Mesabi Group

Commentary

November 10, 2009

Tape Fallacies Exposed — the Future of Tape Is Still Bright

Tape is not going away anytime soon. That may disappoint the "tape is dead" crowd who believe in fallacies such as disk can replace tape for all practical purposes, that disk is cost competitive for all tasks where tape has traditionally been strong, and that tape has significant manageability and usability issues. The fact is that tape shines in areas such as off-line, off-site data protection, tape has dramatic cost and green advantages over disk, and tape is more than acceptable from a manageability and usability perspective. Will the critics never learn?

Tape Is Here to Stay

Tape solutions are not going away. They will continue to play a fundamental role for backup, recovery, compliance, information security, long term data retention and data protection in the IT organizations of a great many enterprises. That assertion should not be controversial, but it is. A number of critics have essentially said that tape is an end-of-life technology.

There are a number of fallacies that these critics espouse that lead them to that conclusion. Let's expose those fallacies for what they are by reference to first principles, best practices, simple logic, and facts.

A number of fallacies exist, but basically they fall into three primary categories:

- **The Business Use Fallacy:** Disk can replace tape for all business purposes except perhaps for deep archiving. This might be called "the anything you can do, I can do better" fallacy. **Reality:** The optimal strategy is a blend of disk and tape to address business, performance and data protection objectives.
- **The Cost Fallacy:** Disk solutions are cost competitive enough so there is no need to keep tape for economic reasons alone. **Reality:** Even if disk and tape were the same cost, tape is needed for off-line data protection best practices. But disk and tape are not the same cost. Native tape in automation is the low cost storage solution by an order of magnitude, is the lowest energy consuming green storage and will continue to maintain that gap with ongoing tape innovations.
- **The Manageability Fallacy:** Relative to disk, tape imposes an administrative burden on IT organizations because tape is unreliable as well as difficult and complicated to use. **Reality:** Today's tape drives and

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

media, like LTO technology, have proven to be highly reliable and advances in tape automation have dramatically simplified the management of tape storage.

These fallacies may be expressed in other words, but that is the essence of them. Let's peel back each of them and expose them for the fallacies that they are.

The Business Uses of Tape

A little background on "copies of data" and "pooling" are necessary before a discussion of where tape and disk are each best used.

Disk and tape are used to house one or more copies of data. A *copy* of data is typically used either for *production* purposes *or* as a *data protection* copy (for purposes of simplicity test data will be ignored).

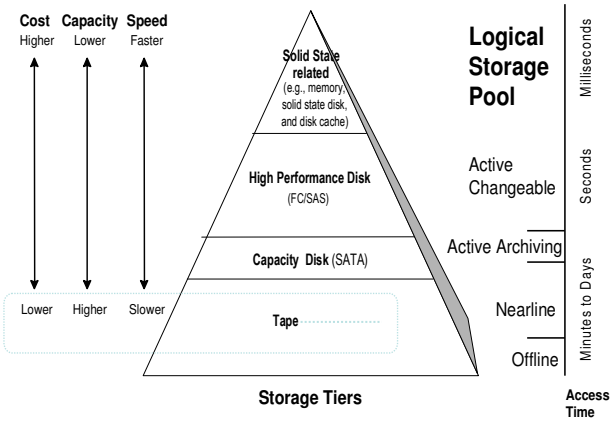
A pool is a collection of data that is managed as a homogeneous whole for quality of service (QoS) purposes. Pools can be incorporated into tiers (Figure 1) (Note: nearline represents data that resides say in a tape library and can be accessed fairly quickly (seconds or minutes) whereas offline is data that takes longer to recover, such as at a disaster recovery site.)

Each of the different storage media has different cost, capacity, and speed characteristics.

Further analysis reveals where disk and tape are each best used (Table 1).

Production data comes in three pools: active changeable, active archive, and deep archive. An active changeable pool contains data that is eligible to be modified or deleted and is therefore

Figure 1: The Storage Pyramid — Tiering and Pooling



Source: Mesabi Group, November 2009

subject to change. An active archive pool contains data that is fixed content data, i.e., data that has been single instanced and is not eligible to change, but still needs to be available to a business user with online or near line service levels. A deep archive pool is fixed content data that a business hopes to never have to access again, but still has an obligation to keep. A deep archive has longer recall service levels so can be stored offline.

From a pragmatic perspective, the majority of data is fixed content data, but many organizations do not have — or are slow in moving data to — an active archive. The active changeable pool may be bloated with data that it does not need, but the impacts of that is a separate discussion. Note though whether or not bloated, the active changeable pool is the traditional focus of the greatest attention in pooling.

