White Paper

Analyzing the Economic Value of LTO Tape for Long-term Data Retention

By Jason Buffington, Senior Analyst – Data Protection, and Adam DeMattia, Research Analyst

February 2016

This ESG White Paper was commissioned by the LTO Program and is distributed under license from ESG.

Contents

Introduction

Executive Summary

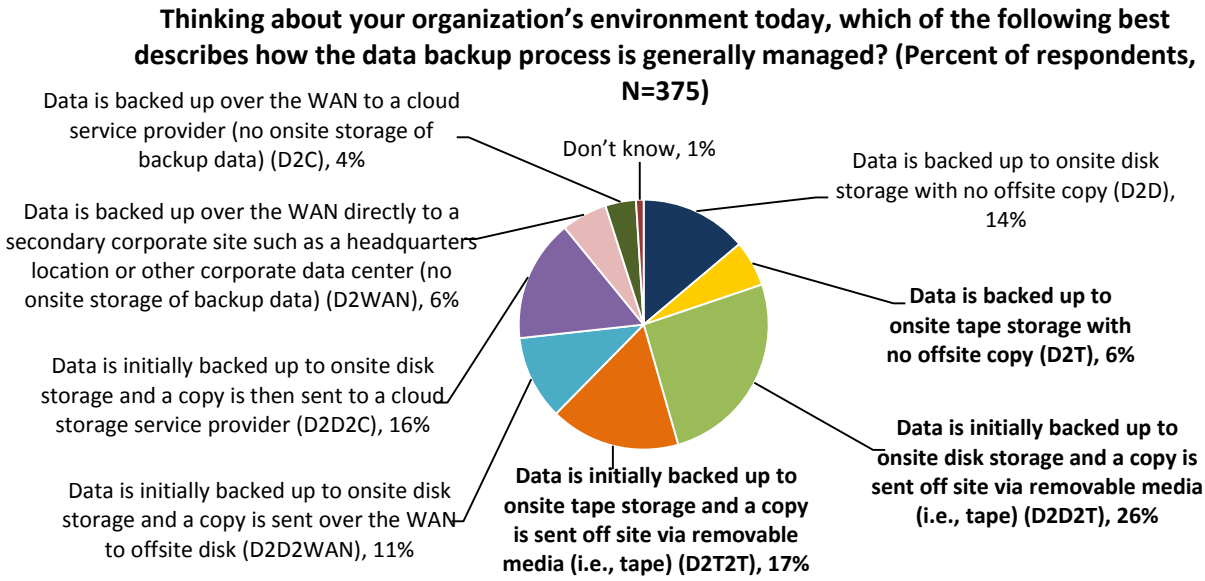
ESG was engaged by the Linear Tape-Open (LTO) Program, a consortium of LTO-based technology providers, to develop a detailed economic analysis of the LTO tape specification, which is the basis for many high-capacity, high-performance, and high-integrity tape storage products in production today. The analysis is designed to help IT organizations compare the fully burdened costs of leveraging modern LTO-based solutions for long-term data retention with a "present mode of operation" (PMO) that represents alternative all-disk storage approaches. This analysis builds upon ESG's evaluation of current-generation LTO products, the forward-looking LTO Ultrium technology roadmap, in-depth interviews with technical stakeholders at The LTO Program's member organizations (HPE, IBM, and Quantum), additional qualitative and quantitative market research conducted by ESG with IT decision makers, and ESG's general familiarity with data storage products and trends.

As discussed in the following pages, LTO solutions offer the opportunity for organizations to drastically reduce cost of ownership, avoiding significant solution capital expenditures, improving IT operational expenses, and improving user productivity compared with disk-based alternatives. **In fact, ESG's analysis of a typical large-scale data retention use case yields nearly $13.5M in estimated cost savings over a ten-year time horizon with an additional $400K in incremental user benefit delivered over and above what is expected with disk-based alternative.** The result is an impressive 577% return on investment over ten years when accounting both for avoided costs and net-new economic improvements. LTO-based solutions yield a compelling outcome for organizations struggling to cope with fast-growing data, reliable retention, and budgetary pressures. The selection of the proper storage platform for its requirements can have a significant impact on the financial success of an organization.

Market Overview

Many IT organizations need to discover what others already have: Modern tape is just that, "modern,"—with speeds, durability, features, and economic considerations that are undeniable. Recent ESG research shows that nearly half (49%) of surveyed organizations report relying on tape as some part of their primary backup ecosystem (see **Error! Reference source not found.**).[1]

Figure 1. How Data Backup Process Is Generally Managed



Source: Enterprise Strategy Group, 2016.

[1] Source: ESG Research Report, *2015 Trends in Data Protection Modernization*, September 2015.

While disk storage continues to be the most commonly leveraged media type when it comes to backup and recovery operations, less than a third of all organizations rely *solely* on disk. Instead, even when disk is meant for initial recoveries, tape is in use in 26% of longer-term retention within those environments. In fact, *tape is the exclusive media for data protection in 23% of surveyed organizations*.

Recognizing that both production and protection storage requirements continue to be a burden to storage and data protection professionals alike as evidenced by the fact that ESG's annual IT spending intentions survey consistently shows *managing data growth* to be among the most cited IT priorities,[2] organizations need to rethink their plan to simply scale what they have been doing for first-tier recoveries. As such, many need to rediscover what modern tape can offer. To help with this rediscovery, ESG Lab recently conducted a hands-on evaluation of modern LTO tape products to assess tape and its suitability in modern IT environments.

