Data Mobility Group

datamobilitygroup.com

White Paper

20 October 2005

Dianne McAdam

Is Tape *Really* Cheaper Than Disk?

 76 Northeastern Blvd. Suite 29A, Nashua NH 03062 Phone: **877.463.6456** Fax: **603.563.8282**

Not since the Y2K phenomenon has the IT industry seen such intensity over technology as it has with the ongoing disk-versus-tape controversy. What started as a one- or two-vendor crusade against tape in the 1990s has continued in the 21st century as a full-on attack by many vendors.

Disk supporters, buoyed by the emergence of Serial ATA disk, point to the declining cost of disk and the increasing reliability of Serial ATA models for enterprise-class applications as proof of their argument that tape is dead, not needed—or at least dying.

Proponents of tape, meanwhile, argue the sequential media's ongoing usefulness in the data center. After all, tape is still less expensive than disk and is transportable for data sharing and optimal data protection.

So who is right? As it turns out, both are. Disk prices are dropping, which makes disk an increasingly appealing option for a number of traditionally tape-based applications, but tape innovation continues to keep tape significantly less expensive than disk, which means its role is also secure as the medium of choice for a variety of applications for which cost-containment is paramount.

Ultimately, the decision to go with disk or tape depends on the application and its specific requirements.

Disk Is.....

Disk is—and will continue to be—the mainstay for online storage in the data center. High-performance disk will always be a requirement for online applications that require fast performance. Disk also plays a critical role as the target device in a synchronous mirroring situation (i.e., data is being replicated between two sites), and it can also be effective as the initial backup target device in a disk-to-disk-to-tape (D2D2T) backup situation. However, disk is still more expensive than tape and tape offers unique attributes.

Tape Is.....

Tape is—and will continue to be—a perfect technology for reading and writing large streams of data, such as those from satellite feeds, due to its fast streaming, sequential nature. But legacy tape does not typically perform nearly as well when dealing with small blocks of data. However, some of the new tape offerings have technology that helps it keep the tape streaming. For example, LTO Ultrium 3 tape technology has a large 128-MB buffer and data speed matching capability that helps keep it streaming with many applications.

The transportability of tape is a key criterion for various applications. Because tape can be easily removed from the data center and locked up in physically secure vaults, it will continue to play an important role for business continuity and archive situations where data needs to be protected, secured, and retained for many years.

The recent natural disasters that have impacted numerous companies underscore the need to store data offsite. Storing data at a remote location is imperative to ensure that the business will survive any disaster.

Today's tape technology is also highly reliable. LTO technology, for instance, has servo tracking mechanisms to help achieve accurate reads and writes of data, immediate read after write verification to help ensure recoverability, and advanced error correction code to detect and correct errors. Tape also has infinite capacity. With tape an organization can simply add more cartridges, not drives, to increase capacity stores.

Tape always has been—and continues to be—less expensive than disk.

Calculating the Costs—Tape vs. Disk

How much does tape or disk storage really cost? To answer this question, IT organizations must first determine how much tape or disk storage they *really* need.

As an example, let's say that an IT organization needs to store 125 TB of backup data and wants to compare the costs of storing this data on both tape and disk over time.

The customer wants to determine the initial cost of storing that data on these media types *as well as* the cost of storing that data over the long term (in this case, seven years).

Let's assume that data volumes for this IT organization are growing at 20% a year, which means that backup volumes are also increasing at about 20% a year.

Raw versus Usable Capacity

Even in the best scenario, disk and tape storage utilization is inherently limited; 100% utilization is not possible. In open-systems environments, a good rule of thumb for IT organizations is 85% utilization for tape media and 70% utilization for disk storage.

Step 1: Calculate Storage Requirements

The first step in determining how much capacity IT organizations really need is to calculate usable, not raw, storage capacities.

Using the 125-TB initial capacity requirement example above, the IT organization would require 150 TB of usable storage capacity at the end of the first year. This translates into 176.5 TB of raw tape capacity or 214.3 TB of raw disk storage, when utilization rates are factored in for both technologies (85% and 70%, respectively).

