Considerations for Archiving in Exchange Environments

An Osterman Research White Paper

Published July 2013

SPONSORED BY

EMC²
EXECUTIVE SUMMARY

Archiving is an essential best practice for any organization in order to satisfy regulatory, legal or functional requirements; and in order to improve the overall performance of the IT infrastructure. Recognizing this, Microsoft added native archiving capabilities in Exchange 2010 and improved these capabilities in Exchange 2013. However, there are inherent limitations in these capabilities that will require many Exchange customers to use third-party archiving solutions.

KEY TAKEAWAYS

- Osterman Research believes that most organizations need to deploy third party archiving tools in order to satisfy all of their archiving and related needs, particularly those that deal with eDiscovery and storage-related matters.

- Most organizations that have deployed Microsoft Exchange have the bulk of their users on Exchange 2010. While most Exchange-enabled organizations have stated their intention of migrating to Exchange 2013 over the next year, we believe that migration will occur more slowly than this. The result is that over the next 12 months many organizations will be left with the more limited feature set available in Exchange 2010.

- Exchange-enabled organizations require a variety of critical email archiving capabilities as shown in the figure below, chief among them those that are focused on eDiscovery, support for large mailboxes and the ability to implement litigation holds.

---

Interest in Various Email Archiving Capabilities

<table>
<thead>
<tr>
<th>Capability</th>
<th>% Responding Interested or Very Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>The ability to search for e-discovery purposes</td>
<td>70%</td>
</tr>
<tr>
<td>Support for very large user mailboxes</td>
<td>65%</td>
</tr>
<tr>
<td>The ability to implement legal holds</td>
<td>62%</td>
</tr>
<tr>
<td>The ability to archive .PST files</td>
<td>56%</td>
</tr>
<tr>
<td>Journaling for compliance purposes</td>
<td>53%</td>
</tr>
<tr>
<td>The ability to selectively dispose of content</td>
<td>52%</td>
</tr>
</tbody>
</table>

---

ABOUT THIS WHITE PAPER

This white paper discusses the archiving capabilities in Microsoft Exchange and why most organizations need to consider deploying third party archiving capabilities to supplement them. We also discuss and present the results of a survey conducted specifically for this white paper. Finally, we present a brief overview of the sponsor of this white paper, EMC, and their relevant archiving solutions.
KEY REASONS TO ARCHIVE CONTENT

There are several reasons to archive email and other electronic content, although the relative importance of these reasons will vary based on a variety of factors, such as the industry served by an organization, its size, the geographic regions in which an organization operates, management’s and legal counsel’s tolerance for risk, etc.

REGULATORY DRIVERS

Akin to the legal drivers (discussed below) for implementing an archiving capability are a number of regulatory drivers. There are thousands of statutory requirements in the United States and other countries to retain business records, including records in email, files and other sources. A small sampling of these requirements includes the following:

• **Energy**
  The Federal Energy Regulatory Commission requires that certain non-public, electronic transmission function information exchanged between transportation and marketing function employees must be retained for a five-year period.

• **Financial services**
  The Securities and Exchange Commission (SEC) and the Financial Industry Regulatory Authority (FINRA) dictate that various securities transaction records must be retained for up to six years. A failure to retain records can result in monetary penalties totaling millions of dollars, as well as censures and cease-and-desist orders.

• **Healthcare**
  The US Department of Health and Human Services (HHS) has expanded the requirements for protecting confidential and sensitive information, expanded the number of organizations that are subject to the Health Insurance Portability and Accountability Act (HIPAA), and can be expected to levy fines and penalties more frequently than has been the case in the past. The Omnibus rule allows HHS to impose fines that range from $100 for a “Did Not Know” breach of Protected Health Information to $50,000 for a single, uncorrected and willful violation. However, fines can reach $1.5 million per year or more.

  Health plans, health plan clearinghouses and healthcare providers (e.g., physicians, nursing homes and clinics) must retain electronic health records for six years from the date of their creation or the date when they last were in effect, whichever is later. Medicare requires that in most cases, clinical records must be retained for up to six years from date of discharge or last entry.

• **Pharmaceuticals**
  The Food and Drug Administration requires that records related to food receipt, release and processing must be kept anywhere from six months to two years. Records related to non-clinical lab studies must be retained from two to five years; and records related to drug receipt, shipment and disposition must be retained for two years after a marketing application is approved for a drug.

• **Publicly held corporations**
  The Sarbanes-Oxley Act of 2002 includes provisions that accountants of publicly held companies must keep certain records and workpapers relevant to the audit or review of such corporations’ financial statements for a period of seven years.