Excerpts of the ESG Lab Review on LTO Tape Technology

Contrary to misconceptions that tape is an unreliable, slow, and antiquated storage technology, ESG is pleased to report that LTO tape is thriving and has a bright future in organizations of all sizes, including some of the largest public cloud providers on the planet. An ESG Lab audit of LTO consortium reliability testing has confirmed that LTO-7 media and drives are orders of magnitude more reliable than enterprise-class disk drives. As a matter of fact, it's more than a billion times more likely that you will get hit by lightning in your lifetime than that you will encounter an unrecoverable LTO-7 media error.

ESG Lab also confirmed that LTO technology can be used to meet the performance needs of the most demanding streaming workloads (e.g., real-time geospatial data archiving and high-definition video production) with native, uncompressed, single tape drive LTO-7 performance of up to 300 MB/sec and the ability to get tens, or even hundreds, of tape drives working in parallel in one or more enterprise-class tape libraries.

While ESG believes that the reliability and tape modernization analyses presented in this report are vital considerations when making a long-term data retention technology decision, *the bottom line for most IT professionals is simple economics*. Nothing is more cost-effective, reliable, or energy-efficient for long-term data retention than a tape in a library slot or on a shelf, and it continues to play a key role for organizations across the globe.

For many IT strategies to be implemented, two equally important questions have to be answered:

- ***Will it work?*** To address that, ESG Lab tested modern LTO and reported on its findings.
- ***Can we afford it?*** That topic is covered in the remainder of this paper.

[2] Source: ESG Research Report, *2015 IT Spending Intentions Survey*, February 2015.

LTO Tape Technologies: Economic Value Analysis

Methodology

For a discussion of the research and modeling methodology ESG adhered to in the process of writing this report, please see Appendix A.

Economic Value Model Overview

As articulated in Appendix A, ESG's economic value analysis compares two scenarios: The first is an organization that elects to use LTO tape products to support its long-term data retention requirements. The second scenario is an all-disk storage approach that many customers currently take to meet their retention requirements. The basic profiles for each scenario follow:

- **LTO tape scenario:** In this scenario, the customer is using a properly sized, high-end tape library, housing tape drives and LTO-6 cartridges at the beginning of the time horizon, and transitions to LTO-7 after five years. The configuration also includes servers to run Linear Tape File System (LTFS) directory services and software costs for LTFS itself.
- **PMO scenario**: In this scenario, the customer is using a comparable set of hardware and software components based on tier-2 disk solutions, including properly sized disk arrays, nearline disks (accounting for RAID overhead), and archive software.

The specific differences ESG's economic value model financially quantified and compared between the two scenarios include:

- Relative costs associated with tape libraries, drives, media, and associated LTFS software versus disk arrays, HDDs, and associated archive software.
- Relative maintenance charges incurred over time as a function of the previous bullet's elements.
- Power and rack space costs of the configurations based on the size and complexity of the solutions required to meet the organization's requirements.
- The relative number and growth of administrators needed to manage the solution over time.
- Data migrations over time to account for the implementation of both new generations of LTO media (a minor transition) and forklift upgrades of disk-based solutions.
- IT administration tasks related to data retrieval requests. For the purposes of this analysis, considered equivalent in the LTO tape scenario and all-disk PMO. When considering other PMOs, like cloud-based data retention, cost improvements for data access and egress are significant for the LTO tape scenario.
- User productivity improvements related to data retrieval requests (e.g., the impact of data availability during retrieval events).
- Note that ESG's model considers data storage requirements, which grow over a ten-year time horizon.

Simply put: ESG's analysis estimates the likely costs and potential benefits of purchasing and operating both an LTO tape-based, large-scale data retention environment and a disk-based alternative.

Cost Categories

This ESG analysis considers eight cost categories: hardware, software, media, infrastructure (i.e., power costs and rack space opportunity costs), maintenance charges, staff costs, data migration costs, and data access costs. The sum of these categories equals the total cost of ownership (TCO) of each solution. See Appendix B for full details.

Benefit Categories

This ESG analysis considers one primary incremental benefit category: user productivity improvements during data retrieval events. Savings offered by the LTO-based solution over and above what is expected in the PMO are categorized as a net benefit.

Default Scenario

ESG developed a baseline profile of a hypothetical enterprise to illustrate the relative costs and benefits of LTO-based data retention compared with the PMO discussed in this report. For the purposes of this analysis, ESG tuned its assumptions to be representative of a growing enterprise with 1 PB of data to be retained and with 10% of retained data necessitating retrieval each month by the end-user community.

Moreover, to illustrate a dynamic environment with changing needs over time, ESG's model assumes the size of data is growing at a rate of 30% annually. This is an important assumption as, in the PMO scenario, a single storage administrator (with an unburdened salary of $65,000, $91,000 fully burdened) is assumed to be capable of managing 200 TB of storage. By comparison, much larger scale tape libraries can be managed by a single administrator. Data growth deepens the financial divide between the solutions over the ten-year time horizon.

To account for data compressibility on LTO tape media, ESG's model assumes a 2:1 compression ratio when sizing the LTO solution. By contrast, data is not assumed to be compressed on the PMO's storage media due to the tier of array likely to be selected in this use case. In fact, the PMO pays a 20% RAID overhead penalty when sized, effectively increasing the required raw capacity by 20%. Additionally, storage media—whether disk or tape—are likely to decrease in costs over time. To account for this fact, a 10% annual cost deflation assumption is built into the model, which assumes that the default cost per-TB for both solutions declines by 10% per year for incremental capacity required.