One or more copies of each pool of production data must be maintained for data protection purposes.

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090

Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

The Active Changeable Pool

The production copy of the active changeable pool is disk. This is a no brainer because of the random access and speed required by this pool.

Table 1: Where Tape and Disk Are Each Best Used

	Production Copy	**Data Protection Copy**
Active Changeable	Disk	*First line of defense (OR)* Disk/tape (depends upon environment and requirements) *Second line of defense (DR)* Disk/tape (depends upon circumstances) *Last line(s) of defense* Tape
Active Archive	Disk (when online access is needed) Tape (where sequential processing is acceptable)	Disk (if fast failover is required) Tape (if slower recovery is acceptable)
Deep Archive	Tape	Tape

Source: Mesabi Group, November 2009

Whether or not data protection copies of data should all be on disk is far from obvious. Data protection copies may reside on disk, tape, or other media. (Even the primary production disk "copy" of the data, where the active changeable production data resides, can be considered a data protection copy if some form of protection, such as RAID, has been added for data protection purposes.) These copies form a number of lines of defense. Each line is a layer or degree of data protection and represents a distinct physical copy of the data. Best practices require a *minimum* of three copies: First, Second and Last Line of Defense.

First Line of Defense

The non-production copy that is the first line of defense is typically for operational recovery (OR). OR tends to be not for recovery from a physical data protection problem, as disk arrays are really very reliable, but rather tends to be for recovery from a logical data protection problem. Logical data protection problems include accidental deletion of a document that needs to be recovered, a virus, hacker, or database corruption; anything where non-physical changes have been made that turn out to be undesirable.

The OR copy needs to be point in time consistent and a full physical copy. These are generally created using backup/restore application software. The critical issue for creation is whether there is enough time in the backup window — the running-out-of-night problem. The critical issues for restoration are selecting what is needed to be restored (as an OR tends to be very selective in what needs to be restored as a single document or the data used by a single application) and the time that is necessary to restore that information.

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090

Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

It is not intuitively obvious as to whether tape or disk is best. Consider the following real case study:

User Story: *Tape Usage at EDPS*
EDPS is the IT department for a trucking company with a number of divisions that has annual revenues of about $1.5 billion. The main data center in Richmond, Virginia runs enterprise-class storage on a SAN for key applications, such as billing, and midrange-class storage for imaging applications, such as delivery receipts. A co-location site in Mesa, Arizona serves as the DR site. EDPS uses asynchronous remote mirroring to keep the Mesa site data up to date. But remote mirroring (either synchronous or asynchronous) only provides physical data protection because any changes that reflect logical data corruption would be quickly propagated from a main data center to a DR site.

For logical data protection and for overall OR data protection, EDPS has long used an open systems LTO tape library solution. A snapshot copy of all production systems is performed daily. A full backup to tape takes place during the day without any impact on production servers. All production data is backed up 365 days a year and sent off-site.

Dick Cosby, EDPS Systems Administrator, is extremely happy with tape especially since EDPS had a true disaster from Hurricane Gaston in 2004 that flooded all systems and the offline/offsite tape was able to recover everything (as the Mesa site was not in operation then).

> "Disk and tape work incredibly well together and both are utilized to achieve our performance and data protection objectives. In my experience, tape is the most reliable data storage on the face of the earth."
>
> — Dick Cosby, EDPS

Don't Automatically Throw Disk at a backup Window Problem
There are cases where dramatic SLA performance requirements may favor disk. In this scenario a disk to disk to tape strategy may be advised. The backup to disk or VTL can provide quick recalls requested by the backup application and the tape copy can provide offline and offsite data protection as well as long term data retention.

But the decision may not just be between tape and disk. Assume for the sake of argument that a VTL using disk can speed up a backup process and thus reduce the length of the backup window. Might that same result have been achieved by moving the fixed content from the active changeable pool to an active archiving pool (while at the same time deleting data that is not needed to be kept)? A smaller active changeable pool may solve the problem. Throwing hardware (in the form of disks) at a problem is not necessarily the best answer.

For example, a study by the University of California at Santa Cruz found that 90% of data stored to disk was never accessed again in a 90 day period. Whether or not this data can be moved to an active archive on disk or a deep archive on tape depends upon a number of factors, but

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090

Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

keeping it on expensive disk may not be the most cost-efficient option.

Bottom line: organizations should not automatically think that disk is the best solution for the first line of data protection defense, but rather think clearly about all their options.

