

WHITE PAPER

Driving to Best Practices in Archiving

Sponsored by: Plasmon

Laura DuBois

March 2007

EXECUTIVE SUMMARY

Archiving, at once a set of policies, processes, and technologies, is increasingly a business imperative in today's legal, regulatory, and IT-complex enterprise environments. While firms face significant data growth year over year, technical administrative resources to manage those increased capacities remain fixed. Challenges such as record retention to ensure regulatory compliance and preservation of information in response to legal holds are causing firms to save more and more data. As a result, disparate archives have appeared. The management of disparate content repositories, archives, and archive technologies while striving for administrative efficiency, investment protection, and disaster recovery can quickly make archiving overwhelming.

As a result of these business and IT challenges, firms are leveraging best practice approaches to archiving. These best practices to archiving include a combination of process, policy, and technology to meet economies of scale, reduce risk, and improve efficiencies. As a first step, firms must identify and classify information assets and assign policies to these assets based on data content, business rules, and IT criteria. Once this inventory is done the automation of defined policies through technology can reduce the cost to manage the storage environment and minimize the level of risk a firm faces. Plasmon's unique, enterprise-class archiving solution, the Ultra Density Optical (UDO) Archive Appliance, combines the performance and simplicity of network-attached RAID cache with the longevity and authenticity of UDO media to provide easy access, read/write performance, long-term retention capability, portability, and flexible management required to optimally align today's archive best practices with a firm's core business objectives.

SITUATION OVERVIEW

Today's business and IT organizations are facing the challenges of the convergence of IT, legal, and regulatory pressures. As more enterprise information is created electronically, the growth and importance of this archive content to a firm increase significantly. Today's firms face the following challenges related to archiving:

Ensure regulatory compliance. Regulatory mandates, such as SEC 17a-4, NASD 3110, HIPAA, CFR Part 11 and SOX, subject companies in many different industries to requirements that specific business records are captured and retained in a way that ensures record integrity, reliability, and accessibility, while some also prescribe the media used for storage be non-rewriteable. These regulations also stipulate information retention, which should encompass the index and storage of information

for specified time periods. Central to satisfying regulatory compliance requirements is the fast provisioning of records in response to an audit, typically within a 24- to 48-hour time period.

Mitigate ediscovery risks. The recent amendments to the U.S. Federal Rules of Civil Procedure have placed more process and control around electronic discovery. Moreover, state and federal court rulings have become more stringent regarding information destruction and tampering. In an increasingly litigious business environment, firms must comply with a duty to preserve electronic information for the purposes of discovery under legal hold orders. During a discovery request, the time and cost of retrieving emails and other documents with traditional serial-based archive technologies, such as tape, can be costly in administration time and lost productivity. Efficient retrieval of authenticated, indexed information in response to electronic discovery is an absolute requirement in today's environment.

Manage disparate repositories, archives, and technologies. Today firms seek to archive content from different applications, both commercial and home-grown applications, into a single-archive architecture. However, it is not uncommon for a firm to have not only different content sources but also disparate and different physical archives and archive technologies. The operating and capital costs are high in management, training, and technology for disparate systems and limit horizontal information management services such as centralized search/retrieval and data security, authentication, and integrity.

Reduce administrative and capital costs. IT organizations seek operational efficiencies while protecting investments in archive infrastructure. The consequences of managing disparate archives are at odds with goals seeking to improve IT operational efficiencies and share common knowledge across the organization. IT organizations want to protect investments in technology, application, and IT strategy. If technology is adopted and rolled out across the enterprise but cannot integrate with other components along the technology stack, it will limit the ability for a firm to maximize its investment. It is important that software, systems, and storage, both legacy and the latest technology, work together to form an integrated, holistic management approach that allows a firm to optimize and continue to expand protection of its previous technology investments.

Ensure disaster recovery. Disaster recovery processes and technologies are critical for business and legally valuable archive data. With regulations focusing on ensuring that corporate information is protected and "safeguarded," a disaster recovery plan and approach for both primary and archive data are important. Archive data must be recoverable in the event of physical disasters, and the security of the data at rest must also be addressed. While some firms have started to rely more upon remote replication as a disaster recovery option, there are many firms that use low-cost removable media solutions for offsite storage.

BEST PRACTICES IN ARCHIVING

Best practices in archiving take into account both the business and technical requirements of the firm. At once, these best practices include both business rules for how a firm wants to manage its critical information assets while also meeting the cost, performance, availability, and security requirements established.

Process Best Practices

Proactive Information Inventory Assessments

As a result of the changes to the rules of U.S. civil procedure, be prepared and understand what information repositories exist. Once an inventory of all information assets has been conducted, firms must establish policies against these information assets. Policies should incorporate access/retrieval requirements, retention and disposition requirements, as well as protect sensitive information about the company, its employees, and partners from compromise.

Policy Development and Documentation

Archiving is a horizontal, cross-functional process that instantiates business rules through technology. The most critical first step in archiving is policy development and documentation. Once technology is used to implement these policies, there should be some means to validate or prove they are in effect.

