Tape is Primed for the Rise of Secondary Storage

Delivering the Most Sustainable, Cost Effective, Highly Scalable Mass Storage Systems

TAPE STORAGE COUNCIL

September 2023

The Tape Storage Council, which includes representatives of BDT, Cozaint, FUJIFILM, GazillaByte, Grau Data, Hewlett Packard Enterprise, IBM, Imagine Products, Insurgo Media, Iron Mountain, MagStor, mLogic, Oracle, Overland-Tandberg, OWC Mercury, PoINT Software & Systems, QStar, Qualstar, Quantum, Spectra Logic, StorageDNA, SullivanStrickler, Symply, Turtle, and XenData has issued this report to highlight the current trends, usages and technology innovations occurring within the tape storage industry.

Introduction

As the widespread impact of Covid-19 eases, new concerns have emerged. These include geo-political tensions, global warming, a fragile supply chain, prolonged inflation, higher interest rates, and rising transportation, health care, food and energy prices. All of which have painted a choppy macro-economic picture. Not surprisingly, these trends have also caused storage demand and forecasts to fluctuate. Current growth projections of at least **25% annually** suggest a doubling of data stored every three years.

Zettabyte era storage demands are driving advanced solutions to effectively contain petascale and exascale requirements. In parallel the unquenchable thirst for IT services has placed energy usage, carbon emissions and the overall environmental impact of data centers squarely in the bull's eye for most organizations to address. Magnetic tape is currently the most economical and environmentally friendly storage solution available, but even with the new compelling advantages of tape, much of the world's lower activity data continues to reside by virtue of inertia and a lack of strategic planning, in the wrong place. Specifically, that's on more costly to operate HDDs (hard disk drives) which require constant power and cooling to maintain reliable operations.

What's Fueling the Shift to Secondary Storage?

Secondary storage is persistent storage designed to keep less critical and less active data on more economical, secure storage mediums that don't need to be accessed as frequently as data on primary storage. TRENDFOCUS projects installed storage capacity across all enterprise data centers is expected to reach ~8.4 ZBs by 2025 with at least 80% (~6.72 ZB) classified as archival or "cold" data making secondary storage by far the largest data storage market. By 2030, total installed capacity is projected to reach ~26.3 ZBs suggesting ~21.04 ZBs of secondary storage requirements. These projections indicate that most of the world's cold data would continue to reside on costly, energy consuming HDDs unless enterprises begin to take advantage of modern tape technology's benefits. Clearly the opportunity for data centers to strategically optimize their storage infrastructure with tape has never been greater.

TRENDFOCUS
Enterprise Storage Installed Capacity - Total Exabytes

Installed Storage Capacity	2025	2030
eSSD Total EB	847	1,956
Captive eSSD Total EB	151	565
Hyperscale Nearline Total EB	4,584	18,181
OEM/Other Nearline HDD Total EB	2,120	4,016
Performance Enterprise HDD Total EB	106	62
Tape Total EB	588	1,517
Total Ent EB Installed	8,396	26,296

TRENDFOCUS, May, 2023

Big Data Trends – 2025 and Beyond

Big Data is having a major impact on all aspects of secondary storage and refers to data sets so large or complex that most traditional data management tools can't store or process them efficiently. Big data applications generate data faster than it can be analyzed, significantly extending data retention timeframes and filling data lakes which ignore any size limits.

Big Data, AI, ML Fuel Shift to Secondary Storage

Key Trends - 2025 and Beyond

Secondary (persistent storage) is designed to keep less critical data on highly economical, secure storage mediums that doesn't need to be accessed as frequently as data in primary storage.

- **~11.5 ZB Stored on SSD, HDD and Tape by 2025 (cagr. 25-30%)**
- **~80% of All Data Stored is Archival (Seldom Accessed, Cold)**
- **Over 80% of Data Created is Unstructured**
- **AI, ML, Big Data Analytics, Edge/IoT Computing, Virtual Reality, Augmented Reality, Gaming and Robotics are Filling the Archives**
- **The Archival Copy is Usually the Only Copy of Data**
- **Retention Periods Over 100 Years are Common**
- **The Active Archive Becomes a De-facto Standard Tier**
- **Tape Becomes Optimal Target for Big Data Repositories**
- **Will Any New Technologies Emerge?**

Source: Horison Information Strategies

It is estimated that ~97% of businesses are investing in AI and ML technology to mine Big Data to gain competitive advantages, deeper customer insights and are the leading sources of secondary storage requirements. Several leading Big Data applications listed below are thriving in the zettabyte era. More intelligent filtering and data reduction techniques are being deployed at the edge for IoT created data, which is sorted, analyzed, and trimmed improving business efficiency as well as reducing unnecessary storage and bandwidth costs. Two applications are highlighted below, video surveillance and social media, which are retaining data for much longer timeframes. Data from both applications typically can reach archival status in a matter of days. Both are redefining their historical data lifecycle and aging profiles.

LEADING BIG DATA APPLICATIONS	
Cloud Storage	Projected to grow from $83.4 B in 2022 to $376.3 B by 2029
Edge Computing storage	Reached $44.7 B in 2022 growing to $101.3 B by 2027
HPC Storage	Projected to grow from $43.6 B in 2023 to $54 B in 2026
Health Care & Life Services Analytics	The life science analytics market size was valued at $9.3 B in 2022 and it is expected to reach around $24.1 B by 2032
Media & Entertainment	The media and entertainment storage market was valued at $6.6 B in 2019 and is expected to grow to $17.8 B by 2027
Sports Archive and Analytics	The global Sports Analytics Market was worth approximately $2.5 billion in 2021 and projects a revenue of $8.4 billion by the end of 2026
Social Media	The global social media market grew from $193.5 billion in 2022 to a projected $434.8 billion in 2027
Video Surveillance	The Video Surveillance storage market is expected to grow from $7.5 B in 2020 to $10.2 B by 2025.

A Closer Look at Video Surveillance Storage Demand

With a rapidly increasing threat landscape, video surveillance is becoming even more important than it was just five years ago. As a result the amount of video surveillance content being stored is accelerating. To deter crime, organizations are installing cameras everywhere from on-premise to the edge having much higher resolution and frame rates, with much longer retention periods. HDDs have historically been the primary storage for surveillance data however increasing retention periods have made storing archival surveillance data on HDDs a costly choice. Coupled with AI, the value of historical video surveillance content is quickly increasing, demanding even longer retention times. Estimates suggest the video surveillance market now consumes ~8% of all HDD capacity shipments. Though most data will reach archival status in 90-180 days, surveillance data can become archival less than a week after creation. Leveraging modern tape coupled with HDDs can reduce cost, power and carbon footprint enabling longer, more secure data retention periods. This scenario is another classic example of an active archive implementation.