These and other key assumptions used in ESG's default scenario are summarized in Table 1.

Table 1. Key Default Scenario Assumptions for Long-term Data Retention

Parameter	Default Assumption
Time horizon of analysis	10 years
Initial data (usable capacity) retained (in TB)	1,000
Annual growth of data	30%
Percentage of retained data to be retrieved each month	10%
Raw disk capacity per HDD in PMO (TB)	1.2
Disk utilization as a percentage of raw after accounting for RAID in PMO	80%
Average cost per TB of nearline disk	$350
Number of TB of disk storage manageable for one administrator in PMO	200
Capacity per LTO-6 cartridge (TB)	2.5
Achievable compression with LTO storage	50%
Average cost per LTO-6 cartridge	$145
Number of IT FTEs required to manage retained data store for LTO environment	1
Annual burdened cost – typical IT administrator	US$91,000
Annual burdened cost – typical employee (application user)	US$140,000

Source: Enterprise Strategy Group, 2016.

Summary of Results

With the model parameters tuned to the default assumptions in Table 1, ESG's analysis concludes that the net benefits (avoided costs plus incremental benefits, less solution costs) of leveraging LTO-based storage solutions to support a large-scale long-term data retention use case far outweigh the costs. Table 2 shows the modeled return on investment (ROI), annual total cost of ownership (TCO), and annual incremental benefit over the time horizon for LTO solutions compared with a similarly sized disk-based alternative approach. The following section details the most compelling findings from this analysis as they relate to both the costs and benefits associated with the LTO solution.

Table 2. Economic Value Summary, LTO Versus Disk-based PMO for Long-term Data Retention

Scenario	Project ROI	Annual TCO (also Avoided Cost Benefit for Alternative Scenario)	Annual Incremental User Benefit
LTO Tape Scenario	577%	$241,110	$40,685
PMO Scenario	N/A	$1,590,960	N/A

Source: Enterprise Strategy Group, 2016.

Annual TCO

Annual TCO is the sum of all the cost categories included in the analysis (as outlined in Table 5), averaged over ten years. As displayed in Table 2, the annual TCO for the LTO tape scenario is estimated as $241,110, a significant 85% savings compared with the PMO.

For the purposes of ROI calculations, the $1,590,960 annual TCO estimated for the PMO scenario is also counted as a cost avoidance benefit for the LTO scenario. If the organization were to forgo an investment in LTO storage, it is feasible to assume an alternative investment—in this case, a disk-based storage solution—would be incurred. Thus, those avoided costs are counted as an additional benefit for the LTO investment scenario.

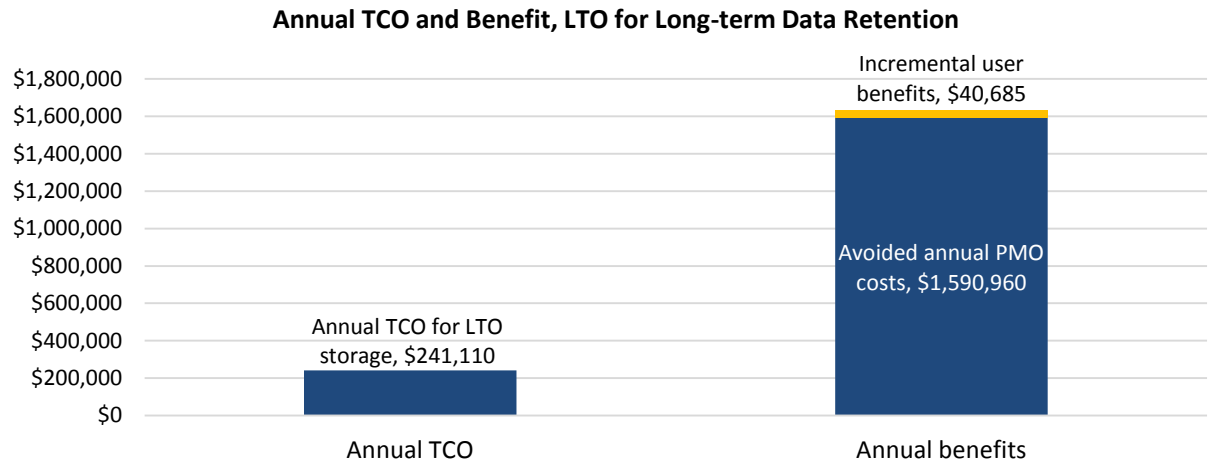
However, TCO should be only one part of the customer consideration when weighing different technology approaches. As shown in Table 2—and discussed in the following report section—the lower costs associated with LTO storage are augmented by benefits in the area of improved user productivity.

Annual Benefit

Although this analysis is dominated by material differences in the TCOs estimated for each solution, there are incremental benefits for the LTO scenario quantified by ESG's model, namely improvements in user productivity during retrieval events. By leveraging LTFS, organizations have the opportunity to allow for self-service retrievals of data, eliminating any back and forth with backup administrators and allowing users to access the data they need faster. To account for this potential benefit, ESG's model looks at the percentage of data to be retrieved each month and assigns a small per-GB charge to the PMO case. This charge represents the lost productivity to the user when submitting a request to an administrator and waiting for the administrator to deliver the data. It is important to note that ESG's model is very conservative: Even as the environment scales to multi-petabytes, with 10% of the data being retrieved monthly, the average value delivered to the organization annually is less than one-half of a fully-burdened user's cost—$40,685. For organizations fully embracing LFTS's capabilities, this value may in fact be much larger.