Assumptions:

- Initial Storage Requirement – 125 TB
- Annual Growth Rate – 20%
- Average Tape Cartridge Utilization – 85%
- Average Tape Compression – 2:1
- Average Disk Utilization – 70%

The following table compares the storage requirements for backing up to both tape and disk over the seven-year period.

Table 1: Calculating Storage Requirements for Disk and Tape

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7
Initial Data Stored	125 TB	150 TB	180 TB	216 TB	259.2 TB	311 TB	373.2 TB
New Data Added	25 TB	30TB	36 TB	43.2 TB	51.8 TB	62.2 TB	74.6 TB
Total Data Stored	150 TB	180 TB	216 TB	259.2 TB	311 TB	373.2 TB	447.8 TB
Disk Storage Requirements	214.3 TB	257.1 TB	308.6 TB	370.3 TB	444.3 TB	533.1 TB	639.7 TB
Tape Storage Requirements	176.5 TB	211.8 TB	254.1 TB	304.9 TB	365.9 TB	439.1 TB	526.8 TB

The conclusion is simple: Disk storage requirements are higher than tape storage requirements due to disk's lower utilization rates.

Step 2: Determine Hardware Configurations

The next step is to calculate the cost of the hardware for the tape and disk configurations.

For this particular example, we decided to calculate the cost of an LTO 3-based tape configuration.

We chose LTO technology, largely due to the breadth of its current acceptance among vendors of tape autoloaders and tape libraries.

For this example, the tape library is configured with four LTO 3 tape drives and 659 LTO generation 3 tape cartridges for a total capacity of about 527 TB compressed. An LTO 3 cartridge can store up to 800 GB of capacity with 2:1 compression.

Table 2: Purchase Price for Tape Solution

Hardware	Cost
1 LTO tape Library	$ 104,000
4 LTO 3 drives ($18,400 per drive)	$ 73,600
659 cartridges ($124.46 per cartridge)	$ 82,020
Total cost	$ 259,620

Note: *List pricing is used for all hardware configurations.*

On the disk side, we chose a readily available Serial ATA disk storage system. The total disk configuration had a total capacity of 644 TB. It consists of 23 disk controllers, all equipped with seven expansion units, for a total of 161 expansion units. Each disk controller (equipped with seven expansion units) had a capacity of 28 TB.

Table 3: Purchase Price for Disk Solution

Hardware	Unit Cost	Total Cost
23 disk controllers	$22,385	$ 514,855
161 expansion units	$10,999	$1,770,839
16 racks	$ 4,850	$ 77,600
Total		$2,363,294

Step 3: Calculate Environmental Costs

The next step is to calculate the environmental costs for the tape and disk configurations.

Floor space and electricity prices can vary widely across geographic regions of the country. For this example, we estimated the floor space cost to be $20/ft^2 per month and the electricity cost to be $0.10/kwh.

Also, since the disk configuration involved multiple expansion units, requiring eight racks of disk space, for the purpose of these calculations we stacked the disks on sixteen 41” x 25.5” racks

Assumptions:

Computer room cost per square foot – $20/ft^2

Cost of electricity – $0.10 kwh

Calculations:

Electricity is required to both power and cool the hardware. To calculate the yearly electrical cost of running the two systems, add the power and cooling requirements for each device.

The yearly electrical cost =

[[Power requirements (kwh) + cooling requirements (kwh] *x* the number of units] *x* 8760 (number of hours in a year) *x* $0.10 (the cost per kwh)

Total environmental cost =
Cost of footprint *x* 12 (number of months in a year) + yearly electrical costs

Table 4: Environmental Costs for All-tape Solution

	Footprint	Yearly cost	Power requirements Per unit	Cooling Requirement per unit	Electrical costs per year	Total Yearly Environmental Costs
LTO tape library	10.4 ft²	$2,496	1.6 kwh	1.6 kwh	$2,803	$5,299
4 LTO3[1] drive	0	$0	.017 kwh	.017 kwh	$119	$ 119
Total yearly costs		$2,496			$2,922	$5,418

Table 5: Environmental Costs for an All-disk Solution

	Footprint (16 racks)	Yearly cost	Power requirements Per unit	Cooling Requirement per unit	Electrical costs per year	Total Yearly Environmental Costs
23 Disk controllers	116.2 ft²	$27,888	.33 kwh	.39 kwh	$14,507	$42,395
161 Expansion units	0	$0	.33 kwh	.39 kwh	$101,546	$101,546
Total yearly costs		$27,888			$116,053	$143,941