The Rise of Social Media Drives Data Creation, Transmission and Storage Demand

As of January 2023, there were 4.76 billion social media users and 5.07 billion internet users worldwide, almost 63.5% of the world's population. The widespread reach of mobile technology has increased the number of Internet users everywhere. The Internet is the superhighway for social media, and these applications currently generate about 5.1% of the entire world's internet traffic volume (bandwidth = PBs transferred). Social media has enabled people to expand their network, add friends, enhance careers, make connections, and find people with scarce skills. The blinding speed that social media platforms can reach 100 million users attests to an insatiable appetite to engage with today's digital universe. ChatGPT, the popular artificial intelligence based chatbot, recently became the fastest social media platform ever to reach 100 million users, just two months after launching and two years two months faster than Instagram.

Months to Reach 100 million Users	
ChatGPT	2 months (AI)
Instagram	2 yrs, 4 months
WhatsApp	3 yrs, 4 months
Snapchat	3 yrs, 9 months
YouTube	4 yrs, 1 month
Facebook	4 yrs, 7 months
Twitter	5 yrs, 2 months
Pinterest	5 yrs, 6 months
LinkedIn	7 yrs, 11 months

Over the past 12 months, the number of active social media users increased by 190 million, growing 4.2% annually and averaging 6 new users joining every second. Social media users spend almost 2.5 hours per day creating, transmitting and storing data. Every "tweet", "swipe left or right", "post", "tag", "friend" or "like" creates some amount of stored data. Consider in just 60 seconds YouTube uploads 500 hours of videos, twitter users write 347,200 tweets, and 16 million texts are sent. That adds up to a lot of data. Though most data will reach archival status in 90-180 days, social media data can become archival less than a week after creation. Like surveillance data, understanding how to properly incorporate tape and HDDs makes this a classic use case for a highly scalable, cost-effective active archive implementation.

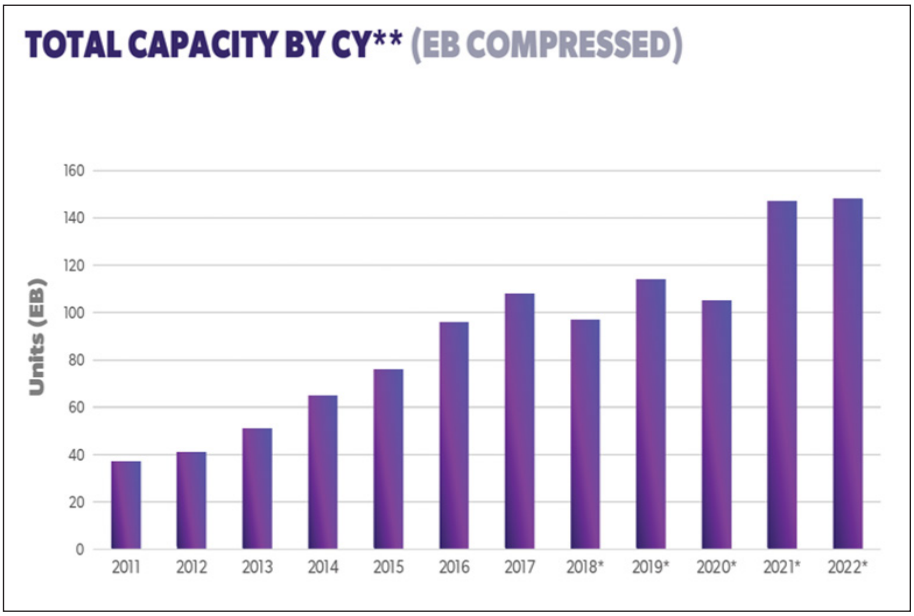
Cloud Repatriation Using On-Premises S3 Compatible Tape Systems Becomes Strategic

Cloud storage repatriation describes a shift away from the cloud and back to an on-prem storage infrastructure. Repatriation is a strategic initiative, gaining momentum and fueling the rise of tape based hybrid clouds. Many enterprises have experienced significantly higher costs associated with cloud storage outsourcing compared to on-premises storage. "On-prem" storage is usually managed by the organization, while the public cloud storage provider is responsible for the management and security of cloud data.

Fortunately enterprises can quickly set up their own private Glacier-like cloud system in their own data center using the familiar AWS S3 Glacier interface already used by many cloud applications. With S3, applications can write directly to tape creating a higher performance, on-prem glacier storage solution. For example, the beginning of data restoration from the AWS Glacier service can range from 4 to 48 hours depending on service pricing levels selected by the customer. With on-premises tape, data restoration can begin within minutes improving SLAs. The hybrid tape cloud provides organizations with much better TCO, higher security, availability and fault tolerance since data is stored both on-premises and in a cloud storage infrastructure increasing the effectiveness of the popular 3-2-1-1 data protection strategy. Leveraging tape significantly enhances the value proposition for cloud repatriation.

LTO Tape Capacity Shipments Reach New Record In 2022

The LTO Program Technology Provider Companies (TPCs), Hewlett Packard Enterprise Company, IBM Corporation and Quantum Corporation, released their annual tape media shipment report, detailing year-over-year shipments through the fourth quarter of 2022. LTO tape capacity shipments outperformed alternative storage technologies that experienced capacity shipment declines. The report reveals 148.3 Exabytes (EB) of total tape capacity (compressed) shipped in 2022. This result was driven by continued hyperscaler and enterprise investment in LTO technology. The strong performance in shipments continues the growth trend following the previous record-breaking 114 EB capacity shipped in 2019, 105 EB of capacity shipped in the pandemic affected year of 2020 and 148 EB shipped in 2021.



