White Paper

Archive TCO: Five-Year TCO Comparing Dell DX to Tape and NAS for Long Term Archive

By Terri McClure

February, 2011

This ESG White Paper was commissioned by Dell and is distributed under license from ESG.

Table of Contents

Introduction: The Data Growth Challenge

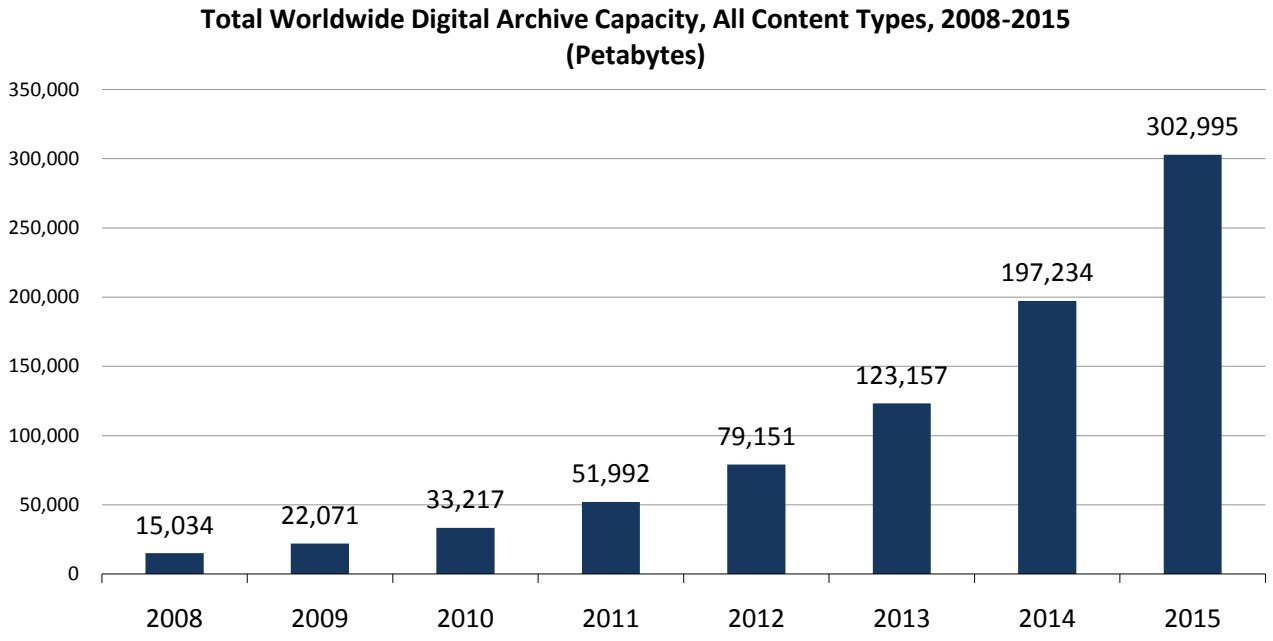
Today's IT managers are facing an inflexible reality: an enormous and still-growing amount of unstructured digital data needs long-term archiving. And this archiving process must be handled with 100% fidelity to legally mandated retention policies, in-house records management schedules, and other strict guidelines.

Of course, not all archival data is the same. Different data must be retained for different lengths of time. But even just establishing and following policies for a single environment, such as e-mail, can bring big administrative headaches and result in a lot of staff overhead.

Those hassles, however, are secondary to a more urgent problem. Archiving can require a big budget. For many enterprises faced with accelerating data growth, keeping so much archival data online is becoming too expensive. In fact, keeping up with data growth is now the top challenge cited by IT managers who have storage planning and buying responsibilities.

Their sentiment is borne out by numbers. According to ESG research, total worldwide archive capacity in the commercial and government sectors will increase from 33, 217 petabytes in 2010 to 302,995 petabytes in 2015.[1] That's nearly 303 exabytes of archival data needing to be stored by 2015, growing at a 56% compound annual rate.

Figure 1. Total Worldwide Digital Archive Capacity, All Content Types, 2010-2015 (Petabytes)



Source: Enterprise Strategy Group, 2010.