The modeled annual costs and benefits for the LTO scenario are depicted graphically in Figure 2.

Figure 2. Annual TCO and Benefit, LTO for Long-term Data Retention



Source: Enterprise Strategy Group, 2016.

ROI

ROI is a profitability ratio for investments. It is calculated by dividing the net benefits of an investment (i.e., the total benefits minus the associated costs) by the total cost of the investment. A positive ROI indicates that total benefits exceed the costs of the investment. As displayed in Table 2, the modeled ROI for utilizing LTO storage to retain a large data store using the inputs defined in Table 1 is 577%.

Quantifying Relevant Cost Differences

Economic models are, by definition, abstractions from reality. In any model, numerous estimates and assumptions must be made. ESG's research methodology leverages rigorous market research and in-depth interviews to estimate material differences between two approaches to satisfying data retention requirements, both in terms of how those solutions would be configured in different environments and the related differences in solution CapEx and OpEx over time. The next two sections discuss important estimates incorporated into ESG's EVV model.

Comparative Cost Analysis

For the hypothetical customer scenario described in Table 1, the estimated ten-year aggregate TCO for the LTO solution—compared with the TCO estimated for the PMO—is displayed in Table 3. As shown, the LTO scenario is expected to yield dramatically lower TCO over the time horizon.

Table 3. Ten-year TCO, LTO Versus Disk-based PMO for Long-term Data Retention

Category	LTO Solution	PMO
Hardware	$496,396	$3,893,830
Software	$210,000	$172,894
Media	$120,710	$2,159,881
Infrastructure	$527,547	$982,403
Maintenance and Support	$85,248	$1,705,537
Staff Personnel	$910,000	$6,779,500
Data Migration Costs	$20,284	$174,640
Data Access Costs	$40,915	$40,915
Total ten-year cost of ownership	**$2,411,099**	**$16,808,264**

Source: Enterprise Strategy Group, 2016.

Key TCO Assumptions Driving Cost Differences between LTO and the PMO

Several of the core cost assumptions ESG utilized in its model and analysis include:

- **Hardware:** Hardware costs are quantified for both scenarios. In the LTO scenario, this cost is representative of the tape library, drives, and associated LTFS hardware required in the scenario. Costs for tape cartridges fall into the media cost category. In the PMO, this cost is representative of the costs for arrays likely to be purchased in the scenario. While the hardware cost includes the drives needed to satisfy the initial storage environment, additional drive costs to accommodate storage growth over time are calculated on an annual basis for this time horizon.

  All costs, whether LTO or PMO, are determined using blended, publicly available pricing from a range of vendors. Additionally, all costs are depreciated using a five-year depreciation cycle applied in an annualized fashion over the full ten-year time horizon to account for technology refreshes likely to occur over such a long time horizon.

  The key driver determining the unit cost of the array or tape library to be configured is the number of drives or cartridges that must be supported in the hardware. For the LTO scenario, the model assumes a 2:1 data compression ratio; essentially, for each TB of data, only 500 GB of raw tape capacity is required. ESG believes this ratio to be typical, though it is important to note that over a ten-year time horizon—and depending on the data stored and the IT organization's ability to limit wastage—this ratio may vary somewhat from organization to organization and use case to use case.

  Additionally, utilizing LTO-6 specifications of 2.5 TB per cartridge for the first five years of the time horizon, and a doubling of per-cartridge capacity for the balance of the time horizon to represent LTO-7 specifications, the model estimates that the tape library and drives bought by the organization will need to accommodate 572 LTO-6 cartridges and 388 LTO-7 cartridges. Considering a tape library of this capacity, and blending costs across a range of LTO technology providers, ESG estimates hardware costs to be in the range of $496,396 over ten years for the LTO hardware to be selected by the organization. These costs also include three LTFS directory servers, one of which is modeled to be purchased for every 1 PB of raw LTO capacity (rounded down) at a unit cost of $16,800 in this model.

  In the PMO scenario, the assumptions are that compression is not able to be utilized on the tier storage system selected for long-term retention and that the required capacity overhead is 20% to account for RAID drive raw capacity requirements much higher than what is observed in the LTO scenario. The logic behind these assumptions should be clear to any storage administrator: RAID or mirroring requires, by definition, more effective space. Additionally, disk storage requires much more work to clean up, and it is usually enough of a hindrance that many users overbuy capacity to accommodate their needs. For a combination of these reasons, ESG's model shows a clear advantage for tape in the area of reducing wasted disk storage and improving usable to raw capacity ratios.

  In total, in excess of 13.2 PB raw capacity is estimated to be required in the PMO case at the end of the ten-year time horizon. This figure, combined with the assumption that each HDD can house 1.2 TB, determines the number of disk arrays the organization is likely to purchase for the data store. In this scenario, the organization is anticipated to purchase several disk arrays, incurring costs of $3,893,830, which is significantly higher than the comparable tape library. This cost also includes a $19,000 cost estimated to cover the backup appliance required in the PMO scenario.

