archival requirements. The chart below identifies key storage issues that are addressed by disk, tape, or a combination of both to optimize the storage infrastructure. Though challenging and becoming increasingly complex as storage requirements grow, re-architecting the data center yields much improved operational efficiencies and cost savings. The time is here for businesses to develop plans to begin re-architecting and optimizing their IT systems before the task becomes overwhelming.

Disk and Tape Combine to Deliver Storage Optimization

Storage Issue	Efficiency Solution	Avg. Improvement %	Technologies Used
Optimizing data placement	Tiered storage, data classification	Capacity: 20%-50% GB	Disk and tape
Too many devices	Consolidation/ virtualization	Reduction varies	Disk and virtual tape
Energy consumption too high	Move less active data from disk to tape	~20-50% of GB can be re-allocated	Disk and tape
Reducing file size	Compression (2:1, 3:1)	Capacity: 50%-66% GB	Tape
Eliminate duplicate data (primarily backup data)	Deduplication	Capacity: 70%- 80% GB	Disk
Over-allocation of storage	Thin provisioning	Capacity: 20%-30% GB	Disk
Backup/restore performance - smaller files, smaller DB	VTL, de-duplication, CDP, Snapshot, Replication	~10-50% Optimal speed for smaller files	VTL
Backup/restore performance – big files, large DB	Streaming data transfer performance	~10-50% (Improves as file size increases)	Tape
Data protection theft/alteration	Encryption/WORM media	Data unusable and not alterable without secure key	Most all midrange and enterprise tape drives, select disk drives

Source: Horison, Inc.

Bottom line: *With disk and tape containing the vast majority of digital data, optimal data management strategies will combine the capabilities of disk and tape to most effectively address the assorted objectives and future storage requirements that storage managers face.*

Dispelling the Myths of Tape Reliability

Questions about the true reliability of tape have persisted for some time. In recent years, significant tape advances for reliability, capacity and performance are being delivered at an accelerating rate. Tape cartridges have been ruggedized for improved reliability, portability while tape drive MTBF (Mean Time Between Failure) has soared from 80,000 hours to 250,000 hours at a 100% duty cycle in the past decade. Today's advanced tape media formulations provide an expected media shelf life range for new tape cartridges up to 30 years or more lending itself effectively to long-term archival storage.

Both disk and tape have made significant reliability improvements in recent years. For tape, reliability progress has been even better than disk comparing the BER (Bit Error Rate), which is quickly becoming a more popular means of measuring reliability. The BER is the percentage of bits that have errors relative to the total number of bits received in a data transfer, usually expressed as 10 to an integral power. A data transmission might have a BER of $1x10E^{8}$ meaning that out of 100,000,000 bits transmitted, one bit was in error. BER is also an indication of how often a data or other packet has to be retransmitted because of an error. Several retries are typically performed by the drive read/write logic until a soft error threshold is exceeded; then it is logged as a permanent I/O error.

Published Values for Bit Error Rate (BER)

Tape drives (Midrange and Enterprise)		
Midrange (typical LTO Ultrium drive)	Hard Error Rate	1 x $10E^{17}$ bits
Enterprise (typical mainframe drive)	Hard Error Rate	1 x $10E^{19}$ bits

Disk (FC, SAS, SATA)		
Enterprise FC/SAS	Hard Read Errors per Bits Read	1 sector per 1 x $10E^{16}$ bits
Enterprise SATA	Hard Read Errors per Bits Read	1 sector per 1 x $10E^{15}$ bits
Desktop SATA	Hard Read Errors per Bits Read	1 sector per 1 x $10E^{14}$ bits

In addition, today's tape technologies like LTO Ultrium drives perform a read after write verification process. That is, a set of read elements read the data immediately after it is written to tape to verify that the data was written accurately. LTO Ultrium tape also includes advanced servo tracking mechanisms to help provide precision tracking of the read/write heads with the tape.

