WHITE PAPER

Why is a Unified Archiving System Essential?

An Osterman Research White Paper

Published October 2013

SPONSORED BY

ZLTECH

Osterman Research, Inc.
P.O. Box 1058 • Black Diamond, Washington • 98010-1058 • USA
Tel: +1 253 630 5839 • Fax: +1 253 458 0934 • info@ostermanresearch.com
www.ostermanresearch.com • twitter.com/mosterman

OSTERMANRESEARCH

EXECUTIVE SUMMARY

There are a number of reasons that any organization should archive its email and other electronic content. These can be distilled into three fundamental drivers:

1. To comply with the regulatory obligations that every organization faces to preserve its business records that are stored in email and other repositories.

2. To make IT staff members more efficient in the performance of their work by improving the efficiency of hardware and software resources.

3. To make employees more productive by enabling them to retrieve business records and other content for as long as they need to do so.

KEY TAKEAWAYS

While roughly one-half of mid-sized and large organizations archive their email and at least some of their other electronic content, there are numerous problems with current archiving technologies and approaches:

- Most archiving solutions store content in information siloes, resulting in significant duplication of data, more difficult coordination when tasks like e-discovery must be performed, and lack of scale.

- There is a lack of assurance on the part of decision makers that every information silo has been discovered and adequately searched for required content.

- An inability to adequately conduct e-discovery or successfully go through a regulatory audit can result in regulatory penalties, legal sanctions, lost revenue, damaging publicity or other consequences.

Consequently, enterprises should implement a truly unified archiving solution that can provide the ability to rapidly search for required content, return all of the required content, and ensure that all relevant content has been identified and appropriately searched.

ABOUT THIS WHITE PAPER

This white paper discusses the reasons to archive email and other electronic content, discusses the importance of unified archiving, and provides a brief overview of the sponsor of this document, ZL Technologies.

Most archiving solutions store content in information siloes, resulting in significant duplication of data, more difficult coordination when tasks like e-discovery must be performed, and lack of scale.

WHY SHOULD YOU ARCHIVE EMAIL AND OTHER CONTENT?

There are three primary reasons to archive email and other electronic content:

- To comply with external obligations imposed by regulators and the courts
- To improve the efficiency of IT operations
- To make employees more productive

Each of these is discussed below.

COMPLIANCE OBLIGATIONS ARE INCREASING AND BECOMING MORE ONEROUS

One of the most important drivers for archiving electronic content is the need to mitigate the risk associated with non-compliance with legal and regulatory requirements to retain relevant electronic content:

- **Legal considerations**
  Email, files and other electronic content repositories contain a growing proportion of the typical organization's business records. Consequently, relevant content from these data stores must be preserved for long periods of time. Because this content is frequently requested during discovery proceedings, it is critical that all relevant electronic content be made available for e-discovery purposes.

  One of the key drivers for e-discovery is the Federal Rules of Civil Procedure (FRCP). Formally enacted in 1975, the FRCP governs court procedures for civil suits filed in the US federal courts. As a result of additional amendments to the FRCP that went into effect in December 2006, discovery of Electronically Stored Information (ESI) – which includes email messages, instant messages, word processing files, spreadsheets, presentations and other content – is now a mandatory point of discussion in civil cases. When subpoenaed for information, the responding party has no more than 30 days to respond according to Rule 34 of the FRCP.

  The current version (2007) of the Rules requires the responding party to "[…] produce documents as they are kept in the ordinary course of business […]" Rule 34: 34(b)(2)(E)(i). This means that if the responding party uses data online and searches it electronically, it cannot supply that data as hard copy. The amendment also requires opposing parties to discuss e-discovery issues within 120 days of a lawsuit's filing.

