The Truth About Tape

FAQ

By Jon Toigo

Digital magnetic tape just celebrated its Diamond Jubilee – 60 years as a stalwart servant in the recording, preservation and protection of electronic data. The appeal of tape technology has had its highs and lows over the decades, with the medium's popularity linked less to its capabilities or limitations than to the level of noise emanating from the marketing campaigns of vendors of competing data storage technologies. However, tape has proved to be a resilient warrior, one that continues to be entrusted with the backup and long-term archiving of nearly 80% of the world's digital information.

In the selection of questions for this FAQ, we have strived to steer clear of the cliché and stilted.

Which technology is best for backup: tape or disk?

Depends on the data, but both if you are smart.

We also sought to avoid old debating points:

Industry analysts claimed back in 1999 that one in 10 tapes fail on restore. How has tape changed to improve its reliability?

That claim, which has been disavowed by the analyst who said it, was bogus at the time it was made. Interesting what a $50 million marketing budget will buy a disk array vendor from its paid mouthpieces in the analyst community.

Instead, this FAQ focuses on what you really need to build a business-savvy business value case for tape:

- The return on investment and cost containment value of tape.
- The risk reduction value of tape as a technology, as a hedge against regulatory noncompliance and as a technology investment safe from early obsolescence.
- The importance of tape as an enabler of greater productivity.

We hope you will find it useful.

FAQ

As a data storage technology, what are the key things that prospective users need to know about tape?

Magnetic tape is one of the oldest and most reliable data recording technologies ever invented. It uses a streaming medium in the form of a spool of linearly oriented tape – comprised of several functional layers, including a magnetic layer, on a flexible mylar plastic support – that passes under a read-write head to record binary bit information into tracks.

Successive generations of tape have brought with them significant improvements in terms of the capacity and the performance of the medium. These gains have resulted both from improvements in the recording media itself and associated tracking and positioning technologies in tape drives and also from the integration of technologies borrowed from the hard disk world, including giant magneto-resistive (GMR) read/write head technologies. In 2010, IBM and Fujifilm demonstrated a new tape media coating technology and magnetic particle, barium ferrite (BaFe), promising a 35 TB capacity in a single cartridge before the end of the decade.

Today, the tape industry produces several types of tape products, distinguished mainly by media and drive characteristics. Linear Tape-Open (LTO) tape has been advanced by a consortium of vendors offering products viewed as price/performance appropriate for applications commonly seen in small, medium and some larger IT environments. In 12 years, over 4 million LTO format tape drives and nearly 200 million cartridges have shipped. Standards for the sixth generation of LTO tape have just been announced, and conformant media and drives providing 6.25 TB of compressed capacity and a transfer rate of 400MB/sec should be available within the year.

In addition to LTO, various proprietary tape formats advanced by longtime tape automation vendors continue to have market share in large to very large IT environments. Significant among these are Oracle's Sun/STK tape drive and cartridge – the T10000C – which is significantly faster and more capacious than LTO tape, and IBM's TS1140, which rivals the Oracle offering in terms of both speed and cost.

Whether one considers LTO or so-called "enterprise-class" tape, the technology's main competitor in recent years has been the disk array. For about a decade, disk-based products have been encroaching on workloads that were once the exclusive purview of tape, including backup and archive. Despite objective financial analyses and operational comparisons between disk and tape-based solutions, which nearly always favor tape, industry analysts and consumers appear to have been taking their guidance from disk vendor marketing campaigns. More recently, however, the combination of a challenging economy (a dynamic that favors less expensive tape) and improvements in tape technology itself (enabling it to move into new roles, including "active archiving") appear to be reinvigorating interest in and adoption of tape technology.

Tape is seeing an increased uptake in industry verticals such as broadcast and surveillance video, media and entertainment, and healthcare, driven by the specialized needs for data retention in these industries, the

FAQ

nature of the data being preserved and other factors. But across all industry segments, tape is enjoying resurgence as a key technology for backup, archive and secondary storage of less frequently accessed files (so-called active archive).

Cost of ownership discussions appear in most tape-versus-disk articles. What are the facts? Is one cheaper than the other?