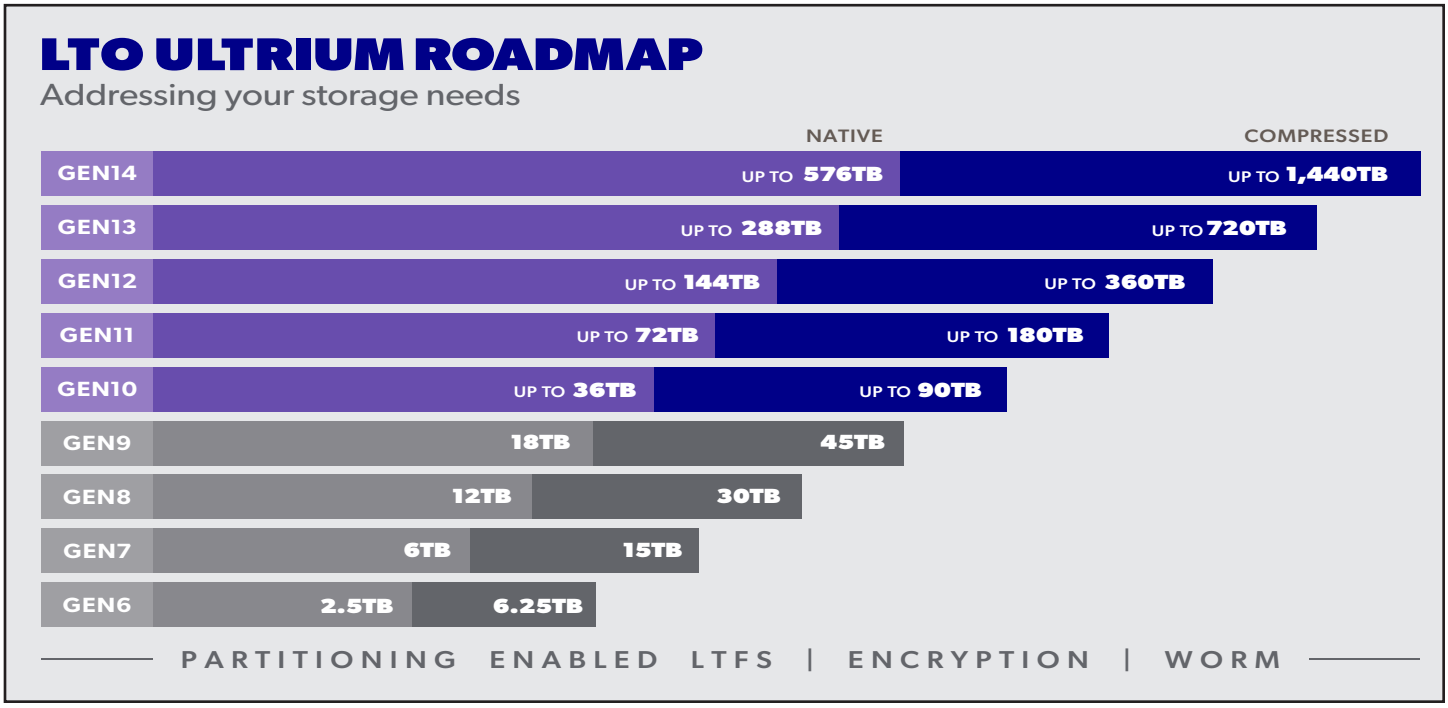
Strontium Ferrite Demo Supports New LTO Roadmap

In Dec. 2020, IBM and Fujifilm demonstrated a record areal density of 318 Gb/in2 on linear tape yielding a native cartridge capacity of 580 TBs (1.45 PB compressed 2.5x) using a new magnetic particle called Strontium Ferrite (SrFe). SrFe has improved magnetic characteristics compared to Barium Ferrite (BaFe). Like BaFe, SrFe is a chemically stable oxide using smaller nano-particles ideally suited for long-term data preservation. Fujifilm is paving the path to 1.0 PB and higher native tape cartridges by developing an advanced technology based on Epsilon Ferrite (-Fe2O3), with even smaller nanoparticles. Scaling tape areal density for the next decade and beyond signals a sustained volumetric and native capacity advantage for tape technology.

LTO Ecosystem Extends Roadmap to Generation 14 to Address Secondary Storage Growth

On Sept. 6, 2022, the LTO Program Technology Provider Companies (TPCs), Hewlett Packard Enterprise Company, International Business Machines Corporation, and Quantum Corporation, announced an updated LTO technology roadmap that extends the LTO Ultrium standard through 14 generations. The roadmap calls for tape capacities to double with each new generation, with LTO-14 delivering up to 1,440 TB (1.44 PB compressed) per tape. Once introduced, LTO-14 could surpass current LTO-9 cartridge capacity by 32 times. Given the rise of exascale secondary storage requirements, the new LTO roadmap extension is more relevant than ever as low-cost and sustainable long-term storage with the added benefit of strengthening cybersecurity. At this point no other storage technologies have revealed a comparable multi-generational, long-term roadmap.

(see LTO Ultrium roadmap below)



LTO-9 Delivers Higher Capacity, Throughput, and Improved Access Times

LTO-9 is the latest generation and has delivered new functionality to tape including higher capacity, data rate, access time and reliability improvements. Containing 3,396' of magnetic tape media, LTO-9 increased the native cartridge capacity of LTO-8 by 50% to 18 TB (45 TB compressed) and increased drive throughput (11%) up to 400 MB/sec enabling a single LTO-9 drive to write up to 1.44 TB/hour. A new feature for the LTO family with LTO-9, oRAO (Open Recommended Access Order) reduces initial file access times to first byte of data. New R/W head and servo technologies that record even narrower data tracks increasing cartridge capacity have also been developed. LTO-9 includes previously introduced features such as multilayer security support via hardware-based encryption. LTO-9 includes full backward read and write compatibility with LTO-8. A native 18 TB LTO- 9 cartridge can hold about

2,340,000 digital photos, 21 months of data from the Hubble Space Telescope, 720 Blu-ray discs, 9,000 hours of movies or approximately 6,480,000 songs. The cartridge capacity increase is especially welcome for hyperscale, cloud, HPC, and large enterprise data centers who have massive secondary storage requirements.

LTO-9 Provides the Highest Storage Device Reliability, with Media Life Reaching 50 Years

Since the introduction of LTO-1 in 2000 with a native capacity of 100 GB, the capacity of LTO cartridges has increased by 180 times and data rates have increased by 20 times. Over the same period, the specified uncorrectable Bit Error Rate (BER) of LTO cartridges has improved by a factor of 1000, three orders of magnitude improvement. LTO-9 provides an industry leading uncorrectable bit error rate of $1x10^{20}$ compared to the highest HDD BER at $1x10^{17}$. A BER of $1x10^{20}$ corresponds to one unrecoverable read error event for every 12.5 exabytes of data read. A tape drive may fail, but if the data on the cartridge can still be read back with another drive, then retrieval is successful. Today, both the latest LTO and enterprise tape products are more reliable than any HDD.