Second Line of Defense

This line of defense tends to be for disaster recovery (DR) purposes. Remote mirroring of disk provides high availability, but only for physical problems. If data corruption occurs (e.g. due to system error, virus, hacker, etc.) then the remote replicated copy will also be corrupted. Electronic off-line vaulting can produce a copy with both physical and logical data protection (assuming that the original production copy was not unknowingly corrupted). Still the availability requirements after a true disaster are relaxed over those of OR (as restoring applications and data rapidly does no good if networks and users are not able to access the data). Organizations have to consider all circumstances before making a decision and that decision may very well be some combination of both disk and tape.

Last Line(s) of Defense

The last line represents the last chance to recover data. Yes, that could be disk, but frankly it had better be tape. Tape is offline and provides I/O isolation. Being offline protects critical data from intentional or unintentional corruption. That means that it is logically protected (i.e., no writes) as well as being a physical copy.

If that tape is offsite at a neutral third party, that is an advantage. The third party has a vested interest in ensuring the physical preservation of the data. Also the data can be sent to any number of locations from which it can be restored using any compatible tape drive. These removability, transportability, and interoperability advantages are characteristics that disk arrays do not possess.

The downside is loss of availability, but, if a rare catastrophic event occurs, data preservation trumps data availability.

The Active Archive Pool

If online access (in the sense of a few seconds) is required, then disk is the choice for the active archive production copy. If sequential processing is used and if the performance is adequate, then tape would be a solution. Note also that online access where longer start times are acceptable (such as a minute for a video) might also use tape (in conjunction with disk as front-end cache).

On the data protection side, disk is a requirement if fast failover is necessary. However, tape may be acceptable if slower recovery is acceptable. Fixed content data follows a long-tail distribution, i.e., the frequency of access is very low. The type of information, its value, and the value of rapid recovery all have to be taken into account.

The Deep Archive Pool

Tape is the preferred solution both for the production copy and as many data protection copies as required. Spinning down selective disks or using MAID is an alternative, but that means that the data would be retained on-site and would be on-line (no I/O isolation). (Removable disk is possible, but not likely to scale and is primarily tape emulation.) Moving data to a third party provides I/O isolation, physical isolation, and relieves

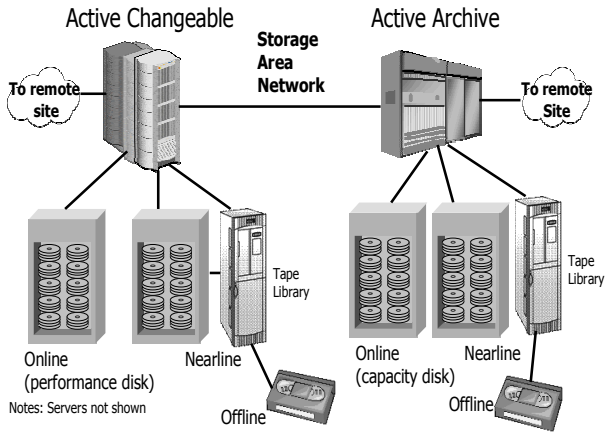
© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

the administrative burden for managing the data in-house. For example, in the case of a disaster, even though this pool of data would have a very low priority, it would still have to be restored to the primary site if the deep archive pool copy is located there.

Summing Up

Figure 2 shows a typical topology in which the different pools may reside. The different layers of data protection defense need to be appropriately spread out over the different physical components of the topology.

Figure 2: A Topological Example



Source: Mesabi Group, November 2009

Dollars and Sense for Tape and Disk

The second great fallacy is the cost fallacy that disk solutions are cost competitive enough so there is no need to keep tape for economic reasons alone. The implied assumption here is that disk has enough competitive differentiation on any valuable dimension so that it is a "superior" product. Even when cost was not taken into consideration, dismissal of the first fallacy shows that not to be the case. However, were it true, a reverse Gresham's Law (named after a 16th century English financier) would take effect. That economic "law" states that an inferior currency tends to drive a superior currency out of circulation. The reverse is that a "superior" product technology would drive an "inferior" product technology out of the market. So that is what probably lies behind the theories of the "tape is dead" crowd.

Why the cost fallacy got started in the first place is open to question. Perhaps the rapid adoption of high-capacity SATA drives and the expectation that the acquisition cost of a tape library and tape drives would somehow be comparable to the acquisition of a disk array and its drives fueled the fallacy.

However, what really happens is that tape has an overwhelming cost advantage over disk. That is clearly illuminated in the February 13, 2008 total cost of ownership paper "Disk and Tape Square Off Again — Tape Remains King of the Hill with LTO-4" by the Clipper Group.

