



# Strategic Snapshot

## Mapping the Business Value of Data Backup and Archival Solutions

By Clay Ryder

The Sageza Group, Inc.  
January 2005

[sageza.com](http://sageza.com)  
[info@sageza.com](mailto:info@sageza.com)

**The Sageza Group, Inc.**  
32108 Alvarado Blvd #354  
Union City, CA 94587  
650-390-0700 fax 650-649-2302  
London +44 (0) 20-7900-2819  
Milan +39 02-9544-1646

# Mapping the Business Value of Data Backup and Archival Solutions

---

## ABSTRACT

***In this first of two papers, we differentiate the roles of backup and archival processes and examine various technologies in the marketplace that seek to help organizations operate effective solutions for each of these two areas. In our next paper we will highlight the business benefits of leveraging both processes in a unified strategy.***

*Over the past year, Information Lifecycle Management (ILM) has become the concept du jour among storage vendors, their customers, and the greater IT industry. Much of ILM's charm lies in the simplicity of the idea that since the value of information naturally declines over time, wise organizations make certain that information resides in tiered storage solutions whose costs and capabilities closely match its own. However, getting from the here of complex, even chaotic enterprise storage to the there of integrated ILM is hardly so simple. While the rote process of such a transition is difficult enough, organizations must contend with a wide range of disparate commercial ILM strategies and technologies. Even worse, vendors who lack a full range of appropriate solutions sometimes trumpet products that they claim can be leveraged, Swiss Army Knife-style, across a host of applications, business needs and time horizons. This problem has become especially severe in the data backup and archival space, two different and distinct areas of particular interest to ILM proponents. While at first glance backup and archival processes might appear similar, they are in reality quite different; as such requiring significantly different solutions to achieve their business goals in an efficient and effective fashion.*

*In this paper, specific to archiving, we examine the capabilities of EMC's Centera storage solution and reflect on how this platform offers unique capabilities in the marketplace.*

# Mapping the Business Value of Data Backup and Archival Solutions

---

## TABLE OF CONTENTS

Backup vs. Archive: Isn't It All the Same? .....	1
What is Backup? .....	1
Understanding Customer Pain: Business Drivers for Backup .....	1
A Quick Look at Other Technologies that Can Be Used for Backup .....	1
What Is an Information Archive? .....	2
Understanding Customer Pain: Business Drivers for Archiving .....	3
A Quick Look at Technologies that Can Be Used for Archiving .....	3
EMC Centera: An Interesting Archive Solution.....	4
Sample Scenarios Where Centera Fits Customers' Needs .....	4
Conclusion .....	5

## Backup vs. Archive: Isn't It All the Same?

While some may interchangeably use the terms backup and archive, the difference between backup and archive processes/solutions surface with respect to the purpose and timeframe driving the need for access to stored information. What follows is a brief definition and comparison of the differences between backup and archival processes.

### What is Backup?

Backups are secondary copies of primary information. They

- ◆ provide short-term protection of production data to ensure business continuity,
- ◆ are generated point-in-time and typically in a periodic automated fashion, and
- ◆ are systematically overwritten.

The key value of backup is the ability to provide a point-in-time copy of information in order to protect critical business processes. For this reason, backup solutions are appropriate solutions for business continuity and disaster recovery.

### Understanding Customer Pain: Business Drivers for Backup

Historically, tape-based storage is the traditional choice for many backup and archive applications. Reasons include a bottom-line focus on tape's lowest initial purchase price and that it is widely embraced as a de facto accepted media for WORM-based compliance solutions.

Despite these benefits, there is a plethora of stories highlighting the challenges of recovering from tape backups. Tape technology for backup falls short in key areas. Specifically, it offers low data access performance – relative to disk-based storage – limiting its effectiveness as a backup technology. In sharp contrast with its lower entry price, tape-based solutions incur higher management and maintenance costs, are prone to mechanical error, and deliver slow recoverability of data. These negatives are compounded by the fact that over time many organizations are left at a dead end with obsolete tape solutions as the technology has undergone many evolutions since first introduced.

### A Quick Look at Other Technologies that Can Be Used for Backup

#### *Optical*

Optical-based storage provides low performance making it an inappropriate choice for backup. Optical-based storage typically incurs higher management and maintenance costs than disk-based storage. Similarly, as with tape, the evolution of optical storage technologies has left many with obsolete optical solutions based upon state-of-the-art technology from not all that long ago.

#### *FC Disk*

Conventional (FC) disk solutions have emerged as appropriate storage solutions for primary backup and business continuity needs. Notably, they offer significantly better performance and reliability than tape- or optical-based storage. However, these solutions also tend to require a dedicated IT staff for management and maintenance, creating considerable long-term cost concerns. They provide speed, but in general with a higher initial cost, and do nothing to lower the total cost of ownership. Customers choosing FC disk solutions for backup often cite the need to replicate information for online offsite access as a deciding factor.

### ATA Disk

ATA disk is a relatively recent alternative that is being implemented in very different ways for backup and archives. ATA-based solutions offer significantly better performance than tape- or optical-based storage and are more cost-effective than FC Disk solutions. Although ATA disk solutions incur slightly higher initial costs than tape- or optical-based solutions, they are notably less expensive to manage and maintain.