This question is framed incorrectly. Cost-of-ownership comparisons make sense when the options being compared are equally appropriate to the task under consideration. That is another way of saying that storage technology must first be best suited to the needs of the business process it serves and the requirements of the data generated and used by the process before entering into analyses of internal rates of return on technology investments. This is easiest to understand from the perspective of data protection or backup.

If an application serving a mission-critical business process must be available continuously, on an uninterrupted basis, the data supporting that application must also be continuously available. Data inherits these protection and availability requirements like so much DNA from the business process and application it serves. Since restoring data from tape can take some time, often some sort of disk-based replication with failover is used to provide a first line of defense for the mission-critical data – delivering swift access to an alternate copy of data in case the original becomes compromised.

With most business processes and applications today, however, adequate "time to data" windows exist for recovery to make tape-based data protection and recovery well suited to data protection and recovery requirements. A good "defense in depth" strategy for mission-critical apps uses both disk- and tape-based replication approaches, with the latter providing an additional hedge or safety net in case of a failure of the disk replication process.

As a rule, tape is generally less expensive to own and operate than disk. Failure rates in tape media are well below comparable failure rates in disk, and the media itself is separate from the drive, so replacement of worn media does not entail the cost or inconvenience of replacing entire kits – as is the case in disk media replacement. And media capacity cost is measured in pennies (or fractions of pennies) per GB with tape versus dollars per GB with disk.

Tape is also an "offline" or "nearline" medium, so the power and cooling costs for tape are significantly lower than those for disk, which is always powered, even when idle. This is increasingly significant as energy costs increase and availability of new power to data centers gets harder to come by owing to utility grid limitations.

Some have tried to argue that tape's dependency on external storage (offsite facilities are commonly used to house data copies that may be needed for restore in the wake of a disaster) and the procedures for moving tapes back and forth to offsite storage facilities make tape more expensive than disk from an operational labor cost perspective. This may be true in some cases, but it is an increasingly specious comparison as more and more companies leverage WAN-based disk-to-disk data replication techniques. Most are discovering that data protection

FAQ

via distributed disk repositories is significantly more expensive to own and operate than is tape backup with offsite storage – by several orders of magnitude of cost.

Again, such cost comparisons are less important, from a planning perspective, than is the question of suitability to task of tape versus disk. In addition to backup, tape technology is increasingly used to support data archiving efforts – and even general-purpose production storage.

Tape has long provided service as an archival medium, a task to which it is well suited both from the standpoint of media durability and reliability as well as automation and drive standards that provide for backward compatibility and "in-library" (background) media management and data migration.
A basic metric for reliability is found in bit error rates – a measure of how often the writing of data to a medium results in errors and data corruption. Recent studies reported by USENIX place bit error rates in SAS and Fibre Channel disk at 1 in 10^{16} – meaning that 1 in every 90 disk drives contain a non-recoverable error. This "silent corruption" can be significant, impairing the ability to open a particular file or rendering an entire RAID disk set inaccessible. With tape, we are looking at a significantly better ratio: 1 in 10^{17}. That's one bad bit in 12.5 petabytes, rather than disk's one bad bit every 67 terabytes. Moreover, in-line error correction and read/write verification technology available in tape provides an early detection capability of bit errors, delivering an undetected error rate of about 1 in 10^{27} – in a word, an infinitesimal impact.

Given these error rates, and the need to store data reliably over time in a deep archive, tape appears to be significantly better suited to archive workloads than is disk. The cost to maintain the repository is also lower, as hardware refresh cycles tend to be spaced further apart in tape than in disk arrays and media migrations between tape generations tend to be as effortless as data migrations between different generations of disk. In fact, given the proprietary nature of value-add software on disk arrays today (de-duplication, compression, etc.) migrating data from one generation of disk to the next may entail the surmounting of increasingly nontrivial technical barriers going forward.

So, tape is likely to be more fit to purpose in the case of deep archive than are disk rivals – and that is before it makes sense to conduct cost-of-ownership comparisons. Recently, tape has started to gain traction once again as a form of primary storage – or at least as a repository for files that are infrequently accessed and consuming the lion's share of space on expensive disk arrays. This is an interesting development.