Comparisons between disk and tape reliability are often debated and much of the debate is the result of past experiences with aging tape technologies, however things have changed. In addition, today's tape drives have a useful life at least twice as long as disk drives. With the help of backward read-compatibility, current tape drives typically read data from a cartridge in its own generation and the two prior generations enabling tape drive upgrades to occur asynchronously with minimal disruption. BER is becoming an increasingly important

consideration when using data reduction features such as encryption, compression, or de-duplication as the loss of a single bit in any of these scenarios will make reconstruction of the data virtually impossible yielding unusable data and the reconstruction process will likely fail.

> ***Bottom line:*** *The belief that tape is less reliable than disk has been dispelled with the latest tape technology advancements, long media life, MTBF ratings, and BER surpassing that of disk. Tape is the premier choice for long-term data preservation and protection.*

Get Greener with Tape Storage

As the cost of electrical power and the amount of tier 3 data continues to increase, the favorable impact of using tape storage on the IT budget will continue to grow. Unlike storage providers, energy providers haven't shown much interest in lowering their rates as average data-center energy costs are growing at 10-20% per year or more. Strategies to reduce energy consumption are needed to address energy expense especially since frequent rate increases can offset reductions in data center consumption.

Compounding the energy dynamic is the fact that power density is going up for most IT equipment at a rate of 20% or more per year. This has the following domino effects; 1) more power needs to be supplied to each square foot of a data center 2) more power is required to cool hotter equipment and 3) more heat extraction equipment is needed to support each square foot of a data center. The limits of utility company power distribution in many data centers is being approached, forcing organizations to begin exploring new cooling techniques such as water-cooled racks, outdoor cooling or in some cases, building another data center. Building another data center is normally a last resort and is extremely expensive mandating that energy consumption be properly managed. Average IT electrical consumption percentages for support infrastructure and IT hardware are summarized in the chart below. Tapes have compelling low power and cooling requirements and can be used to increase the amount of data stored without increasing environmental requirements.

Where Does the Energy Go?

Average Electrical Power Usage for Data Centers	
Chillers, cooling, pumps, air-conditioning	24%
Uninterruptible power supply	8%
Air movement, circulation, fans etc.	10%
Misc. lighting, security, perimeter surveillance	3%
Total support infrastructure – external consumption	**45%**
Servers	30%
Disk drives, control units	12%
Tape drives, robotic libraries	3%
Network gear, SAN switches and other devices…	10%
Total IT hardware – internal consumption	**55%**

Source: Horison, Inc. and estimates/averages from various industry sources.

Average electrical consumption within a data center (not including external HVAC) indicates that tape typically uses about one-fourth of the electricity that disk does. In addition, heat is the enemy of all IT technology regarding reliability and maintaining recommended environmental requirements is a critical factor. Utility companies in certain areas are restricting the amount of power some businesses can use.

> ***Bottom line:*** *Data stored on tape will require less energy than data stored on disk. Tape has the lowest energy consumption per gigabyte of any tier enabling the elusive IT goal that "data that isn't accessed shouldn't consume energy" to become a reality.*

Tape TCO Improves Storage Economics

Storage managers identify security, high availability, data classification, virtualization, tiered storage, improved utilization, and managing big data archives as their more pressing storage priorities. Businesses are re-architecting their storage systems with a variety of new capabilities that improve storage efficiency by reducing the number of devices, reducing energy consumption, improving I/O performance, increasing utilization levels, and adding stronger security measures. Though the challenges for today's IT executives may seem endless, cost reduction is almost always at or near the top of everyone's list.

Several TCO studies have been published recently for backup and for archiving applications. Using hardware and software acquisition costs, management costs, energy, facilities, and personnel costs, the studies conclude that the TCO (Total Cost of Ownership) for disk is significantly higher than tape for archive by 10x or more. Analyst firm ESG (Enterprise Strategies Group) TCO study comparing an LTO-5 tape library system with a SATA disk VTL system using de-duplication shows the disk VTL system having a 2-4x higher TCO than the tape system for backup in several scenarios. The TCO advantage for tape is expected to increase for the foreseeable future. [1]

In addition to significantly lower costs, the following points will enable storage managers to better identify and plan for the role tape can play in solving many data center storage problems.

