Aberdeen Group
A Harte-Hanks Company

Archive Agility

Building Business Resilience through Active Archiving

July 2011

Dick Csaplar

Aberdeen Group
A Harte-Hanks Company

Executive Summary

In May, 2011 Aberdeen conducted a survey on data archiving, exploring the challenges faced, strategies employed, and ultimately the tangible business and operational benefits organizations have achieved. Representatives from a diverse group of 113 organizations responded.

One of the key data management challenges organizations often face is how to keep their archived data accessible and active, without spending the time and resources associated with primary storage. The amount of data in the archives can range from one half to 10 times the amount of data actively managed in primary storage. How can end-users gain access to historical files in a reasonable amount of time without pulling IT employees from higher priority projects? Aberdeen's research found the answer in the technologies and processes that comprise active archiving.

Best-in-Class Performance

Aberdeen used the following key performance criteria to distinguish Best-in-Class companies:

- Amount of time to recover an archived file
- Number of business interruptions in the last 12 months
- Longest period of time of a business interruption in the last 12 months

Competitive Maturity Assessment

Survey results show that the firms enjoying Best-in-Class archiving performance shared several common characteristics, including:

- 75% have deployed archive management software
- 70% have IT trained in new archiving tools
- 58% utilize open archive standards
- 45% have a formal process to ensure redundant data is not archived.

Required Actions

In the Aberdeen report *Storage Virtualization: Experience Begets Benefits*, companies reported on average that their primary storage requirements were growing at about 30% per year. In addition to the strain that this data explosion currently places on the datacenter, a similar impact will eventually be felt in the data archives.

Best-in-Class companies in this report clearly understand how to manage archived data, as they reported their archives to be just 50% larger than their primary storage capacity. Furthermore, even though the size of their archives outranked other organizations in terms of overall bytes, their end-users could access archived files the fastest. To achieve this success, they employ active archiving strategies as well as corporate programs to manage the data.

Research Benchmark

Aberdeen's Research Benchmarks provide an in-depth and comprehensive look into process, procedure, methodologies, and technologies with best practice identification and actionable recommendations

Aberdeen Group
A Harte-Hanks Company

Table of Contents

Figures

Tables

Chapter One: Benchmarking the Best-in-Class

Business Context

Computer systems that support business processes such as order fulfillment, inventory, and customer relationships are critical to the financial performance of an organization. Every moment these systems are collecting and creating data valuable to the survival of the company. This report will focus on the long term management, protection and recovery of this critical business data.

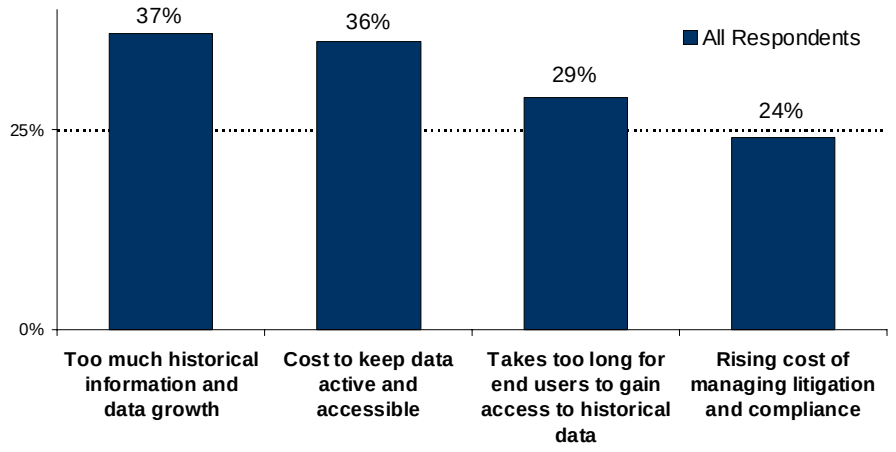
Archiving was traditionally based on hierarchical storage management, or the movement of large amounts of data from disk to low cost mass storage such as tapes or optical devices. The storage devices were then shipped offsite, often never to be seen again unless the company got sued, their taxes were questioned or a governmental agency wanted to review the historical data.

Today organizations understand there is still knowledge and business value in that legacy information and keeping it accessible allows it to be exploited for the corporate good. An active archiving process, where end users can independently search and find files and initiate a restore, is how Best-in-Class organizations now manage their archived data.

Business Drivers to Archiving

During May of 2011 Aberdeen surveyed 113 end-user organizations on the subject of archiving and the technologies they use to manage and retrieve legacy data. Respondents reported that the top pressure causing them to deploy an archiving solution was they had too much historical information - and how fast it keeps growing.

Figure 1: Top Pressures Driving Companies to Archive their Data



Percentage of Respondents n = 113

Source: Aberdeen Group, May 2011

Definitions

- √ **Primary Storage** – Also called production storage, is where data is held for the fastest access by end-users or applications
- √ **Backup** – A second copy of data that is used in case the primary copy is deleted or corrupted
- √ **Remote Storage** – A separate copy of the data held at an off-site location in case the primary datacenter and its contents are destroyed
- √ **Archiving** – Data help for long term storage to meet legal, tax or other compliance reasons

The cost of keeping all of this data active and on primary storage devices such as traditional SAN and NAS arrays would be excessive, hence the second most cited pressure for archiving, the cost of keeping the data active and accessible. These two pressures are really highly correlated as they were cited by virtually the same number of organizations, 37% and 36% respectively.

The third most stated pressure, the time it takes end users to gain access to historical data will be addressed in great detail later in the report. Best-in-Class organizations have processes in place to eliminate this issue.

The fourth most common archiving pressure has to do with litigation. While important, the practice of e-discovery, or managing corporate data requested in a lawsuit, will not be extensively examined in the confines of this report. The legal requirements for discovery, file management, and lawyer involvement is significantly unique that Aberdeen will cover the topic separately in later research.