File-based content such as office documents, web pages, digital images, and audio and video files now represent 76% of archival data. Those file types, along with e-mail archives, will continue to constitute the vast majority of archived digital assets for the foreseeable future. And they can be expensive to store using traditional methods.

Last year, however, Dell unveiled a product called the Dell DX Object Storage Platform that is designed to intelligently access, store, protect, and distribute fixed digital content. This self-managing, self-healing archiving system is built with commodity x86 hardware and integrated software. The company positions it as a way to help users cut storage costs (versus traditional performance storage methods) by affordably storing and allowing access to digital content without being locked into a costly, inflexible strategy.

[1] See: ESG Research Report, *Digital Archive Market Forecast 2010-2015*, July 2010

This paper will compare the projected five-year cost of ownership of the Dell DX platform against tape and network attached storage alternatives, and will examine many of the other important considerations that must go into any smart archiving platform purchasing decision.

What can Happen When Archiving isn't a Core Practice

As expensive as it can be to retain historical digital assets properly, not having an intelligent retention and deletion policy can be far more costly. For example, in the U.S., the Public Company Accounting Oversight Board (PCAOB) will impose fines for noncompliance with Sarbanes-Oxley data retention regulations for financial records. The PCAOB may even bring federal court proceedings against corporations that haven't complied. And the U.S. SEC may penalize a company by halting trade in shares of its stock without advance warning. Such actions can result in long-lasting damage to a business's reputation and financial heath.

If archiving non-compliance relates to medical records, outcomes can be just as bad. Electronic patient records must be protected against loss for certain legally mandated periods of time. The retention rules imposed by the Health Insurance Portability and Accountability Act (HIPAA) are covered under U.S. federal law; violating these rules constitutes a criminal felony and can result in large civil fines and penalties.

Challenges Associated With Archival Data Growth

The stakes are high. What are the options for IT managers tasked with protecting archival data in today's budget-tight and regulation-heavy landscape?

Typically, three alternatives are available:

- Archiving to **tape**, which involves storing digital information on magnetic drives, autoloaders, libraries, and similar media.
- Archiving to **optical media**, which typically involves archiving digital information using 5.25-inch cartridges (ultra-density optical or magneto optical cartridges), CDs, or DVDs.
- Archiving to **disk**, which involves storing information on internal disks within a server or on external devices such as direct-attached storage arrays, network-attached storage appliances, storage area network devices, or a specialized archive solution such as the Dell DX Object Storage Platform.

ESG research has determined that by 2015 users will be storing most of their archival data on external disk. This external disk-based archiving is expected to grow at a 61% CAGR from 2010 through 2015—a growth rate higher than that of the general disk storage market.

The increase in external disk-based archiving appears to correlate to reductions in other modes of data archiving. For instance, ESG forecasts that traditional tape-based storage will likely decline from 40% in 2010 to just 27% in 2015. Internal server-based archive storage will likely decline from 29% in 2010 to 25% in 2015.

One reason for the shift is already clear: preserving archival data is a perfect use-case for cloud storage and cost-efficient cloud environments rely on external networked storage, not DAS, tape, or optical media. Rather than continue to engage in onsite disk archiving and offsite tape-shipping, IT organizations are beginning to store archive data in locations operated by third-party managed service providers. These service providers offer easy-access, are guaranteed regulatory-compliant, and provide fully secure archiving to the cloud.

A Balancing Act

IT organizations need to perform a balancing act, making sure they can:

- Access archived data in a timely manner to meet regulatory guidelines.
- Opt for a platform capable of scaling up to massive capacity without becoming burdensome to manage.

- Stay "future-proof" by choosing an archiving system that will work with legacy systems, lend itself to upgrades, and evolve online as new technology advances become available.
- Rein in costs, regardless of how much archival data must be stored.
- Establish protection levels corresponding to the data's value, leveraging the most dense and least expensive medium that still provides timely data access.
- Incorporate efficiency features to reduce the data center's footprint and energy consumption.