- Cartridge tape capacities and data rates will continue on an unprecedented growth pace.
- The average price per gigabyte for automated tape library storage is expected to remain far below that of magnetic disk storage for the foreseeable future.
- The TCO for tape systems should remain significantly below that of online disk storage as people, facilities and energy costs rise and make tape even more appealing.
  Tape bit error rates have improved dramatically having surpassed disk. These reliability improvements favor tape for critical applications especially those requiring encryption

[1] A Comparative TCO Study: VTLs and Physical Tape, By Mark Peters, ESG, Feb., 2011

and compression as the loss of a single bit can render disk data unrecoverable.

- Tape media life has improved to 30 years or more favoring it for Tier 3 archive applications over other technologies.
- The long-life media eliminates the need to frequently migrate data to new media further reducing labor intensive conversion costs.
- Moving stagnant or unchanging data from the backup pool to a tape archive can reduce the size of the backup pool thus reducing backup time and costs.

Bottom line: *The many developments for tape including higher capacities, low energy costs, significantly improved reliability and a TCO far lower than disk have made tape the most cost-effective archive and backup solution.*

What Does the Future Hold for Tape Technology?

Tape usage is shifting as disk slowly encroaches on tape's traditional backup/recovery market while tape is positioning itself for the exploding tier 3 demand drivers such as fixed content, compliance, archive and potentially cloud storage archiving. In addition, tape as secondary backup is normally employed to provide off-line protection for critical backup data that could be at risk with an "on-line only" backup solution. For the foreseeable future, the primary tape market segments encompass all mainframe environments and the midrange and SMB (Small to Medium Business) markets. These usage and market segments will benefit from the significant tape enhancements that lie ahead.

A paper comparing areal density titled Tape Based Magnetic Recording: Technology Landscape Comparisons with Hard Disk Drive and Flash Roadmaps from the IBM Systems Technology Group at the Almaden Research Center in San Jose, CA, describes areal density growth rate scenarios for three SCM (Storage Class Memory) technologies, tape, HDD and NAND Solid State FLASH Drives (SSD). These scenarios 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 HDD and NAND annual growth rates will not maintain their traditional 40% values and will rather slow to 20% values. Specifically, the tape bit cell is 300-X to 500-X larger than the HDD and NAND bit cells and is scalable to smaller areas without being impacted by nano-technology issues related to lithography and bit stability. The net result of these areal density scenarios is a sustained volumetric and total capacity storage advantages for tape technology over HDD and NAND technologies. This also signals a lower total cost per gigabyte for tape than other technologies for the foreseeable future.

Another 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 tape developed by Fujifilm Corporation at a density of 29.5 billion bits per square inch that could produce a native

capacity tape cartridge of 35 TB. Several new tape technologies were developed as a result 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.

Why Tape is Important in the Storage Hierarchy?

Tier 3 Capability	Tape	Disk
Long-life media	✓ Yes, 30 years or more on all new media	~4-5 years for most HDDs before upgrade or replacement, 7-10 years for tape drives
Reliability	✓ Tape BER has surpassed disk since 2005	Disk BER still not improving as fast as tape
Portability	✓ Yes, media completely removable and easily transported	Disks are difficult to physically remove and to safely transport
Move data to remote location for DR with or without electricity	✓ Yes, can move remotely with or without electricity. Natural disasters can force physical media movement	Difficult to move disk data to remote location for DR without 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, except in the case of MAID or "spin-up, spin-down" disks
Hardware encryption for highest security level and performance	✓ Yes, encryption capability available on essentially all midrange and enterprise tape drives	Becoming available on selected disk products, PCs and personal appliances
Capacity growth rates	✓ Roadmaps favor tape over disk with 35 TB capability demonstrated	Continued steady capacity growth but roadmaps project disk to lag tape
TCO	✓ Favors tape for backup (4:1) and archive (15:1)	Higher TCO, more frequent conversions and upgrades

Source: Horison, Inc.

