



**Email Archiving: the definitive checklist -  
33 crucial questions you must get answered  
before selecting a solution.**

[www.re-soft.com](http://www.re-soft.com)  
203 972 8462  
[info@re-soft.com](mailto:info@re-soft.com)

This page left intentionally blank

## 2012 Update

This document was first published in 2007 as the demand for email archiving solutions was growing significantly. Today, we see the majority of organizations have now implemented some form of email archiving system, typically driven by one or more of the key factors - regulatory capture, e-discovery or mailbox size management.

In 2012, the market has matured to where an incumbent solution now typically exists. So we have a situation where some organizations are looking at email archiving from a greenfields scenario while others will, as with most technologies, be reviewing whether they should continue investing with the existing supplier.

Hindsight dictates how well a strategic decision an organization made as to the promise and longevity of their selected solution. As a consequence of the shakeout after too many vendors entered the email archiving market, some have successfully maintained their position, some have been acquired and then reacquired, some have just ceased to exist.

Most archive solutions have been architected to lock you into their proprietary formats and do not make it easy to move to a different solution. For early adopters the number of messages and volume of data under management is now significant. So any plans to modernize from a poorly-chosen legacy system require considerations that will not have been anticipated.

This has generated a need for migration tools, some of which we offer as solutions, to ease the migration from old legacy systems to the latest.

So whether you are seeking to replace an existing archive, wanting to validate your original decision or are one of those new to the email archiving world in 2012, the issues highlighted in this paper are still valid and worthy of consideration.

## The Challenge of Selecting an Email Archiving Solution

Before you start reading this paper, go to Google and search on 'email archiving'. The number of results you get is staggering; there is no shortage of potential solutions. But who do you trust? How do you choose the solution that is right for you today and into the future?

This paper is designed to provide invaluable planning information to you along with Best Practice Advice built upon significant experience gained by ReSoft International in assisting its clients in implementing Email Archiving solutions.

As you research the market for appropriate solutions you will find there is a lot of hype, and it is difficult to distinguish between fact and scare tactics. The questions that follow (and we will come back to the Google search later) will aid in making those critical decisions.

This paper has been revised since its first introduction on 2007, when we had 29 Critical Questions. Interestingly, there have been some additional questions such as using a Cloud-based solution but many of the planning issues have not changed significantly.

## What is driving the need to archive....?

**Regulations** such as HIPAA, SEC rule 17a-4 and NASD rule 3010 have illustrated the need for many organizations to commit to implementing long-term indexing and archiving of Email and Instant messaging conversations. Important to these regulations is the timely extraction of messages that meet the criteria of a Regulator.

Similarly, any organization who has received a **Legal Discovery** request against their email and IM files knows the importance of timely extraction. In the past this has involved installing software that gradually crawls through users' mailboxes, restoring backup files by rebuilding email servers from backup tapes, and then hoping that you have been able to find all of the personal archives sitting out on user desktops to have the confidence to certify that you are compliant with the Discovery request.

And finally, **Mailbox Size**. The user-friendliness and ubiquity of the email client means it is all too easy for users to use the email system as their personal electronic document filing system. It is functional, easy-to-use and simple to find items. But without some kind of retention strategy the cost of storing all this information becomes prohibitive, mail-files become corrupted, backups and restores take longer.

### **Proactive or Reactive? Which is best?**

Clearly, being prepared to respond in advance of an email retrieval request is the most effective approach, although there are investment costs involved. If there is no business driver such as Regulatory compliance or a pending litigation, it may be difficult to financially justify an email archiving system for this purpose. Often, a legal incident or a pending regulatory request will help change budget priorities. Similarly, offsetting an imminent hardware or storage upgrade caused by email disk-space consumption can release funding.

If you are unable to build a convincing business case, there are still tactical tools that will streamline the otherwise manual process of collecting data from mail files and backup tapes so you can respond in a reasonable manner to search requests.

If you are able to fund an email archiving project, determine which of these questions are important to your archiving strategy. You will save considerable time and remove the 'gotchas' that you may not see as an issue until post-implementation.

