White Paper

A Comparative TCO Study: VTLs and Physical Tape

With a Focus on Deduplication and LTO-5 Technology

By Mark Peters

February, 2011

Contents

Contents 2
LTO Ultrium Generation 5 and Deduplication TCO Analysis 3
VTL with Deduplication Compared with LTO Ultrium Generation 5 Technology 3
Market Overview 3
LTO-5 Technology and VTL with Deduplication in this TCO Study 4
Scenario 1: Comparing a "Deduped VTL" with an LTO-5 Tape Library 5
Scenario 2 Suite: Disaster Recovery Scenarios 5
Scenario 2a 6
Scenario 2b 7
Scenario 2c 8
Summary Comparative Costs for Year 1 9
The Impact of Future LTO Generations 10
The Bigger Truth 10
Appendix 11
Assumptions for TCO Analysis 11
General and Hardware 11
A Note on the Extent and Frequency of Backups 11
Software and Maintenance 12
Personnel 12

LTO Ultrium Generation 5 and Deduplication TCO Analysis

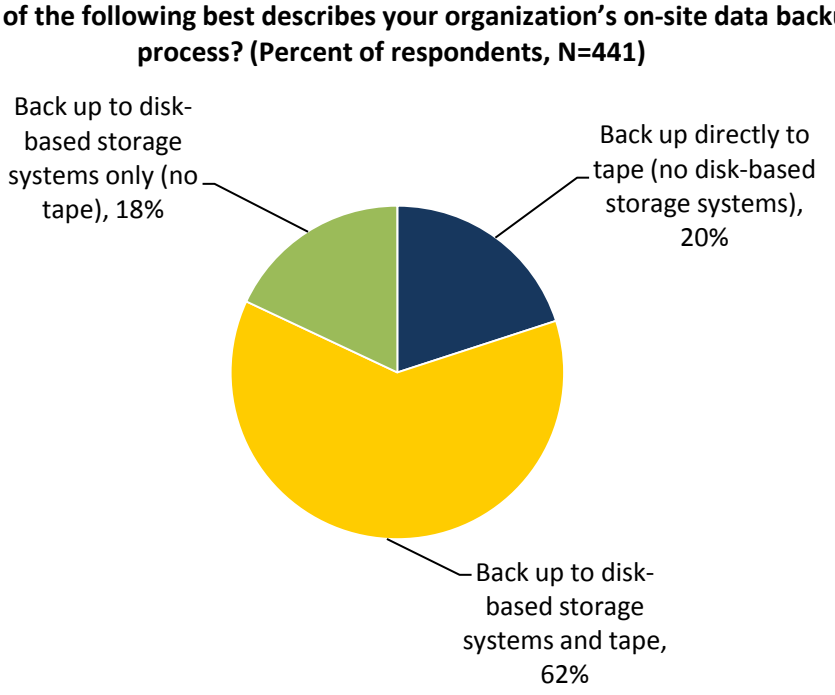
VTL with Deduplication Compared with LTO Ultrium Generation 5 Technology

Market Overview

The magnetic tape industry is changing. In recent years, tape has been focused on the data center and the middle-to-high end of the SMB (small to medium business) market with decreasing penetration below this level. Disk and tape markets have been shifting roles as disk encroaches on tape's traditional backup/recovery role, first by using faster disk arrays that appear as tape libraries—virtual tape libraries (VTLs)—and then with disk-based deduplication to reduce the amount of storage needed. Tape is still utilized by many businesses for backup and low access applications and is positioning itself to compete in the fast growing tier-3 storage opportunities that include fixed content, compliance, and long-life archival applications.