In 2019, Fujifilm and JEITA (Japan Electronics and Information Technology Industries Association) officially confirmed the longevity of Barium Ferrite magnetic signal strength to be stable for at least 50 years based on studies of LTO-7 tapes. Prior to this confirmation, the number of years for LTO tape longevity had been rated at 30 years. HDDs typically offer a five-year warranty and, in most cases, will be replaced at that time. While modern tape media can last 50 years or more, tape drives are replaced approximately every seven to ten years driven by the technological, economic, and operational benefits of newer drives and media. Knowing the data stored on tape can be relied upon for several decades gives confidence in the integrity of the archives. Extended media life effectively reduces the frequency of media remastering and conversions for tape systems.

IBM and FUJIFILM Announce TS1170 Hi-Density Tape System

In August of 2023, Fujifilm and IBM announced a new ultra-high-density tape drive with a native storage capacity of 50TB in a single cartridge and capacities up to 150 TB per cartridge with 3:1 compression. The IBM TS1170 storage system represents the world's highest cartridge capacity ever announced and enables data intensive secondary storage applications like big data, archiving, cloud computing, AI, and analytics to significantly reduce the total cost of ownership.

Fujifilm has succeeded in achieving this advanced cartridge capacity by evolving the technologies developed in previous tape generations. The improved areal recording density (the amount of data that can be recorded per square inch) and the overall recording area using a 15% longer tape cartridge enable the higher cartridge capacity. Fine hybrid magnetic nanoparticles have been developed by combining the technologies used in the next-generation Strontium Ferrite (SrFe) magnetic particles and the Barium Ferrite (BaFe) particles that are currently used in high-capacity data storage tapes. Reduction in the size of the magnetic particles and enhancement in their magnetic properties significantly improves the areal recording density.

With a 250% increase in capacity as compared to the previous IBM TS1160 tape drive, the IBM TS1170 tape drives consist of two new models: the TS1170 Model 70F with a dual-port 16 Gb Fibre Channel interface, and the TS1170 Model 70S with a dual-port 12 Gb SAS interface. Other features include RAO (Recommended Access Order) which improves recall time and time to first byte and the IBM Storage Archive (LTFS format) for direct, intuitive, and graphical access to data. Investments in tape library automation are protected by offering compatibility with existing tape library solutions. The TS1170 announcement demonstrates IBM's commitment to continue the evolution of tape drives.

Data Center Sustainability Benefits From Tape

As the world faces a growing climate crisis, many data centers are establishing aggressive sustainability projects to reduce energy consumption and CO2 emissions. By 2027, 75% of organizations are predicted to have a data center sustainability program in place, according to Gartner research. Analysts expect that leap (from less than 5% in 2022) will be fueled by a combination of stakeholder pressures and the desire to optimize costs. Data centers are estimated to be responsible for up to 3% of global electricity consumption today and are projected to reach 4% by 2030. Servers and HDDs are the two largest consumers of data center energy. Tape cartridges containing low activity data spend most of their lifetime residing in library slots or offline racks consuming energy only when they are mounted in a tape drive.

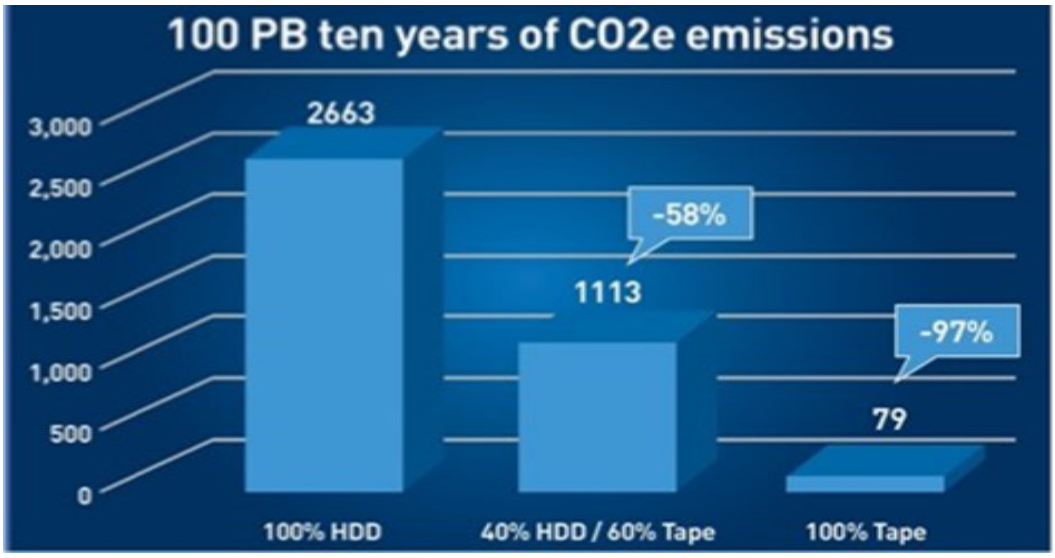
The average hyperscale facility consumes 20-50MW annually, theoretically enough electricity to power up to ~37,000

homes. Hyperscale data centers store over 50% of all enterprise data (see TRENDFOCUS chart on p1) and are the largest generators of CO2 emissions and carbon footprints. However as aggregators of an enormous compute and storage footprint, hyperscale data centers are the most efficient of all data centers at reducing energy and carbon footprint and have the best data center power usage effectiveness (PUE) ratings. While estimates vary, some research estimates suggest the entire IT industry could use as much as 20% of all electricity produced and emit up to 5.5% of the world's carbon emissions. The overall environmental impact from reducing data center carbon emissions and eWaste is enormous.

