2020 Technology Update Series

Tape Performance *Accelerates*

Access Time and Throughput Take Off

HORISON Information Strategies

Fred Moore
President
Horison Information Strategies
Horison.com

Introduction

The tape industry is making significant performance strides by delivering much faster initial access times and throughput levels with the arrival of the Active Archive, RAO, TAOS, LTFS, much faster data rates, and RAIT capabilities. For all the amazing technological progress made for the traditional data centers, Hyperscale and cloud computing, the fundamental challenges of reliably transferring large files and bulk data volumes at high speeds to different geographic locations continue to be problematic. Moving large amounts of streaming data, archives, images, video/audio streams, tele-conferencing, tele-medicine, scientific data capture, and large-scale disaster recovery scenarios is performed much faster on high speed modern tape than other digital technologies. Cloud egress and ingress is slow using network bandwidth and can become cost prohibitive, taking days or even weeks compared to moving the same amount of data on removeable tape media via truck or airplane. Since tape media is readily portable, using tape for cloud storage can be highly advantageous if a CSP (Cloud Service Provider) shuts down or should you want to quickly move your entire digital archive media to another provider. These trends and scenarios present a growing mass transit problem for bulk data movement - if it weren't for the tape industry's renewed focus on performance and throughput. This report will examine several new performance capabilities for improving tape access and data transfer times.

The Anatomy of Tape Performance

HDDs are online and continually spinning having access times to first byte of data in the 5-10 ms range. By comparison, tape performance requires three steps before data transfer can begin increasing the time to the first byte of data: 1) robotic cartridge access 2) drive load time 3) drive/file access time 4) throughput/data transfer. Once data transfer begins, tape data rates are much faster than HDD rates.

Tape Performance Accelerates

New Access Time and Throughput Improvements

Robotics	Drive Load	Drive/File Access	Total Access Time To 1st Byte of File	Throughput/Data Transfer
Tape Library Mount Time	Tape Drive Load Time	File Access Time Locate the File	Σ Robotic+Load+Access Time	Sequential Data Transfer Rates LTO-8 360 MB/sec. TS1160 400 MB/sec. HDDs 160-220 MB/sec
4 - 10 secs	Up to 11 sec	10 - 100 secs	25 – 121 secs	

Access Time Improvements	
Active Archive (ms)	Provides HDD-like Cache Access Time to 1st Byte of Tape Files (cache hit ratio ~60-90%)
RAO – Recommended Access Order (Enterprise Tape)	Order Tape Requests to Optimize Tape Movement Time to 1st Byte. Reduces Drive and Media Wear.
TAOS - Time-based Access Order System (LTO)	Order Tape Requests to Optimize Tape Movement Time to 1st Byte. Reduces Drive and Media Wear.
LTFS	Partitioned Self-describing File System to Drag and Drop Files for faster Access
Faster Robotics	Sorting Move Commands and Optimizing Move Sequence-Based on Robot Location. Faster Robotic Movement, Multi-Media Support, Improved Reliability.

Data Rate and Throughput Improvements	
Data Rates *Today*	Tape Data Rates are 2-3x Faster Than HDDs.
Faster Data Rates *Tomorrow*	Projected to be ~5X Greater than HDD by 2025.
RAIT	Striping Multiplies Tape Drive Data Rates. Increases Availability With Fault-tolerance.

Robotics are transporters that move a tape cartridge from a library slot to the tape drive. A combination of faster hardware and intelligent library software features have reduced the time required (robotic movements) to locate a tape cartridge and place it in a tape drive while improving library reliability. Robotic mount times typically range from 4 – 10 seconds.