Bottom line: Tape densities will continue to grow and costs will decline, while disk drive capacity growth is leveling off. The future opportunities for tape storage solutions have grown considerably and are being fueled by significant advancements in tape technology.

Tape Gets Smarter

Tape innovation continues to march forward. Beginning with LTO Ultrium generation 3, WORM tape was introduced helping to address data retention and compliance needs. Tape drive encryption was added starting with LTO generation 4 to help provide additional security to protect sensitive data. LTO Ultrium generation 5 changes the rules for tape access with the ability to create two partitions on tape enabling a new file system called Linear Tape File System (LTFS). The LTFS specification was created by IBM and is available as an open standard for standalone tape drives as part of the LTO program. LTFS tapes have two partitions and are self-describing enabled by storing metadata information, a hierarchical directory structure, and fast search indexes in the first partition and file content in the second partition. The LTFS software allows the formatting and mounting of LTFS tapes, and access of the files directly on tape, including the ability to drag and drop files to and from the tape.

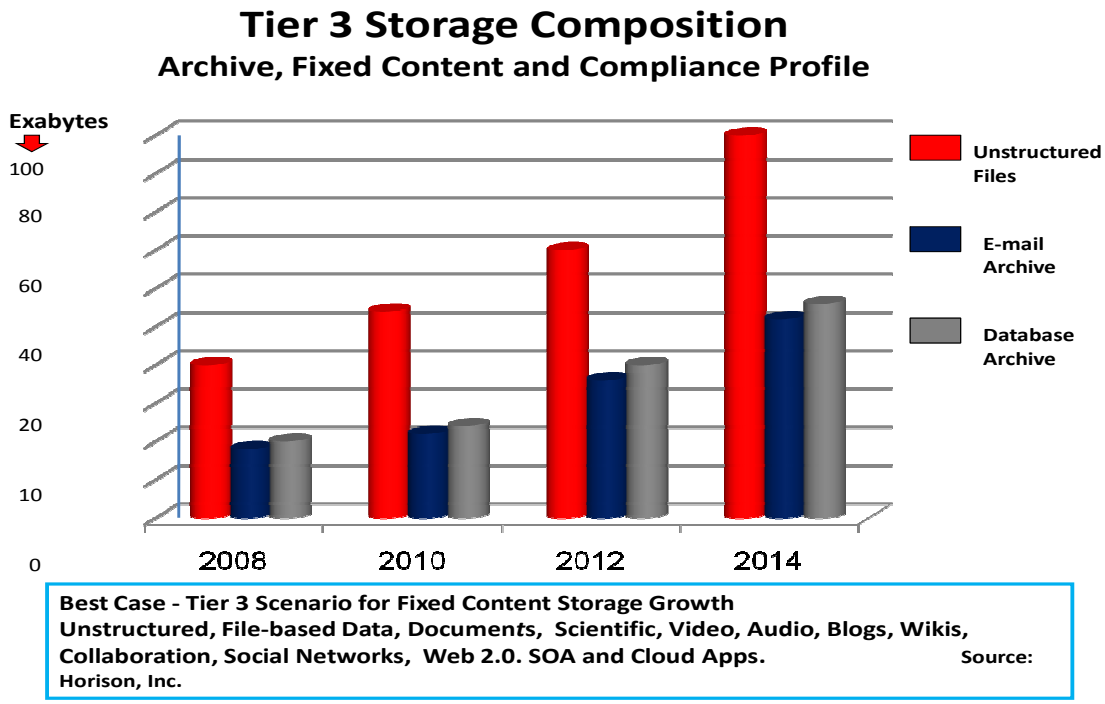
Applications, file browsers, image viewers and media players can directly browse and access files on tape. LTFS opens the door to a whole new world of possibilities and users are finding new applications as a result. Any file type can be stored to an LTFS formatted LTO-5 tape cartridge as LTO-5 can address long-term archive strategies by creating a self describing tape that has no application dependencies. LTFS can be readily utilized for media and entertainment video archive and workflow applications as well as in related rich media industries such as digital video surveillance, medical imaging, engineering drawings, and in the cloud. A variety of offerings based on the LTFS format are available from a number of vendors with more in the works every day. LTFS for automated tape library environments is also becoming available with the first offering recently announced by IBM called LTFS LE (Library Edition) helping to make LTFS an even more viable option for archival storage. The LTFS LE software allows the user to have direct access to all library files and directories stored in the LTFS Volumes.