- **Software:** Software costs required to support the data store over time are included within the scope of ESG's economic value model. In the LTO scenario, this cost is representative of licensing costs required to run LTFS directory services in the environment. LTFS is included in the model as a means of easing the operational costs associated with the LTO solution. Not only does this file system allow the tape environment to be managed at greater scale than the PMO, but its drag and drop interface and self-service capabilities also drive assumptions in the model that data retrieval requests will be slightly easier and faster to satisfy. Costs are estimated as a function of capacity, with the base software charge being equal to

$156,000 and incremental licenses, of which three are required, being purchased at a cost $18,000. In total, the organization is estimated to spend $210,000 on software over the time horizon.

In the PMO, software costs estimated are representative of more traditional disk-to-disk archive management tools. These costs are estimated as $172,894 per backup appliance required in the PMO. In the scenario described in Table 1, a single backup appliance is required. Thus, the unit cost covers the entirety of software expenses in the PMO.

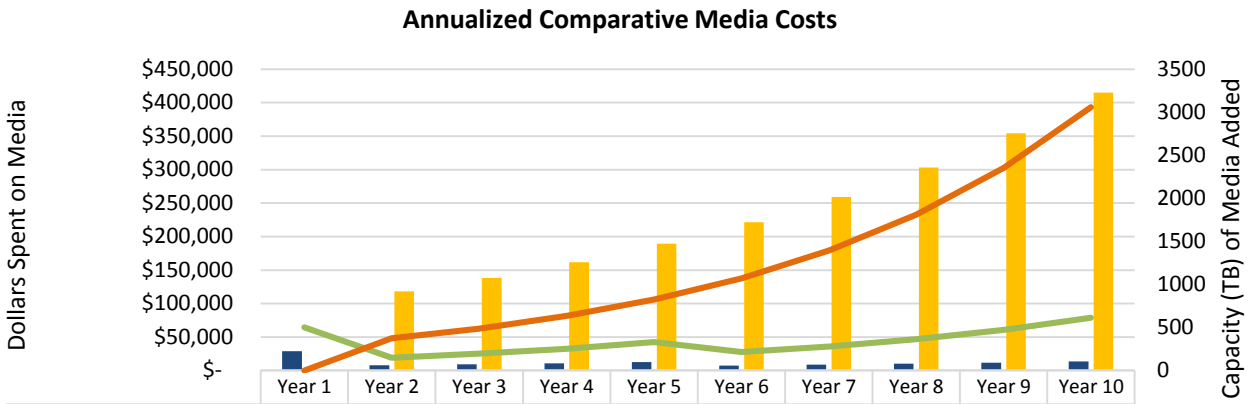
- **Media:** Media costs for both scenarios are broken out from hardware costs as they make up a major cost component, and delta, between the two technology approaches. For both scenarios, media costs are driven by the raw capacity requirements of the environment as well as the estimated per-TB cost of media.

  For the LTO scenario, it is important to bear in mind that it is assumed that data can be compressed by a 2:1 ratio, effectively halving the amount of raw capacity needed. As such, to support the initial 1 PB data store, only 500 TB of cartridges are required. Utilizing the assumptions that each LTO cartridge costs $145 and can retain 2.5 TB of data, ESG's model assigns a per-TB cost in the LTO scenario of $58. Thus, in the first year of the time horizon, the organization can expect to spend $29,000 ($58 x 500) on tape media. Over the course of the time horizon, the per-TB cost of media is anticipated to decrease by 10%. As a result, at the end of the time horizon, the organization is estimated to pay $22 per TB for LTO media. However, the compounded growth of the environment results in over 612 TB of raw tape capacity being added in the final year of the time horizon. In total, ESG's model estimated that the organization can expect to purchase $120,710 worth of tape cartridges over the full time horizon.

  Media costs estimated in the PMO scenario are markedly higher than in the LTO scenario, driven by several factors. First, data is assumed to not be compressed, based on the array capabilities likely to be present in the disk arrays selected for a long-term data retention use case. The result is that much more raw capacity is required to store the same amount of data. Second, exacerbating the previous point, in a disk-based data store, it is assumed RAID will be utilized for data parity. To account for this fact, ESG's model includes an assumption that raw capacity will be overprovisioned by 20% to account for parity drives. Finally, ESG's model uses an initial assumption that on a per-TB basis, disk capacity can be procured for $350 at the beginning of the time horizon. While this per-TB cost is modeled to decline over time at the same 10% rate as is assumed in the LTO scenario, the starting point is much higher, resulting in much higher media costs over time.

  The arrays present in the all-disk scenario are sized and configured with the disk capacity required to accommodate the initial environment's size and growth anticipated in the first year (and included in the hardware costs estimated by the model). In years 2-10 of the time horizon, the organization is modeled to purchase capacity incrementally. For example, in the final year of the time horizon, the organization is estimated to require over 3 PB of raw disk storage. Even with declining disk costs, the organization is still estimated to spend $414,790 in just that year on disk capacity. See Figure 3 for a graphical comparison of the annual media spend modeled in both the LTO scenario and the PMO scenario.

Figure 3. Annualized Comparative Media Costs, LTO for Long-term Data Retention Versus Disk-based PMO



	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10
LTO Media Cost	$29,000	$7,830	$9,161	$10,718	$12,541	$7,336	$8,583	$10,043	$11,750	$13,747
Disk Media Cost	$-	$118,125	$138,206	$161,701	$189,191	$221,353	$258,983	$303,010	$354,522	$414,790
Tape Capacity Added	500	150	195	254	330	214	278	362	471	612
Disk Capacity Added	0	375	488	634	824	1071	1392	1810	2353	3059

Source: Enterprise Strategy Group, 2016.