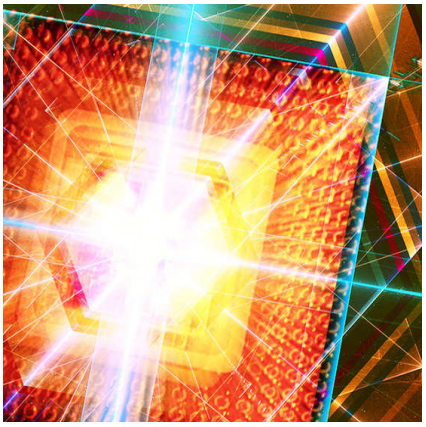
Tape Reduces Carbon Emissions and eWaste

Improving Information Technology Sustainability with Modern Tape Storage, a research paper issued by Brad John's Consulting, compared an all data on HDD solution to an all tape solution and to an active archive that moved 60% of the HDD resident (low activity) data to tape. Moving 60% of HDD data to tape for 10 years reduced carbon emissions by 58% and electronic waste was reduced by 53%. For eWaste (electronic waste), storing all 100 PBs on HDDs and refreshing them after 5 years generates 8.2 tons of eWaste, while storing 60% of the data on tape created 4 tons of eWaste, a 51% reduction. The research paper analyzes the CO2e generated due to energy consumption and emissions associated with the acquisition of raw materials, manufacturing and the final disposal of storage media.

According to the report, tape produces 97% less CO2e compared to an all-HDD storage strategy!
With insatiable secondary storage demand ahead, expect tape to play a pivotal role in data center sustainability initiatives as moving low activity data from HDDs to tape pays enormous dividends. Storing low activity and archival data on spinning HDDs is a strategy, just not a cost-effective or eco-friendly one.



Will A New Secondary Storage Tier Emerge?



Momentum continues to build for a new secondary storage tier sometimes referred to as the deep archive tier. Any new technology must be easily scalable and require minimal remastering (long media life) cycles and can be the home for the "golden copy", an immutable master copy of the most critical data. The optimal design will have a robust roadmap, minimal environmental/carbon footprint, realistic access times and a price point below all other competitive offerings. The latest LTO roadmap has relatively few areal density limitations and is currently the best positioned technology to address a new deep archive tier.

Other large-scale, potential deep archive storage technologies remain under various stages of development including, photonics, holographic, glass media, ceramics and DNA. Some of these technologies have been under development for decades and any timeframe for commercial availability remains unknown.

Tape Air Gap Data Provides Cybercrime Protection Capability

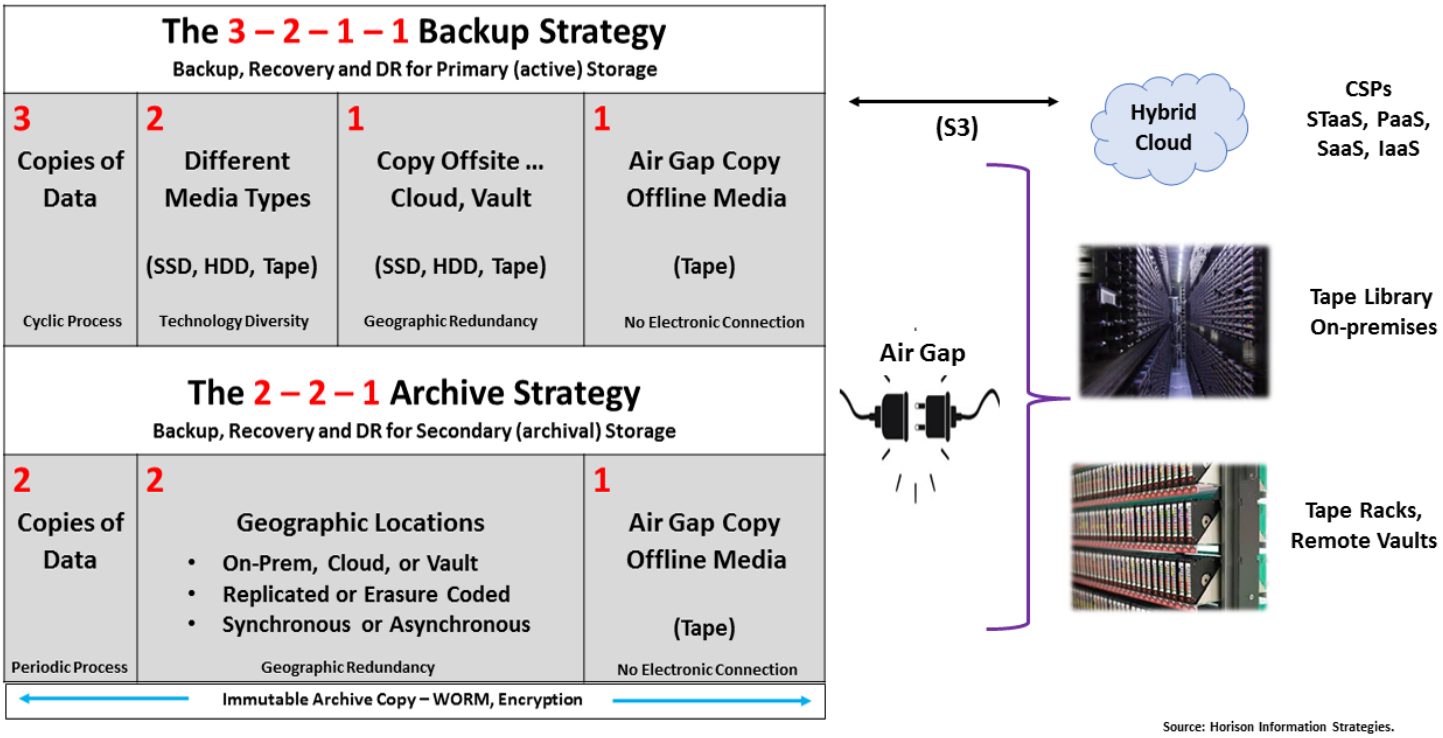
Tape is the only true air-gapped data center storage technology. The “tape air gap” means there is no electronic connection to data stored on a removeable tape cartridge without robotic or human intervention preventing a direct malware attack (hacking) of data on tape. HDD and SSD systems are connected and online 7x24x365 are always vulnerable to a cybercrime attack making tape the only viable data center air-gapped storage solution currently available, a key component of any data protection strategy. Cartridges located in robotic tape library slots or racks are always protected by the air gap.

Backup was the original data protection strategy, but having one backup copy is no longer sufficient. The popular and genetically diverse 3-2-1-1 Backup Strategy states that enterprises should have three copies of backup data on two different media types, one copy offsite and one air gap copy. Combining the air gap copy with encryption and WORM, tape strengthens any data center cyber resiliency strategy. Recovery time for backup data on tape has been significantly improved with features such as oRAO with LTO-9 which can reduce the time needed for file access to recover backup files and archive data from tape by as much as 73%. Remember – backup is important but recovery is everything.

