

Best Practices for Managing Email Archiving

An Osterman Research White Paper

SPONSORED BY



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

EXECUTIVE SUMMARY

There are four key issues to consider for any organization in the context of managing its electronic content:

1. All organizations should archive their relevant electronic information in order to satisfy their legal, regulatory and other external retention requirements, as well as to maintain good server performance and the ability for their end users to find important, older content.
2. A failure to archive content can result in a variety of consequences, some of which can be quite serious and costly on both a legal and regulatory level, as well as for overall system and application performance.
3. While email has been the traditional starting point for archiving business records and other relevant content, the practice of archiving is evolving and will require organizations to manage a growing number of different content types, including social media content, files, voicemails, etc.
4. A key element of good archive planning is the ability to migrate between archiving providers when necessary, and to migrate archives to the cloud if management decides that doing so will reduce costs or provide other benefits. Organizations that do not plan on a migration to a new archiving platform run the serious risk of being unable to satisfy their long-term retention needs in a cost-effective way.

ABOUT THIS WHITE PAPER

This white paper discusses the key drivers for archiving electronic content and provides recommendations on the important issues that decision makers should consider as they plan their archiving strategy. The document also provides a brief overview of Micro Focus, the sponsor of this white paper, and their relevant archiving and related solutions.

THE FIVE DRIVERS FOR ARCHIVING

There are a variety of reasons that organizations should archive their email, although the specific reasons to do so – and their relative importance – will sometimes vary significantly from one organization to the next based on a number of factors. These factors include the regulatory environment in which an organization operates, the legal risks an organization faces, the overall level of risk that senior decision makers are willing to accept, the willingness of IT to implement user-focused capabilities for recovering email, etc. That said, there are five primary drivers that decision makers should consider as they plan their archiving strategy, as discussed below.

LITIGATION

For most organizations, the most important reason to implement email archiving will be the litigation environment that an organization faces. Specifically, litigation-related drivers focus on three basic capabilities:

- **Early case assessment**

Many decision makers, when they believe that their organization might somehow be involved in a legal action, will want to conduct an early case assessment to determine the risk they face in either defending themselves or prosecuting the case. For example, if an organization believes that a former employee may sue for wrongful termination, legal counsel and others may want to review relevant emails and other documents from key employees, managers, etc. to determine if the former employee might be successful in his or her case. An email archive will enable decision makers to extract the relevant information to help them make that decision quickly.

There are a variety of reasons that organizations should archive their email, although the specific reasons to do so – and their relative importance – will sometimes vary significantly.

- **Legal holds**

Legal, or litigation, holds involve retaining all relevant content when decision makers reasonably believe that litigation might be forthcoming, even if no formal legal action has yet occurred. For example, in the wrongful termination case example noted above, if decision makers reasonably believe that such a case will be initiated, all relevant emails and other documents must be preserved until the legal action has been settled. An archiving system, in conjunction with appropriate retention policies, is extremely useful in enabling legal holds.

- **eDiscovery**

Electronic discovery is the formal process of searching for, extracting, producing and reviewing relevant information from email or other data stores based on court orders and/or other rules established during the process of litigating a case. An archiving system that can index, store and enable search of relevant content – all while protecting the integrity and the authenticity of the stored data – is a key enabler of eDiscovery.

REGULATORY COMPLIANCE

Another important driver for archiving that is somewhat akin to the litigation driver is regulatory compliance. While organizations in all industries face some level of regulatory compliance obligation that requires retention and production of business records, some industries face strict and sometimes onerous regulations. These include the financial services industry, healthcare, pharmaceuticals, energy and certain others.

For example, in the United States the communications of registered representatives and others that sell securities or provide financial advice must comply with Securities and Exchange Commission (SEC) and Financial Industry Regulatory Authority (FINRA) requirements to retain relevant emails, instant messages and social media posts. Healthcare organizations must comply with various US government rules focused on retaining patient records. Energy-related companies must comply with retention obligations imposed by the Federal Energy Regulatory Commission.

It is important to note that even supposedly “non-regulated” industries like retail, manufacturing, etc. face some level of regulatory obligation to protect business records. Consequently, decision makers in all industries must consider regulatory compliance as a reason for implementing an archiving capability.