The Maturity Class Framework

Aberdeen used three key performance criteria to distinguish the Best-in-Class from Industry Average and Laggard organizations in archiving practices:

- Amount of time to recover an archived file
- Number of business interruptions in the last 12 months
- Longest period of time of a business interruption in the last 12 months

Table 1: Top Performers Earn Best-in-Class Status

Definition of Maturity Class	Mean Class Performance
Best-in-Class: Top 20% of aggregate performance scorers	▪ 20 minutes to recover an archived file ▪ .5 business interruptions experienced in the last 12 months ▪ 36 minutes - longest period of downtime over the last 12 months
Industry Average: Middle 50% of aggregate performance scorers	▪ 12.4 hours to recover an archived file ▪ 2.1 business interruptions experienced in the last 12 months ▪ 5.7 hours - longest period of downtime over the last 12 months
Laggard: Bottom 30% of aggregate performance scorers	▪ 44 hours to recover an archived file ▪ 3.3 business interruptions experienced in the last 12 months ▪ 11 hours - longest period of downtime over the last 12 months

Source: Aberdeen Group, May 2011

According to Aberdeen's research, Best-in-Class companies can recover an archived file 8 to 150 times faster than Industry Average or Laggard organizations. The time of 44 hours to recover a Laggard's archived file reflects that their archives are most likely on a hierarchical-organized tape located off site. The 44 hours would include identifying the required file, requesting its delivery from the off-site location and then mounting the entire tape content for the targeted file recovery.

Aberdeen included the number of business interruptions as a key performance metric in order to identify those superior IT organizations that were better able to keep their environments up and running despite adversity.

Aberdeen also asked for companies to identify their longest downtime events as those are most likely to require recovering data from archives. Those with their archived data live and easily recoverable should have the shortest downtime events. Best-in-Class organizations reported their longest downtime events, on average were just 36 minutes vs. 6 or 11 hours of lower performing IT organizations.

The Best-in-Class PACE Model

Defining and deploying a successful strategy for archiving and preserving historical corporate data requires a set of strategic actions, organizational capabilities, and enabling technologies. Table 2 summarizes the characteristics of those leading companies. Additional information about the Best-in-Class PACE Framework can be found in Appendix A.

Table 2: The Best-in-Class PACE Framework

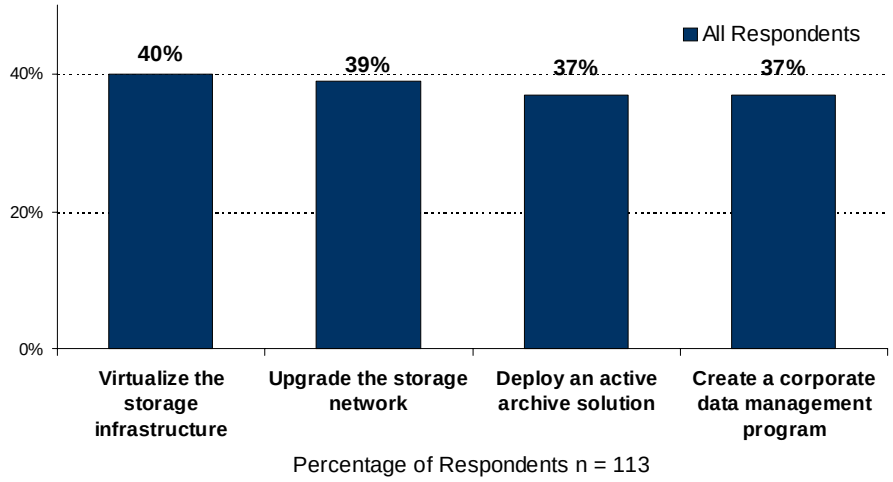
Pressures	Actions	Capabilities	Enablers
▪ Takes too long for end users to gain access to archived data	▪ Deploy active archiving solution ▪ Outsource archiving function to third-party ▪ Outsource some or all of storage to the Cloud	▪ Formal archiving training for IT staff ▪ Creation of cross-functional Information Lifecycle teams ▪ File recovery times measured and reviewed ▪ Process to ensure that redundant data is not archived	▪ Archive management software (75% Best-in-Class Adoption) ▪ Information life cycle management software (55% Best-in-Class adoption) ▪ Open archive standards (58% Best-in-Class adoption) ▪ Data deduplication (72% Best-in-Class adoption) ▪ End User initiated archiving (60% Best-in-Class adoption)

Source: Aberdeen Group, May 2011

Best-in-Class Strategies

Most reporting organizations are utilizing similar strategies when addressing the pressures listed in Figure 1. The planned strategic actions can be divided into two classes of response (Figure 2).

Figure 2: Strategic Actions Taken in Response to Pressures



Percentage of Respondents n = 113

Source: Aberdeen Group, May 2011

The first two actions, virtualizing the storage infrastructure (40%) and upgrading the storage network (39%) deal with improving the primary storage capabilities of the organization. These do not encompass traditional archiving. However, as vendors of low cost disk-based storage devices and traditional data backup suppliers have targeted the archiving market, the line between storage and archiving is becoming blurry. Organizations that are storing more legacy information in their primary storage will understandably see better performance after taking actions to upgrade those areas of their information infrastructure.

The second two actions, deploying an active archiving solution (37%) and creating a corporate data management program (37%) are more relevant to traditional archiving programs. It is good to note that 37% of the responding organizations recognized data management as an important part of a successful archiving program. Data management is an important component for Best-in-Class archiving programs (see Knowledge Management Capabilities in Chapter Two).

Archiving Metrics

After identifying companies with Best-in-Class archiving programs, Aberdeen looked at their archiving practices to measure their success. Table 3 displays a series of archive program metrics for the three maturity classes.

Definitions

√ **Metadata** – "Data about the data" includes attributes like date created, size, author, last touched, etc.

√ **Storage Tiering** – The concept of moving data to slower or less reliable storage technologies as it ages and its value diminishes.

"You can replace destroyed hardware, but you cannot replace destroyed critical business data. Protect it adequately. We use LTO Tape. That's how we recovered after a flood destroyed all of our hardware. Our tapes were safely stored off-site."