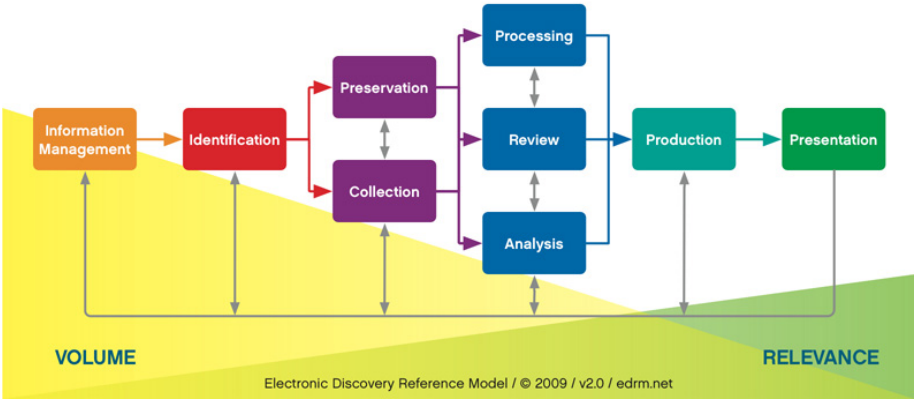
  When a hold on data is required, it is critical that an organization immediately be able to preserve all relevant data, such as all email sent from senior managers to specific individuals or clients, documents that may contain corporate policy statements, spreadsheets with auditors' opinions, and so on. An archiving system allows organizations to immediately place a hold on data when requested by a court or on the advice of legal counsel.

  Those that fail to preserve electronic content properly are subject to a wide variety of consequences, including brand damage, additional costs for third-parties to review or search for data, court sanctions, directed verdicts or instructions to a jury that it can view a defendant's failure to produce data as evidence of culpability.

  Another benefit of an archiving system is that it allows an organization to perform either formal or informal early case assessment activities. For example, if a terminated employee has threatened to sue his or her former employer in a wrongful termination action, senior managers can search the archive for information that will help them determine the potential liability they face. If the assessment results in a determination that the company was indeed wrong in firing the employee, they can instruct legal counsel to pursue a quick settlement. If, on the other hand, the assessment results in information that supports the appropriateness of the company's decision, that information can also be used to convince the ex-employee to drop the case or it can help win the case if it goes to trial. In either case, the archiving system can help the organization to understand its position early on, either avoiding unnecessary legal fees or an adverse judgment, or reducing its costs by proving the sufficiency of its case.

- **The Electronic Discovery Reference Model**
  The Electronic Discovery Reference Model (EDRM) Project, diagrammed in the following figure, was a response to the relatively few standards and lack of generally accepted guidelines for the process of e-discovery that existed prior to its development. The team that developed the EDRM was facilitated by George Socha (Socha Consulting LLC) and Tom Gelbmann (Gelbmann & Associates), and included 62 organizations, among whom were software developers, law firms, consulting firms, professional organizations and large corporations.

When a hold on data is required, it is critical that an organization immediately be able to preserve all relevant data.



The development of the EDRM was important because it represented a major step forward in the standardization of the e-discovery process. Standardization will become increasingly important for e-discovery for several reasons, most notably because of the growth in quantity and diversity of ESI and the large number of entities that will need to process this data (internal and external legal counsel, senior managers, archiving solution vendors, cloud-based IT managed services, outside forensics firms and others).

- **Regulatory compliance**
  There are a large and growing number of regulatory obligations to preserve email and other forms of ESI. Among these requirements are:

  - *Securities and Exchange Commission Rules*
    Members of national securities exchanges, brokers and dealers are obliged to preserve all records for a minimum of six years, the first two years in an easily accessible place (SEC Rule 17a-4). The affected records are broad and encompass originals of communications generated and received by individuals within financial institutions, including inter-office memoranda and internal audit working papers. Also included are automated messages sent to all customers, which could include email blasts. The records may be "immediately produced or reproduced on 'micrographic media' [microfilm, microfiche or similar] or by means of 'electronic storage media'.

  - *Financial Industry Regulatory Authority (FINRA)*
    FINRA is a non-governmental regulator formed in 2007 by the merger of various functions of the New York Stock Exchange and the National Association of Securities Dealers. FINRA manages a wide variety of rules that are imposed upon the more than 5,000 brokerage firms and nearly 675,000 registered representatives it oversees.

  - *Federal Energy Regulatory Commission (FERC)*
    FERC 18 CFR Parts 35 and 284, issued in May 2006, changes the record retention requirement from three to five years for various types of communications and content, including market-based rate authorizations for electricity and related products, and for transactions related to the sales and marketing of natural gas products. This requirement imposes significantly enhanced archiving requirements on firms impacted by the ruling.