For many data centers, the archive copy is often the only copy of archival data exposing it in case of any data loss event. Since the business value of untapped archival data is increasing, creating a second, secure air-gapped copy in a different geographic location will become a strategic data protection strategy. For higher availability, the archive copy can be protected by implementing a 2-2-1 Archive Strategy creating a second copy of archival data at a different physical location.

Data Protection Strategies

Tape Plays Key Role to Protect *All* Data In The Security Ecosystem

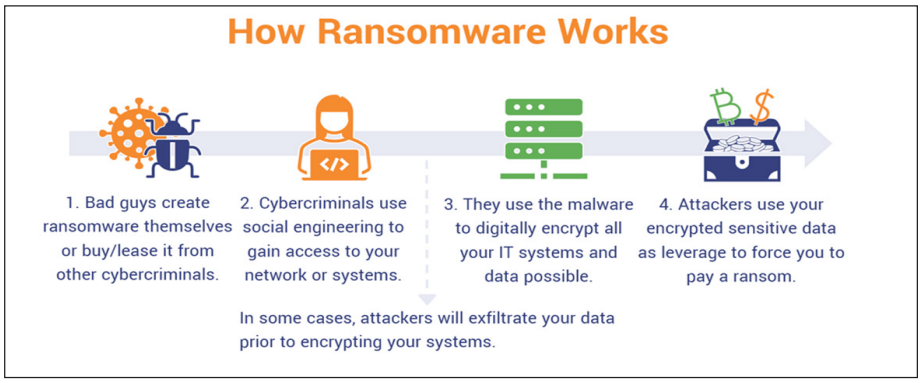


Source: Horison Information Strategies.

Several tape library vendors are adding additional air gap protection capabilities by offering libraries with managed partitions consisting of dedicated slots which are invisible to external applications. These isolated partitions contain no drives creating a secure air-gapped storage location within the library providing an additional barrier to access. The partitions are solely configured by the tape library administrator who can create, modify, delete, or reconfigure partitions to meet any size required. Since partitioned tapes remain in the library, physical media handling is avoided. Modern tape provides immutable air gapping at a favorable price point at scale for backup and archive data and should be an integral part of any data protection strategy whether on-premises, in the cloud, or a hybrid cloud.

Ransomware Targets Backup Repositories

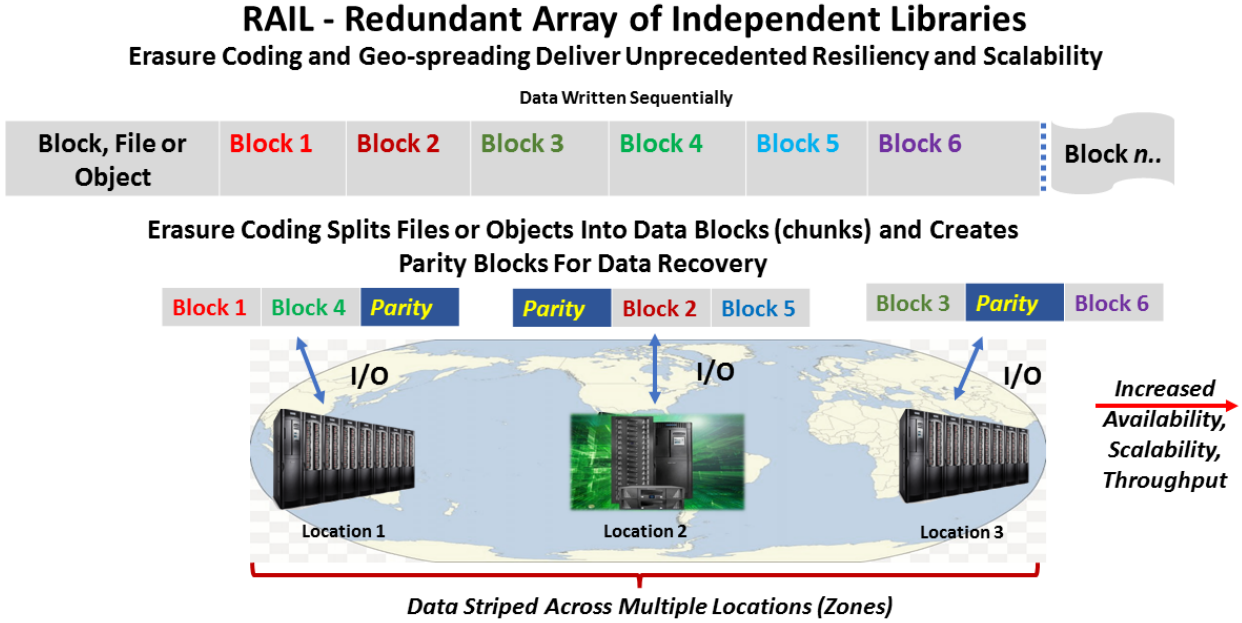
Ransomware is malicious software preventing access to computer files, systems, or networks and demands you pay a ransom for their return. The average cost of ransomware in 2022 was $4.54 million, slightly higher than a data breach at $4.35 million while the total cost of cybercrime is projected to reach $10.5 trillion by 2025. According to the 2023 report from VEEAM, 93% of worldwide ransomware attacks attempted to infect backup repositories.



Encryption protects data if stolen, changed, or compromised and works by scrambling data into a secret code that can only be unlocked with a unique digital key. Cybercriminals cannot read or utilize the stolen encrypted data. Enterprises can build an even stronger cyber defense by implementing the 3-2-1-1 backup strategy with encrypted, air-gapped tape and off-site backups.

RAIL Provides Geographic Resiliency for Large-scale Secondary Storage Systems

Geo-spreading is the distribution of mission-critical IT components or data across multiple data centers that reside in different geographic locations. Geo-spreading acts as a safety net in case the primary site fails, in the event of an outage or a disaster that impacts an entire region. RAIL (Redundant Arrays of Independent Libraries) provides high availability for large-scale secondary storage and hyperscale requirements.