~ IT Director, Large Transportation Company, US

Table 3: Archiving Metrics

Metrics	Best-in-Class	Average	Laggards
Quantity of active data in storage	160TB	64TB	10TB
Quantity of archived data	244TB	110TB	63TB
Archive data retention plan	10.8 Years	11.6 Years	10.6 Years
Year over year archive data growth	+27%	+24%	+24%
Amount of time knowledge worker spends looking for data each week	1.7 Hours	2.5 Hours	3 Hours

Source: Aberdeen Group, May 2011

Archive Volumes

Aberdeen's research shows that Best-in-Class organizations are managing the largest amounts of both active and archived data. The Best-in-Class designation did not automatically go to large enterprises, but to those with the best archiving programs. In fact, large companies are represented in all three maturity classes, including Laggards. Small companies reported an average of just over 100TB of archived data, mid-sized organizations indicated they manage about 285TB of archives and large enterprises said they have archives with an average of 170TB. Clearly being the largest does not automatically mean being the best.

The ratio of archives to active information is also a clear indicator of superior archiving programs. The Best-in-Class reported that their archives were only 50% larger than their active storage, Industry Average companies had archives about 70% larger, and the Laggards reported their archives were up to 6 times the size of their active storage. Clearly the Best-in-Class know how to limit the amount of data they archive, delete it when it reaches the end of its life, and ensure that multiple copies of the same data are not archived. Laggard companies seem to be practicing archiving by making full, bulk copies of their data at regular intervals rather than selectively archiving corporately targeted data only. See the Process Capabilities in Chapter Two for more information on the practice of ensuring redundant data is not archived.

Definitions

- √ **Small Company** – Less than $50M in revenue per year
- √ **Mid-sized Company** – Between $50M and $1B in revenue per year
- √ **Large Enterprise** – Revenue greater than $1B per year

Archiving Retention and Growth Rates

All three maturity classes reported that they plan to hold their archived data for about 10 to 11 years. By US law, tax documents and legal records need to be kept for 7 years but other records such as medical information and pharmaceutical test results need to be retained much longer. HIPPA law in the US requires that medical records be kept for the life of the patient, potentially a century or more. Certain verticals such as gas and oil exploration plan to keep their survey data in perpetuity, as it contains information on underground formations that can be used to look for new elements far into the future. Given the multiple regulations that might impact data storage, corporations need to create a retention schedule that includes the required archive times for each class of corporate data.

It is interesting to note that all three groups reported keeping their archives for a similar period, as all classes of companies feel the same legal and tax pressures (the Best-in-Class, Industry Average and Laggards are a broad mix of industries and company sizes).

Similarly all three maturity classes reported a similar growth in their archive capacities: year-over-year rates of about 25% to 27%. This is slower than the reported growth of primary storage, which stands at 30% to 40% on average (see Required Actions in the Executive Summary). These archive growth rates will continue to increase for all organizations as the current data explosion ages to the point of requiring archiving. New sources of corporate data, including video, tweets, blogs, web content and other social media, will require archive space as their importance to the enterprise increases.

Information Search Times

The only way that archived information can continue to remain valuable to an organization is if it can be accessed by the right people in an acceptable amount of time. Aberdeen asked survey respondents to report how much time the average knowledge worker spends each week looking for information. This information search includes looking up multiple data formats, such as information on the public internet, data stored internally to the organization and archived historical data. There is a very strong correlation between having a highly organized archive repository and the amount of time knowledge workers spend seeking information. Best-in-Class organizations reported their workers spend about half as much time each week as the Laggard companies looking for information.

"Old data can only be used against you. End user retention requirements should cite a law or regulation to back up their request. Purging when that period ends is the best policy."

~ IT Staff, Very Large Shipping Company, US

How to Gain the Benefits of Active Archiving

Aberdeen asked all respondents to identify how they are able to deliver the performance reported in the metrics seen in Table 3. The Best-in-Class are gaining operational advantages from their archiving programs (including the ability to find information quickly and managing the size of their archives from overwhelming their storage capacity) and Table 4 shows several areas where they are investing in improving their archived data.

Table 4: Active Archiving

Metrics	Best-in-Class	Average	Laggards
Ability to search archived data	85%	52%	0%
Self Service file restores	33%	21%	0%
Partial retrieval of archived files	79%	43%	25%
What percentage of your archived data has live access?	56%	38%	15%
What percentage of your archived data has metadata tagging?	44%	32%	15%

Source: Aberdeen Group, May 2011

File Searching

Metadata tagging, or creating data about the data file, creates a relatively small, searchable database of the files in the archive. The metadata file is kept live and searchable by authorized individuals to assist them in finding the exact historical file they seek. Metadata tagging can be done manually or automatically depending on the type of technology an organization selects. Today's advanced tape libraries are often linked to a server hosting the metadata database. Once the targeted file is identified, the tape library can automatically select and load the correct tape cartridge, spin to the exact location on the tape where the data resides, and restore only the requested file. Disk-based archiving tools provide similar search and restore capabilities.

Best-in-Class organizations report that 44% of their archive files have metadata tagging. This percentage will grow, as current archives include files created long before automatic metadata tagging was available. As new archived files are created and the old ones deleted this percentage will rise, as will the percentage of files with live access.

The ability to partially retrieve a file is very important to companies with large files, especially rich media files such as video or audio. These organizations have the ability, for example, to restore only the first 10 seconds of a video rather than having to retrieve the entire file. Partial restores are also important for email files. Partial restore means that a single targeted email can be recovered rather than having to restore the entire individual's mailbox or the entire user group volume. Seventy nine percent (79%) of Best-in-Class companies can do partial restores of their archived data.

File Retrieval

In large organizations files are required to be restored or, more accurately "reacquired" on a daily basis. Sixty nine percent (69%) of large enterprises need to recover files on a daily or weekly basis. If IT has to be involved in every restore then they are limited in their ability to work on higher profile projects. One third of Best-in-Class organizations (33%) have the capability to perform self service file restores. This allows end users or designated individuals outside of IT to restore files, leaving valuable IT personnel out of the process.