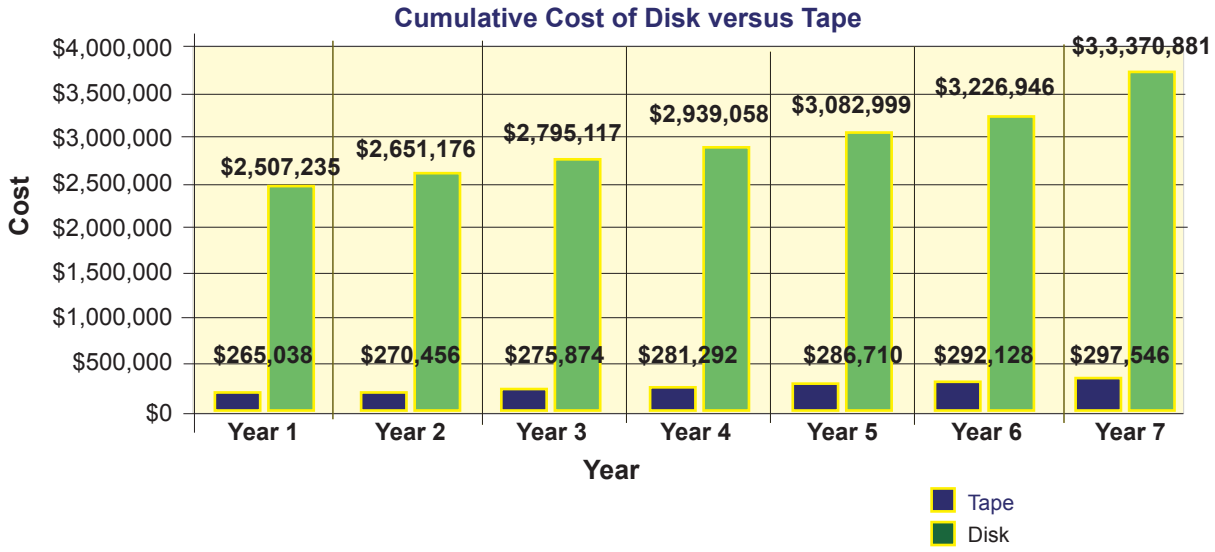
Step 4: Calculate the Seven-Year Total

Table 6: Yearly Cost for Disk and Tape Solutions for the First Seven Years

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Total
Tape	$ 265,038	$ 5,418	$ 5,418	$ 5,418	$ 5,418	$ 5,418	$ 5,418	$ 297,546
Disk	$2,507,235	$143,941	$143,941	$143,941	$143,941	$143,941	$143,941	$ 3,370,881

[1] LTO 3 drives consume 29 watts of power when streaming (writing) data and 13 watts when idle. For this exercise, we assumed the drives were streaming 25% of the time and idle the rest of the time.

Chart 1: The Cumulative Cost for Disk and Tape Solutions over Seven Years



Some Caveats

The cost calculations above represent a simplified view of the total cost of a seven-year disk or tape solution. However, one of the purposes of the exercise is to highlight the effects of environmental costs on long-term storage solutions.

A more complete calculation would also include:

- The cost of maintenance after the warranty period expires. Warranty periods differ among vendors and across products.
- The cost to replace disk and tape after several years (if that was the typical buying pattern of the IT organization). Some IT organizations replace disk every two to three years and tape drives every five to seven years.
- The cost of phasing in equipment purchases. It is not necessary to purchase all equipment in the first year. Not all of the disk controllers or tape drives and cartridges would need to be purchased the first year.
- The cost associated with backing up data to local or remote tape or disk. This exercise assumes that only one copy of backup data exists and that additional backup copies are not stored at remote locations for disaster-recovery purposes.
- The additional cost of mirroring disk (RAID-1). Mirroring disk doubles the cost of the disk solution. Other forms of RAID protection, such as RAID-5, can provide protection against drive failures and require some additional disk capacity without doubling the cost of storage.

Backing Up Data—Which Media to Choose?