It's a lot of pressure to be under, prompting users to consider object-based storage as an alternative to traditional archiving to tape or NAS systems. Significant savings are associated with such object stores, compared with tape or disk alternatives.

By the Numbers

Before going further into a discussion of some of the "softer" factors that play into an archiving purchasing decision, let's examine total cost of ownership (TCO) numbers for three archiving-platform alternatives over the long haul.

ESG looked quantitatively at the cost of managed terabytes over time, comparing the long term ownership costs of:

- Dell DX versus tape archiving
- Dell DX versus highly scalable NAS

The long term cost differences are dramatic.

ESG's Baseline Environment Scenario

ESG made its best effort to attain and assign pricing for the components of each solution—tape, NAS, and the Dell DX—and leveraged its knowledge of platforms and industry-accepted best practices to assign labor time and cost. Relying on Dell for assistance with the DX model, ESG looked at the five-year total cost of ownership for each platform based on the following assumptions:

- The environment's initial raw storage capacity is 35 TB.
- The archive environment is growing at 20% yearly.
- Floor space costs are $75.00/square foot.
- The organization paid list price minus 30% to acquire the platforms.

Table 1. Five-Year Total Cost of Ownership – Tape Archiving

i. Tapes Transported Offsite via Truck	Year 1	Year 2	Year 3	Year 4	Year 5	
Acquisition costs (CAPEX)	$179,104	$260	$317	$375	$461	
Maintenance	$19,639	$19,639	$19,639	$26,676	$26,676	
OPEX (floor space, personnel, media storage and transportation)	$42,491	$41,777	$58,528	$58,562	$87,480	
TOTAL:	**$241,234**	**$61,676**	**$78,484**	**$85,613**	**$114,617**	**$581, 624**

Table 2. Five-Year Total Cost of Ownership – Scale-out High Density NAS

ii. Scale-out High Density NAS	Year 1	Year 2	Year 3	Year 4	Year 5	
Acquisition costs (CAPEX)	$108,000	$0	$0	$36,000	$36,000	
Maintenance	$16,200	$16,200	$16,200	$21,600	$27,000	
OPEX (floor space, personnel)	$32,950	$32,950	$32,950	$32,950	$32,950	
TOTAL:	**$157,150**	**$49,150**	**$49,150**	**$90,550**	**$95,950**	**$441,950**

A note on NAS pricing: a number of scale-out vendors have all-inclusive software pricing and this model is built to reflect that. *But many vendors do charge additional fees for add-on software and software such as protocol, archive, or snap licenses can add significant up-front capital and maintenance costs.* Minimum configuration for most scale-out NAS systems is three nodes; this chart is based on a three-node entry point with twelve 2 TB drives in each node. All backup and NAS costs in the model were blended costs and are used for illustrative purposes only.

Table 3. Five-Year Total Cost of Ownership – Dell DX Object Storage Platform

iii. Dell DX	Year 1	Year 2	Year 3	Year 4	Year 5	
Acquisition costs (CAPEX)	$96,837	$ 0	$17,282	$26,291	$26,291	
Maintenance	$19,367	$19,367	$22,824	$28,082	$33,340	
OPEX (floor space, personnel)	$13,450	$13,450	$13,450	$13,450	$13,450	
TOTAL:	**$129,665**	**$32,817**	**$53,555**	**$67,823**	**$73,081**	**$356,931**

After reviewing these numbers, ESG can conclude that the Dell DX Object Storage Platform is a long term archive that is massively scalable on a very dense and inexpensive commodity-based hardware platform. It is simply going to be—especially from a hardware standpoint because of its density—less expensive over time.

Of course, mileage may vary and every environment is different. Multiple factors influence TCO and it is impossible to create a model that works for every environment. This model is a scenario that outlines a new archive environment built from scratch using tape, scalable NAS or Dell DX. It is important that users understand the nuances of their environments and perform their own due diligence. However, in this five-year model showing a 35 TB archive environment growing at 20% yearly, ESG determined that archiving using Dell DX would cost the organization **$224,693 less** than it would have spent using tape and **$85,019 less** than using NAS.