Table 5: How Often Do You Have to Recover Data?

Metric	Small Company	Mid-Sized	Large Enterprise
Daily	10%	17%	46%
Weekly	17%	30%	23%
Monthly/Semi-monthly	22%	30%	16%
Quarterly/Yearly/Rarely	49%	20%	16%

Source: Aberdeen Group, May 2011

In summary, Best-in-Class companies keep their archives searchable by either metadata or other searchable formats. However there is more to a successful archiving program than just the technology. In Chapter Two we will examine how Best-in-Class organizations surround the archives with business practices and policies that enable success.

Aberdeen Insights — Strategy

Data backup and archiving are not the same. They serve two different purposes, treat the data differently and therefore, need to be managed separately.

Backup is a store of data used for disaster recovery. It is a (remote) copy of the primary data, which changes often, and therefore requires multiple backups (at least daily). The goal of a good backup program is fast and efficient operational recovery. Decisions on how and where you store backup data need to keep this objective in mind.

Archive data is just the opposite. Archives should contain only one copy of data, as one of the goals of archiving is to get inactive data off of the primary storage devices and free up space for newer, more strategic data. The data is static, meaning that it stopped changing, and is stored for legal, historical and compliance purposes.

Choosing the right technology and storage medium for each task is important. The benefits of disk-based storage fit well to the changing and dynamic aspects of backup. Tape, with its low cost per TB, little management overhead, and long storage capabilities fit well to archiving.

With the longer retention periods now being mandated by HIPAA, legal holds and tax authorities, the differences between backup and archiving are growing, not disappearing. Both backup and archiving are important programs to protect corporate data; don't try and cut corners by making one serve the purposes of the other.

Chapter Two: Benchmarking Requirements for Success

Aberdeen asked a mid-sized organization how they have deployed archiving within their company.

Case Study

The TV and broadcast industry manage some of the largest archives in existence. They need to protect thousands of hours of shows and clips but still have them available for rebroadcast or repurposing for the web and mobile devices.

This was the challenge for Trinity Broadcast Network, the world's largest religious network. "We needed a digital asset management solution that could store our video files indefinitely while still allowing for quick, easy access when needed," said Ben Weber, IT Engineer, Trinity Broadcasting Network in Costa Mesa, California.

At first TBN attempted to manage their broadcast portfolio with tapes. They would record a live or edited broadcast to tape. To save space on their video server they would delete the show off of the server's hard drive and just depend on the tape to save the content. When the show was scheduled for rebroadcast they would copy the tape to the server and then, again delete it off the server hard drives when finished. "This was a very manual process and took a lot of time and planning," reported Weber.

TBN then purchased a DVD changer which kept the shows recorded on DVD disks. This solution was very expensive and slow. A single show would encompass multiple disks and a disk failure required them to go back to the archive tape for recovery. The system could only accommodate about 15TB and that capacity was soon exceeded.

Now Trinity relies on an LTO5-based tape library product that supports active archiving. The library management software can be programmed to automatically load the correct shows onto the video server when required, greatly reducing the manual work to support the tape inventory. The total capacity of the library exceeds 1PB and has upward capacity built right is as LTO5 is an open standard and will support LTO6, the next generation tape format.

"The active archive provides fast, online access to our video files and is extremely dependable. Implementing this active archive solution to manage our assets was an easy decision," said Weber.

Fast Facts

We asked organizations to tell us how much data they are currently managing in their active files today:

√ Small Enterprise – 57TB

√ Mid-Sized Enterprise – 132TB

√ Large Enterprise – 193TB

Competitive Assessment

Companies must recognize that archiving is a survival strategy as well as a legal and tax requirement. However, this process must be part of a larger

effort to ensure that data is managed correctly as part of the organization's business process.

Aberdeen Group analyzed the aggregated metrics of surveyed companies to determine whether their performance ranked as Best-in-Class, Industry Average, or Laggard. In addition to having common performance levels, each class also shared characteristics in five key categories: (1) **Process** (the approaches they take to execute daily operations); (2) **Organization** (corporate focus and collaboration among stakeholders); (3) **Knowledge management** (contextualizing data and exposing it to key stakeholders); (4) **Technology** (the selection of the appropriate tools and the effective deployment of those tools); and (5) **Performance management** (the ability of the organization to measure its results to improve its business). These characteristics (identified in Table 6) serve as a guideline for best practices, and correlate directly with Best-in-Class performance across the key metrics.

Table 6: The Competitive Framework

	Best-in-Class	Average	Laggards
Process	Process for Handling Corrupted Data		
	65%	32%	0%
	Process to Ensure Redundant Data is not Archived		
	45%	25%	22%
Organization	Senior Executive in Charge of Archiving Program		
	60%	42%	0%
	IT Staff Trained in New Archiving Tools		
	70%	39%	25%
Knowledge	Formal Archiving Performance Metrics		
	60%	19%	0%
	Master Data Management Program		
	45%	31%	0%
Technology	Archive Management Software		
	75%	23%	20%
	Informational Lifecycle Management Software		
	55%	16%	0%
	Open Archive Standards		
	58%	30%	0%
Performance	IT Archiving Services Measured Against those provided by External Resources		
	37%	23%	0%
	Operational Cost Objectives Established for Archiving Program		
	50%	30%	0%