    FERC Order No. 717 requires that communications between transmission and marketing employees – ranging from instant messages to paper correspondence – be retained for five years.

  - FERC Part 125 published under the Federal Power Act and Natural Gas Act, requires specific retention periods for the records maintained by public utilities and their affiliated companies. For example, procurement

There are a large and growing number of regulatory obligations to preserve email and other forms of ESI.

agreements must be maintained for six years, stockholder-related meeting minutes must be kept for five years, and plant ledgers must be kept for 25 years.

- *Sarbanes-Oxley Act of 2002*
  The Sarbanes-Oxley Act of 2002 requires all public companies and their auditors to preserve relevant records like audit workpapers, memoranda, correspondence and electronic records – including email – for a period of seven years. Company officers are obligated to report internal controls and procedures for financial reporting and auditors are required to test the internal control structures. Businesses have to ensure employees preserve information -- whether paper- or electronic-based -- that would be relevant to the company's financial reporting.

- *Health Insurance Portability and Accountability Act of 1996 (HIPAA)*
  All organizations operating in the healthcare field need to comply with HIPAA to ensure the safety of Protected Health Information. Organizations are required to protect the data from unauthorized users, as well as to retain for six years a broad range of documentation regarding their compliance.

  The provisions of HIPAA were expanded as part of the American Recovery and Reinvestment Act of 2009 (ARRA). A key element of ARRA is the Health Information Technology for Economic and Clinical Health Act (HITECH). Now, business partners of entities already covered by HIPAA, such as pharmacies, healthcare providers and others, are required to comply with HIPAA provisions. This includes attorneys, accounting firms, external billing companies and others that do business with covered entities. While these business associates were accountable to the covered entities with which they did business under the old HIPAA, these associates are now liable for governmental penalties under the new law.

  Related to the point above is that penalties for HIPAA violations have been expanded dramatically. For example, if a covered entity or one of their business associates loses 500 or more patient records, they must notify HHS and a "prominent media outlet" to let them know what has occurred. Fines for violations can now reach as high as $1.5 million per calendar year.

- *Model Requirements for the Management of Electronic Records (MoReq)*
  Originally developed in 2001, MoReq is a specification that defines the functional requirements for the manner in which electronic records are managed in an Electronic Records Management System. MoReq has been used widely in Europe and has been updated with MoReq2.

- A small sampling of the many other requirements for data retention are FINRA 3010, the Investment Advisors Act of 1940 (hedge funds), the Gramm-Leach-Bliley Act, IDA 29.7, FDA 21 CFR Part 11, OCC Advisory, the Financial Modernization Act 1999, Medicare Conditions of Participation, the Fair Labor Standards Act, the Americans with Disabilities Act, the Toxic Substances Control Act, the UK Companies Act, the UK Company Law Reform Bill - Electronic Communications, the UK Combined Code on Corporate Governance 2003, the UK Human Rights Act, Basel II, and the Markets in Financial Instruments Directive.

These regulations are but a very small sample of those that are focused on data retention that impact archiving requirements and practices.

- **Records Management**
  As critical as ESI preservation is for legal and regulatory purposes, it is equally important to consistently manage retention policies and disposal so that data doesn't linger once required lifecycles have ended. This process has often been

As critical as ESI preservation is for legal and regulatory purposes, it is equally important to consistently manage retention policies and disposal so that data doesn't linger once required lifecycles have ended.

handled by stand-alone enterprise content management (ECM) systems, where records managers declare documents as records and categorize them based on policies. However, records professionals currently face growing challenges in the management of this data due to several factors:

- *Variety and volume and of records*
  The definition of a business "record" has expanded greatly, and now often includes ESI such as email and social media that traditional systems were never meant to handle. This, combined with increased data volume, means it has become extremely difficult to manage all items that may have legal relevance as records.

- *Duplicate material*
  Multiple types of systems for unstructured information have caused overlap and duplication of material. File-shares, ECM solutions, email archives, and other specific-use systems often house the same data that is stored elsewhere. This duplication requires additional storage space, and creates immense difficulty in managing the lifecycles of single documents that may exist in multiple disparate locations.