• **State and local requirements**
  A variety of state, provincial and local governments have requirements to retain public records under “sunshine law”, Freedom of Information, or related requirements. In the event of a request for public information, these various open records laws may require production of relevant information within a given time frame.
LEGAL DRIVERS
Legal drivers are among the most critical considerations for deploying archiving solutions in many organizations, particularly among organizations that are not as heavily regulated as those in the financial services, energy, pharma or healthcare industries. Legal drivers for deploying archiving solutions can be broken down into three primary focus areas:

• **Litigation holds**
  Litigation holds are a key driver for deploying archiving, since relevant information must be retained for long periods after a legal action has been initiated, or when decision makers can reasonably expect that such an action is likely to occur.

• **eDiscovery**
  The search for electronic data when seeking relevant information during a legal action, the extraction of this data for analysis by attorneys or paralegals, and the presentation of information to various parties, is the leading driver for archiving in many organizations. Because large organizations, as well as many smaller ones, must produce information from their email and other archives on a regular basis, many organizations use eDiscovery as a key driver for the deployment of an archiving solution.

• **Early case assessments**
  Whether a formal or informal inquiry, the ability to perform early case assessments is another key driver for archiving. The ability to search an archive of email, file servers or other electronic content stores can be extremely valuable in helping senior managers, legal staff or outside attorneys to perform assessments of an organization’s legal position before a legal action has commenced or during its early stages.

FUNCTIONAL DRIVERS
Another reason to archive email and other content forms applies to all organizations, namely the functional benefits that email archiving can provide, which are several:

• **Email server storage management**
  An archiving solution, by migrating email and files from servers to archival storage, can drastically reduce the amount of content stored on these servers. The result is that backups are shorter and can be completed within the allotted backup window, restores are much faster, and server performance improves.

  Osterman Research has found in numerous surveys over the past several years that many of the top ten problems in managing email servers are related to excessive storage, a problem that archiving addresses directly by migrating data to archival storage. This minimizes the overall cost of managing storage and it reduces IT’s storage-related costs by delaying or eliminating altogether the deployment of high-performance, primary storage.

• **Employee productivity**
  An archiving system can allow end users to access their own archived content, enabling them to retrieve missing, deleted or older emails and attachments without having to ask IT to do this for them. Not only does this greatly reduce IT’s workload or force IT to deny the request for lack of manpower, it makes employees better able to be productive because they have access to more information.

  An archiving solution can minimize the time employees spend on managing their mailbox content. In the absence of an archiving system, employees must spend time filing, deleting or otherwise managing their emails and other content to stay under the mailbox size quota that most organizations impose on their users. However, with an archiving system users have a seemingly unlimited mailbox because content is automatically migrated out of mailboxes to the archive.
CURRENT AND PLANNED ADOPTION OF EXCHANGE

Our research found that the vast majority of Exchange-enabled users are served by Exchange 2010, as shown in the following figure. Another 20% are still on Exchange 2007, with the balance on Exchange 2003 or earlier versions. In a separate survey conducted in early May 2013, we found that among that are using Exchange exclusively, 59% of users are expected to be served by Exchange 2013 as of May 2014. However, we believe this figure is too optimistic:

- Exchange 2013 is a relatively new offering, available only in November 2012, but it wasn’t until April 2013 that Exchange 2013 could be deployed in Exchange 2010 or 2007 environment (via Cumulative Update 1). Consequently, if we consider that Exchange 2013 became a production system in early April 2013, achieving 59% penetration in just a year’s time seems a bit ambitious.

- Our research has found that many organizations skip an Exchange generation because of the cost and effort required to deploy a new version of Exchange. Given the large proportion of organizations that are currently using Exchange 2010, we anticipate that adoption of Exchange 2013 will be slower than many anticipate. This means that many organizations will be left with only the more limited archiving capabilities available in Exchange 2010.

WHAT ARE ORGANIZATIONS LOOKING FOR IN AN ARCHIVING SOLUTION?