In the example above, the tape solution turns out to cost about 9% of the comparable-capacity disk solution, or to say it another way, the disk system is over 11 times more expensive than the tape solution, factoring in the initial purchase of the systems as well as environmental costs.

Therefore, if low cost is the IT organization's highest priority, tape is the logical choice. However, if quick access is more important than the overall cost of the solution, disk may be the better choice.

Organizations today are implementing long term data archive policies to address internal and external data retention requirements. As this analysis shows, to keep data stored on disk for long periods of time could be quite costly and could raise the risk of data protection exposures. An increasingly popular option would be to implement a hybrid disk/tape solution. In this scenario, still using the same example, 25% of the backup data would be stored on the higher-performing Serial ATA disk over the seven year period and the remaining 75% would be kept on the LTO 3 tape. As the data on disk ages and is accessed less frequently it could then be moved to tape storage to take advantage of tapes lower costs and data protection attributes.

Table 7: Disk and Tape Storage Capacities

	Raw	Useable
Disk capacity	160 TB	112 TB at 70% utilization
Tape capacity	395.3 TB	336 TB at 85% utilization
Total capacity	555.3 TB	448 TB

Table 8: Cost of 25% Disk Storage

Hardware	Unit Cost	Total Cost
6 Disk controllers	$22,385	$ 134,310
40 Expansion units	$10,999	$ 439,960
4 Racks	$ 4,850	$ 19,400
Total		$ 593,670

Table 9: Cost of LTO tape library

Hardware	Cost
1 LTO tape library	$104,000
2 LTO3 drives	$ 36,800
494 Cartridges	$ 61,483
Total cost	$202,283

 76 Northeastern Blvd. Suite 29A, Nashua NH 03062 Phone: **877.463.6456** Fax: **603.563.8282**

Table 10: Environmental Costs for 25% Disk Solution

	Footprint (4 racks)	Yearly cost	Power requirements Per unit	Cooling Requirement per unit	Electrical costs per year	Total Yearly Environmental Costs
6 Disk Controllers	29 ft^2	$6,960	.33 kwh	.39 kwh	$3,784	$10,744
40 Expansion Units	0	$0	.33 kwh	.39 kwh	$25,229	$25,229
Total yearly costs		$6,960			$29,013	$35,973

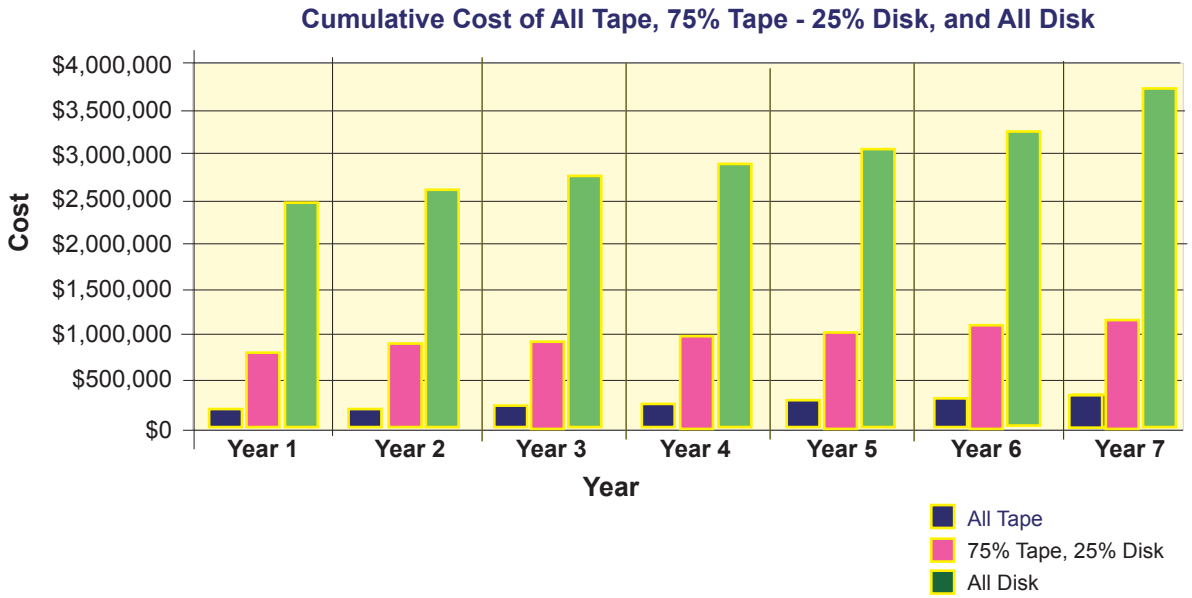
Table 11: Environmental Costs for 75% Tape/25% Disk Solution