- **Infrastructure:** ESG's model quantifies the likely infrastructure costs incurred in both solution scenarios, made up of both power costs and a measure of the opportunity cost incurred as a result of the data center real estate being occupied by the two solutions.

  With respect to power costs, ESG's model leverages some common assumptions. First, ESG's model assumes both solutions are operated in 24x7 environments and thus need to be powered for the entirety of each year in the time horizon. Similarly, ESG's model uses a common utility cost of $.10 per kWh for both solutions.

  With regard to power costs specific to the LTO solution in place, ESG's model uses the assumption—based on an average of properly sized tape libraries in production today—that the tape library utilized by the end customer in the LTO scenario will draw nearly 300 watts when in operation. Thus the annual power cost for that tape library is estimated as $252. ESG's model also considers the servers running the LTFS software in place in the solution configuration, using generic server wattage estimates of 600 watts per server. The result is an additional $2,102 per year in utility costs expected for the organization.

  With regard to the power costs in the PMO, there are two material components: The power costs required for the three disk arrays purchased to store data and the power costs required to run the backup appliance in the environment. Two of the three disk arrays configured by the model are very large and—including both controllers and disk shelves—are estimated to draw 11,680 watts each. The third disk system is considerably smaller, requiring only 1,230 watts to power. Finally, ESG's model assumes the backup appliance in the environment will draw 300 watts per hour. Using the same uptime requirements and power costs as the LTO scenario, the total data center power costs in the PMO scenario are modeled at $21,804 annually, over an order of magnitude more than in the LTO scenario.

  In addition to power costs, ESG's model includes a data center floor space cost within the scope of its analysis. For many organizations, this represents the hard cost paid to a colocation provider for data center space. For other organizations, this is representative of the opportunity cost of utilizing a precious commodity—space on the data center floor—for their long-term retention needs as opposed to some other project. In either case, there is a cost to the organization associated with deploying larger hardware solutions. ESG's model quantifies this cost by estimating the total rack units (u) needed to house each solution and assigning a per-u, per-month cost to the solution of $50. In the LTO scenario, the tape library as well as the associated hardware required to run the LTFS environment are estimated to occupy 84 rack units, or approximately two full racks. Per the cost assumptions in ESG's model, the real estate cost to the

organization is estimated as $50,400. By contrast, the storage arrays and backup appliance configured in the PMO scenario are estimated to occupy 92 u. The result is an estimated annual real estate cost of $55,200.

In total, considering both power and data center real estate costs, the LTO solution is estimated to be 46% less costly over a ten-year time horizon for the environment described in Table **1** ($527,547 versus $982,403).

- **Maintenance and Support:** ESG's model includes estimated maintenance and support costs paid to solution vendors over time. These costs are estimated as functions of hardware and software capital expenditures over time and ESG leveraged published list prices as the basis for its model assumptions.

  In its research, ESG found that LTO technologies were often maintained at significantly lower rates than their disk array counterparts. ESG's model utilizes the assumption that maintenance on the tape library is approximately 4% of list pricing and applies this fee on an annual basis. By contrast, ESG's model utilizes the assumption that maintenance on disk arrays is approximately 8% of list pricing and applies this fee on an annual basis. The increased maintenance rate assumption, along with the much higher estimated capital expenditure associated with disk arrays, yields a significant $1.6M in lower maintenance costs anticipated in the LTO scenario compared with the PMO.

- **Staff Personnel:** Staff costs are a material consideration in ESG's model. ESG utilizes per-TB staffing considerations in the PMO to quantify the value of freeing up staff to focus on other initiatives when comparing staffing requirements in an LTO solution. In the PMO scenario, ESG's model assumes that one full time equivalent (FTE) is required for every 200 TB of data to be retained in the data store. With a data store of 1 PB, growing at an annual rate of 30%, it is easy to see how staffing costs can quickly escalate. In fact, at the conclusion of the time horizon, the organization is modeled to have allocated ~10.5 FTEs to manage the data store in the PMO scenario. By multiplying the FTEs required by the annual fully burdened FTE cost of $91,000 for each year, ESG's model estimates a total staff cost over ten years of ~6.8M.

  In contrast to the PMO, an LTO-based data store powered by the usability of LTFS is a far more scalable solution. Many multi-PB customer environments are typically managed by a single FTE or less. ESG's model utilizes this single FTE assumption to calculate staffing costs over a ten-year time horizon, thus the total staff allocation is estimated as $910,000 in the LTO scenario.

- **Data Migration Costs:** As any storage administrator can attest, data migrations are a painful process. Within an LTO environment, potentially painful forklift upgrades and costly data migrations are eliminated. LTO drives are open: Different vendors' solutions are compatible with one another and future generations of LTO drives are compatible with the previous generation. In this model, the organization utilizes a mix of LTO-6 and LTO-7 cartridges over time and thus the risk of a major migration, which may be required due to compatibility, is marginalized and there is no risk of vendor lock-in.

  By contrast, disk-based solutions are fraught with the risk of a painful data migration. An organization may outgrow an array or change vendors over time, either of which comes with an expensive and painful data migration. While estimating these costs over a ten-year time horizon is difficult to do with certainty, ESG has attempted to quantify this cost in its value model by utilizing capacity as a driver of migration costs, with larger environments being more costly to migrate than smaller. These costs coupled with an assumed likelihood of a migration, estimated as 33% (the annual likelihood of an array upgrade using a three-year replacement cycle), yield the annual expected cost of data migrations in the PMO scenario on an annualized basis. In total over the time horizon, the PMO is estimated to incur in excess of $150K more in data migration costs over the time horizon compared with the LTO solution.