That fact-driven analysis is very clear in demonstrating the dramatic cost superiority of tape over disk for a traditional backup application (and that includes taking into account data deduplication for disk). The Clipper analysts are very fair in saying that a disk to disk to tape solution is probably the best overall combination, but that the total elimination of tape would not be in the best interests of the enterprise. That blend would enable organizations to achieve performance and other key business goals while at the same time being very conscious of cost.

Costs are typically looked at on two levels — acquisition costs and total cost of

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

ownership (TCO). Acquisition costs are typically CAPEX (capital expenditure) costs and come out of a capital budget. However, looking at acquisition costs only ignores ongoing costs and could distort an economic business decision. A TCO analysis corrects this deficiency. Typically, a TCO analysis is run on a 3 or 5 year basis.

The Clipper Group did a five year analysis and tape came out ahead with lower acquisition costs and lower operating costs, such as floor space and energy costs. So tape wins the economic battle on both the acquisition cost and TCO fronts.

Rhetorically, why ever buy disk? The first answer is that disk can do a "job" (i.e., perform a necessary business function) that tape simply cannot do. Using primary disk for online transaction processing applications is an example. In that case, no cost comparison between disk and tape has any validity.

However, there are cases where both disk and tape can perform the same job, but one has an advantage over the other. Say that disk can shave hours off a backup job that typically runs on tape. When is this a must have (say when backup jobs are not able to complete on a regular basis) versus a nice-to-have (say when fewer hours for backup provides time for a rerun of a backup job if it should initially fail).

So what is the dollar value of that time shaving? How does an organization measure indirect economic benefits in such a way that accounting will find it acceptable in an economic justification analysis? That burden falls upon disk to prove as tape otherwise wins on an economic basis. And that can be a difficult challenge for disk as will be seen in examining the third and final big fallacy.

Managing Tape Effectively

That third fallacy states: relative to disk, tape imposes an administrative burden on IT organizations because tape is unreliable as well as difficult and complicated to use. That fallacy is actually composed of a number of smaller fallacies — such as tape is unreliable — which summed up challenges the usability and manageability of tape.

Dick Cosby of EDPS certainly does not feel that tape is unreliable or is difficult and complicated to use. He feels that "reliability is incredible," "retrieves couldn't be simpler," and "almost zero training."

That indicates that tape does not intrinsically have to be an administrative burden. However, other IT organizations have different workload mixes and business requirements and may not be as sanguine about tape solutions.

Tape is different from disk and does have to be managed differently. Understanding the impact of those differences from a technology perspective, from the perspective of applications that use tape, and from the perspective of the processes and people needed to use a tape solution is important. Let's start with tape technology.

The Management of Tape Technology Continues to Get Easier

Of course, tape technology consists of three basic elements: tape automation, tape drives, and tape media cartridges. The interplay of these three elements is actually fascinating to watch. However,

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

historically, these three elements did not have much information that helped with their management.

Potentially any electromechanical (or electronic) component can suffer failures. Failures can be catastrophic (in the sense of total and immediate), but often are the result of degrading performance over time. Failure of a tape media could result in the inability to read the cartridge; failure of a drive would be its inability to ingest or eject a cartridge; and failure of a tape library could be the failure of the robotics to move a tape.

Rather than wait until an actual failure occurs (reactive management), having the information and analytical capabilities available to take corrective action prior to a service-level-impacting event actually occurring (proactive management) is essential. And proactive management is now available for the use of tape technology.

The Management of Tape Media Inside the "Box"(i.e., within the Tape System) The worry that some organizations may have about the ability to read media is probably a major contributor to the concerns about the reliability of tape.

That concern should now be greatly alleviated. For instance, LTO technology has a number of advancements that contribute to high reliability and data integrity including: LTO drives perform a "read after write verification" of all data written to tape. That is, the drive immediately reads what was written to tape and compares it to what is in the drives memory buffer. If it is not correct it writes it again until it is correctly validated helping to ensure the data is readable when retrieved. In addition, LTO drives use an advanced servo tracking mechanism that keeps heads and media aligned for accurate recording and reading.

In addition, each piece of media in LTO tape technology contains a cartridge memory (CM) chip that contains a variety of information including historical information on the cartridge. That information includes the numbers and types of records with errors, numbers of bytes read/written, numbers of loads, and the age of the cartridge. Analyzer tools can examine tapes and give tape health scores, which can lead to a recommendation to replace a piece of tape media before a service-level impacting event occurs.