Source: Aberdeen Group, May 2011

Capabilities and Enablers

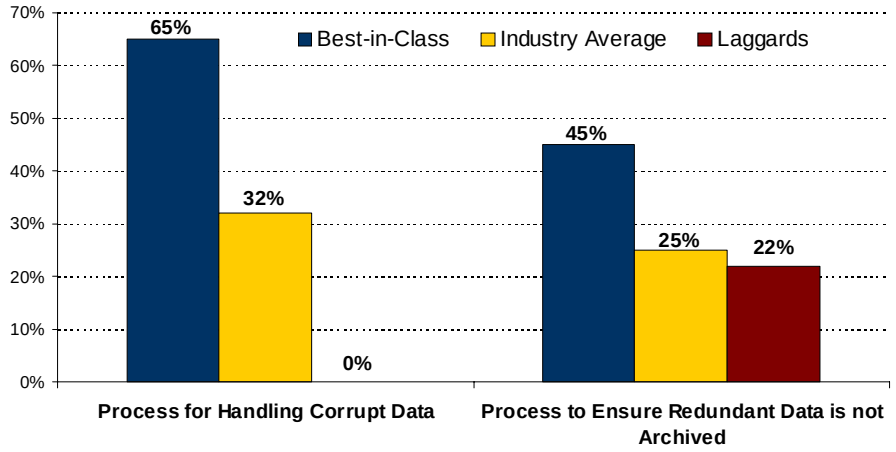
Aberdeen's analysis reveals the most important factors for Best-in-Class archiving programs.

Process Capabilities

Data corruption can be caused in many ways. When files are transmitted, created or edited they can become unreadable by the application or end-users. End-users can overwrite data with wrong attributes, software applications can be poorly coded and place data in the wrong fields, and files can be mislabeled. No matter what the cause, archiving corrupted data is at best a waste of space and money. At worst, it can cause serious problems when, years down the road, the only copy of a suddenly important piece of information is discovered to be unusable. Sixty-five percent (65%) of Best-in-Class organizations reported they have either a manual or an automatic process in place to resolve corrupted data before it reaches the point of being archived.

Archiving redundant data inflates the size of the archive without improving the effectiveness. In Table 3, Best-in-Class organizations were shown to have managed archives just 50% larger than their active data. Achieving this level of efficiency is not possible if copies of the same data are stored multiple times. Forty five percent (45%) of Best-in-Class companies report having a process to ensure that redundant data is not archived.

Figure 3: Best-in-Class Process Capabilities



Percentage of Respondents n = 113

Source: Aberdeen Group, May 2011

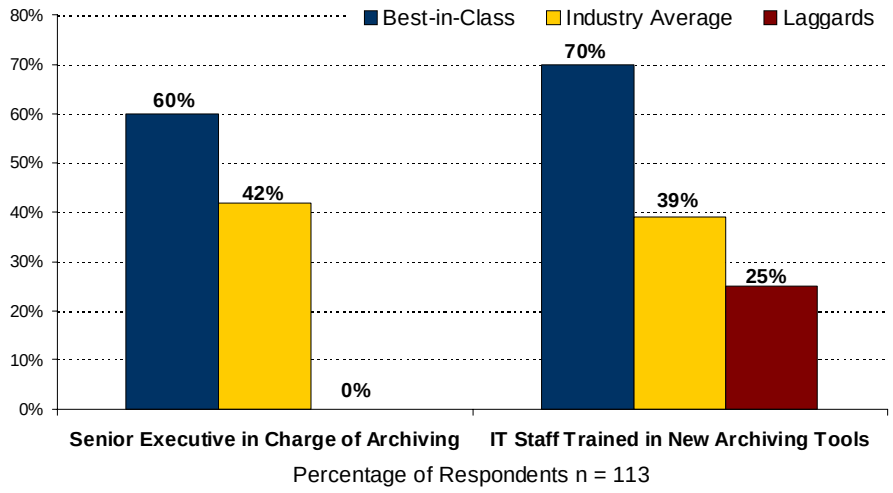
Organizational Capabilities

Sixty percent (60%) of Best-in-Class organizations reported that they have a senior executive in charge of their archiving program versus just 42% for Industry Average companies. Time and again, Aberdeen surveys show that having a designated senior manager in charge of a program is found to be a

best practice. A senior manager can keep focus on a project and ensure it stays top of mind to those charged with its execution. Senior level interest keeps the archiving program from becoming just another task on the long IT work list.

Another process that almost always is found as being a best practice is a formal training program for IT personnel on new products or technologies. Seventy percent (70%) of Best-in-Class organizations report having their IT staff trained in new archiving tools. There is nothing intuitive to understanding archiving or the software tools that enable it. Best-in-Class organizations are almost three-times more likely than Laggard companies to have a formal archiving training for their IT employees. Most vendors offer formal training classes on their products and this should be part of the educational process. Training needs to be on-going and provided in a just-in-time schedule for new employees or those assigned to new programs.

Figure 4: Best-in-Class Organizational Capabilities



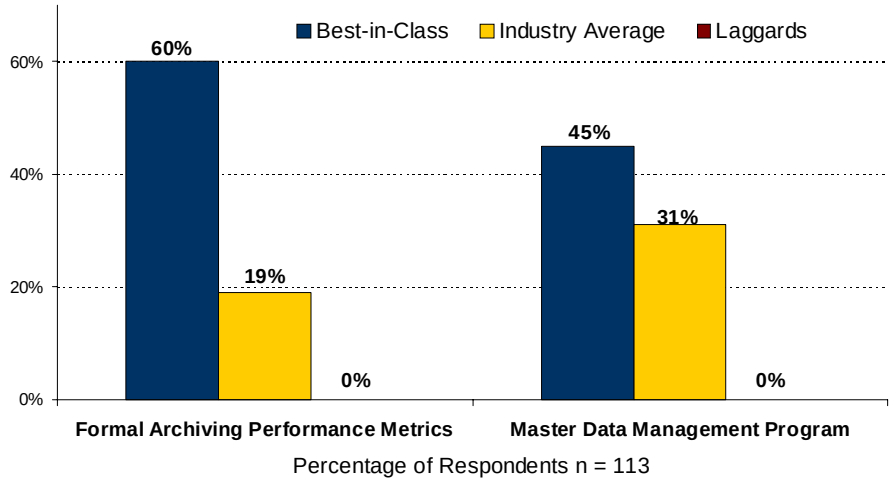
Source: Aberdeen Group, May 2011

Knowledge Management Capabilities

With no previous work to clearly define goals for implementing an IT project, the strategy for implementing archiving technology can be characterized as "ready, fire, aim." Whether driven by word of mouth, hopping on the buzzword bandwagon, or through management decree, programs are often deployed without clearly defined metrics to demonstrate project success, which means executing with no direction. For a company to understand how well it is doing, it first needs to define success and then prepare to measure their performance against it. Sixty percent (60%) of Best-in-Class organizations report establishing formal archiving performance goals for their company versus just 19% for Industry Average and 0% of Laggard companies.