Historically, tape began life as general-purpose storage for data in mainframe shops, where it was substantially less expensive than drum or core memory for data storage and less cumbersome than punch cards or punch tape. However, the medium lost its luster as a general-purpose data repository in the early 1970s as rotational disk became cheaper and more capacious. Over time, tape was relegated to a much narrower set of roles: mainly backup and archive.

Today, however, the use of tape in a "production storage role" has recently returned to the forefront of IT infrastructure discussions with the appearance of a general-purpose file system for tape, the Linear Tape File System (LTFS) from IBM. An enabling

FAQ

technology for bridging tape storage and file systems, the technology is seen as paving the way to a network-attached storage solution providing extraordinary capacity and significantly lower cost than disk – just in time for the data deluge. LTFS isn't the only file system for tape, but its arrival, in concert with the first generation of LTO tape featuring "partitioning" (a capability required for LTFS to work), is exciting planners who are (1) prohibited from deleting data, (2) confronting unprecedented levels of capacity demand growth and (3) hamstrung by budgetary constraints from scaling existing storage capacity to meet new demand. An LTFS-enabled "Tape NAS" would provide a low-power, low-cost, network-accessible storage platform for the up to 40% of file data currently stored on disk that is seldom, if ever, accessed.

Bottom line: Whether you are looking at data protection, data archiving or capacity production storage, the selection of tape is not simply a matter of cost comparisons with disk – though tape tends to win such comparisons handily. The choice of tape is first and foremost about the suitability of the storage technology to the needs of the workload and, perhaps, to the skills and knowledge of the staff. With respect to the second criterion, staff knowledge, tape backup is a familiar meme, with most issues having less to do with the hardware infrastructure and/or media than with the vicissitudes of backup software functionality.

In general, tape is already easier to operate than disk. It is "pre-virtualized" and completely agnostic regarding the workload or the control software (backup, archive, etc.) with which the tape library is used. Disk, by comparison, is highly variable given the propensity of vendors to add value in the form of specialty software joined to proprietary controllers on arrays. As a result, management complexity of disk infrastructure is much greater than tape, resource allocation and utilization efficiencies are much lower, and labor costs are much higher.

From a media cost perspective, recent reports place the cost of tape at about 10 to 20 cents per GB (without data compression), while the per GB cost of disk ranges from 70 cents to upwards of $6.00. Both disk and tape are poised to see huge capacity improvements, and compression technologies for squeezing more data into the same capacity are improving, both of which will improve the cost efficiency of the media at the media level. However, the argument can still be made that falling media costs are not passed along to consumers as efficiently in disk systems as they are in tape libraries. The cost of a disk storage system is often 100x greater than the cost of the disk drives inside the box, adding significantly to the cost of disk storage. A tape library system does not accelerate the cost of tape media by such a significant rate, in part because upgrading to new generations of tape drives and media rarely require a "forklift upgrade" of the entire kit. Hence, even at the system level, tape continues to show greater cost efficiencies than disk.

Does your comparison of tape and disk include a fair extension of total cost of ownership (TCO) elements: labor, external services, etc.?

An extended TCO model would still work to the advantage of tape. All storage technology has two dimensions of cost: CAPEX (the cost of acquisition) and OPEX (the cost of administration). From a capacity perspective,

FAQ

the CAPEX cost of tape is significantly lower than disk. Moreover, risk of vendor lock-in, at the media, drive or automation levels in LTO tape, at least, are mitigated by de facto standards from the LTO Consortium, reducing investment friction concerns.

The rest of the cost of either tape or disk storage has to do with administration and environmental costs. Environmental costs include cost of energy, heating and cooling. From this perspective, the creation of a large, multi-petabyte repository for data using tape requires far less power for both equipment operation and for HVAC than does a comparable repository built on disk – even when parts of the disk infrastructure are operated in reduced power states or contain de-duplicated or compressed data so that fewer drives are used to hold the same quantity of data.