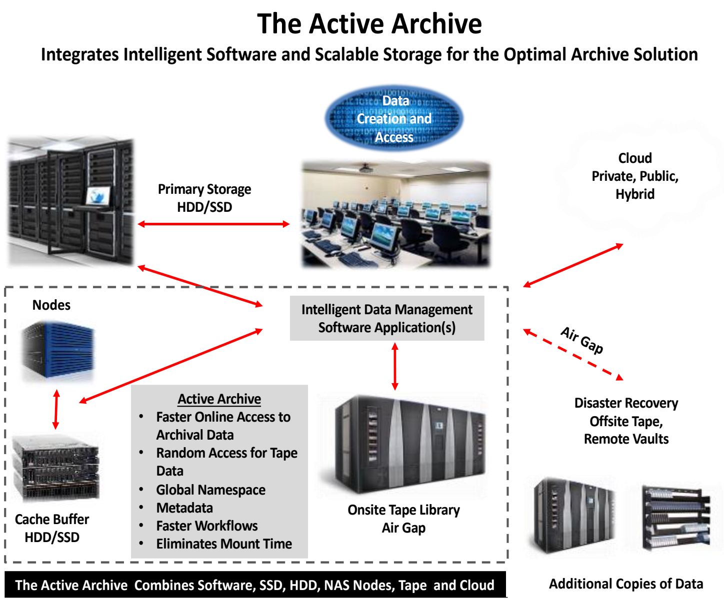
Drive (media) load time is the amount of time between cartridge insertion by the robotic arm and the drive becoming ready for host system commands**.** In all tape drive systems, the tape is pulled from the cartridge and guided through the tape path across the read-write heads to the appropriate file's load point where data placement begins. Drive load times typically take ~11 seconds.

Drive (file) access time is the time from when the drive receives a host-system command to read data and the time when the drive *begins* to read the data. File access times are typically expressed as averages (time to tape midpoint), since the requested file might be located in the beginning, middle or at the end of the tape. The typical file access (lodcate) time can range from 10 – 100 seconds.

Data Rate/Throughput is the speed at which data is written to tape from the drive's internal buffer and is usually measured in MB/sec. Tape drives de-compress data that is stored compressed before transfer to the server. Tape data rates are much faster than HDDs and RAIT multiplies tape data rates by striping data across multiple drives for transferring data in parallel.

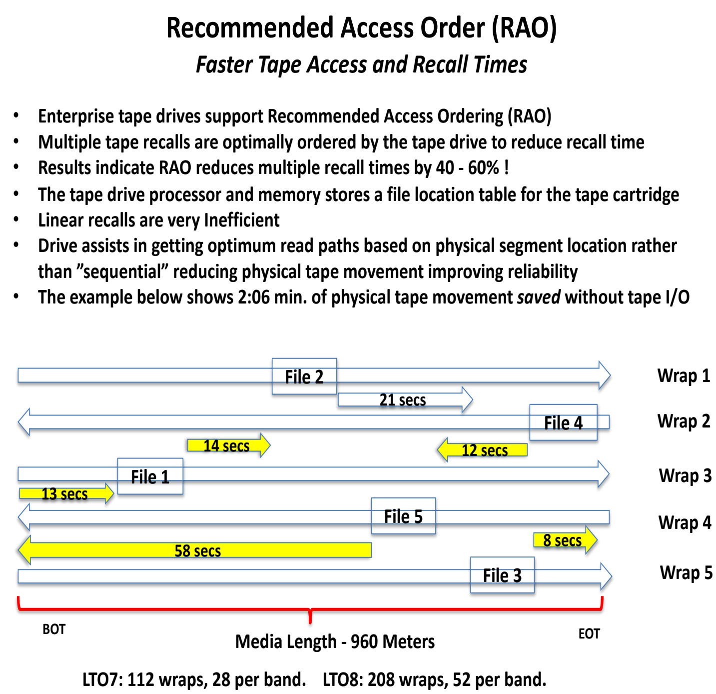
Active Archive

An Active Archive integrates SSD, HDD, tape, and cloud storage (public, private or hybrid) making it a special case of the popular tiered storage model that is dedicated to improving archive performance. The active archive greatly improves tape access time by using HDD or SSD as a cache buffer for a tape library enabling a high percentage of accesses to the tape subsystem to be satisfied online from SSD or HDDs (the cache hit ratio) significantly improving access time to first byte of data. Intelligent management software migrates data between tiers based on user policies and in some cases can create metadata and global namespaces to facilitate access to unstructured data. LTFS is often used as the standard open tape file system for an Active Archive.