STORAGE MANAGEMENT

Another important reason to implement an archiving system focuses on the functional benefit of reducing storage on “live” servers. For example, if an organization allows each user to have a one-gigabyte mailbox and each email server supports 1,000 users, that means that each email server will potentially need to store one terabyte of mailbox data. Backing up such a large amount of data can take a substantial length of time, not to mention the performance degradation that might result from managing so much content. Moreover, if a server crashes because of a failed hard drive, a patch that does not work as it should or some other fault, restoring up to one terabyte of data onto the server will take a substantial length of time and can result in an unnecessarily long system downtime.

Alternatively, if an archiving system is implemented that will automatically migrate data to archival storage once a mailbox reaches a certain size, it can do two things:

1. It can implement much smaller mailbox quotas (perhaps 250 megabytes instead of one gigabyte); and
2. It can give users access to their archived content directly in the mailbox or through a web-based interface. This will not only improve email server performance, but will also dramatically shorten backup and restore downtimes, with important and positive implications for disaster recovery and business

It is important to note that even supposedly “non-regulated” industries...face some level of regulatory obligation to protect business records.

continuity.

KNOWLEDGE MANAGEMENT

Another important reason to implement an archiving capability is to preserve the content that organizations pay employees to produce. An email system, for example, contains a record of communications with customers, prospects, business partners and others; it contains records of appointments, contacts and tasks generated by each user; and it contains a record of an organization's "digital heritage". Consequently, organizations should implement an archiving capability to preserve this information for use in the future.

It is important to note that few organizations will ever implement an email archiving system with knowledge management as the primary driver in the decision process. However, it is an important secondary or tertiary driver and something of a "nice to have" if implementing archiving for regulatory compliance or legal reasons, for example.

END USER AND IT PRODUCTIVITY

Finally, another important reason to implement an archiving capability is simply to improve the productivity of both end users and IT. Organizations that implement archiving and then give their users access to archived content, enable them to recover their own missing, deleted or older email without having to ask IT to do this for them. This enables individuals not only to recover their own email more quickly, but it makes IT more efficient because they do not have to respond to these types of recovery requests or tell users they don't have the time or resources to do so.

This is a key consideration in the selection of an archiving capability, since most users employ email more than any other application. For example, Osterman Research has found that the typical email user sends and receives 110 emails per day. That means that in an organization of 2,500 email users, roughly 17.9 million emails will be sent or received every three months. If even a small portion of this enormous amount of content must be recovered periodically, an archiving capability can make both end users and IT much more productive by speeding access to this material.

THE EVOLUTION OF ARCHIVING

THE TRADITIONAL FOCUS OF ARCHIVING

The focus of archiving has traditionally been email archiving – and for good reason:

- Email is where the typical information worker spends roughly 30% of his or her time. For example, an Osterman Research survey of end users found that the typical email user spends 149 minutes per workday doing something in email, making it the primary communications platform for most users¹.
- Email is the primary file transport mechanism in most organizations and contains a large proportion of the typical organization's intellectual property and other content. The study cited above found that 25% of all emails contain attachments of some kind, 3% of them that are larger than five megabytes. The result is that about 98% of the bits that flow through email systems are actually attachments, not the emails themselves.
- Email is the system of record for how employees spend their day, containing information on appointments, contacts, tasks, interactions with clients and business partners, and a wealth of other information that can be critical for legal or regulatory purposes.

¹ Results of a Survey With Email Users, April 2013; Osterman Research, Inc.

Another important reason to implement an archiving capability is simply to improve the productivity of both end users and IT.

Consequently, most organizations begin their archiving journey by retaining email-based content because of its unique importance and utility for a wide range of purposes. In short, email archiving has traditionally represented the “low hanging fruit” of electronic content archives – the content in email systems – and you’ve addressed most of your retention requirements.

THE NEED TO PRODUCE CONTENT

Osterman Research has found that most organizations have been required to produce various types of electronically stored information as part of legal actions and regulatory audits, with employee email being the most commonly requested type of information. Organizations have also made use of their own archives and backup tapes to determine which course of action to take in legal cases.