While tape- and optical-based storage solutions are appealing for their lower entry price point, their initial costs can be dwarfed by high ongoing management costs. Over the long term, there are simply fewer mechanics involved in disk-based solutions than in tape-based solutions (e.g., tape libraries struggle to offer swift access to data). The ongoing management benefits of disk-based solutions are demonstrable and offer a more reliable platform on which to build backup or archives.

The introduction of ATA-based disk solutions focused on Disk-to-Disk backup has organizations poised to embrace these new cost-effective online solutions. However, although ATA offers great advantages as a storage medium, when looking to use it for an archive (where information is moved from transactional storage to the archive) extensive software capabilities are required to align it fully with the various needs of archiving. Needs such as long-term retention and protection, assured authenticity, single-instance store, ensuring application policies are enforced intrinsic at the storage layer, and so forth.

### What Is an Information Archive?

Information Archives are the primary copies of information. The characteristics of archived information are:

- ◆ It is valuable and retained for future reference. For this reason, information authenticity must be assured.
- ◆ It is typically in its final form, and subject to limited or no modification.
- ◆ As the sole copy of information, archived information must be retained for long-term periods (e.g. months, years, or decades). Think of a thirty-year mortgage, medical images, or product warranty information
- ◆ Archives focus on access and retrieval of a specific piece of information rather than all the content (a.k.a. backup).
- ◆ Archival processes often include specific ILM time frames including deletion.

Desired features include:

- ◆ *Fast retrieval access* to content with assured information authenticity for the life of the information;
- ◆ The ability at the storage layer to enforce an application's retention and disposition policies on a piece of information;
- ◆ Even in the largest of archives with the longest of retention periods, the ability to eliminate the management requirement for backup due to the capability to safeguard information at its primary location and replicate to another location for disaster recovery.

Whether for day-to-day business, governance, or regulatory purposes, in today's business climate it is an archive's ability to deliver on both the base attributes and desired features that creates maximum information return.

## Understanding Customer Pain: Business Drivers for Archiving

Traditionally, the range of solutions has offered some but not all the features needed for an effective archive. For example, high-cost disk-based solutions do not offer content authenticity, the more-affordable optical-based solutions sacrifice access speed for authenticity, and tape falls short on both fronts. To maximize the business value of their content, organizations need fast online standards-based access to content from any application on any platform. Organizations need to be able to assure that content has not been modified. These are requirements for day-to-day business, internal governance policies, compliance with regulatory requirements.

Ongoing cost control remains a concern and areas where costs can be reduced include not storing duplicates of original content, requiring minimal systems management, and moving inactive or final form content off of primary storage and into an archive. An ideal archive solution offloads a production environment of its final form information without negatively impacting retrieval time should the information be needed. To accomplish this task an effective archive must be online and able to scale massively (to petabytes of storage, or billions of items if needed) while serving multiple applications in an architecture designed for the future evolution of business storage needs.

## A Quick Look at Technologies that Can Be Used for Archiving

### *Tape*

As mentioned earlier, tape has been widely embraced for archiving because it has the lowest initial purchase and is a de facto accepted media for WORM-based storage. Despite these benefits, tape is a backup medium at best: it provides no simultaneous access to information, and it is best for serial reading of an entire tape rather than random read-and-write requests for various pieces of information. It is prone to mechanical error, which often results in the policy of keeping many multiple copies of the same information, generating substantial management costs when information needs to be deleted. It is not a technology that provides fast recovery for business continuance or disaster recovery. In addition, tape has a history of technology obsolescence. All of these facts add up to slow access to information, high management costs, high maintenance costs, and a high overall total cost of ownership. These limitations are diametrically opposed to the needs of today's Internet world and the ability to get maximum business value from archived information.

### *Optical*

Optical-based storage is the traditional choice for many archive applications in large part due to its acceptance for WORM-based compliance solutions. Its competitive entry price and reasonable scalability make it initially appealing for enterprise archival storage, but it suffers from the majority of limitations that tape does. Optical access times to information are slow; content is managed by optical platter and information retention policies at a platter level. Similarly, as with optical for backup, the road is littered with customers that have dead-end, technology-obsolete solutions they must live with.

### *FC Disk Using a File System's Architecture*

A conventional (FC) disk solution incurs higher initial and upgrade costs, making it an inappropriate choice for a long-term archive. FC disk storage solutions also tend to require a dedicated IT staff for management and maintenance, creating considerable long-term cost exposure. For compliance-focused shops, FC disk is not acceptable as a WORM-based compliance solution.

### *ATA Disk*

Organizations could benefit from deploying an online ATA disk storage solution that is cost-effective, but it must be application- and information-aware. It must have a rich software layer that, combined with the hardware, meets today's needs for archives. Software that has the built-in intelligence to eliminates the negative impact of file systems, keep just one copy of a unique piece of information, and be able to enforce an application's policies for retention and deletion at the storage layer. These are but a few of the features that an intelligent, purpose-built ATA hardware and software solution must have to be appropriate for archive needs.