Tape remains a viable economic part of the storage hierarchy due in part to its lower cost per GB, lower operating expenses, and lower energy costs. It provides a greener storage solution than disk drives and is suitable for on- and off-site storage of several generations of backup and archival data. Tape technology has added security features including encryption, WORM, and greatly enhanced media life and reliability, with media expectations of up to 30 years or more. The notion that "tape is dead" ignores the substantial evidence that favors tape as a lower cost, environmentally friendly removable medium well suited for offline data protection and high growth compliance, fixed content, and archiving applications. Indeed, recent ESG research demonstrates that tape is still a significant component in the overall data storage universe: Figure 1 shows that despite broad adoption of disk-to-disk as a part of the backup process, tape is still involved in the vast majority (82%) of onsite backup approaches. Nor is this involvement declining as rapidly as many commentators would suggest: Figure 2 shows the percentages of all backup data stored on various media. For tape (on and offsite) plus offsite third parties (which can reasonably be assumed to involve tape to some degree), the sum total is 36% for 2010. By 2012, users expect the share for the same combination of media to remain at 34%.[1]

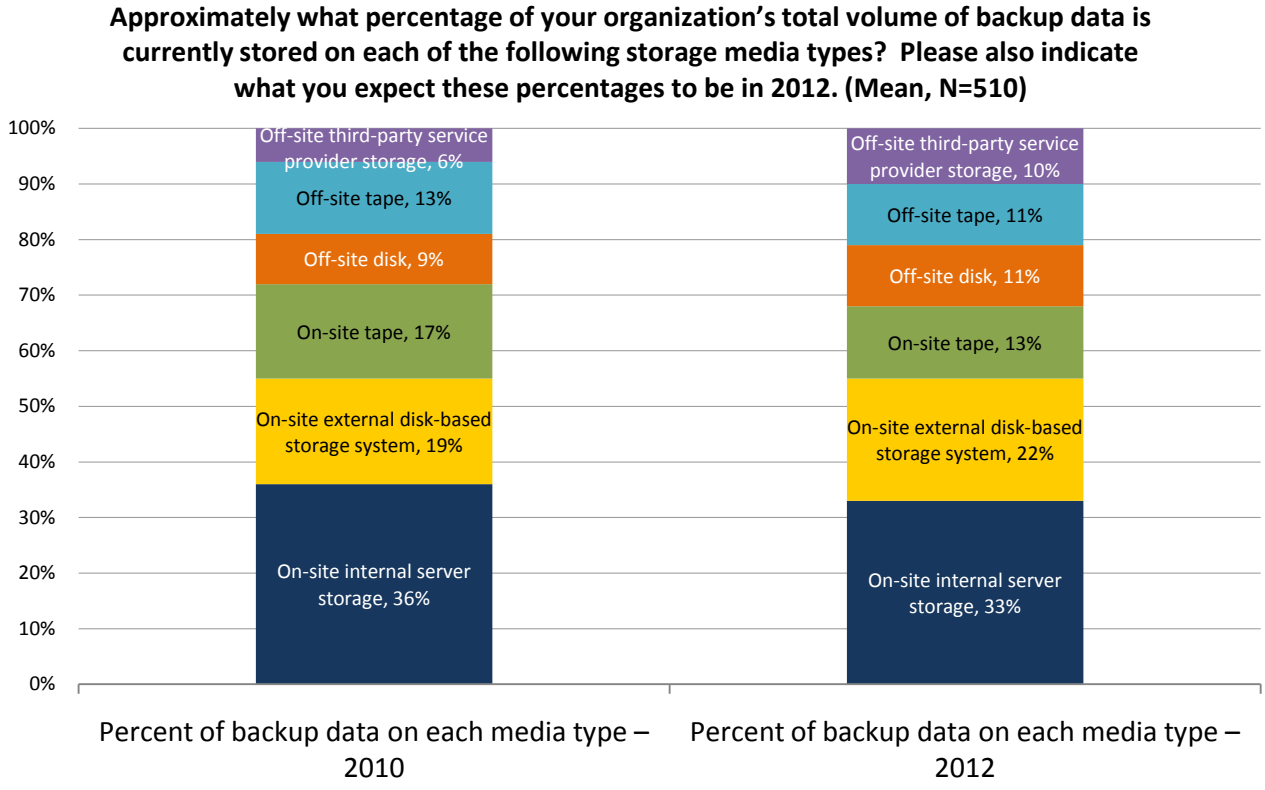
Figure 1. The Majority of On-site Data Backup Processes Involve Tape



Source: Enterprise Strategy Group, 2010.

[1] Source: ESG Research Report, *2010 Data Protection Trends*, April 2010.