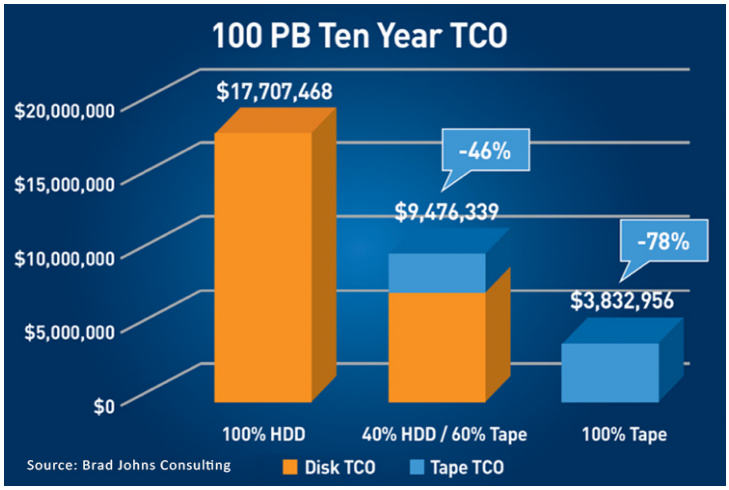
- **Erasure Coding Distributes Files/Objects Across Drives in Multiple Libraries Providing Parallel Access**
- **Erasure Codes Provide ~50% Space Saving Versus RAID, Ideal for Hyperscale and Large Cloud Object Archives**

Source: Horison Information Strategies

RAIL is a component of geo-spreading and like RAIT (Redundant Arrays of Independent Tape), stripes data across different tape cartridges - but in different physical libraries that can be in different geographic locations. RAIL uses erasure coding for higher availability to provide increased geographic resiliency in case of a device, full data center outage or natural disaster.

Tiered Storage Effectively Addresses the Secondary Storage Challenge

Effectively addressing the storage optimization challenge of "getting the right data, in the right place, at the right time and at the right cost" presents the biggest total returns for storage managers. To meet this challenge, the strengths of tiered storage combining intelligent software, SSDs, HDDs, and tape must be leveraged. As most data ages, access frequency drops off rapidly and data typically reaches archival status between 90 to 180 days eventually becoming cold data. Much archival data continues to live on HDDs long after it reaches archival status, an expensive residence for cold data. The TCO for an HDD solution is compared to a tape solution using LTO-9 tape media and drives. The estimated ten- year TCO for keeping all 100 PB on HDDs is $17,707,468. With the tiered approach moving 60% of HDD data to tape, the ten-year TCO drops to $9,476,339, a 46% savings. Moving all data to tape results in a 78% cost reduction.



The greatest economic benefits of tiered storage are realized when the tape tier is used. Intelligent data management software that moves data between tiers based on policies is a key component for an optimized tiered storage infrastructure. Look for AI to play a more significant role in data management decisions going forward. In reality, adding disk is tactical, adding tape is strategic.

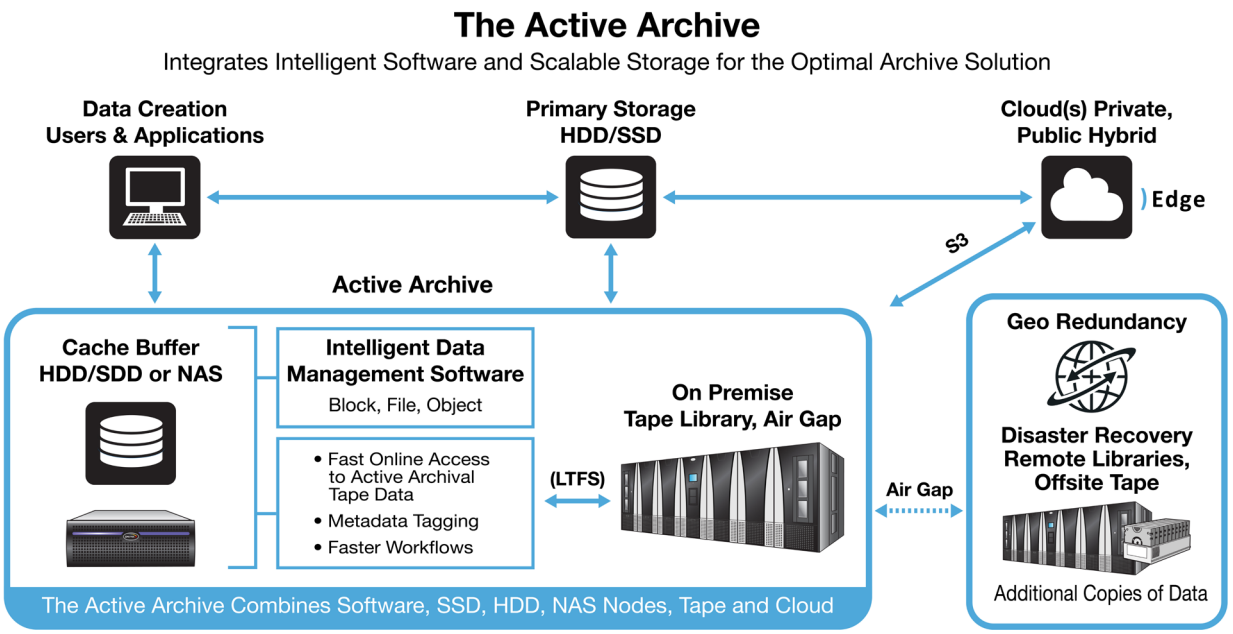
Market Dynamics Favorable for Flape (Flash plus Tape)

The concept of "flape", that is to say flash + tape, is not new, but is benefiting from current market dynamics as the cost of flash/SSDs comes closer to HDD and the value proposition of tape (low cost, low energy, capacity/scalability, long archival life, air gap, etc.) becomes more relevant. Flash and tape systems are complementary and organizations can eliminate the middle tier of costly, energy intensive HDD and simply rely on a flash layer for its outstanding performance gains, micro latency, macro efficiency and enterprise reliability. Tape is then used for best data streaming performance, unequaled cost-effectiveness, high capacity, scalability and low power consumption and footprint.

Leveraging the benefits of flash + tape was first identified by the analyst firm Wikibon in 2012. The concept behind the flape architecture is to place the most active data as well as the metadata on flash and the rest of the cold data on tape. A flape architecture helps meet the performance requirements of live, active data (best $/IOP) and the budget requirements for long-term data retention (best $/TB). Given the investments being made in flash/SSD, tape and HDD and the costs of each solution over time, a flape architecture makes sense, especially when looking at the manageability and environmentals (power, cooling, floor space) of the solution. Flape solutions can enable IT to turn unmanageable, infrequently accessed big data into a true business asset.