How the Financial Picture Changes When the Cloud is Introduced

This report has outlined in detail the financial comparison of long-term data retention on tape and all-disk media. However, many organizations are beginning to consider the public cloud as a potential locale for data. ESG's model includes the flexibility to compare tape retention with a hybrid disk and cloud retention model.

In order to accommodate hybrid PMO scenarios, assumptions must be made for the percent of data stored externally in a public cloud versus internally on disk-based storage. For the purposes of this paper, ESG's model is tuned to assume that 20% of retained data resides in a public cloud. Next, costs for the cloud service must be incorporated. ESG's model assumes a charge per upload request of $.055 and assumes that every GB of data stored in the cloud carries with it a $.007 cost per month. In addition to loading data into the cloud, egress costs are considered. ESG's model assumes a charge per retrieval request of $.055 and assumes that each GB of data retrieved carries with it a $.08 charge. Finally, ESG's model assumes that data is uploaded and retrieved from the cloud in chunks which average 4 TB.

These cloud-centric assumptions are added to the use case assumptions in shown in Table 1 so that aggregate costs can be compared amongst all data retention scenarios. Table 4 shows the TCO comparison over ten years.

Table 4. Ten-year TCO, LTO Versus Disk-based PMO Versus Hybrid PMO for Long-term Data Retention

Category	LTO Solution	All-disk PMO	Disk/Cloud Hybrid PMO
Hardware	$496,396	$3,893,830	$2,858,598
Software	$210,000	$172,894	$172,894
Media	$120,710	$2,159,881	$1,727,905
Cloud Services	$0	$0	$356,248
Infrastructure	$527,547	$982,403	$787,429
Maintenance and Support	$85,248	$1,705,537	$1,291,441
Staff Personnel	$910,000	$6,779,500	$5,960,500
Data Migration Costs	$20,284	$174,640	$139,712
Data Access Costs	$40,915	$40,915	$606,028
Total ten-year cost of ownership	**$2,411,099**	**$16,808,264**	**$13,900,747**

Source: Enterprise Strategy Group, 2016.

How Cloud Retention Shifts Costs in a Hybrid Model

As the hypothetical organization incorporates cloud storage for data retention, several costs are diminished while others are increased. First, the organization requires less on-premises hardware to host the data since a portion resides in the cloud. In this scenario, the result is fewer drives and smaller, less expensive storage systems. In ESG's analysis, the hardware and media costs are reduced by 27% and 20% respectively. Similarly, the smaller hardware footprint results in commensurately less OpEx over time dedicated to support and maintenance costs, data center power and cooling, and staff management requirements.

However, these cost reductions do not come without a price. The costs to upload, retain, and retrieve data over time must be considered. Adding data to the cloud is the least-costly component of this equation. As data grows in the environment, ESG's model estimates that uploads to the cloud (in 4 TB chunks) will grow from 50 initially up to 122 events in the tenth year of the time horizon. In total the organization will incur a trivial ~$350 in services costs tied to upload requests.

While getting all that data into the cloud is a relatively cheap proposition, retaining it there results in significant costs over time. There are two components to the retention calculation that matter: the unit cost to keep data in the cloud and the amount of data residing there. As described previously, the monthly cost/GB of data in the cloud at the beginning of the time horizon is $.007 resulting in an annual cost of $84/TB/year. ESG's model assumes that, much like on-premises storage capacity, cloud capacity will continue to decline in cost over time. At the end of ten

years the cost/TB/year for cloud retention is estimated to be $40/TB/year. When turning to the amount of data residing in the cloud, we return to the assumption that 20% of the initial PB of data will be housed in the cloud. As that PB grows at a 30% annual clip, additional data is modeled to be cloud-sourced. By the end of the ten-year time horizon, cloud-based capacity is estimated to have grown to over 2.1 TBs, resulting in an annual cost of $65,547 in that year alone. In total, ESG's model estimates expenditures for retaining data in the cloud to exceed $350K.

Finally, in this dynamic scenario, data is not simply sent to the cloud never to be seen again. The 10% monthly retrieval rate is applied to this data store. For example, in the first year of this analysis, with an estimated 200 TB of data residing in the cloud, and 10% of that data retrieved per month (20 TBs) at a per-TB cost of $55 ($.055/GB X 1,000), the organization can anticipate egress charges in excess of $13K in that first 12 months. As the data stored in the cloud continues to grow, so too do egress charges incurred. In the tenth year of the time horizon, these costs have grown to over $142K.

In this analysis, the incorporation of the cloud for data retention helps reduce overall costs compared to an all-disk scenario. However, these net reductions still leave a deep financial divide between the higher costs of a disk/cloud hybrid approach and an all LTO-tape approach.

The Bigger Truth

As evidenced in ESG research, IT organizations are broadly utilizing tape solutions in their data protection processes.[3] ESG believes these decisions are made based on a combination of the functionality of modern tape (i.e., its speed and reliability characteristics) and its compelling economics compared with disk-based alternatives and even disk/cloud hybrid approaches. While many data stores may benefit from retention architectures that take advantage of multiple technology platforms and the incremental performance offered by disk solutions in the right use cases, for organizations with large data stores that must be maintained over extended time horizons, tape's economics are hard to argue against.