Figure 2. Percentages of Total Organizational Backup Data Stored on Various Media Types



Source: Enterprise Strategy Group, 2010.

LTO-5 Technology and VTL with Deduplication in this TCO Study

The LTO-5 technology specifications were announced on January 19, 2010 and raised the bar for tape as a viable backup and archival solution by offering native data rates of up to 140 MB/sec. and a native capacity of 1.5 TB. LTO-5 tape also enhances tape's ability to meet the needs of the archive, disaster recovery, and business continuity markets with the announcement of LTFS (Linear Tape File System), which will work on LTO generation 5 technology. LTFS is a true file system for tape, allowing tape data to be accessed in a manner like disk or removable media, including directory tree structures and drag-and-drop capability.

In recent years, deduplication has become a popular backup methodology as it eliminates redundant (duplicate) data and can significantly shrink storage requirements and improve bandwidth efficiency. Because the price of primary disk storage steadily decreases over time, businesses frequently store multiple versions of the same information without deleting redundant data. In particular, the backup application stores extremely redundant information—much of the exact same data that is backed up today has already been backed up previously. Several deduplication products are available today and many are integrated with a VTL, a disk array that appears to the system as a tape library. This report will use standard industry specifications to describe the deduplication system and the tape library with LTO-5 drives in all scenarios. Like deduplication, VTLs can complement tape in the backup role.

This report will address the TCO (total cost of ownership) aspects of the newly announced LTO Ultrium Generation 5 (LTO-5) tape drives and tape libraries compared to a selection of deduplication scenarios over a five-year timeframe. The initial amount of disk capacity at the primary site assumption will start at 35 TB; 52% (the industry average allocation for non-mainframe systems) is allocated to live data, leaving 18 TB of data to physically be backed up. List prices with a 30% discount are used for consistency to determine the ASPs for tape and deduplication hardware.

Scenario 1: Comparing a "Deduped VTL" with an LTO-5 Tape Library

This scenario (see Table 1) compares the five-year TCO of a deduplication appliance with an LTO-5 compatible tape library; it assumes that they are both in a primary data center and used for backup.

Table 1. TCO Details – Deduped VTL Compared with an LTO-5 Tape Library

Scenario 1							
LTO-5	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$60,157	$26,744	$31,444	$37,472	$46,072	$201,889	35.9%
Software	$130,924	$53,664	$53,664	$53,664	$53,664	$345,580	61.4%
Personnel	$3,125	$3,125	$3,125	$3,125	$3,125	$15,625	2.8%
One-time Charge	$0	$0	$0	$0	$0	$0	0.0%
Total	$194,206	$83,533	$88,233	$94,261	$102,861	$563,094	100.0%
VTL/Deduplication	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$326,287	$52,070	$109,088	$60,522	$60,552	$608,519	60.9%
Software	$146,325	$53,664	$53,664	$53,664	$53,664	$360,981	36.1%
Personnel	$6,094	$6,094	$6,094	$6,094	$6,094	$30,470	3.0%
One-time Charge	$0	$0	$0	$0	$0	$0	0.0%
Total	$478,706	$111,828	$168,846	$120,280	$120,310	$999,970	100.0%
Cost Difference	($284,500)	($28,295)	($80,613)	($26,019)	($17,449)	($436,876)	

Summary

The TCO for the VTL with deduplication is 2.46 times greater than the LTO-5 tape library in year one. By year five, the TCO for the VTL deduplication solution is 1.78 times greater than the tape library. This relative difference is because the tape library will have been upgraded with expansion modules (each containing 92 additional tape slots with added media) and two additional LTO-5 drives are added during the period. The VTL with deduplication is upgraded once in this period (in year three) to support a higher raw capacity for the data ingest.

Bottom Line

In a head-to-head comparison, the VTL with deduplication TCO ranges from 1.78 to 2.46 times higher than the LTO-5 tape library TCO. In all cases, the LTO-5 library has a lower TCO and is $436,876 less expensive than the VTL deduplication appliance over five years.

Scenario 2 Suite: Disaster Recovery Scenarios

Scenario 2 adds various remote disaster recovery components to the basic VTL and tape library comparison and is split into three possible implementations: a, b, and c.