- *Coordination of legal holds*
  When the legal duty to preserve is triggered, a record that might be relevant needs to be "frozen" in its lifecycle so that it doesn't continue the path to disposal. If legal actions and record retention are managed by separate systems, there is no guarantee that legal holds will actually be applied to ESI lifecycles without manual intervention. This creates needless difficulty and opens the door for legal sanctions and other repercussions.

- *Classification*
  The growing volume of ESI that can fall under the "record" definition has created a problem of scale. Manual classification of every single item is no longer physically possible, and fully accurate automatic classification is not yet a reality. Records managers are stuck looking for options that can balance the two approaches to effectively manage all information.

- *Security Requirements*
  Records managers must often meet high benchmarks for data security, including the industry gold standard of DoD 5015 certification. With non-traditional files now common as records, it is difficult to find systems that can offer full DoD 5015 security for a broad variety of file types. ,

The growing volume of ESI that can fall under the "record" definition has created a problem of scale.

TWO IMPORTANT TRENDS

It is important to note that two important trends are occurring in the context of managing electronic content:

- Governance over this content is becoming more difficult for a variety of reasons, including the use of a growing number of data siloes that include email, SharePoint and other types of document repositories, CRM systems, ERP systems, ECM systems, social media systems, and instant messaging systems; as well as devices like smartphones, tablets and employee-owned equipment.

- Governance is becoming more important as a result of more oversight of business activities by regulators and the courts. With regard to the latter, the courts are becoming savvier with regard to archiving technology and so are less amenable to arguments that data is not accessible or cannot be produced in a timely manner.

FUNCTIONAL CONSIDERATIONS ARE ALSO IMPORTANT

Another important set of reasons to archive email and other electronic content are focused on archiving's functional benefits – i.e., its ability to improve the operation of

email and other servers, and its ability to improve the efficiency of IT operations. For example:

- An archiving system can dramatically reduce the amount of server-side storage by migrating older content to an archive instead of retaining it on the server. This can significantly improve the performance of servers, such as shortening message delivery time and improving server reliability.

- Archiving can also reduce the impact of storage growth, a set of problems that continues to vex email administrators and others charged with maintaining email servers. Osterman Research has found that of the top ten problems in managing email systems, five are related to excess storage, large attachments and other storage-related issues. An archiving system can dramatically reduce storage requirements on "live" servers and solve one of IT's most time-consuming and expensive problems.

- IT also benefits from archiving in the context of normal backup and restore operations. Because content is automatically migrated from live servers to archival storage, the reduced amount of content to back up shortens backup windows. Correspondingly, when servers need to be restored, such as after a crash or a faulty patch, the restoration period is also much shorter.

OTHER ISSUES

There are other reasons to archive email and other forms of ESI, including:

- An archiving system permits users to gain access to their own content, such as missing or deleted emails, rather than asking an IT staff member to recover the content for them. This can significantly reduce the amount of time and effort required to retrieve older information.

- An archiving system permits analytics to be performed on stored content. Analysis of content can provide useful insights into how an organization is operating and can help decision makers to better understand an organization's relationships with customers, prospects, business partners and others.

While all organizations should implement an archiving system to retain necessary data, not all archiving systems provide the same level of functionality or protection from the risks and problems.

PROBLEMS WITH MANY CURRENT ARCHIVING SOLUTIONS

While all organizations should implement an archiving system to retain necessary data, not all archiving systems provide the same level of functionality or protection from the risks and problems discussed above. As a result, there is not only a risk from not archiving content at all, but also a risk of inadequate archiving. These risks focus on three primary areas that will be explored below:

- Maintaining content in information siloes
- Not archiving enough content
- Not archiving the right content

CONTENT IS STORED IN INFORMATION SILOES

One of the fundamental problems with many archiving solutions is that they store archived content in information siloes – separate repositories of content that are managed individually, such as one for email, another for SharePoint, another for records, another for files, etc. This creates a number of problems both when managing the archive as well as when performing searches for data during e-discovery or a regulatory audit. These problems include:

- Significant duplication of data, such as the same attachment stored in an email repository and a SharePoint repository, as well as the same attachment stored in multiple emails (e.g., an email that is sent to a distribution list and then

archived). This not only adds significantly to storage costs, but also requires more time to search across the archives and it requires culling of duplicate data during e-discovery and related activities.