ESG's research and financial modeling conducted for LTO-based solutions shows that dramatic savings in hardware, media, staff, and maintenance costs can be achieved by leveraging tape over disk.

For IT organizations looking to increase their efficiency in the realm of capital expenditure, while boosting their ability to be an exceptional service bureau for the organization by freeing up highly skilled administrators to focus on other endeavors, LTO tape solutions warrant close consideration.

[3] Source: ESG Research Report, *2015 Trends in Data Protection Modernization*, September 2015.

Appendix A

For this project, ESG adhered to the following research and modeling methodology:

- ESG conducted initial market research across LTO providers (including HPE, IBM, and Quantum) to assess current market trends, vendor value claims associated with LTO tape, and the purchase considerations that are most important to existing and prospective customers researching long-term data retention solutions.
- Based on the results of this initial research, ESG subsequently identified a "present mode of operation" or PMO—effectively, the alternative approach that customers are likely to take to meet their long-term data retention requirements—against which the costs and benefits of utilizing LTO technologies was to be compared. A conventional disk-based infrastructure is used.
- ESG then developed a comprehensive financial model designed to qualify and quantify the potential costs and benefits of utilizing LTO technology compared with the PMO.
- Next, ESG conducted a series of in-depth interviews with systems engineering, service and support, and technical marketing representatives from LTO providers. The data collected in these interviews was used to refine assumptions built into the model related to current customer environments and the direct and indirect costs and benefits attributable to both LTO technology and the PMO. Product demonstrations, ESG Lab validations, and vendor case studies were also used to identify specific IT tasks and the labor burden (in both time and cost) associated with those tasks. These findings were augmented by findings from ESG's qualitative and quantitative market research in the data protection coverage area. The totality of this research helped to inform ESG's understanding and analysis of storage solutions for long-term data retention in terms of adoption drivers, usage trends, and the technical, operational, and financial benefits that have been realized by customers.
- Once the economic model was finalized and all validation was complete, ESG modeled a default scenario that was designed to demonstrate the relative costs and benefits of LTO technologies in a large-scale environment. Those results were then compared with model outcomes for a similar-scale disk-based solution. The results for this default scenario are described in the body of this paper.

Please note that the data and conclusions presented in this report regarding the costs and benefits associated with implementing LTO technologies compared with disk-based storage solutions reflect the output of ESG's economic value analysis based on the specific use case and default scenario assumptions modeled for this report. ESG acknowledges that changes to these assumptions will lead to a different set of results and as such, advises IT professionals to use this report as one validation point in a comprehensive financial analysis process prior to making a purchase decision. A composite of current standard pricing and configuration information for LTO technologies was provided to ESG by the LTO solution providers engaged in this project. Other IT equipment and labor cost assumptions were obtained from publicly available sources such as IT vendor and channel partner websites and published price lists.

Appendix B

Table 5. Cost Categories in the Scope of the Analysis

Category	Description
Hardware	• For the LTO scenario, this includes capital costs associated with library(s), drives, and associated LTFS hardware required in the scenario. Tape cartridges are included in the media cost category. • For the PMO scenario, this cost is representative of the costs for arrays likely to be purchased in the scenario. While the hardware cost includes the drives needed to satisfy the initial storage environment, additional drive costs to accommodate storage growth over time are calculated on an annual basis and allocated in the media cost category. Additional hardware costs to account for a backup appliance are also included. ○ Note that ESG used annualized depreciation costs when extrapolating hardware costs over the full ten-year time horizon examined.
Media	• For the LTO scenario, this cost is representative of the incremental tape media (in the form of either LTO-6 or LTO-7 cartridges) required to accommodate the environment over time as data grows. • For the PMO scenario, this cost is representative of the added disk capacity required (excluding the initial capacity procured, which is sized to manage start-up and first-year capacity growth) to accommodate the environment over time as data grows.
Software	• For the LTO scenario, this cost is representative of licensing costs required to run LTFS directory services in the environment described. • For the PMO scenario, software costs estimated are representative of more traditional disk-to-disk archive management tools.
Infrastructure	• For both scenarios, this includes costs related to data center infrastructure, such as power and rack space. Since the infrastructure build in the LTO scenario is assumed to be considerably smaller and less power-intensive, material differences are observed.
Maintenance and support	• For both scenarios, support and maintenance costs include vendor charges for the solutions procured. Maintenance charges are annualized and likely relative maintenance rates are utilized based on publicly available pricing data from both LTO and disk-based solution providers.
Staff	• For both scenarios, this includes IT staffing costs associated with the management and operation of the solutions deployed. • All position-specific salary information is based on publicly available median average salary data. For all employees, the model uses a fully burdened rate (i.e., including the cost of benefits, payroll taxes, etc.) of 40%.
Data migration costs	• For both scenarios, this includes costs associated with data migrations occurring over time. While data migration costs are largely eliminated in the LTO scenario due to LTO technology's open format and backward compatibility, significant migration costs for the PMO are estimated based on both potential vendor changes and array upgrades likely to occur in the environment over a protracted time horizon.
Data access costs	• Data access costs are marginalized in the scenarios described in this paper. This cost category is largely represented by significant vendor costs associated with data egress from a cloud storage provider. While data access is not "free" in an on-premises solution—due to the time and effort required of IT staff to retrieve data—when comparing two modern data storage solutions, costs are normalized.

Source: Enterprise Strategy Group, 2016.

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

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