- 2a: Replicated VTLs compared to tape library.
- 2b: Replicated VTLs compared to tape library with tapes transported offsite by truck.
- 2c: VTL replicated to offsite tape library via WAN compared to tape library with tapes transported offsite by truck.

Scenario 2a

This scenario (see Table 2) is an analysis of a VTL with a second replicated VTL compared to a tape library—the local VTL deduplication appliance will use the remote replication feature to connect to another VTL over a WAN for DR instead of a tape library. This scenario using two VTLs is then compared to the tape library as in Scenario 1.

Table 2. TCO Details – Replicated VTLs Compared to Tape Library

Scenario 2a							
LTO-5 Library	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$60,157	$26,744	$31,444	$37,472	$46,072	$201,889	35.9%
Software	$130,924	$53,664	$53,664	$53,664	$53,664	$345,580	61.4%
Personnel	$3,125	$3,125	$3,125	$3,125	$3,125	$15,625	2.8%
One-time Charge	$0	$0	$0	$0	$0	$0	0.0%
Total	$194,206	$83,533	$88,233	$94,261	$102,861	$563,094	100.0%
VTL-VTL WAN	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$656,571	$108,137	$222,173	$125,102	$125,102	$1,237,085	76.0%
Software	$146,325	$53,664	$53,664	$53,664	$53,664	$360,981	22.2%
Personnel	$6,094	$6,094	$6,094	$6,094	$6,094	$30,470	1.9%
One-time Charge	$0	$0	$0	$0	$0	$0	0.0%
Total	$808,990	$167,895	$281,931	$184,860	$184,860	$1,628,536	100.0%
Cost Difference	($614,784)	($84,362)	($193,698)	($90,599)	($81,999)	($1,065,442)	

Summary

The TCO for the VTL to VTL connection over a WAN is 4.16 times more expensive than the TCO for the LTO-5 library in year one and 2.89 times greater over the five year period. This scenario uses two VTLs, which of course adds significantly to the VTL cost.

Bottom Line

Because of the significantly higher initial hardware costs associated with purchasing a second VTL for DR, the TCO for the LTO-5 library approach is significantly less than adding a second VTL.

Scenario 2b

In this implementation, there is a replicated VTL compared to a tape library with truck—the oft-quoted "PTAM" or "Pickup Truck Access Method"! The local VTL deduplication appliance will use the remote replication feature to connect to another VTL for DR instead of using a tape library. This scenario (shown in Table 3) is then compared to using the tape library as in Scenario 1, but also using a truck to move tapes to an offsite location.

Table 3. TCO Details – Replicated VTLs Compared to Tape Library with Tapes Transported Offsite by Truck

Scenario 2b							
LTO-5 Truck	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$60,157	$26,744	$31,444	$37,472	$46,072	$201,889	35.9%
Software	$130,924	$53,664	$53,664	$53,664	$53,664	$345,580	61.4%
Personnel	$3,125	$3,125	$3,125	$3,125	$3,125	$15,625	2.8%
One-time Charge	$4,600	$4,600	$4,600	$4,600	$4,600	$23,000	4.1%
Total	$198,806	$88,133	$92,833	$98,861	$107,461	$586,094	104.1%
VTL-VTL WAN	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$656,571	$108,137	$222,173	$125,102	$125,102	$1,237,085	76.0%
Software	$146,325	$53,664	$53,664	$53,664	$53,664	$360,981	22.2%
Personnel	$6,094	$6,094	$6,094	$6,094	$6,094	$30,470	1.9%
One-time Charge	$0	$0	$0	$0	$0	$0	0.0%
Total	$808,990	$167,895	$281,931	$184,860	$184,860	$1,628,536	100.0%
Cost Difference	($610,184)	($79,762)	($189,098)	($85,999)	($77,399)	($1,042,442)	

Summary

The TCO for the VTL to VTL connection over a WAN is 4.06 times more expensive than the TCO for the LTO-5 library that uses a truck to move backup tapes offsite in year one. The offsite truck and storage services used are priced at $450 per month (see appendix for all assumptions). Over the five year period, the TCO for the VTL to VTL connection over a WAN is 2.78 times greater than the combination of a LTO-5 tape library and truck to offsite location.