RAO (Recommended Access Order) Reduces Tape File Access Times for Enterprise Tape

The RAO capability of HPSS (High Performance Storage System) is available on enterprise tape drives for improving tape file access and recall times. Modern tape drives write data to tape in a serpentine manner in different directions on alternating wraps. However, files are most often accessed (reading data) in random order. As tape capacities and the number files on a cartridge increase, file access times will increase as the probability of multiple file requests for the same cartridge rises. The RAO determination is performed by the drive producing an optimized list called "best access order" which determines the least amount of time needed to locate and read all concurrently requested tape files. RAO can improve the seek-time between files on tape by 40% to 60%, a major access time reduction. HPSS also performs offset ordered tape recalls (like RAO) for LTO. Offset ordered recalls improve the efficiency of recalling collocated files. Because of the serpentine nature of data recorded on tape, LTO tape offset ordering does not minimize seek-times for files that are not collocated.



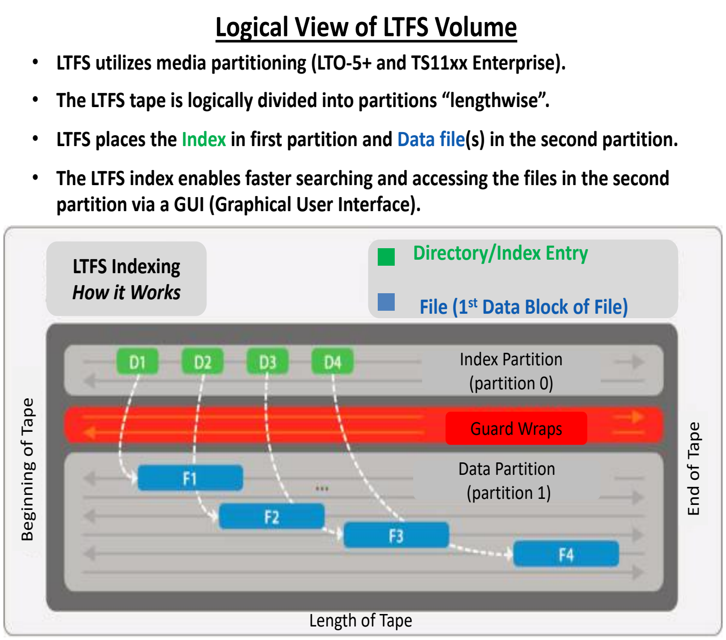
TAOS (Time-based Access Order System) Reduces Tape File Access Times for LTO

TAOS does for LTO what IBM's RAO feature does for enterprise tape drives. TAOS is a tape control unit feature from [Spectra] that provides up to a 4 times improvement in overall access time and up to a 13 times reduction in physical tape movement across the drive heads. Tape files are written on tape in sequential order but are most often accessed (reading data) in random order. LTO files are laid out in wraps that follow a serpentine pattern starting at the physical beginning of the tape, traversing all the way to the end of the tape, then back to the beginning (see chart above). Less tape movement reduces tape media and drive wear improving overall tape reliability and delivering faster performance.

The TAOS algorithm runs on the Spectra tape robotic library processor and orders read requests to minimize tape movement. TAOS uses a "Nearest Neighbor" algorithm to create an optimized best access order list. Retrieving non-consecutive files from an LTO tape 960 meters long can result in inefficient seek times between file reads. An optimized best access order list, similar to RAO, is generated based on the least amount of time needed to locate and read all concurrently requested files on a tape. From the list of files to recall, every file's start and end position is translated to a physical position on tape. An estimate in milliseconds between each file's end position and every other file's beginning locate position determines the optimal order. TAOS support is presently available for LTO-7, LTO-8, and future LTO drives.