From an administrative perspective, a good tape management application and process will require the cycling of tape copies to an offsite storage facility as a hedge against a facility or equipment room disaster, which is an additional cost. Accomplishing the same level of protection for data stored to disk requires a high-speed, high-capacity WAN link between two pre-designated and identically configured sites, replication software or identical storage kits with compatible value-add replication software, and a potentially labor intensive and operationally disruptive testing regime to confirm the proper operation of replication processes. To test a mirror, typically applications must be quiesced, caches flushed, replication processes stopped and comparisons made between source and target disks to confirm that data is being successfully copied and within acceptable deltas. By comparison, performing a read/verify on a tape storage at an offsite facility does not interfere with day-to-day processes at all.

In the final analysis, the comparison of disk and tape TCO, assuming that both are minimally acceptable to the workload requirements they are tasked to support, usually result in better outcomes for tape. Where tape may lag behind disk is in I/O performance, though this is a highly nuanced area of discussion.

Since we are there, let's talk performance. Clearly, disk wins this war, right?

Actually, we need to define win. A lot of performance is relative and contextual – related to the application that is being served – for example, backup and restore.

Disk backup – or replication over distance using a WAN – is considered to be the analog of tape backup and removal to offsite storage. Is WAN-based replication faster than tape backup and offsite storage? That depends on several factors, including the amount of data that must be moved across the WAN, the distance that data must be moved and the overall throughput efficiency of the WAN link. Truth be told, WAN-based data replication isn't that efficient: An OC192 link, for example, would take about four hours under ideal conditions to move 10 TB of data from point A to point B, while a T1 would take over a year to move the same volume of data. And that is considering only link speed, not distance.

When moving data over distance, you need to consider distance-induced latency. The rule of thumb here is that for every 100 kilometers data travels, the target disk is at

Toigo Partners International

least 12 SCSI read/write operations behind the primary disk. That's called a delta, and it gets worse the further that data needs to travel over an interconnect.

And latency is only part of the problem. You also need to concern yourself with jitter – stumbling blocks such as queuing and processing delays that occur inside the WAN infrastructure that accumulate to add more delay to the transfer of data between two points. This is where the real problems start to show up. WANs, like LANs, are not optimized to provide the most direct or shortest path between two points, but rather the smallest number of router hops. So, to use the example of airline travel, your flight may route way out of its way to move you between two airports that are geographically fairly proximate to each other. The fastest way to move data between geographically dispersed locations is via passenger pigeon – not kidding, there is a protocol for it called IP over Avian Carrier.

Some vendors say that "virtual WANs" will help by combining the Internet and WAN services. They might, but there is no service-level guarantee on the resulting transport facility. Other vendors claim that MPLS networks solve the problem, and they do, provided that you are only replicating data over the common span of an MPLS WAN (usually less than 50 kilometers). Unfortunately, that may not be distant enough to insulate data copies from the natural or man-made disaster that impacts the original data.

From a performance perspective, then, tape backup may be very much in line with WAN-based data protection, but without the cost, the hassles of delta management and so forth. So, performance is a bit more nuanced.

Restore is another matter. Today's LTO 5 tape drives, practically speaking, deliver about .675 TB/hour/drive in read/write speed. Data must be transferred from tape to disk to make it useful. Simply put, this is not required if data is already instantiated on disk. If instantaneous failover is the requirement, disk may well be the better choice for data restore. However, even this rule of thumb has a few issues. As noted above, the copy of data at the remote site may be quite different from the original data, owing to transitory deltas and jitter. So, there may well be some delay waiting for the recovery disk to synchronize with the original, or to resolve issues with applications resulting from missing data. As a measure of simple throughput or time to data, disk may be the same speed or slower than tape as a recovery mechanism.

Another interesting dimension to performance comparison is revealed when you look at tape NAS solutions that are beginning to come to market. The disk vendors have been quick to point out that retrieving a specific file from a tape-based NAS running LTFS or some other tape file system will be considerably slower than the speed of retrieving a file from disk. On its face, this comparison makes sense, but not necessarily in reality.

First, LTFS tape NAS uses a front-end controller and a back-end tape library. The front-end controller, like a controller on a disk array, uses buffering to expedite performance. One tape NAS head in the market stores the first 2 MB of every file stored to tape on disk so that, when the file is requested, it can begin loading from the disk while the full file is sought on the appropriate

FAQ

tape volume. This bit of "spoofing" may well conceal the "time to first byte" delay that accrues to tape NAS.