In practice, Master Data Management (MDM) is all about having a single, authoritative repository for critical information, or "one version of the truth." While an MDM initiative is less a strategy and more of an ultimate goal, the importance of such a program cannot be overstated. Such a program eliminates common human errors through data normalization, supplements incomplete records through data enrichment, and allows separate data silos to be integrated into a seamless whole. In the end, a successful MDM program enables accurate, complete, high quality data to fuel the applications and processes that drive business operations. Forty five percent (45%) of Best-in-Class organizations have an MDM program.

Figure 5: Best-in-Class Knowledge Management Practices



Source: Aberdeen Group, May 2011

Technology Capabilities

Archive management software provides a single point of management for control and allows companies to meet the tight requirements of regulatory agencies such as HIPAA, SEC, Sarbanes-Oxley and other government mandates. Archive management software sets schedules for creating archives, tracks the archive assets and manages the recovery of required assets. The management software should be based on open standards to ensure access to different archiving technologies as they evolve over time. Clearly this is a Best-in-Class tool as 75% of the leading archiving companies are using this technology. Only about 25% of other companies have invested in this important capability.

Information Lifecycle Management (ILM) is a comprehensive approach to managing data from creation and initial storage to when it becomes obsolete and is archived. ILM involves all aspects of dealing with data, not just date of creation and last accessed. ILM allows companies to assign a value to data based on its usefulness to the company and then track it through its lifecycle. If MDM is about maximizing the use of data that is active and important right now, ILM takes a high level approach to corporate data

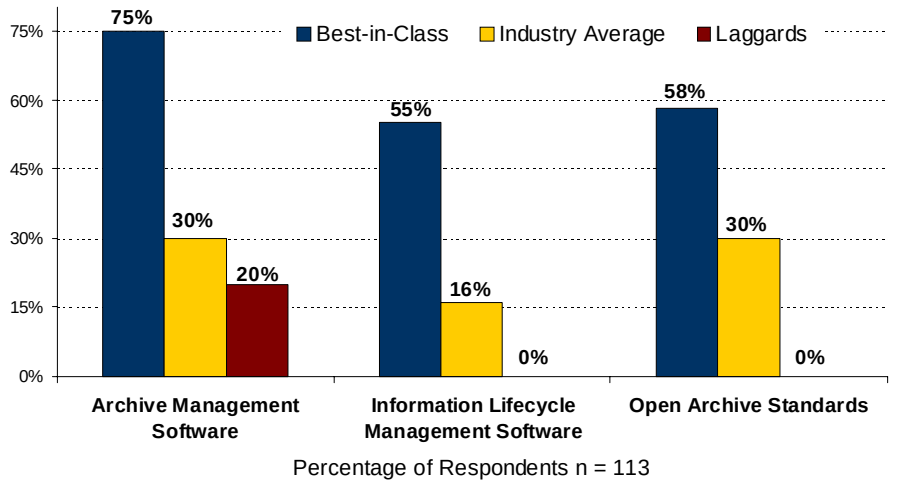
management. ILM covers the entirety of a record's existence, from creation, through any changes, moves, edits and daily use, to final archiving and deletion. Long after a record has been designated obsolete and removed from an MDM system, ILM policies will dictate where it is stored and when it will finally be retired. Fifty-five percent (55%) of Best-in-Class organizations employ some form of ILM software to help ensure they are archiving all critical data.

With the retention plans discussed in Chapter One, (see Archiving Metrics) archives that are saved for multiple decades will likely be required to migrate their technology base at least several times. As archive tools become obsolete, archived data will be required to be reformatted to meet new storage technologies. If an open archive standard has not been selected, then companies with a proprietary technology are in danger of being orphaned by their vendor. By selecting an open archiving standard, multiple vendors will offer support and migration services to the new open standard.

The following are two organizations that are promoting open archiving standards. These groups are composed of multiple vendors of archiving products that have come together to promote a standard which is supported by their products:

- www.trustlto.com – This organization is made up of nine vendors of tape-based storage products and supports the use of LTO5 tape format.
- www.activearchive.com – This alliance represents nine vendors of various hardware and software products promoting a view of the archived data, making it easier to access it whenever needed.

Figure 6: Best-in-Class Technology Enablers



Percentage of Respondents n = 113

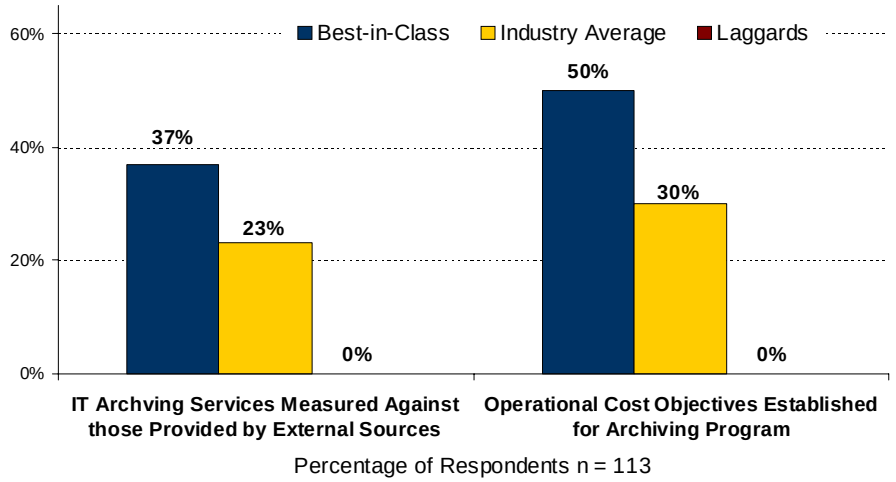
Source: Aberdeen Group, May 2011

Performance Management Capabilities

One needs only to search Google for "archiving services" or "archive as a service" to see there are hundreds of companies that will provide this service for a fee. Understandably, they have to offer a competitive product as they compete with each other for enterprise business. However, the internal IT organization should also "compete" to ensure that the services they provide and their costs are in-line with what is offered by the market. Costs well in excess of what the market offers would make the archiving program a candidate for outsourcing. Just over one third, 37% of Best-in-Class organizations evaluate their own IT performance against those offered by the market.