Establishing Comprehensive Policies

Policies are really business rules that should take into account regulatory (retention, integrity, availability), legal (preservation), and business disposition and reuse requirements across systems. Increasingly, firms are taking into account not only external requirements but also internal IT governance for how it wants to manage its valuable information assets, such as financial, employee and customer data, as well as intellectual property.

Taking a Team-Based Approach

Archiving needs to take into consideration the different needs of key technical and business stakeholders. Increasingly, firms are establishing tech-based approaches to information management including archiving. Key stakeholder organizations on an archive team typically include finance, risk management, legal, compliance, security, records management, IT, and line of business managers.

Technical Best Practices

Move Reference Data From Primary Production Storage Systems While Still Providing Access

IDC refers to two general classes of data used by applications: reference data and transactional data. Transactional data is data created and managed during business transactions and is related to specific business processes. Reference data is typically fixed or static data that describes facts about the business that are referenced during the course of business but have no specific process-determined point of termination.

The amount of non-transactional reference data continues as a percentage of overall storage and is typically over 50% depending on the internal systems installed, business processes performed, and vertical industry in which the firm operates. The migration of reference data to lower-cost media while still providing accessibility improves production system performance and lowers production storage costs while compressing production backup windows.

Retain Redundant Copies of Reference Data

Commonly, firms are retaining two copies of reference data in multiple locations to enable disaster recovery and reduce the risk of data compromise. For heavily regulated firms, often the disaster recovery copy is moved to a vault or bunker, which is fireproof and has limited physical access. As a result, firms often rely upon removable media.

Provide Fast, Random Access to Archive Content

Firms are increasingly using fast, random access media for archive content to streamline and optimize location of relevant content as a result of a regulatory audit, discovery request, or business need. IDC finds that response times to digital content are only compressing. Regulatory audits, legal discovery, and business needs require that information be produced quickly and accurately — in hours, not days.

Ensuring Tamper-Proof, Record Authenticity

Some of the most rigorous regulations require that data be stored in a non-rewriteable format. The most sensitive and risk-averse environments want this level of WORM protection to be done at the hardware level. Firms require that data stored in an archive does not change over time, cannot be modified or deleted during its retention period, and rely upon technology to verify record authenticity for legal and regulatory purposes.

Archive Longevity Commensurate with Data Life

Many firms today face retention periods that are permanent with infinite archive of things such as financial records and customer contracts. Other firms retain relevant patient or financial records for between 50 to 100 years. These retention periods of 50 years to forever require longevity in the media and the applications that store the data.

Design for Architectural Scalability

While a current archive might range from 50–100TB today, depending on data growth rates as well as retention, duty to preserve and legal corporate requirements, this archive could grow significantly in a 5-year period. Firms are designing their archives to scale to growth 5 to 10 years out with architectures that can migrate to next-generation technology in a non-disruptive, online manner.

Ensure Multiple Copies for Recovery

Many firms want multiple archive copies in place, one local copy for access and another as a backup safety measure for disaster recovery. With archive content being critical to business environments, firms must have a copy that is available an adequate distance to survive a regional disaster.

Use Standards-Based Interfaces

Most organizations have had poor experiences with custom API integration or lack of integration with internally developed applications. Increasingly, best practices mean use of open standards, such as CIFS, NFS, WBEM, and developing standards, such as XAM. This standards-based support makes application integration easier to implement and maintain.

PLASMON'S APPROACH TO ARCHIVING

Founded in 1984 and a pioneer in archival storage systems, Plasmon has over two decades of expertise in developing fast, secure, reliable archiving solutions for long-term storage and data retention. Plasmon technology is installed in firms worldwide across a broad range of industries, including healthcare, financial services, manufacturing, law enforcement, and military and government agencies. Plasmon solutions meet all current quality standards and industry certification, including ISO 9001:2000, and are frequently cited as approved media sources in government regulations. Plasmon focuses exclusively on archive solutions, and its offerings include archive systems based on UDO, the first storage technology specifically designed for long-term professional data archive requirements.

In 2003, Plasmon responded to customer demand for long-term archival solutions that combined the strengths of both disk cache, UDO media, and policy management. The result, The UDO Archive Appliance, offers the only enterprise solution that ensures data authenticity, access, longevity, and portability that are required by new regulations.

By eliminating reference data from production systems and permanently archiving critical data to a UDO Archive Appliance, IT executives can remove archive data from the backup window. UDO media assures data is safe, secure, and can be readily accessed for long periods without the need for backups. Plasmon solutions improve operational efficiencies, decrease costs, and reduce business risks.