The Active Archive Provides Fast Access to Secondary Storage

Secondary storage receives a big performance boost from an active archive architecture. An active archive integrates two or more storage technologies (SSD, HDD, tape, and cloud storage) behind a file system providing a seamless means to manage archive data in a single virtualized storage pool. SSDs or HDDs serve as a cache buffer for archival data stored on secondary storage (tape) providing faster access to first byte of data, higher IOPs, and random access for more interactive archival data. Many data management products now support tape as an object storage target using the S3 interface. Combining the open tape file system LTFS with tape partitioning, data mover software (HSM, etc.), an HDD array or NAS in front of a tape library creates an active archive. Using an SSD for the cache creates an instant archive. LTFS currently has 34 implementers, and it is expected that an increasing number of ISVs (Independent Software Vendors) will exploit LTFS in the future. The active archive is supported by the Active Archive Alliance. See the Active Archive conceptual view below.



Tape Performance Improves – Faster Access Times and Throughput

HDDs and SSDs have much faster access times than tape to the first byte of data. For large files, tape systems have faster access times to the last byte of data. Previously available for IBM enterprise tape drives, oRAO (Open Recommended Access Order) was made available with LTO-9 full-high tape drives. oRAO reduces initial file access time serving as a data retrieval accelerator enabling applications to retrieve non-consecutive (random) files from tape by optimizing physical seek times between files. oRAO can improve random access time to data segments on tape by as much as 73%, when compared with retrievals of the same data segments linearly, while significantly reducing physical tape movement and drive wear by creating an optimally ordered list of files on a cartridge. In addition, LTO-9 increases drive data rate to 400 MB/sec. making it ideal for data streaming and large file transfers. RAIT serves as a data rate multiplier enabling parallel data transfer from an array of tape drives.

Tape Access Time Improvements (Time to 1st Byte)	
Active Archive	Active archive greatly improves access time to file and object tape data by using high-capacity HDDs or SSDs as a cache buffer in conjunction with a tape library.
oRAO	oRAO (Open Recommended Access Order) arrived with LTO-9 and produces an optimized list called “best access order” enabling applications to retrieve non-consecutive tape files by minimizing physical seek times (~73%) between files.
LTFS	LTFS provides access to files directly without the application that wrote the data.
Faster and Smarter Library Robotics	Intelligent robotics optimize robotic movements reducing mount and access times while improving reliability. Ransomware-free partitions boost security.
Tape Throughput Improvements (Data Transfer Rate)	
Fastest Data Rates	The LTO-9 and TS1160 enterprise drives each have a data transfer rate of 400 MB/sec. This compares to the 7,200 RPM HDDs ranging between 160 – 260 MB/sec.
RAIT	RAIT (Redundant Arrays of Independent Tape) stripes data across multiple tape drives in parallel significantly increasing throughput and provides parity for data reconstruction like RAID does for HDDs.
RAIL	RAIL (Redundant Arrays of Independent Libraries) stripes data across tape cartridges but in different libraries which may be in different geographic locations.

LTO Delivers a Compelling Value Proposition for Secondary Storage

The current state of the tape industry is highlighted with continued development and investment in smart libraries, new drives, advanced media, and intelligent management software. The tape value proposition below is compelling and addresses the relentless demand for higher reliability, higher capacity, better power efficiency, ease of use and the lowest $/TB and TCO of any available storage solution.

Tape Function	Benefits Summary (Tape Re-enters Growth Phase)
Price/TCO	Tape Has the Lowest Acquisition Price $/TB, Lowest TCO.
Energy, CO2 Sustainability	Tape Uses Much Less Energy and Has Much Lower Carbon Footprint Than HDDs (~97% Lower).
Performance (Access time)	Much Improved Access Times - Active Archive, Fastest Data Rates, Smarter and Faster Robotics, RAIT, RAIL, New Time to 1st Byte Features (oRAO), Re-writable.
Capacity	LTO-9 Cartridge Capacity @18 TB (45 TB compressed) with 400 MB/sec Data Rate. Smart Zone, Exabyte+ Capacity Libraries are Available. Lab Demos Reach 580 TBs per Cartridge.
Scalability	Tape Easily Scales Capacity (PBs to EBs) by Adding Media/Racks Without Adding Energy Consumption, HDDs Scale Capacity by Adding Drives and Adding Energy Consumption.
Portability	Tape Media Easily Portable in Case of Disaster, HDDs More Difficult to Physically Move.
Cybersecurity	Air Gap, WORM and Encryption Options Protect Against Malware Attacks, Providing Immutability.
Durability/Media	LTO Reliability BER ($1x10^{20}$) Surpassed HDDs ($1x10^{16}$), Media Life >50 Years for all Modern Tape.
Recording Limits	HDDs Facing Areal Density and Performance (IOPs) Limits. Tape Has a Well-Defined Roadmap extending through LTO-14.
Open Standards	LTO and LTFS Provide Open Standard File Interface, APIs. SW (S3 API) Support for Tape Object Storage.
Tape and Cloud Ecosystem	Tape Interfaces Seamlessly with Clouds Using Industry Standard API's. Native Cloud Applications Can Write To and Read From Tape. Hot and Cold clouds.

CONCLUSION

Digital data creation continues to grow at 25% or more per year, at least 80% of the worlds digital data is lower activity data optimally suited for secondary storage. In response to this challenge, the tape ecosystem has significantly expanded its capabilities in recent years. Evolving from its original roles as a backup and then an archive target, modern tape supports numerous Big Data applications that have historically been anchored in the domain of costly HDDs. Tape has also become the leading secure, pure long-term storage solution to defend against cybercrime by seamlessly integrating air gap, encryption and WORM capabilities. Roadmaps signal that the trend of steady tape innovation will continue well into the future.

Tape is the greenest storage technology and can significantly reduce carbon emissions and eWaste from data center operations. Leading edge organizations determined to contain their infrastructure costs and improve sustainability metrics will be motivated to rethink existing data storage practices and take advantage of advanced magnetic tape. Combined with improved access times, faster data rates, a 50-year media life, the lowest TCO, the highest device reliability, and huge sustainability benefits, modern tape has the greatest potential to address the massive demands of the zettabyte era.

It's no coincidence that the rise of many new tape advancements corresponds with the rise in secondary storage demand. After carefully considering the alternatives for the foreseeable future - tape has clearly positioned itself as the primary choice for secondary storage.

For more information on ransomware-resilient, long-term data archiving and tape storage from the Tape Storage Council, visit: https://tapestorage.org

TAPE STORAGE COUNCIL

V1-091923