Bottom Line

Adding remote tape storage via a truck to transport tapes to a remote storage facility adds relatively little to the TCO for the tape library solution. The LTO-5 library solution remains significantly less than the VTL to VTL deduplication solution in this scenario.

Scenario 2c

This scenario (shown in Table 4) will compare a VTL combined with a remote offsite tape library over a WAN to an LTO-5 library using an offsite LTO cartridge storage facility with no tape hardware installed at the storage facility (as in Scenario 2b). In the event of a recovery action that requires offsite services in this scenario, the VTL approach will use a WAN to deliver data back to the primary site. The offsite tape storage facility will need to use a truck to deliver tapes back to the primary site.

Table 4. TCO Details – VTL Replicated to Offsite Tape Library via WAN Compared to Tape Library with Tapes Transported Offsite by Truck

Scenario 2c							
LTO-5 Truck	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$60,157	$26,744	$31,444	$37,472	$46,072	$201,889	35.9%
Software	$130,924	$53,664	$53,664	$53,664	$53,664	$345,580	61.4%
Personnel	$3,125	$3,125	$3,125	$3,125	$3,125	$15,625	2.8%
One-time Charge	$4,600	$4,600	$4,600	$4,600	$4,600	$23,000	4.1%
Total	$198,806	$88,133	$92,833	$98,861	$107,461	$586,094	104.1%
VTL-WAN-Tape	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**	**%**
Hardware	$390,442	$82,812	$137,126	$102,022	$110,622	$823,024	50.5%
Software	$146,325	$53,664	$53,664	$53,664	$53,664	$360,981	22.2%
Personnel	$9,219	$9,219	$9,219	$9,219	$9,219	$46,095	2.8%
One-time Charge	$0	$0	$0	$0	$0	$0	0.0%
Total	$545,986	$145,695	$200,009	$164,905	$173,505	$1,230,100	75.5%
Cost Difference	($347,180)	($57,562)	($107,176)	($66,044)	($66,044)	($644,006)	

Summary

The TCO for the VTL with tape library connection over a WAN is 2.75 times more expensive than the TCO for the LTO-5 library using a truck to move backup tapes offsite in year one. Offsite truck and storage services used are priced, as before, at $450 per month. Over the five year period, the TCO for the VTL to the LTO-5 tape library connection over a WAN is 2.08 times greater than the LTO-5 tape library and truck to offsite location combination.

Bottom Line

Replacing the remote VTL via WAN with an LTO-5 tape library improves the TCO by 32%, but still leaves the VTL/tape solution 2.76 times more expensive in year one and 2.1 times more by year five. Replacing the remote VTL with the tape library saves nearly $400K in remote VTL costs over the five-year period. Adding remote tape storage via a truck to transport the tapes to a remote storage facility adds relatively little to the TCO for the tape library solution.

Summary Comparative Costs for Year 1[2]

The table below is the "technology cost summary" for the hardware, maintenance, environmental, software, personnel, and one-time charges for the first year, including list price comparisons. A year-two summary would have the maintenance and operations costs, but not the acquisition costs.