Consider a law firm that archives its emails to just ATA disk and cannot prove their authenticity, or has documents associated with multiple lawsuits stored multiple times where the process can not assure that content is retained or deleted appropriately because it does not know how many copies are being stored.

Intelligent software that works with and at the storage layer is a necessity to get maximum business value from archived information.

## EMC Centera: An Interesting Archive Solution

One vendor that has risen to meet today's enterprise archiving storage needs is EMC with its Centera storage solution. As a combined software and hardware solution, Centera's ATA-based content addressed storage (CAS) approach is a well-established solution that offers capabilities desired by archival customers heretofore not available in a unified offering. It delivers enhanced speed of access to information compared with tape- or optical-based storage as well as assured content authenticity. Centera's use of digital fingerprinting to create location-independent content addresses for unique content reduces management overhead and results in single-instance storing which dramatically improves storage efficiency. It guarantees integrity and, by utilizing a metadata layer, enforces any application policies intrinsic to storage at the storage layer.

Centera addresses concerns over long-term management and maintenance through its self-healing and self-managing capabilities that increase ATA disk stability and availability while lowering the TCO. Organizations concerned with governance or regulatory compliance should take note of Centera's Compliance Edition, which has been used by many organizations for governance, regulatory-sensitive, and compliance-focused applications. With better price and performance than traditional tape- and optical-based offerings, Centera unites the best of many storage technologies into a solution.

The ideal archive storage solution should possess a feature set that stands alone yet is flexible enough to work with complementary offerings be they primary, backup or software solutions. Centera has over 160 integrated partner offerings from vendors including KVS, Mobius, Connected, and EMC's own Documentum and Legato. Areas of application concentration include email archiving, ECM, workflow management, medical imaging, mainframe output archiving, and others. Through Centera Universal Access software, non-integrated applications can store and retrieve content from Centera via standards including NFS, CIFS, FTP, and HTTP.

## Sample Scenarios Where Centera Fits Customers' Needs

Some companies that have deployed an array of disk-based storage may give in to the temptation to drop in another ten drives to deal with their archive storage needs rather than discussing how to achieve greater benefits through implementation of a purpose-built solution. Scenarios where customers could deploy Centera for archiving include:

- ◆ Customers needing a unified single solution: cost effective, fast access with assured information authenticity as a means to get maximum business value from their archived information, whether it be to create higher SLAs or to create new service offerings.
- ◆ Customers seeking to migrate all or most of their final form content from tape- or optical-based media to achieve better performance and better reliability, and to assure content authenticity and reduce management costs.
- ◆ Customers seeking, for all of their archiving needs, a future-proof archiving architecture and standard regardless of application platform or access method.
- ◆ Customers seeking to shift from using expensive primary disk for archiving to achieve low-cost, fast access with assured content authenticity
- ◆ Customers seeking the most cost-effective backup strategy to deploy. How do you make backup 30% faster and more cost-effective? Answer: Archive 30% of your information.

Centera is uniquely positioned to address archival needs. It supports features that should appeal to IT staffs, including single-instant store, self-healing, self-management, and self-configuring capabilities. Collectively, these help ease long-term storage management.

Enterprises and other organizations require an archival solution that can be used by all applications for all their information types. Centera meets this need. As a solution offering multi-application support, Centera can centrally store and share business content while preventing multiple copies of content from being stored or creating information silos. This obviates the need to purchase and deploy separate storage solutions for each application (e.g., one for email, one for content management), uniting storage across applications. Centera's location-independent content addressing enables the sharing of information across applications.

## Conclusion

In today's business climate, IT attitudes toward enterprise storage are increasingly turning strategic with a specific focus on ILM. Enterprises seeking a successful approach to ILM must embrace a strategy that appropriately addresses the specific distinct requirements of backup and archiving. While both processes may seem similar conceptually, their respective roles relating to preserving and adding value to enterprise information assets cannot be misunderstood or inappropriately deployed. Each requires a significantly different solution to work effectively. Enterprises that evolve their storage accordingly are most likely to achieve success in ILM.

A range of solutions are available, from tape- and optical-based storage to FC-disk and ATA disk. Each of these requires organizations to consider fit, cost, and ongoing operational expense. Overall, businesses are best served by choosing and working with IT vendors whose backup and archive solutions are specifically designed to meet the needs of each market.

EMC's Centera archiving storage solution provides notably price and performance benefits over tape- or optical-based storage. Centera also possesses a unique feature set that affords enterprises the opportunity to deploy one solution across a range of applications. Any organization with a large repository of critical information that must be stored for reference will appreciate Centera's ability to provide online access with assured content authenticity at a tape TCO.

Enterprises with adequate storage solutions should take a moment to reassess their current deployment and ask if there are opportunities to add value, achieve greater storage efficiency, and improve the ROI from their information assets. Those that are serious about their ILM strategy would be well advised to consider EMC's Centera as a way to maximize the value from their archived information. As the needs of enterprise archiving continue to rise, from our perspective EMC has risen to meet them with Centera.