LTFS Enables Faster File Access

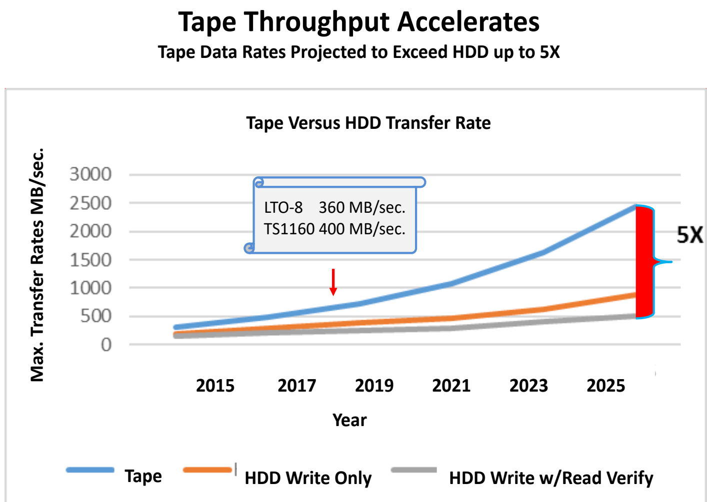
To improve the access and interchange capabilities of tape, a new, long awaited open standard file system specification for LTO called LTFS (Linear Tape File System) was developed by IBM and became available with LTO-5 in 2010 proving an easier way to access tape without the need for another software product. With the new dual partitioned tape functionality of LTFS, one partition holds the index and the other contains the content, allowing the tape to be self-describing. The metadata of each cartridge, once mounted, is cached in server memory and operations, such as browsing directory tree structures and file-name searches are performed quickly in server memory and do not require physical tape movement.



LTFS provides connection with OpenStack Swift to enable movement of cold (archive) data for large object storage to more economical tape and cloud storage. The faster file access capability provided by LTFS becomes more important as tape capacities continue to increase and the number of files stored per tape steadily increases.

Tape Delivers the Fastest Data Rates

Tape capacities and data rates (tape throughput) are growing faster than all other storage technologies. Comparing native data rates, the enterprise TS1160 at 400 MB/sec. and LTO-8 at 360 MB/sec. both transfer data much faster than the typical 7,200 RPM HDD at 160 - 220 MB/sec. The write verification process is a key storage requirement to ensure data integrity. Multi-bump tape heads enable tape to immediately read the just-written data, performing on-the-fly verification without compromising any of the transfer rate. For disk systems, the host must request a full read process, resulting in an overall reduction of the data transfer rate when writing data. INSIC has projected steady increases with data rates ~ 5x faster than HDDs by 2025. When architected with RAIT, the aggregate throughput of a tape subsystem is unmatched.

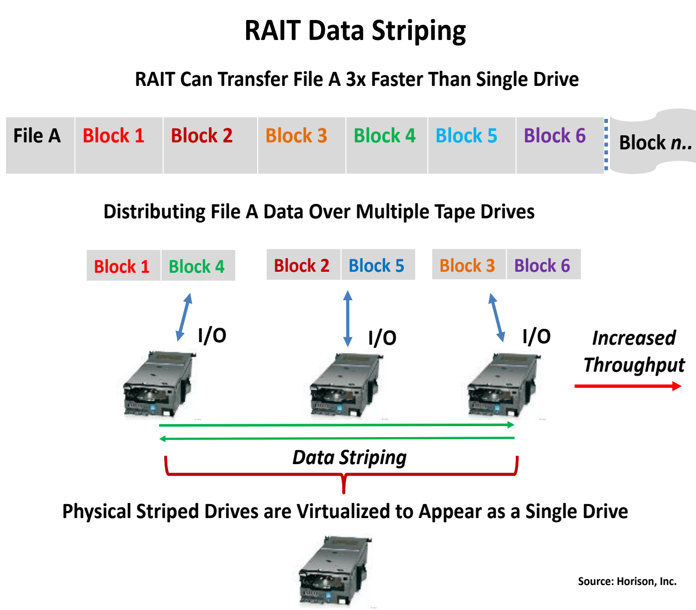


Note: Multi-channel head technology enables faster transfer rates in tape drives. Disk drives use a single active channel.