Table 5. Comparative Technology Cost Summary for Year 1

	LTO-5		VTL/Deduplication		LTO-5 Truck		VTL-VTL WAN	
Component	Annual Cost	Percent	Annual Cost	Percent	Annual Cost	Percent	Annual Cost	Percent
Hardware								
1. Product Acquisition	$ 46,913	78.0%	$274,217	84.0%	$ 46,913	78.0%	$552,432	84.1%
2. Annual Maintenance	$ 11,098	18.4%	$ 49,359	15.1%	$ 11,098	18.4%	$ 98,718	15.0%
3. Environment								
a. Floor Space	$ 591	1.0%	$ 554	0.2%	$ 591	1.0%	$ 1,108	0.2%
b. Power & Cooling	$ 1,556	2.6%	$ 2,156	0.7%	$ 1,556	2.6%	$ 4,313	0.7%
Total	$ 60,157	31.0%	$326,287	68.2%	$ 60,157	30.3%	$656,571	81.2%
Software								
1. Product Acquisition	$ 130,924	100.0%	$146,325	100.0%	$130,924	100.0%	$146,325	100.0%
2. Annual Maintenance	$ -	0.0%	$ -	0.0%	$ -	0.0%	$ -	0.0%
Total	$ 130,924	67.4%	$146,325	30.6%	$130,924	65.9%	$146,325	18.1%
Personnel								
1. Operations	$ 3,125	100.0%	$ -	0.0%	$ 3,125	100.0%	$ -	0.0%
2. Storage Management	$ -	0.0%	$ 6,094	100.0%	$ -	0.0%	$ 6,094	100.0%
2. Network Management	$ -	0.0%	$ -	0.0%	$ -	0.0%	$ -	0.0%
Total	$ 3,125	1.6%	$ 6,094	1.3%	$ 3,125	1.6%	$ 6,094	0.8%
One-Time-Charges								
1. Off Site Storage Fee	$ -	0.0%	$ -	0.0%	$ -	0.0%	$ -	0.0%
2. Storage Media Purchase	$ -	0.0%	$ -	0.0%	$ -	0.0%	$ -	0.0%
2. Other Fees	$ -	0.0%	$ -	0.0%	$ 4,600	100.0%	$ -	0.0%
Total	$ -	0.0%	$ -	0.0%	$ 4,600	2.3%	$ -	0.0%
Grand Total	**$ 194,206**		**$478,705**		**$198,806**		**$808,990**	**100.0%**

[2] Includes product list prices.

The Impact of Future LTO Generations

Although outside the strict remit of this paper, it is worth pointing out that LTO technology has had a consistent history of improvement in its specifications and capabilities. There have been new generations—and hence capacities and per GB costs—roughly every 2-3 years. During the timescales evaluated in this paper, LTO-6 technology should become available. Information made public by the LTO program states that its intended key specifications will include

- Native capacity of 3.2 TB compared to LTO-5's 1.5 TB.
- A compressed capacity of 8 TB (nearly 3X LTO-5's 3 TB). The assumed compression ratio increases from 2:1 to 2.5:1 due to a larger compression history buffer.
- Data rates of up to 210 MB/sec (native) and 525 MB/sec (compressed).
- Generation 7 and 8 of LTO have also been outlined, so the technology clearly has "longevity."

Clearly, LTO-6 technology (making some reasonable assumptions about pricing based on historical precedence) would represent a reduced $/TB. The "march of progress" would make overall LTO adoption and TCO even more attractive for new users at that stage or existing LTO users that want to add to their infrastructure at some point. Obviously, there would be no obligation for existing users to move up to LTO-6 products, and the broader costs of such a move are out of scope in this paper—the bottom line is further TCO improvement.

The Bigger Truth

The recently-released LTO-5 tape drive technology has increased tape capacity, reliability, and performance, taking tape to levels it has never attained before. That said, tape technology is faced with ever-increasing competition from disk-based solutions such as VTLs and deduplication for storage capacity optimization in backup applications. Such competition is often based on perceptions and entrenched opinions as much as on fact; as we have examined the costs of ownership of these two architectures in four different scenarios over a five-year timeframe in our model, the TCO of VTL with deduplication ranges from 1.78 to 4.16 times higher than the various implementations of the LTO-5 tape library. *This makes the tape option the least costly backup solution (TCO) to own over a five-year period by a substantial margin*. Of course, TCO isn't everything; other issues to consider (that are not part of this report) include the impact of each solution on the backup window and recovery time objectives. Like so many aspects of storage management, the decisions boil down to business needs: these needs typically include addressing performance, compliance, security, energy consumption, archive, data protection, and costs. A combination of disk and tape storage may be the optimal strategy for many users to address these varied needs. A VTL or disk-based backup can provide the performance needed for the recall of files for high access applications. As data backed up to disk becomes infrequently or never accessed, it should be moved to tape for long term retention. Tape technology can provide data security, compliance, and offline protection (against viruses, hackers, system errors, and so on) and a long term, low cost archive repository.

Appendix

Assumptions for TCO Analysis

As with any TCO model, the assumptions are crucial for users to be able to judge the accuracy and relevance of the model with respect to their particular situation. We believe our assumptions are "reasonable," but the full disclosure below will allow individual situations to be judged and users to make adjustments to their specific inputs and opportunities.