Content Management and Production Incidents That Have Ever Occurred

Capability	%
We have been ordered, as part of a legal action, to produce employee email	65%
We have used archived content for pre-discovery purposes (i.e. to determine in advance whether or not to settle or fight a lawsuit)	53%
We have referred back to our content archive or backup tapes to support our innocence in a legal case	43%
We have been ordered, as part of a regulatory audit or similar event, to produce employee email	33%
We have been ordered, as part of a legal action, to produce employee instant messages	14%
We have been ordered, as part of a legal action, to produce employee social networking content (e.g., employee Twitter or Facebook posts)	13%
We have been ordered, as part of a regulatory audit or similar event, to produce employee instant messages	13%
We have been ordered, as part of a regulatory audit or similar event, to produce employee social media content (e.g., employee Twitter or Facebook posts)	10%

NEW CONTENT TYPES THAT MUST BE ARCHIVED

The traditional archiving paradigm is changing. Partly as a result of amendments to the Federal Rules of Civil Procedure in 2006 which expanded the envelope of what could be included in eDiscovery, and also because of the growing number of non-email tools available for information workers to share communications and content, the concept of Electronically Stored Information (ESI) has expanded to include a wide range of data types, such as instant messages, files, social media content and even voicemail. Because email is but one type of ESI, but by no means the only type of electronic content that might be relevant in legal or regulatory actions or that might be useful for other purposes, other electronic content should also be retained.

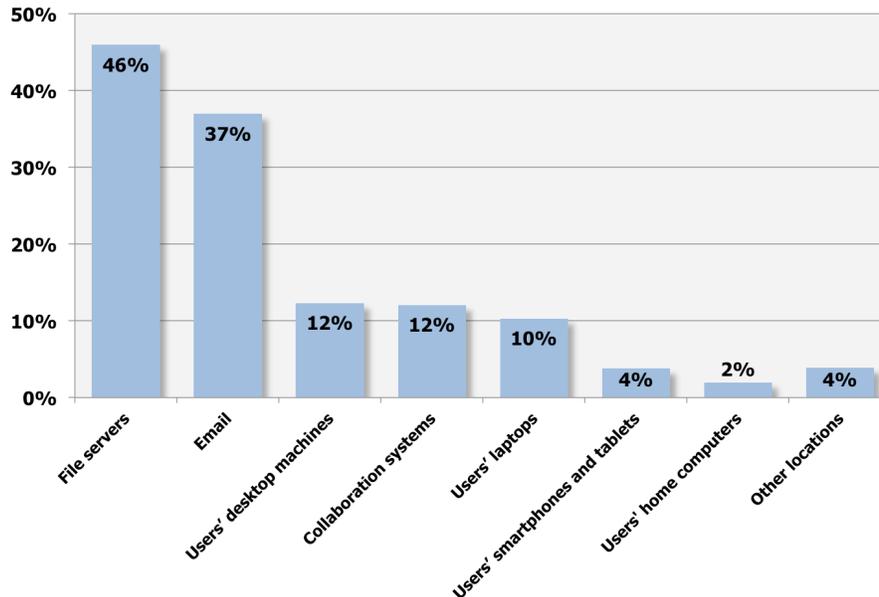
To be sure, email content continues to remain the most important single type of content that must be retained in most organizations. However, financial services firms have had to archive instant messages since 2003 and social media content since 2010 and this trend is slowly expanding to include other industries, as well.

THE FUTURE OF ARCHIVING

Every firm, regardless of the industry in which it is involved, should seriously consider how their archiving policies and platforms should be expanded to include non-traditional content moving forward. While the specific types of content to archive will vary based on a number of factors like legal rulings, advice from corporate counsel, regulatory changes, the various systems that users employ, and senior management’s

tolerance for risk, every organization should consider how to retain and archive its relevant electronic content. For example, as shown in the following figure, email systems are the primary starting point for archiving because of the ubiquity of email and the variety of content it contains. However, email systems are actually not the most common location of electronic content in most organizations.

Proportion of Electronic Content Stored in Various Locations



Add to these venues for relevant electronic content the growing importance of archiving information from various applications like social media (Twitter, Facebook, LinkedIn, et al.), file systems, SharePoint, various collaboration tools, voicemail systems and any other repository of electronic business records that might contain discoverable content.