	Footprint	Yearly cost	Power requirements Per unit	Cooling Requirement per unit	Electrical costs per year	Total Yearly Environmental Costs
LTO tape library	10.4 ft^2	$2,496	1.6 kwh	1.6 kwh	$2,803	$5,299
2 LTO 3 drives	0	$0	.017 kwh	.017 kwh	$60	$60
Total yearly costs		$2,496			$2,863	$5,359

Table 12: Seven-Year Cost of Disk/Tape Solution

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Total
Tape	$207,642	$ 5,359	$ 5,359	$ 5,359	$ 5,359	$ 5,359	$ 5,359	$239,796
Disk	$629,643	$35,973	$35,973	$35,973	$35,973	$35,973	$35,973	$845,481
Total	$837,285	$41,332	$41,332	$41,332	$41,332	$41,332	$41,332	$1,085,277

Chart 2: Cumulative Cost for Three Solutions
All-tape, 75% Tape and 25 % Disk, and All-disk—over Seven Years



The total cost of the hybrid disk/tape solution, which stores 25% of the backup data on disk and the remaining 75% on tape, is significantly less than the all-disk option, and it has room to grow. The tape library can accommodate an additional 218 cartridges as users' capacity requirements increase.

However, it is important to note that the caveats explained above also apply to these calculations. This example is only meant to give users a general idea of the long-term costs of storing data using three different configurations.

What If Electrical Costs Increase?

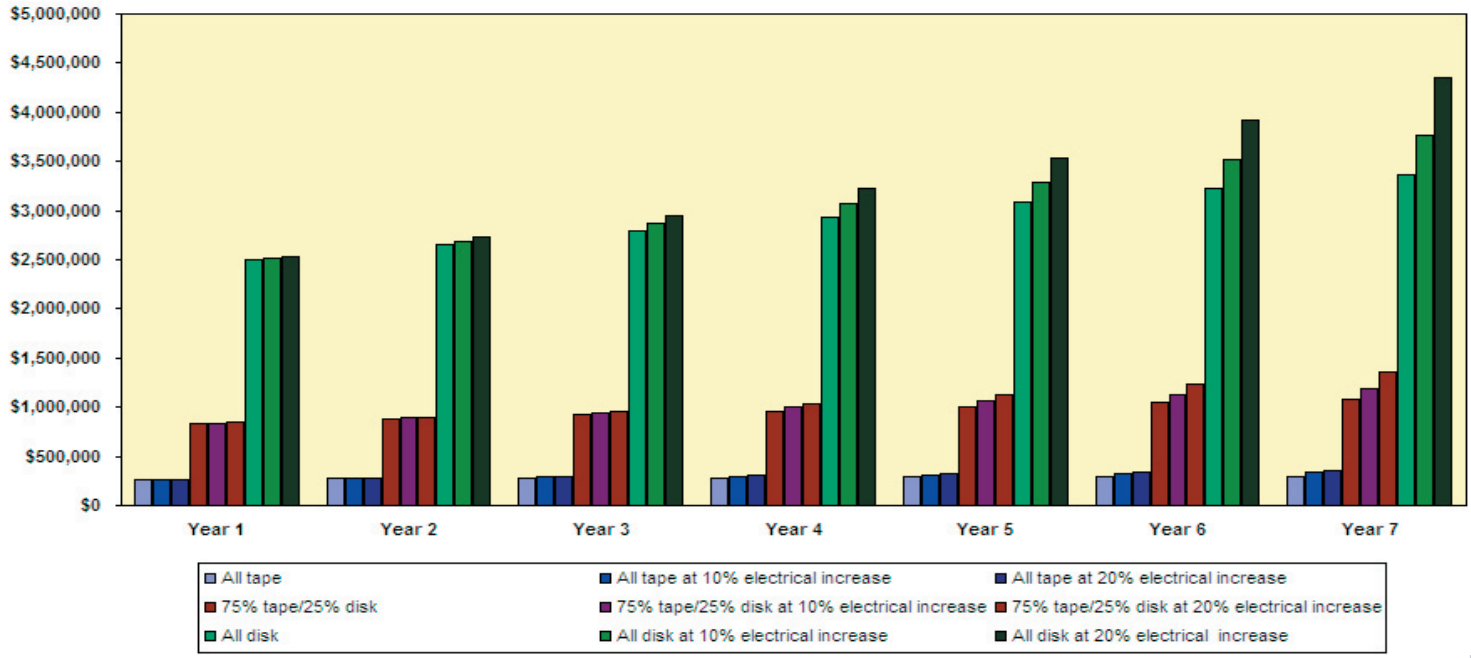
To calculate the cost of cooling and powering the equipment, we used a conservative figure of ten cents per kilowatt hour. However, in certain areas of the United States, the average cost per kilowatt hour is higher. In fact, according to a recent report by the United States Energy Information Administration, commercial facilities in New England are paying about 11.57 cents in 2005, a ten percent increase over 2004. Companies in Hawaii have a higher burden—their cost of electricity in 2005 was a whopping 17.57 cents, again more than a 10% increase over the previous year.