General and Hardware

- Initial raw storage capacity = 35 TB (average allocation level of 52% = 18 TB).
- A 35% compounded annual storage growth rate for the primary 18 TB occurrence.
- A full system backup to tape occurs every night (18 TB).
- Fill six (6) tape cartridges/night.
- Tapes are kept for two weeks before recycling.
- LTO-5 tape cartridge capacity = 3 TB at 2:1 compression, 140 MB/sec data rate.
- A total of five user tape pools and one scratch pool are used, no cartridge will be allocated over 70% (2.1 TB for LTO-5).
- Floor space cost = $75.00 per sq. ft.
- Power costs = 0.10 / KWH.
- ASPs (average selling prices) are calculated using 30% off list for the VTL deduplication appliance and 25% for the LTO-5 drives and tape library hardware.
- The LTO-5 tape library can scale from up to 18 Ultrium LTO-5 tape drive(s) and to just over 400 tape cartridges at full capacity.
- The model assumes that the media is not upgrade or replaced over the five-year timescale with LTO-5 being used throughout. This is good news for users as it illustrates the long life and extensibility of LTO tape before requiring added cost upgrades. Note: also see the "The Impact of Future LTO Generations" in the main document for information about LTO 6 and beyond.
- The VTL with deduplication can scale up to 96 TB of raw capacity and uses a 15:1 data reduction ratio.
- WAN pricing is based on half of a dedicated T3 line.

A Note on the Extent and Frequency of Backups

As mentioned above, this study assumes full nightly backups. Clearly, other assumptions could have been used, such as a 20% daily incremental change and a full system backup weekly. A nightly full system backup assumption is used because that is a very common practice for mid-sized and non-mainframe environments. Budget and staffing constraints have only encouraged this straightforward approach; incremental/differential backup techniques can often be viewed as double-edged-swords—on the one hand, reducing the backup load, but on the other hand, making recoveries more complex. Some users tend to want to keep things as simple as possible and are often hesitant to exploit backup or HSM type approaches to their optimum capabilities and fullest advantage. Overall, the financial impact of the chosen assumption is, in any case, limited; incremental/differential backups would use less tapes and would cause a lower impact (lower reduction ratios) for deduplication, meaning more disk would be needed. Again, as with all the assumptions in this report, the extent and frequency of backup is fully explained and open; any potential VTL or tape users reading this can make an adjustment for their specific situation.

Operational Considerations

- There is a noticeable jump in hardware costs for the VTL scenarios in year three: this is because the VTL was sized optimally without too much extra capacity in the beginning, but an expansion unit (more disk) was ultimately needed to accommodate the 35% annual storage growth.
- The tape hardware costs rise each year from years three to five, which is due to library expansion (more slots) for the 35% annual storage growth rate. Cartridges are purchased every year as needed.
- VTL power and cooling costs are 38% higher than the tape library; this is lower than in a non-dedupe VTL configuration because there are fewer spinning disks.

Software and Maintenance

- All backup software products were purchased during the first year.
- VTL hardware costs include the VTL-dedupe software costs as follows: basic dedupe software costs are included in the dedupe engine price; however, the specific replication software is only added to the cases where remote replication comes into play.
- Maintenance price, if not provided, is calculated at 15% of list price and is applied starting at the first year.

Personnel

- Two skill sets used in cost analysis (industry averages)
    1. Storage Administrator at annual salary of $65,000 averages 7.5% of time at storage management.
    2. Tape Operator at annual salary of $55,000 averages 5% of time at physical tape management.
- Corporate burden rate of 25%
- It is assumed that the VTL-dedupe system requires a slightly higher skilled person for storage management than the tape system because a disk storage administrator has a “somewhat higher skill level” than a tape operator (which is normally an entry level job) especially when dealing with recovery and BUR activities for more critical data. There isn't that much to do with automated tape libraries these days unless there are lots of enters/ejects or special tape media collection from a variety of users/clients/service bureaus.
- Offsite charges to remote storage facility
- Tape transit and offsite storage = $450 per month service fee.

Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel:508.482.0188 Fax: 508.482.0218 | www.enterprisestrategygroup.com