Surveying the Landscape

Of course, beyond reviewing cost-of-ownership considerations, all IT users must examine these archiving choices—object storage, tape, and NAS—from the standpoint of everything else they may require from an archive appliance.

Regulatory-driven Needs

IT managers should consider their organization's regulatory-driven responsibilities. Most IT organizations must follow formal data retention policies matched to the information being archived. Knowing the official retention periods for various records helps an IT manager make smarter archiving decisions and ensure compliance with the operational and legal requirements that dictate how long certain types of data must be saved and when information must eventually be deleted.

IT managers also must examine how the archive data is used. Is it important to ensure that even "dormant" archived data can be retrieved in a timely manner? Will multiple users or business units be accessing the archived data periodically? These are important questions to ask when selecting an archiving platform. Some data is kept solely for possible audit compliance and will never again be viewed unless an audit occurs. Conversely, other archived data, such as older medical images of patients, may still need to be quickly accessible to medical professionals if those patients experience a new health problem. That type of data isn't suited for offsite tape archiving.

A write once/read many (WORM) format may be the best option if, perhaps, the data is not mission-critical to business operations yet its integrity must be maintained. The fact is, a large percentage of archival data must be immutable and unchanging, but some of it also must be quickly retrievable.

Efficiency-driven Needs

If the prime goal is efficiency, an IT manager's focus might be on operational savings. A SAN device would likely be too cost-inefficient for some information that needs to be kept pristine yet may never actually be needed (for example, the scanned blueprints of a 30-year-old factory).

Still other efficiency-related factors pertain to:

Metadata: There's often a need for an archiving system to provide an IT manager with insight into the context and content of archived data to drive policy-based management.

If this is the case, it's vitally important to have a system that can archive not just files, but also rich metadata (additional identifying information) about those stored objects. Object storage enables attachment of metadata to stored files.

For example, a laboratory's collection of millions of spectrometer readings may need to have metadata attached to each one of those files, showing the name of the spectrum that was measured, the time the reading was taken, and the spectrometer settings used. Metadata may also identify a file's pre-established lifecycle and rules for its automatic distribution and eventual deletion. An object storage system such as the Dell DX holds each object in one enormous, flat address space—with its associated metadata—to enable fast identification and retrieval of the correct information.

Metadata cannot become disassociated from its file during storage, deduplication, backup, or retrieval. Again, an object storage system like the Dell DX will ensure that disassociation problems won't occur. This benefit of object storage is one of the factors leading to wide adoption of the technology for storing fixed digital content.

Density: Because archival data is growing exponentially, the systems storing it must be able to hold a tremendous amount of information in a reasonable footprint. The storage nodes in the Dell DX system, for instance, are based on Dell's PowerEdge R510 chassis and are optimized for storage density.

A basic DX configuration consists of either 2Ux4 drive or density-optimized 2Ux12 drive x86-based storage nodes (as used for the cost model) and one out-of-band cluster services node (Each storage node can contain between 1 to 24 TB of raw data). As a result, the platform can scale to handle billions of objects—petabytes of data accumulated over time.

Power-saving Technology: Highly dense archiving systems could (not surprisingly) become highly power-hungry, too. To overcome this dilemma, IT managers are actively expressing interest in solutions such as Dell's Darkive Adaptive Power Conservation technology, a feature of the Dell DX Object Storage Platform.

Darkive caps power consumption and sets a percentage of maximum potential power consumption for hardware proactively and without human intervention. Specifically, it spins down disks and reduces CPU utilization based on predefined policies or predetermined levels of inactivity. The reduced CPU utilization in turn reduces the cost of running and cooling the platform.

Space-saving Technology: Although a good archiving system must be able to hold billions of files, that doesn't mean it has to have a large footprint. After all, a large platform brings large management headaches. IT managers are trying to be more efficient and are tired of paying for and deploying more and more storage.

Reducing a storage footprint reduces costs, improves service levels, and mitigates risks. Some archiving technologies do accomplish this space-saving feat: for example, the DX, based on a grid architecture, requires few floor tiles and rack space.