Of course, not all states have such high electrical rates. Idaho, for example, has one of the lowest rates—5.3 cents. Even those states with lower energy costs experienced an average increase of five percent over last year's costs.

Most energy experts would agree that energy costs will continue to increase over the next few years. These increased energy costs will increase the costs of keeping equipment powered and cooled in the data center.

So what happens if electrical costs continue to rise? Disk drives consume more power than tape drives for one simple reason—they are continuously powered up, spinning and generating heat. Tape cartridges, on the other hand, have no special power requirements since they remain idle in the library most of the time. Rising energy costs will make the costs of tape *even more attractive* than disks.

Chart 3: Cumulative Cost for Three Solutions—All-tape, 75% Tape and 25 % Disk, and All-disk Solutions—over Seven Years with 10% and 20% Increases in Electrical Costs



In the preceding chart, we calculated the cost of the three solutions—the all-tape solution, the all-disk solution and the 75% tape/25% disk solution—and then calculated the effect of 10% and 20% yearly increases in electrical costs over seven years.

If energy costs increase about 10% each year for the next seven years, then the all-tape solution will cost about $36,000 more than originally planned. The 75% tape/25% disk solution will cost about $55,000 more. The all-disk solution will cost almost $400,000 more! If energy costs increase at 20% per year, then the all-tape solution will cost over $63,000 more than originally estimated, the hybrid tape/disk solution will cost about $271,000 more, while the all-disk solution costs increase by almost one million dollars ($986,344 to be exact).

If energy costs escalate to a 20% annual increase, then placing 75% of the data on tape can ease the bite of high electrical bills. That solution could save over $714,000 in electrical costs over an all-disk solution.

While no one looks forward to rising electrical costs, their impact on the budget must be considered when evaluating storage purchases.

The Bottom Line

Which configuration you choose—all tape, all disk, or a disk/tape hybrid—really depends on access requirements (i.e., how quickly you want to retrieve data), the amount of data backed up annually, business continuity/data protection strategies, and budgetary restrictions (how much you're willing to spend on backup media).

If business service level agreements (SLAs) state that all data must be restorable within minutes, then disk may be your preferred option (doing replication between two disk systems). However, if the majority of your data has more lenient recovery requirements, then implementing a costly pure-disk system probably doesn't make sense.

A better option might be a tiered storage hybrid/tape solution. It allows for fast retrieval of some data from disk. As the data on disk ages and is accessed less frequently it could then be moved to tape storage keeping overall storage costs down by utilizing tape while also taking advantage of the data protection benefits inherent in tape. In a sense, it may offer the best of both worlds—fast retrieval from disk and cost-effective data protection with tape storage. When all is said and done, tape and disk are complementary. Tape storage is not dying. Tape is alive and well and is a vital part of a balanced storage hierarchy.■