UDO Archive Appliance

The UDO Archive Appliance is designed with a fully integrated architecture that leverages several components: a NAS server as the network access, RAID cache for fast read/writes, archive software to control data migration within the appliance, and UDO media for long-term data retention. The UDO Archive Appliance NAS interface presents a standard file system supporting NFS and CIFS protocols. The NAS interface simplifies installation and configuration. Any application that supports NFS and CIFS can easily be integrated with the UDO Archive Appliance including email archive, medical imaging, and content management applications. As archived data is written to the UDO Archive Appliance, data is cached on RAID for rapid access as well as immediately committed to UDO for long-term retention. Optionally, IT administrators can use the archive software to set policies to control the archive process based on capacity, time, or event. Off-line replication and management features also allow UDO media to be removed from the system for archive expansion and low-cost disaster recovery. As the industry standard for data archiving, UDO provides a level of media longevity and record authenticity that cannot be matched by magnetic technologies with a life of over 50 years. The capacity of the UDO Archive Appliance scales from 1TB to over 19TB using field upgradeable capacity options. Available now with new 60GB UDO 2.0 drives and media, the capacity can be expanded to 76TB. Features and benefits of the UDO Archive Appliance include:

- ☒ **Integrated tiered storage.** The value of and access to data varies by application and business need. As data growth continues and access patterns vary, firms cannot afford to keep everything on the same media. UDO Archive Appliance is an integrated, fully tested storage appliance that meets organizational cost and access requirements.
- ☒ **Quick, easy implementation.** IT organizations cannot afford lengthy installation, configuration and API integration. UDO Archive presents standard interfaces and appears as a network share to other applications, making integration easy and time to use quick.
- ☒ **Fast, random access to data.** Regulators, litigators, and internal customers require that records be retrieved quickly. Firms cannot tolerate days or weeks to wait for production of requested data. UDO Archive Appliance offers fast cache-based access to data and rapid seek times on UDO media.
- ☒ **Compliant hardware WORM.** For regulations or fixed content requiring non-rewriteable media, UDO provides the gold standard. UDO Archive Appliance provides true WORM at the hardware level (which is physically unalterable) and acts as a guarantee that data cannot be tampered with, deleted, modified, or changed in any way. UDO Archive Appliance provides data authenticity and risk mitigation.

- ☒ **Permanence/longevity of media and data.** UDO media has been certified for a 50-year-plus lifespan, minimizing the need for data migration. Data with long-term retention can be retained on UDO media in the UDO Archive Appliance with uninterrupted chain of custody and without requiring forklift upgrades. Additionally, UDO drives are backward-read compatible with prior generation media; there is a confirmed road map to 240GB, so investments are protected.
- ☒ **Secure record authenticity.** Application-enabled encryption, key management, and crypto shredding are offered for data security and retention management. Unauthorized access to removable media is blocked with UDO Guard's media lock capability. New UDO2 drives include imbedded AES-256 hardware encryption technology.
- ☒ **Low-cost ownership.** An environmentally responsible solution with significantly lower costs for power, cooling, and maintenance than alternative solutions. Additionally, Plasmon archive solutions are substantially less expensive than comparable disk-archiving solutions.

ONGOING MARKET CHALLENGES

Driven by trends — including electronic discovery, regulatory compliance, and storage performance and cost optimization — firms are deploying archiving solutions. Trends such as the explosive growth of unstructured data types, NAS consolidation, tiered storage, the migration of physical and tape archives to digital archives are also major factors driving archiving implementations. Over the next five years, IDC expects archiving to move beyond regulatory and legal motivations, and active, online digital archives will increasingly be used for knowledge management and competitive advantage by mining data in the archive for trends, patterns, and information. Content sources to be archived will move beyond email to file systems, applications, databases, and eventually the desktop. However, with this growth in archiving, challenges still remain in many firms.

Confusion Over Archive and Backup Workloads

IDC finds many organizations still confused over the differences between archive and backup. Backup serves as a process to provide for restore and recovery in the event of corruption of primary system or data. Backups are typically done periodically and media used is rotated, and older backups are destroyed. Conversely, archive is the explicit function of preserving and ensuring retention of specified records for the purpose of regulatory compliance, discovery, and general business use. The hardware and software used to archive and discover information is very different from those used for traditional backup and provide policies for retention, preservation, and disposition. Archive technologies that cannot ensure permanence or accessibility necessitate backup of the archive, sacrificing the cost reduction benefits associated with reference data archives.

Data May Still Outlive the Technology

Many organizations have infinite or permanent retention of critical business or financial records. However, the application that was used to create the data may not be around to support recall of the data if it is required 100 years from now. Moreover, the media on which the data lives 100 years from now may no longer be interoperable with the application on which the data was originally written. The archive industry is still in a state of standards development, and standards are needed to ensure adequate levels of interoperability between the technology of the data and the data created 100 years prior.

SUMMARY

Archiving best practices start with archive and information management policy development and documentation. Firms currently on the leading edge of archiving realize that archiving is a combination of process, policy, and technology to leverage economies of scale, reduce risk, and improve efficiencies. Plasmon's unique, integrated UDO Archive Appliance is an example of a technology that can be used by firms seeking to make use of today's archive best practices.

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2007 IDC. Reproduction without written permission is completely forbidden.