CONSIDERING ARCHIVING FOR THE LONG TERM

MIGRATING FROM ARCHIVING SYSTEM TO ANOTHER

Another important consideration in the development of archiving is the periodic need to migrate from one on-premise archiving system to another, from an on-premise system to archiving in the cloud, or even from one cloud provider to another. Osterman Research believes that the replacement market for archiving will continue to grow as earlier generation archiving platforms are migrated to higher capacity, more scalable or higher performance systems to meet more stringent archiving requirements.

KEY ISSUES TO CONSIDER

As organizations consider their long-term archiving requirements, there are four key issues that they must evaluate, both for their current platform(s) and for any replacement or substitute platforms they might be considering:

- **Performance**
Performance is key in any archiving solution, particularly for very large archives. Going back to our 2,500-user company that generates 110 emails per user per day, let's say that only 15% of those emails contain business records or other

Email systems are the primary starting point for archiving because of the ubiquity of email and the variety of content it contains.

relevant information that need to be archived. Further, let's assume that the organization needs to retain its information in the archive for an average of seven years.

At the end of seven years, that organization would have 75.1 million emails in its archive, assuming that it had an email archiving system that provided single-instance storage and thereby eliminated duplicate emails and files. Moreover, because the average email is 0.42 megabytes², the total volume of content archived would exceed 31.5 terabytes. If 20% of emails required retention, these figures would jump to 100.1 million emails and 42.0 terabytes of storage.

What this means is that archiving systems need to be evaluated based on their performance – their ability to search through very large data stores and return relevant results in a short amount of time. Some older archiving solutions are not capable of delivering this type of performance, and so increase the cost of eDiscovery and other searches for content in the archive.

- **Scalability**

Closely related to the performance of an archiving solution is its scalability – the ability of the solution to manage increasing quantities of content and provide the needed performance. This is particularly important as organizations add not only more email content to the archive, but files, social media posts, voicemails and other content that will need to be retained for long period. An archiving solution that is not scalable will simply not meet the long term content retention requirements of an organization and could increase the level of risk that an organization faces by not archiving all necessary content.

- **Flexibility**

Flexibility is another key decision point for any archiving solution, since as new data types must be archived, the solution must be designed to retain them and allow complex searches to be conducted on their content. Moreover, it is essential that content is stored in a meaningful way – storing social media posts in flat text file, for example, instead of in their native format may not satisfy every organization's social media retention obligations.

- **Cost**

Finally, but certainly not least, is the cost of the solution. While long-term storage costs for archived content may not be a critical consideration for on-premises systems, it can be an issue for content stored in the cloud if per-gigabyte pricing schemes are employed. This is a particularly important consideration given that many organizations do not have detailed and thorough content deletion policies, and so content tends to be retained past the time that is required by legal precedent or regulatory requirement.

An archiving solution that is not scalable will simply not meet the long-term content retention requirements of an organization.

ABOUT MICRO FOCUS

Micro Focus is a global software company with 40 years of experience in delivering and supporting enterprise software solutions that help customers innovate faster with lower risk. By applying proven expertise in software and security, we enable customers to utilize new technology solutions while maximizing the value of their investments in critical IT infrastructure and business applications. As a result, they can build, operate, and secure the IT systems that bring together existing business logic and applications with emerging technologies—in essence, bridging the old and the new—to meet their increasingly complex business demands.

² Based on calculations from an Osterman Research survey conducted in early 2013.

© Osterman Research, Inc. All rights reserved.

No part of this document may be reproduced in any form by any means, nor may it be distributed without the permission of Osterman Research, Inc., nor may it be resold or distributed by any entity other than Osterman Research, Inc., without prior written authorization of Osterman Research, Inc.

Osterman Research, Inc. does not provide legal advice. Nothing in this document constitutes legal advice, nor shall this document or any software product or other offering referenced herein serve as a substitute for the reader's compliance with any laws (including but not limited to any act, statute, regulation, rule, directive, administrative order, executive order, etc. (collectively, "Laws")) referenced in this document. If necessary, the reader should consult with competent legal counsel regarding any Laws referenced herein. Osterman Research, Inc. makes no representation or warranty regarding the completeness or accuracy of the information contained in this document.

THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. ALL EXPRESS OR IMPLIED REPRESENTATIONS, CONDITIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE DETERMINED TO BE ILLEGAL.