It is also worth noting that the streaming rate of tape is much faster than that of disk. Hence, for long block files such as video, once the first byte of the requested file is found on the tape, the transfer of the full file is much faster streaming off of a tape cartridge than it would be coming from a RAID array. This is one reason why a cloud storage service utilizing tape may actually outperform one based on disk.

You mentioned tape's new (old) role as a capacity storage repository using a tape file system. Please explain that more in detail.

With the introduction of the latest tape file system, LTFS, new attention is being paid to building a capacity NAS platform that uses tape for storage. tape NAS is actually an integration of several components:

- Front-end processor (or "head") – a commodity server running an operating system, a complement of disk for buffering and local storage, network file system software (NFS, CIFS/SMB, HTTP, etc.), LTFS and any hardware required to connect to a LAN and to a back-end tape library.
- Back-end storage – the tape library itself, complete with one or more robots, drives and shelves for media.
- Management software – some means for monitoring, configuring, tuning and troubleshooting the overall kit.

In truth, LTFS by itself is only a kind of middleware or glue between the tape NAS head and the back-end physical library. It does not currently present an intuitive file system view of the contents of the library. However, third-party developers are mapping their wares – traditional file systems, archive management systems, media management systems, etc. – to LTFS to make the middleware glue useful in practical applications. Crossroads Systems, for example, has an early lead in the pre-integrated tape NAS head space with its StrongBox™ platform. Users see file listings in the same format as what a typical disk NAS platform presents.

It is also worth noting that LTFS only works with partitioned tape media – media that can be segregated to enable file markers and metadata to be stored in one partition and the actual data comprising a file in the adjacent partition. In the LTO family of tape, this is only possible with LTO 5 or better generations of media. For enterprise tape, both IBM's and Oracle's current products are partitionable.

Retrieval speeds for files, with no buffering or spoofing whatsoever, depend on where data is located. If a requested file is already on the media in a tape drive, its retrieval time can be very short. If the file is located on a tape that must be retrieved from a shelf by a robot, spun up in the tape drive and then accessed, the time to first byte may be greater. The range commonly used to describe tape NAS performance is between 2 seconds to 2 minutes – approximately the same as the "World Wide Wait" experienced when retrieving a rarely accessed PDF file from a website.

As stated above, some disk-based buffering can be used to spoof time to first byte, but the real performance advantage accrues

FAQ

to tape's superior streaming rate once data transfer is under way. The streaming rate of tape is significantly better than the throughput of a disk drive or disk array.

Are there industry verticals that are adopting tape NAS more quickly than others?

Healthcare, media and entertainment, and broadcast and surveillance video, but also storage clouds.

The industry verticals healthcare, M&E and video all confront at least one similar challenge: the need to store growing numbers of "long block" files. These are predominantly rich media files (graphics-, video- or audio-intensive) that can actually take advantage of the superior streaming rate of tape. In the case of healthcare and M&E, and in some video applications, the applications also present a requirement for long-term storage, whether for operational reasons or to comply with regulatory requirements. So, tape's resiliency characteristics have appeal.

The streaming rate and resiliency of tape also factor into the suitability of the technology as the storage platform for cloud storage. Despite the hype, cloud-based storage isn't a general-purpose resource. Wide-area networking imposes delay and jitter on data transfers, limiting the utility of a "great disk drive in the sky." However, the placement of tape in a cloud makes a great deal of sense for certain applications, especially archive and active archive (production files with low rates of re-reference). Permivault™ from Fujifilm is a good example of a tape-based cloud service that is paving the way for a new category of on-demand archive and tape NAS services.

About the Author

Jon Toigo is a 30+ year IT veteran who now serves as Managing Principal of Toigo Partners International LLC, a technology consumer advocacy and vendor watchdog, and Chairman of the Data Management Institute. He is the author of 17 books and thousands of articles and columns for the trade press. He writes a monthly column for TechTarget's STORAGE magazine, entitled Storage Revolution, where he frequently challenges both the common sense and business savvy of architectural and technological marketing memes of the IT industry. He maintains a popular blog at DrunkenData.com.