Our research found that organizations have a strong interest in archiving capabilities for email management. The following figure shows that eDiscovery related features rank the highest followed by storage management features related to mailbox capacity and .PST upload.
WHY THIRD PARTY ARCHIVING IS STILL NEEDED

Microsoft has provided a basic archiving capability, but it will not satisfy the requirements of all Exchange-enabled organizations, as discussed next. Consequently, these limitations in Exchange 2010 and 2013 will necessitate the use of third party archiving systems:

- **eDiscovery lacks some important features**
  Exchange 2013 eDiscovery lacks important features like hit-highlighting and adding comments or tags to individual search results – functions needed for rapid legal review. The search and review workflow process itself is more cumbersome than for many third party solutions and jumps back and forth between locations, adding complexity. The Exchange multi-mailbox search is best suited for basic search and exporting the search results to a third-party eDiscovery solution for detailed legal review and analysis.

- **No support for role-based search**
  Exchange does not support role-based search, meaning the multi-mailbox search commands access to all mailboxes and cannot be tailored to specific groups or departments. This limits the ability for an organization to manage legal discovery securely with multiple individuals – a common requirement for all but the smallest organizations.

- **Personal archives remain on the Exchange Server**
  The Personal Archives are a separate mailbox that is accessed by users in Outlook or the Outlook Web App. Mailbox contents in the Personal Archives remains on the Exchange Server forever, increasing the total storage load (and cost) on Exchange Server and significantly impacting Exchange Server recovery time in the case of a failure or disaster. When considering the need to retain email on legal hold and email for ex-employees, sometimes for many years, the potential impact on Exchange can be considerable.

- **No single-instance storage**
  Single-instance storage (SIS) in Exchange was eliminated in Exchange 2010 and remains absent in Exchange 2013. While there are some good reasons for doing so, such as improving the performance of Exchange servers because of falling storage prices, the use of single-instance storage is an important benefit available with many
Considerations for Archiving in Exchange Environments

third-party archiving systems and one that many IT administrators will find beneficial. Not only are storage requirements in Exchange significantly greater without single-instance storage, Database Availability Groups further exacerbate the storage problem when used for high availability.

Without single-instance storage, migrating personal folders (.PSTs) to Exchange will also increase storage overhead. This data is replicated inside the Database Availability Group and require large amounts of storage for what is largely redundant data. Also, as .PSTs are pulled into the Personal Archive, users lose the offline capability to access that data offline. This data would need to be connected to Exchange in order to access the data after migration.

- **Exchange Servers do not experience load reduction**
  The archiving functionality in Exchange 2010 and 2013 does not reduce the load on Exchange servers because content is not moved to a separate archive infrastructure. This negates a key advantage that is available with some third-party archiving solutions. Consequently, the Exchange infrastructure must support email for its entire lifecycle, including email for all current and ex-employees and email that is held on legal hold.

- **Multi-content eDiscovery workflow is complex**
  Exchange 2010 and 2013 support only eDiscovery for Exchange Server mailbox content. For eDiscovery of Microsoft SharePoint and File Shares, SharePoint 2013 eDiscovery Center is installed. SharePoint 2013 eDiscovery Center works only with 2013 versions of Exchange and SharePoint. This could complicate and increase the cost of eDiscovery and other litigation support functions, so many organizations will opt for a single archiving solution that enables policy management and search from a single interface.

- **Retention controls have limitations**
  Other than for mailboxes that are on legal hold, users are primarily in charge of their own retention management. This can result in the deletion of content from an Exchange mailbox that should be kept in compliance with corporate policies. Many third-party solutions offer much more robust controls over content retention.

- **eDiscovery capabilities are only basic**
  The eDiscovery capabilities built into Exchange 2013 provide some useful e-discovery functions; however, these capabilities are unlikely to satisfy some of the more sophisticated eDiscovery requirements that an organization might have. For example, Exchange 2013 delivers basic search of mailbox contents, but there is no “hit-highlighting” of the search results. The review of hundreds or thousands of items becomes a very difficult task when the reviewer must read each item without the aid of hit-highlights.

  Exchange 2013 offers the ability to place a hold on an entire mailbox or a query-based search; however, a litigation hold can be applied only to data indexed by Exchange. Since Exchange is limited in the number of different file types it can index, third party solutions will still be required in order to manage file types that cannot be indexed by Exchange.

  Exchange 2013 does not support important eDiscovery capabilities that many third party solutions provide, such as role-based search, results analysis and tagging, among other capabilities. Lack of support for these capabilities creates problems for organizations that want support for tagging content, as shown in the following figure.
• **Very large mailboxes create backup problems**
  The 25-gigabyte mailbox size in Exchange, while beneficial for users who want to store large amount of information in their mailbox, can create problems. For example, as shown in the following figure, 27% of decision makers believe that such a large mailbox would create backup and restore problems, while another third of respondents are not yet sure of its ramifications.