Ease of Use: The right archiving system should offer a simple, self-managing, expandable, and cost-effective architecture. An easy-to-use archiving solution is one that:

- Allows IT managers to identify and retrieve the right information quickly by using metadata.
- Automatically manages data and meets data-governance demands by following automated, policy-based retention rules.
- Can scale out to billions of files and petabytes of storage easily without adding a lot of additional management overhead.
- Is self-healing, uses wizard-based setups, and does not require LUNs or RAID groups.
- Avoids the complexity of forklift upgrades, relying instead on a scalable architecture based on industry-standard hardware and software.

Future-proof Architecture: Speaking of forklift upgrades, a smart object storage investment must be able to work with obsolete/legacy platforms, seamlessly incorporate new technology, and ensure that its archived files will be retrievable and readable years into the future.

The IT industry is fast-moving. New approaches to storage are unveiled frequently. IT organizations responsible for protecting digital assets for the long term must ensure that the archival platforms they are using today will keep files safe for as long as that information is required to be retained (sometimes decades). They need a solution that keeps the archived content available and retains all of its original characteristics.

Generally speaking, a future-proof content archiving system is one based not on proprietary hardware components, but instead on a standard platform. The Dell DX system, for example, is based on x86 standards-based rack server platforms from Dell, adding a great deal of assurance that these systems (and their contents) will remain manageable far into the future.

Intelligent Object Replication: Lastly, IT managers need to be able to intelligently and automatically replicate content. They may be doing this to ensure disaster recovery, to move content closer to certain users, or to perform other administrative tasks.

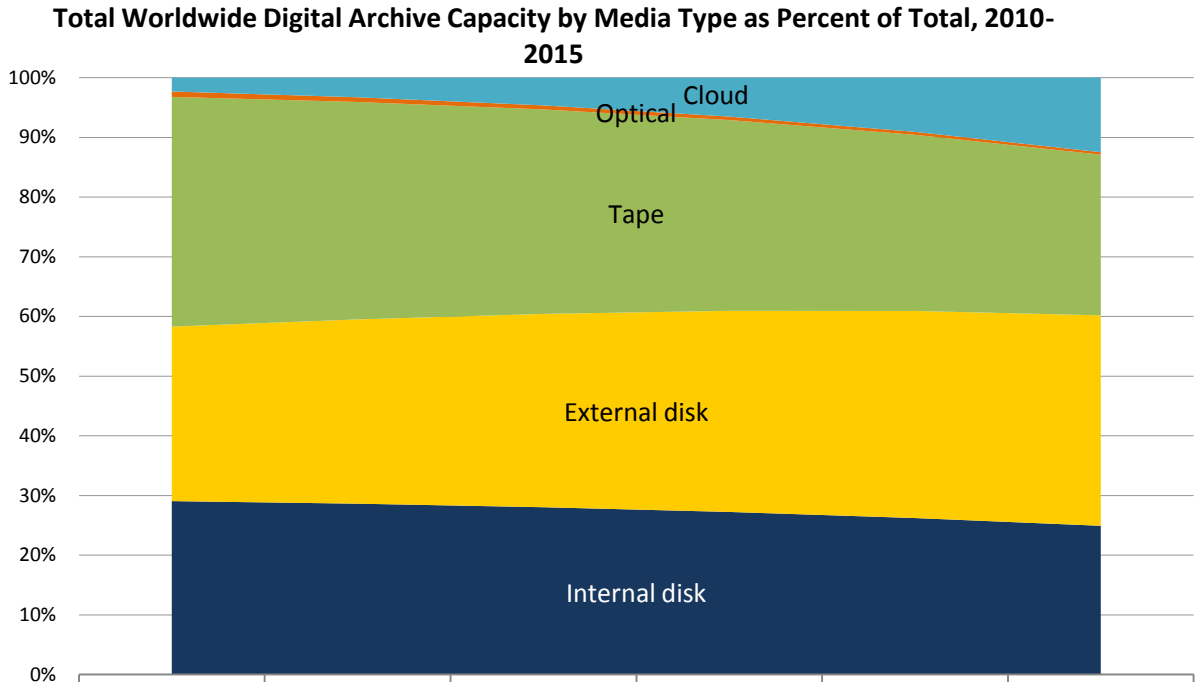
Intelligent object replication is desirable. It is a capability that will automatically move and replicate content to multiple locations using a rules-based workflow engine driven by metadata.