The Management of Tape Media Outside the "Box" (i.e., outside the Tape System) Although a tape cartridge could theoretically spend its entire lifetime safely nestled within a tape library, many tapes need to be exported. Exporting expels them from the automated world to the manual world — which can include moving to a remote site. A good degree of concern has been expressed that tapes transported to a remote site can be lost or stolen, which in some cases could lead to identity theft or the loss of valuable proprietary information.

If confidentiality is the goal in question, encryption is the answer. Tape drive hardware encryption is now manageable, and cost-effective when needed. Encryption of information sent to a third party for custodial care is a best practice if an enterprise does not want information exposed to an unauthorized user.

And there is even a GPS tracking capability available from at least one tape media vendor for those who simply must

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

keep track of their data. Who says tape is not innovative!

The Management of Tape Libraries and Tape Drives

For those concerned about issues, such as making sure tape drives are being used in a balanced manner and determining whether or not alerts should result in corrective action, now have help available. Some software (including one in the form of an appliance) is now available from at least two suppliers and tape library automation firmware can provide tape drive allocation and library partitioning helping to make tape storage management a simple task.

Backup/Restore Software — the Company Tape Systems Keep

Backup/restore software has a tightly coupled relationship with a tape system. Large scale deployments can run into problems where tape may get the blame (such as slow backup performance because the existing tape drives are not properly utilized or the available pieces of tape media run out). Existing backup/restore software is much better today than in the past in dealing with the overall process and special data protection management programs that help with both root cause and predictive analysis. These applications can either nip a small problem in the bud before it becomes serious or prevent one in the first place.

People and Processes

Tape and disk solutions require people and processes. Human error or poor processes can lead to problems such as a failure to restore properly. Sufficient effort has to be made so that the right people with the right training and the right processes have to be put in place.

Performing a necessary function is not a burden unless all business functions that cost time and money are considered a burden or unless an excessive amount of time or cost is involved. Managing tape and disk effectively should not be considered any differently than any other business function.

Conclusions

If, as the fallacies claim, disk can do everything that tape can do, if disk is reasonably price competitive with tape, and if disk is much easier to manage than tape, why is there any reason to maintain tape solutions? Those assertions have been exposed as the fallacies that they are (Table 2).

On the data protection side, some organizations can put together a strong case for a tape only solution. Other organizations can put together a strong case for a blended disk-tape solution. Finally, some organizations have been able to put together a case for a disk-only solution, but considering everything, including cost and off-line/off-site data protection, that would seem to be a real challenge. Perhaps that is why the recent Fleishman-Hillard survey (December 2008) revealed that 61% of the users of disk-only systems plan to start using tape.

The recent Fleishman-Hillard survey (December 2008) revealed that 61% of the users of disk-only systems plan to start using tape

© 2009 Mesabi Group LLC
26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.mesabigroup.com

Perhaps the fallacy should be: disk can replace tape in its entirety. That sums up the implied objective of the three fallacies — business use, cost, and management — in full. Separately, and collectively, those fallacies are false so the hidden fallacy of disk is always the best choice is also false.

Organizations should focus on a blend of disk and tape solutions that best meets their needs rather than be mislead by fallacies that contort decision-making and could lead to IT infrastructure decisions that are not in the best interest of the organization.

David Hill

Table 2: Fallacies Exposed in a Nutshell

Fallacy Name	Fallacy	Reality
Business Use	Disk can replace tape	A blend of disk and tape typically works best to optimize performance and data protection
Cost	Disk is cost competitive with tape	A tape solution typically has a huge cost advantage over disk especially as capacity begins to scale.
Manageability	Tape is an administrative burden	Advances in tape drive, media and automation solutions have greatly improved reliability and manageability

Source: Mesabi Group, November 2009

Analyst Name: David Hill
Topic Area: Tape Infrastructure

Mesabi Group LLC
26 Country Lane
Westwood, MA 02090
www.mesabigroup.com

Mesabi Group LLC is an affiliate of Valley View Ventures that aims to provide thought leadership and sound advice to both vendors and users of information technology. .

Phone: (781) 326-0038
email the author: davidhill@mesabigroup.com
Author of "Data Protection: Governance, Risk Management, and Compliance"

The information contained in this publication has been obtained from sources Mesabi Group LLC believes to be reliable, but is not warranted by Mesabi Group LLC. Commentary opinions reflect the analyst's judgment at the time and are subject to change without notice. Unless otherwise noted, the entire contents of this publication are copyrighted by Mesabi Group LLC, and may not be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written consent by Mesabi Group LLC