> ***Bottom line:*** *LTFS allows tape to be used in a fashion like disk or other removable media making tape an even more effective archival data storage solution and future tape developments indicate this trend will continue for the foreseeable future. LTFS changes the rules for tape access!*

Massive Growth in Digital Data

The era of big data is here. Digital data is growing at over 40% annually and are now being generated by billions of people, not just by data centers as in the past, mandating the emergence of an ever smarter information infrastructure. Smart archiving is emerging as a key requirement as the need to migrate data to a repository that makes it easier to find while lowering costs is rapidly increasing. The growth in business analytics has significantly increased the value of information stored for long periods of time. While some backup applications are moving from tape to disk given faster restore times, tier 3 applications which include less active data, digital archives, fixed content, multi-media, broadcast, social networks, and compliance are being addressed by more cost effective tape storage. Much of this data is unstructured data and has no identifiable structure. Unstructured data typically includes bitmap images/objects,

text, photos, video and other data types that are not part of a database. Most enterprise data today can actually be considered unstructured– see below – making it ideally suited for the latest tape solutions. A growing list of government compliance and legal regulations worldwide now describe the way data should be managed, protected and how long data should be stored throughout its lifetime. For many applications and files, the lifetime for data preservation has become infinite, further increasing the security, size and management requirements of the digital archive.



Best Case - Tier 3 Scenario for Fixed Content Storage Growth
Unstructured, File-based Data, Documents, Scientific, Video, Audio, Blogs, Wikis, Collaboration, Social Networks, Web 2.0. SOA and Cloud Apps. **Source: Horison, Inc.**

Bottom line: *Archive, fixed content, entertainment, scientific, social networks, compliance and unstructured data requirements generate much of today's tier 3 storage demand and have become the primary drivers for future tape storage demand.*

Executive Summary

The general perception of tape is often outdated and today's reality is that the magnetic tape industry has made considerable progress in the past 10 years. Tape has improved its position as a valuable tier in the storage hierarchy and is complementary with disk to optimally address the many data storage objectives in the data center due to its lower total cost of ownership, improved reliability, data protection attributes, fast streaming performance 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 high growth tier 3 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 tier 3 archival opportunities that lie ahead, whether on-site or in the cloud. The tape renaissance is well underway as it readies itself to become the digital archive of the future. For tape, it's a new game with new rules, for users, tape is a valuable part of a data protection strategy. Disk and tape are and will continue to be complementary and when combined in a blended implementation, they can address the most critical storage objectives found in the data center.

End of report

This document was developed with IBM funding. Although the document may utilize publicly available material from various vendors, including IBM, it does not necessarily reflect the positions of such vendors on the issues addressed in this document.

About the author: In 1998, **Fred Moore** founded Horison Information Strategies, a data storage analyst and consulting firm in Boulder, CO. Prior to Horison, Fred began a 21 year career at StorageTek as the first systems engineer and concluded as corporate vice president of Strategy and Marketing. He is a sought-after IT speaker worldwide and has published numerous reports and publications for the data storage industry.
Email: fmoore@horison.com