The Dell DX, for instance, allows for this type of intelligent replication by enabling applications to write data to a local DX cluster and replicate it over a standard IP network to a disaster recovery cluster synchronously, within a single request.

The Alternatives for Long Term Archiving

Archiving to tape; archiving to optical media; archiving to NAS, SAN, or object storage: which option is likely to serve IT users best as they deal with the data explosion?

Figure 2. Total Worldwide Digital Archive Capacity, by Media Type as Percent of Total, 2010-2015



	2010	2011	2012	2013	2014	2015
Cloud	2%	3%	5%	7%	9%	12%
Optical	1%	1%	1%	1%	0%	0%
Tape	38%	36%	34%	32%	30%	27%
External disk	29%	31%	32%	34%	35%	35%
Internal disk	29%	29%	28%	27%	26%	25%

Source: Enterprise Strategy Group, 2010.

The landscape is—and will continue to be—a dynamic one, with users likely to shift their choices in the coming years. According to ESG research, IT users are looking to deploy economical offsite archive-tier storage. This growth in disk-based cloud archiving will continue and it will come at the expense of traditional archive alternatives.

Tape

Tape has played a starring role in backup and archive environments for decades. In fact, it has been deployed as an archive medium since the 1960s. Traditionally, IT organizations used tape for long-term storage because, compared with disk storage, tape libraries were more cost effective. But as seen in the TCO comparison, that's no longer the case. In this data-deluged world, an exponential growth-rate in tape libraries (and the time required to manage those environments) makes any cost advantage of tape seem far less evident. Additionally, tape is not the best choice for IT managers trying to adhere to today's gamut of data-retention regulations. Nor is it the best option for IT managers looking for the kind of space-saving, easy-to-use, future-proof, replication-intelligent system discussed in the previous section.

Some other considerations: data stored on tape is accessed sequentially, so finding a specific file can be a slow process compared with disk archiving alternatives. Tapes are also prone to loss when being transported to or from offsite archiving locations (and that can be a very costly situation). And, tapes tend to be one of the less-appealing options in terms of reliability. If they break, the files they hold may be rendered inaccessible.

Table 4. How Tape Stacks up Versus Dell DX Archives, Based on Regulatory and Efficiency Requirements

Requirement	Tape	DX
Data can be retrieved in a timely manner		X
Offers write-once/read-many (WORM) format	X	X
Immutability	X	X
Provides sufficient knowledge of data context and content to drive policy-based management		X
Stores not just archival data, but also rich metadata about the stored object		X
Offers dense architecture		X
Is a power-saving technology	X	X
Is a space-saving technology		X
Is easy to use		X
Is future-proof		X
Offers intelligent object replication		X

Source: Enterprise Strategy Group, 2011.

Network Attached Storage

Why use scale-out NAS as an alternative to tape for archive? Accessibility of data is certainly faster than tape storage. And one of the hallmarks of NAS is its expandability: adding more capacity is fairly easy. NAS systems are also fault-tolerant, possessing many RAID protection features that tape libraries just don't have.

But NAS systems were not purpose-built to hold and protect long-term archives. They are meant for online storage of primary file data. As the number of managed NAS systems increases, the amount of time and energy needed to manage them also increases—that's a real concern for IT managers struggling with extremely large archive stores.

NAS systems also may not offer the data-immutability assurance that organizations such as hospitals and banks need as they strive to adhere to data-integrity regulations.