Source: TSC State of the Tape Industry Memo 2019, INSIC

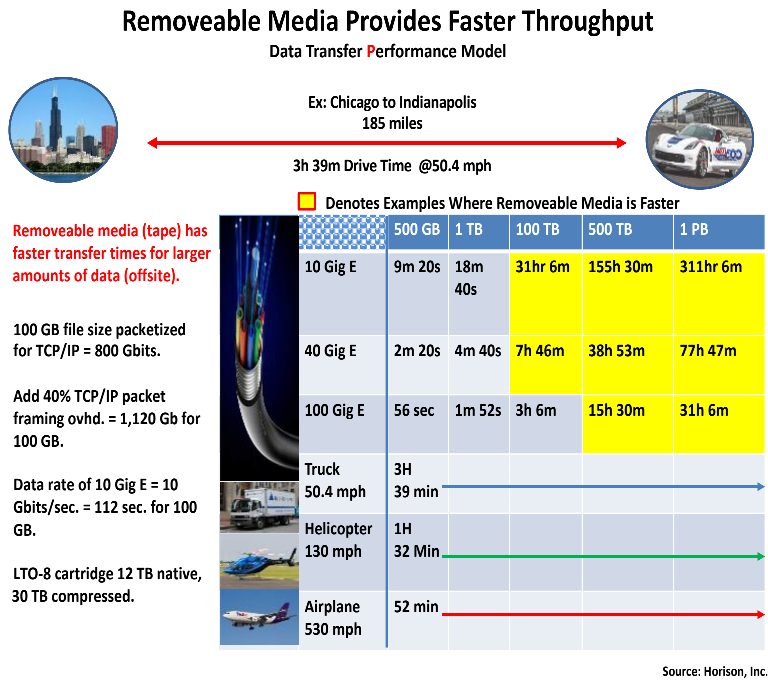
RAIT Improves Tape Throughput and Offers Fault Tolerance

HPSS RAIT (Redundant Arrays of Independent Tapes) is an HPSS software feature that capitalizes on faster tape data rates significantly increasing the throughput of large sequential files by creating multiple parallel data lanes into the tape subsystem. RAIT levels are implemented in HPSS and the stripe width is the number of drives, typically 3, 4, 5, or 6 needed to maximize tape performance. RAIT can provide fault tolerance using parity, much like RAID for HDDs, by ensuring that if a tape drive fails, the application can still operate on the remaining drives without impacting availability. The only extra RAIT overhead cost is the amount of space used for parity. With increasing emphasis on hyperscale with large geographically dispersed tape storage pools, look for **RAIL** (Redundant Arrays of Independent Libraries) systems to evolve providing major reductions in data transfer time and higher tape availability for massive amounts of data. Like RAIT, with RAIL data is striped across tape drives, but each drive is in a different robotic library. The libraries may be located in different geographic locations for higher degrees of fault-tolerance. The stripe width for RAIL is the number of libraries in the stripe.



Data Transfer Performance Model

Removeable tape media has taken a new role for transferring large amounts of data given the speed limitations of modern network bandwidth. The example below compares the time to transfer data loads ranging from 500 GB to 1 PB on standard [GigE network](#) speeds versus moving those same data loads if stored on removeable tape media using a truck, helicopter or an airplane from Chicago to Indianapolis, a drivable distance of 185 miles. The areas in yellow indicate which data loads are faster using removeable media than network bandwidth. As data loads and distance increase, removeable data transfer becomes faster than network transmission. For example, the flight time from Boston to Los Angeles is 6 hours, 21 minutes and the flight distance is 2611 miles. A truck or helicopter wouldn't be feasible. An airplane would still transfer the data loads of 500 TB or more faster than any network transfer time.



	500 GB	1 TB	100 TB	500 TB	1 PB
10 Gig E	9m 20s	18m 40s	31hr 6m	155h 30m	311hr 6m
40 Gig E	2m 20s	4m 40s	7h 46m	38h 53m	77h 47m
100 Gig E	56 sec	1m 52s	3h 6m	15h 30m	31h 6m
Truck 50.4 mph	3H 39 min				
Helicopter 130 mph	1H 32 Min				
Airplane 530 mph	52 min				

Summary

Tape performance has improved for access times and throughput. Though tape won't be a solution for random access, data base, and high IOPs applications like SSD or HDDs, recent performance developments are positioning tape as a mass transit system capable of moving massive amounts of data that will need to be accessed, protected and stored indefinitely. As the information age embraces larger throughput loads from archives, streaming media, cloud services, big data, DR and the IoT, the performance gains tape has and will continue to make are becoming increasingly important. Tape has now emerged as the top choice for both archive data storage and transferring large amounts of data. Future roadmaps indicate tape will make major strides in capacity, reliability, total cost of ownership and overall performance, with relatively few limits in sight.