### Views on the 25-Gigabyte Mailbox in Exchange

- **We are not yet sure** 33%
- **This size mailbox would work well for us** 40%
- **This would create backup and restore problems** 27%
- **Other than for mailboxes that are on legal hold, users are primarily in charge of their own retention tagging. This can result in the deletion of content from an Exchange mailbox.**
NEXT STEPS
Any organization that is considering deployment of archiving capabilities in Exchange 2010 or 2013 should consider a five-step approach to evaluating their archiving requirements:

1. UNDERSTAND THE NEED TO ARCHIVE
Organizations need to appreciate that archiving is fundamentally about retaining corporate records as a best practice, not an option that applies only to “regulated” firms like broker-dealers, pharmaceutical companies or healthcare providers. Although regulatory requirements to archive content will vary based on a number of factors, virtually every organization has legal obligations to retain records for purposes of litigation or regulatory compliance.

Organizations need to examine their needs for in-house legal discovery. Does it rely heavily on outside legal support for eDiscovery? What is the frequency of legal disputes that require outside legal support? What is the related cost of such eDiscovery? Depending on these answers, the more cost-effective approach maybe to increase in-house eDiscovery capability with third party archiving.

2. DEVELOP POLICIES
All organizations should implement policies that are focused on retaining important content in email and other electronic data stores. However, many organizations do not have email retention policies and, among those that do, these policies often are not well defined. Many decision makers view business records in email incorrectly. For example, in a 2012 survey conducted by Osterman Research, we discovered that senior management in nearly one in five organizations surveyed viewed email content as “transitory” and not worthy of long-term retention. Nearly one-half more viewed records in email as important, but subject to retention only by employees. Only about one-third of the organizations surveyed hold the view that records in email are important and should be managed by IT according to corporate policies.

3. CAREFULLY DESIGN EXCHANGE FOR STORAGE MANAGEMENT
Decision makers should carefully design the new Exchange 2013 architecture with respect to per mailbox storage capacity and its impact on backup and recovery and total storage cost. Most organizations will need third-party solutions in order to manage Exchange total storage capacity with a centralized email archive for cost-effective long-term retention of email information, including single instance storage.

4. PERFORM DUE DILIGENCE ON EXCHANGE eDISCOVERY
It is recommended that those individuals in the organization responsible for legal discovery perform due diligence on Exchange eDiscovery capabilities. The basic eDiscovery steps include simple and advanced search, review and culling, saving search results and exporting search results. Observe overall search performance and the ability to perform review considering the lack of key features like hit-highlighting and role-based review. Organizations that routinely perform in-house legal discovery will need third-party solutions to manage their caseload in a manner they are accustomed to.

5. CONSIDER EMAIL CLASSIFICATION AND TAGGING
It is also essential to consider the classification of email, using the retention policy, default policy or retention tags in Exchange; or a third-party classification solution. Tagging content can make it easier to identify information that should be retained and also to assign unique retention periods.
**CONCLUSION**

Exchange 2013 is unlikely to be sufficient for organizations that are concerned about litigation and who have invested in in-house legal staff to perform early case assessment and sophisticated review of content before exporting such content to external legal counsel. Moreover, Exchange 2013 archiving will not meet the needs of organizations that require maximum storage efficiency through the use of single-instance storage. Third-party solutions can provide those eDiscovery capabilities not addressed by Exchange, as well as provide improved storage efficiency. Organizations should diligently review eDiscovery functions of Exchange side-by-side third-party solutions before finalizing their decision.

**SPONSOR OF THIS WHITE PAPER**

EMC Corporation (NYSE: EMC) is the world’s leading developer and provider of information infrastructure technology and solutions that enable organizations of all sizes to transform the way they compete and create value from their information.

EMC SourceOne is an archiving family of products and solutions that allows organizations to manage the lifecycle of corporate information, according to consistent policies based on the content’s business value. EMC SourceOne solutions are designed to scale to meet large enterprise needs, yet offering a simple footprint for mid-sized customers. The solutions provide customers with the ability to reduce costs, ensure compliance to internal, external and government regulations, and to mitigate risks associated with non-compliance and eDiscovery. Coupled with EMC hardware and other EMC technologies, including backup & recovery, deduplication, security and virtualization, corporations can build a complete information governance solution from a single vendor.

Additional information about EMC can be found at [www.EMC.com](http://www.EMC.com).

---

Third-party solutions can provide those eDiscovery capabilities not addressed by Exchange, as well as provide improved storage efficiency.