- Coordination for tasks like e-discovery becomes much more difficult and harder to coordinate because each silo of information requires an individual search and extraction of required content.

- Even when siloes of archived content are integrated using APIs, they cannot scale sufficiently to address enterprise-level archiving because of the extremely large archiving volumes that are created, and because at these high volumes the APIs cannot handle the job of transferring, updating and syncing across siloes. In other words, these solutions simply cannot scale appropriately to meet the archiving demands for very large enterprises.

THERE IS A LACK OF COHERENCY

Another critical problem with many current archiving solutions is a lack of coherency. This results in a lack of assurance that a search across all of the archived data stores has produced the content that will be required for e-discovery, regulatory audits, early case assessments or other purposes. More fundamentally, however, this results in a lack of assurance that every data repository has been discovered and searched for necessary content.

The consequences of this problem can be damaging to an organization. For example, in the case of *Pension Committee of University of Montreal Pension Plan v. Banc of America Securities, LLC*[i], the Court issued sanctions against parties that did not adequately preserve ESI, citing the "gross negligence" of their actions. This ruling was made even though the judge found that there was no evidence of bad faith on the part of those who did not preserve the required ESI. In *Keithley v. Homestore, Inc.*[ii], Keithley won on summary judgment, but was still required by the Court to pay $283,000 in fees for failing to preserve and produce required electronic evidence.

In short, an organization's archiving capability must be designed so that decision makers, legal counsel and others have access to every data repository and every data type, coupled with the assurance that they will be able to identify and extract all of the required content.

As part of a sound archiving solution, an enterprise must have access to all of their ESI.

WHY UNIFIED ARCHIVING IS ESSENTIAL

ARCHIVING BEST PRACTICE IS EVOLVING TOWARD *CONTENT* ARCHIVING, NOT JUST *EMAIL* ARCHIVING

Many organizations begin their archiving journey by implementing an email archiving solution. This is a logical first step given that for most organizations email is the largest single data repository and the most important source of relevant business records. However, archiving best practice is evolving toward *content* archiving that encompasses a much broader range of archivable information than email archiving systems are designed to manage. Email is just one data type and must be archived – email archiving is only a start toward a complete archiving solution.

As part of a sound archiving solution, an enterprise must have access to all of their ESI, including not only email, but also files, content in SharePoint repositories, instant messages, social media posts and other relevant forms of ESI. This information must be accessible across the entire enterprise and available across any storage medium.

It is also important for decision makers to note that best practice is rapidly evolving toward search across all of the data stores in an enterprise and not just by individual mailbox. Because of massive data volumes that may reach into the billions of records in a large enterprise, efficiency of search must be a key criterion on which archiving

systems are evaluated.

IMPORTANT CONSIDERATIONS FOR A UNIFIED ARCHIVING SOLUTION

When evaluating archiving solutions, there are four important factors to keep in mind:

- **Rapid search is essential**
  Legal discovery, one of the most common use cases for an archiving system, is typically an iterative process that requires multiple searches. As a result, the rapidity with which a search can be completed is extremely important. For example, if it takes 30 minutes for an attorney to complete a single search, and 10 searches will be required to winnow down the results to the desired content, then five hours will be required to complete the search. However, it is essential that multiple searches be accomplished in near real time for the sake of efficiency and lower costs.

  Faster search capabilities reduce wasted staff or legal counsel time, they make more time available to satisfy "meet and confer" deadlines, they enable highly complex searches to be conducted more easily, and they enable significantly reduced costs that might be incurred for external and internal legal counsel.

- **Concept searching is becoming more important**
  Concept search is another important element of a unified archiving system because of the need to search unstructured data in archives, such as emails, social media posts, instant messaging conversations and the like. Concept searching can provide more accurate results than keyword searches alone because of the inaccuracies inherent in the latter when applied to unstructured data.