Similar to above, the archiving program must be managed like all other IT projects – with deliverables and within budget. Fifty percent (50%) of Best-in-Class organizations have set operational cost objectives for their archiving program. This helps ensure that the program is correctly sized and remains so through the course of the budget cycle.

Figure 7: Best-in-Class Performance Measurement



Source: Aberdeen Group, May 2011

Aberdeen also found a significant percentage of organizations that have yet to launch a formal archiving program. By following the Best-in-Class practices outlined in this chapter, enterprises will have a strong start to implementing a top performing, cost effective program for retaining and protecting important corporate historical data.

Aberdeen Insights - Technology

Aberdeen asked how often companies are required to recover files. The responses were very different, depending on the size of the organization responding.

Table 7: About how often do you need to recover data?

	Small	Mid-Sized	Large
Daily	10%	17%	**46%**
Weekly	17%	30%	**23%**
Monthly	15%	13%	8%
Semi-monthly	5%	17%	8%
Quarterly	**12%**	0%	8%
Rarely	**37%**	17%	8%

Source: Aberdeen Group, May 2011

More than two-thirds of large enterprises (69%) need to recover data on a weekly or daily basis. Almost half of small companies (49%) reported they must recover data only quarterly or rarely.

There is an important distinction that needs to be made between different types of data recovery. Daily recovery is most likely an individual employee looking to get a file or report reinstated. A volume recovery is a much rarer occurrence and happens only after the loss of an entire computing device (disk or array) or the loss of an entire datacenter. Efficient file level recovery requires different technologies, business processes and flexibility than volume based requests.

When looking at archiving tools and technologies it is important to understand the type of recovery that needs to be supported.

Chapter Three: Required Actions

Whether a company is trying to move its performance in archiving from Laggard to Industry Average, or Industry Average to Best-in-Class, the following actions will help spur the necessary performance improvements:

Laggard Steps to Success

- **Put a senior executive in charge of the archiving program.** A finding consistent with Aberdeen's research over the years is that designating a senior manager to be in charge of a program greatly increases the likelihood that the program will become Best-in-Class. When a senior manager's name is attached to an outcome, they are much more likely to ensure that the details of the program are thought through and the results get measured. This is as true for archiving as any other IT program. Sixty percent (60%) of Best-in-Class organizations have a senior executive in charge of their archiving program while only 38% of other organizations can claim the same.
- **Deploy an Information Management Lifecycle (ILM) program**. Fifty five percent (55%) of Best-in-Class organizations use ILM while only 15% of other companies do so. All data is not equal. Some has much more importance and value to an enterprise, and that information must be afforded the best protection. What data is crucial to protect is too important a decision to be left to IT. A company wide information management team must decide the correct level of importance to assign to each set of data in the company's possession.
- **Select archive solutions based on open standards**. Respondents to this survey cited that they intend to keep their archives for over 10 years. In that time, many IT technologies will be developed, deployed and retired. If a proprietary technology is chosen to support the archiving program there is a greater likelihood that it could be discontinued, with the worst case being that archived files could not be restored. Open standards are supported by many vendors and are unlikely to be abandoned with no fallback solution supported in the market.

Industry Average Steps to Success

- **Develop an archiving training program for IT**. It is surprising to see that the majority of organizations (56%) report that they have no formal training program for their IT employees on new archiving tools. There is nothing intuitive or obvious about any IT product, how to deploy it or establish the appropriate practices and policies. Formally training your IT employees may be one of the

Fast Facts

We asked organizations to tell us what percentage of their archived data is stored at remote locations:

√ Small Enterprises - 52%

√ Mid-Sized Enterprises - 54%

√ Large Enterprises - 50%

easiest and most cost effective ways to ensure a new tool deployment with superior returns.

- **Define a process for ensuring redundant data is not archived**. Aberdeen has shown that Best-in-Class organizations have archives half the size of their active data. Part of the reason for this is that they ensure that they are not archiving the same piece of information multiple times. Forty five percent (45%) of Best-in-Class organizations have this sort of process while only 23% of other companies do so.
- **Set operational cost objectives for the archiving program**. No corporate program should have a blank check. We all live within a budget, and the corporate archiving program should as well. Prior to investing in any archiving tools IT should propose an appropriately sized archiving program and the budget required to manage it successfully. They should then be held accountable for operating within that budget. Fully half (50%) of Best-in-Class organizations have set operational cost objectives for their archiving programs while only 27% of other organizations have such a practice.

Best-in-Class Steps to Success

- **Create a process for self service restores.** Table 5 highlights that 69% of large companies are required to recover archived files either daily or weekly. Tasking IT resources with this amount of work will keep them from higher priority tasks. Creating a process with appropriate management approvals for self initiated restores puts the work where it should be, on the person making the request. Only 33% of Best-in-Class organizations have the capability today.
- **Increase the percentage of the archives with metadata tagging.** Metadata tagging is a technology that allows for quick searches of archived data. However, the study found that on average only 44% of Best-in-Class archives have files tagged with metadata. This percentage will grow as new archives are created with today's technology and old archives are destroyed. Increasing this number will make a higher percentage of a company's archives active and available to end users.
- **Measure an internal archiving program against that which can be supplied by external sources**. One popular method of right sizing any IT program is to see how it compares to the cost of outsourcing the program. If the program is costing a company $X and third-party suppliers can do if for half that amount, the program is a strong candidate for outsourcing. Thirty seven percent (37%) of Best-in-Class organizations already compare their archiving program costs to those charged by external sources, but every organization

should do this research, as it forces internal IT groups to be as efficient as they can.