Table 5. How Scale-out NAS Stacks up Versus Dell DX Archives, Based on Regulatory and Efficiency Requirements

Requirement	NAS	DX
Data can be retrieved in a timely manner	X	X
Offers write-once/read-many (WORM) format	X*	X
Immutability	X*	X
Provides sufficient knowledge of data context and content to drive policy-based management		X
Stores not just archival data, but also rich metadata about the stored object		X
Offers dense architecture	X	X
Is a power-saving technology	X	X
Is a space-saving technology	X	X
Is easy to use	X	X
Is future-proof	X	X
Offers intelligent object replication		X

**For an additional fee*

Source: Enterprise Strategy Group, 2011.

Object Storage

There's a reason why object storage systems are considered to be the next generation of digital archiving. These platforms offer the access, manageability, and recovery assurance missing from tape. And in large archive environments composed of infrequently accessed files, they definitely offer a better cost equation than NAS.

Table 6. Advantages and Disadvantages of Object Archives, Based on Regulatory and Efficiency Requirements

Requirement	Yes	No
Data can be retrieved in a timely manner	X	
Offers write-once/read-many (WORM) format	X	
Immutability	X	
Provides sufficient knowledge of data context and content to drive policy-based management	X	
Stores not just archival data, but also rich metadata about the stored object	X	
Offers dense architecture	X	
Is a power-saving technology	X	
Is a space-saving technology	X	
Is easy to use	X	
Is future-proof	X	
Offers intelligent object replication	X	

Source: Enterprise Strategy Group, 2011.

It's worth reiterating a few additional points of comparison that are not as straightforward in a head-to-head analysis of the Dell DX system in particular. These points relate primarily to the value of implementing a future-proof archiving architecture, especially in relation to tape. For example:

- With Dell DX, there's no need to re-host and copy an archive to roll over to newer platforms.
- The system offers just-in-time scalability in a clustered architecture that can grow as storage requirements grow.
- There's inherent efficiency in building an archive on a commodity hardware layer that scales in a granular fashion. The approach allows users to ride the commodity cost-curve advantages over time.

The Dell DX cluster doesn't need to be backed up and upgrades to new storage nodes can be done without impacting applications. With this system, adding new nodes and retiring old ones is an extremely simple task. An administrator adds new capacity to the cluster by deploying a new storage node with 12 drives. A software license from Dell for this capacity can cost less than overall cluster capacity and can be licensed when additional space is needed.

The Bigger Truth

Of course, some cost-related issues beyond those discussed in this paper also can impact the TCO of an archive platform. But ESG did not consider them in its quantitative comparison model due to the far-flung variances possible. These issues include service-level violations, the cost of data loss or inaccessibility, the cost of planned and unplanned outages (including the costs of personnel idle time), and the fact that personnel costs can vary widely.

Again, in ESG's TCO analysis, the assumptions were:

- Initial raw storage capacity = 35 TB
- Growth rate = 20%
- Floorspace = $75.00/square foot
- Power costs = $0.10/KwH
- List prices minus 30%

And these are the TCO numbers ESG arrived at for the three archive alternatives:

- Cost of tape archiving over five years: **$581,624**
- Cost of scale-out high-density NAS archiving over five years: **$441,950**
- Cost of Dell DX archiving over five years: **$356,931**

In this five-year model, archiving using Dell DX would cost **$224,693 less** than tape and **$85,019 less** than NAS. Additionally, there are a large number of variable and soft costs specific to an organization than can increase costs significantly and must be considered in any TCO calculation, such as additional add-on software licensing common to many NAS systems (all software licensing cost are included in the DX model), the cost of one lost or damaged tape, or inability to meet regulatory compliance guidelines. All of these issues must be considered in any cost comparison exercise.

The cost difference may be an eye-opener for any organization looking for more cost-effective storage. CFOs are finding themselves paying more for storage all the time. They want to be more efficient while remaining compliant with data-retention regulations. CIOs and IT managers are worried about the cost of resources to manage all that storage over time. They are ready to adopt more sophisticated and economical archiving tools.

Simply put, an IT organization that consolidates ten petabytes of archive data on a single platform for long-term storage is going to see a lower TCO.

If IT is to cost-effectively manage the deluge of archive data headed its way, then object storage, like the Dell DX platform, is a logical choice for its manageability, longevity, density, and regulatory advantages.

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