- **Language independence is important**
  Language independence is another important element for any archiving system used in any large, multinational organization, and even those used in smaller ones, as well. Given that most large organizations will produce content in languages other than English, the ability to search across multiple languages is essential. Support for double-byte languages – such as Chinese and Japanese – is also an essential element for an archiving solution.

- **Performance, functions and features can vary widely**
  it is also important to keep in mind that archiving system performance, functions and features will vary from one solution to another. For example, some archiving systems will not archive every message that is sent or received. Some will not track every response to a message. Some will ingest legacy mail stores better than others. Some will support multiple email systems while others support only a single platform. As noted above, speed of search can vary widely from one archiving system to another. Some will offer more flexible litigation hold capabilities than others. Finally, some will scale much better than others.

  Consequently, due diligence in evaluating archiving platforms, coupled with realistic estimates of the total volume of email and other content that will eventually be archived, is essential.

Concept search is an important element of a unified archiving system because of the need to search unstructured data in archives.

THE CONSEQUENCES OF NOT ARCHIVING IN A UNIFIED WAY

There is a significant difference between unified and integrated archiving solutions. In a truly unified archiving system:

- All applications share the same data schema instead of each silo of archived data using a different schema. The latter results in a significant processing and transmission burden each time there is an addition, change or deletion of data.
- A single code base is employed as opposed to the use of third-party code.
- All applications share a single end-user profile.
- There is a single backup and restore capability for all applications

In short, a unified archiving architecture will provide significant advantages in lowering the cost of an archiving solution and increasing its performance. Moreover, the ability to coordinate retention policies across an enterprise and eliminating data movements will result in much greater efficiency and faster processing of data. When processing massive data volumes, such as are common in large enterprises that have been archiving data for many years, a unified architecture is the only viable approach because it enables one copy of data across the enterprise, it enables the use of a single system with all application functions available to it, and it eliminates wasted data movements.

WHAT HAPPENS WHEN ARCHIVING IS NOT UNIFIED

When archiving is not unified, there can be a number of serious consequences that are felt most acutely in large enterprises with enormous volumes of archived data:

- Searches are much more time-consuming in terms of total elapsed time for search. This requires more IT staff time to retrieve data and it results in higher legal costs.
- Management of retention for individual items becomes difficult, making it very likely that duplicate copies linger even after being scheduled for disposal.
- Decision makers have less confidence that all relevant and required information has been discovered, increasing the likelihood of spoliation of evidence.
- If spoliation has occurred, this can result in legal sanctions, adverse inference instructions, regulatory penalties, negative publicity, brand damage, lost revenue and other consequences.

When archiving is not unified, there can be a number of serious consequences that are felt most acutely in large enterprises with enormous volumes of archived data.

ABOUT ZL TECHNOLOGIES

Founded in 1999, ZL Technologies has proven itself as the specialized provider of electronic content archiving software for the most demanding large enterprise environments. The award-winning ZL Unified Archive® addresses E-Discovery, compliance, records management and storage optimization.

Built upon the industry's most scalable platform, ZL offers today's leading organizations the ability to comprehensively manage the entirety of their digital assets. To accomplish this challenge, ZL engineered a number of complex technologies into one seamless solution in order to manage billions of documents from a consolidated point of control.

At ZL, we believe in satisfying customers. Everything else follows.

This tenet defines ZL's strategy. Having never lost an enterprise customer, ZL is focused on driving long-term customer satisfaction and value. ZL has generated consistent growth through customer revenues, enabled by listening closely to customer needs and adeptly crafting customized solutions when customers face challenges. This has resulted in a profitable and sustainable business model over the past 12 years, even through the worst of economic times.

ZL is employee-owned and controlled, free from the short-term focus of investors that regularly results in premature product deployments and other shortcuts. The freedom to focus on the long-term gives ZL the flexibility to make prudent decisions for its customers. ZL's long-term outlook has culminated in a clear differentiation in product quality, a point consistently echoed by ZL's customers.

With reliable products and services, talented people, and constant collaboration between partners and customers, ZL has created a profitable and sustainable business model; one that is taking ZL in imaginative new directions.

[i] 2010 WL 184312 (S.D.N.Y. Jan. 15, 2010)
[ii] 2008 U.S. Dist. LEXIS 61741 (August 12, 2008)