Aberdeen Insight - Archiving Overview Options

Archiving has changed considerably in the last several decades. Historically, the most utilized option for preserving historical records was to put physical documents and tapes of digital records into large white boxes and ship them offsite to a storage facility.

Today networks, clouds, tape libraries and high volume disks are providing options for users to select to closely align to their archiving storage to their recovery needs. The tapes are faster and hold higher volumes, the wide area networks are faster and disks can be put into warm states; all of which help keep the cost of archiving in control and allow them to be part of an overall archiving strategy.

With choice comes the need for end users to do their homework. Investigate and select the products that cost effectively allow for archived searches and fast recovery – how ever you define those parameters.

Appendix A: Research Methodology

In March 2011, Aberdeen examined the use, the experiences, and the intentions of 113 enterprises and their archiving programs.

Aberdeen supplemented this online survey effort with interviews with select survey respondents, gathering additional information on archive strategies, experiences, and results.

Responding enterprises included the following:

- *Job title:* The research sample included respondents with the following job titles: CEO / President (22%); EVP / SVP / VP (15%); Director (14%); Manager (22%); and Other (27%).
- *Department / function:* The research sample included respondents from the following departments or functions: Information Technology (49%); Business Development (9%); Corporate Management (8%); Product Management (7%); Operations (5%); and other (22%).
- *Industry:* The research sample included respondents from the following industries: IT Consulting Services (23%); Software (16%); Financial Services (15%); Government (9%); Education (6%); Computer Equipment (5%); Telecomm (5%); and others (21%).
- *Geography:* The majority of respondents (70%) were from North America. Remaining respondents were from the Asia-Pacific region (13%), and Europe/Middle East/Africa (17%).
- *Company size:* Twenty percent (20%) of respondents were from large enterprises (annual revenues above US $1 billion); 27% were from midsize enterprises (annual revenues between $50 million and $1 billion); and 53% of respondents were from small businesses (annual revenues of $50 million or less).
- *Headcount:* Thirty-three percent (33%) of respondents were from large enterprises (headcount greater than 1,000 employees); 25% were from midsize enterprises (headcount between 100 and 999 employees); and 42% of respondents were from small businesses (headcount between 1 and 99 employees).

Study Focus

Responding executives completed an online survey that included questions designed to determine the following:

√ The size of their archives

√ How long they keep archived data

√ Supporting technologies

√ Business processes that ensure Best-in-Class performance

The study aimed to identify archiving best practices and to provide a framework by which readers could assess their own long term data storage strategies.

Aberdeen Group
A Harte-Hanks Company

Table 8: The PACE Framework Key

Overview
Aberdeen applies a methodology to benchmark research that evaluates the business pressures, actions, capabilities, and enablers (PACE) that indicate corporate behavior in specific business processes. These terms are defined as follows: **Pressures —** external forces that impact an organization's market position, competitiveness, or business operations (e.g., economic, political and regulatory, technology, changing customer preferences, competitive) **Actions —** the strategic approaches that an organization takes in response to industry pressures (e.g., align the corporate business model to leverage industry opportunities, such as product / service strategy, target markets, financial strategy, go-to-market, and sales strategy) **Capabilities —** the business process competencies required to execute corporate strategy (e.g., skilled people, brand, market positioning, viable products / services, ecosystem partners, financing) **Enablers —** the key functionality of technology solutions required to support the organization's enabling business practices (e.g., development platform, applications, network connectivity, user interface, training and support, partner interfaces, data cleansing, and management)

Source: Aberdeen Group, July 2011

Table 9: The Competitive Framework Key

Overview	
The Aberdeen Competitive Framework defines enterprises as falling into one of the following three levels of practices and performance: **Best-in-Class (20%) —** Practices that are the best currently being employed and are significantly superior to the Industry Average, and result in the top industry performance. **Industry Average (50%) —** Practices that represent the average or norm, and result in average industry performance. **Laggards (30%) —** Practices that are significantly behind the average of the industry, and result in below average performance.	In the following categories: **Process —** What is the scope of process standardization? What is the efficiency and effectiveness of this process? **Organization —** How is your company currently organized to manage and optimize this particular process? **Knowledge —** What visibility do you have into key data and intelligence required to manage this process? **Technology —** What level of automation have you used to support this process? How is this automation integrated and aligned? **Performance —** What do you measure? How frequently? What's your actual performance?

Source: Aberdeen Group, July 2011

Table 10: The Relationship Between PACE and the Competitive Framework

PACE and the Competitive Framework – How They Interact
Aberdeen research indicates that companies that identify the most influential pressures and take the most transformational and effective actions are most likely to achieve superior performance. The level of competitive performance that a company achieves is strongly determined by the PACE choices that they make and how well they execute those decisions.

Source: Aberdeen Group, July 2011

Appendix B:
Related Aberdeen Research

Related Aberdeen research that forms a companion or reference to this report includes:

- *Disaster Avoidance and Disaster Recovery: Making your Datacenter Disaster Resilient*; May 2010
- *Datacenter Downtime: How much does it really Cost?*; June 2010
- *"Bare it" or Dare it: The Benefits of Image-based Recovery*; June 2010
- *Off-site Storage and Computing: Keys to Successful Disaster Recovery*; September 2010
- *Managing Virtualized Applications: Optimizing Dynamic Infrastructures*; April 2011
- *Reduce the Cost of Downtime Protection: Tier your Virtualized Applications*; April 2011
- *High Availability for Virtualized Applications: Protecting Against Unplanned Downtime*; May 2011

Information on these and any other Aberdeen publications can be found at www.aberdeen.com.

Author: Dick Csaplar, Senior Research Analyst, Virtualization and Storage (richard.csaplar@aberdeen.com)