*Best Practices suggest maintaining an archive, regardless of size or industry to enable a considered, timely response to legal discovery requests, an effective way to provide storage relief to the email system and an effective mechanism to meet regulatory demands. The benefits typically attributed to this application are compelling:*

- *Meet Regulatory Retention Requirements*
- *Timely response to Legal Discovery Requests*
- *Reduction in space consumed by email*
- *Ease of restoring lost or corrupted messages and folders*
- *Reduction in administrative effort in managing the email system*

## **The Questions:**

### **1. Archive Collection Policies**

#### **1.1 How will the right solution support Regulatory and/or Mailbox Size and/or E-Discovery needs?**

Email archiving tends to be driven by one these factors. Regulatory and E-Discovery needs tend to favor collection of all email and this has historically been addressed by enabling the journaling capability of the native email system. Journaling guarantees that every single message traveling through the email network is captured, regardless of whether the user has deleted it from their inbox. But this puts a huge and potentially unnecessary burden on the storage and indexing systems by including non-relevant email in the archive and requires close

management of the journal mailboxes lest they get corrupted and mail is lost.

The other driver is that of Mailbox Sizes. In this case, it is more appropriate to collect email and the attachments from the mailboxes on a scheduled basis and leave behind a link in the mailbox folder from where the message originated. This should dramatically reduce mailbox size.

But consider that capturing email from the mailboxes may be an effective collection strategy to meet Regulatory requirements as the email system can be configured to capture email that the user thought had been physically deleted.

## **1.2. Which classification options should a solution support?**

If you elect not to use the journaling capabilities, you may wish to implement a programmatic filter to determine what gets captured. You may elect only to capture messages older than a certain age, a particular size, those contained in a particular folder, or containing certain word-content within the email. Or perhaps only begin to archive email once a user's mailbox gets to a particular size.

Alternatively, consider a User-Initiated collection policy. This is where a user can nominate which messages are archived (Business Record) and which get purged. As organizations consider how to meet the need for archiving email and instant messaging conversations, many are implementing a policy of archiving all such email on a just-in-case basis. The reality of this approach, though, is to unnecessarily duplicate much of the traffic and maintain copies of non-business and non-essential email for many years, increasing storage and performance issues as well as increasing the amount of clutter being returned in responsive searches. By spending a little more time on defining policies, organizations can set up a mechanism to capture only relevant email. User-initiated archiving goes a step beyond programmatic archiving. It moves the responsibility to the individual user for what should and should not be sent to the archive, based upon a set of clear guidelines as to what is determined to be business relevant email.

## **1.3. How should a solution maintain an exact replica of the original for Legal purposes?**

A number of potential solutions render the original email message into an SQL-or-similar database format. By doing this the message is no longer in its true original state.

## **1.4. How will the solution allow the capture of other classes of message?**

A users' mailbox contains more than emailed messages. There are calendar items, memos, notes, and importantly draft messages, some of which will not have been

sent through the email system but may be relevant to an archiving strategy.

## **1.5. How will the solution allow the capture of users' personal archives from the desktop?**

Absent a strategic email archiving system, there are many instances where users will have already archived email to their local desktops. This creates inherent problems in that the organization will not necessarily know what information has been saved to that desktop archive – information that is perhaps critical in meeting a legal discovery request. Is it acceptable to continue this practice or do you wish to eliminate it and move the messages to the central archive? Make sure any solution can find those personal archives and import the data into the central archive.

**1.6. How does the solution allow the capture of email from existing email backup files?** Depending on your retention policies and any existing subpoenas requiring you to maintain backup tapes, you may have a significant investment in backed-up email files. Do you need to import this data into the archive so you have a complete picture?

## **1.7 How is access to the Archived email made secure?**

What encryption techniques are applied to protect the archive store from unauthorized access or is the solution reliant upon the Operating System providing access control?

## **1.8 How are Blind Copies (BCC) and Distribution List Members stored in the Archive?**

Ensure that the sender & recipients of any email sent to a bcc or a member of a distribution list can be identified in the archive.

## **1.9 What is the source of the archived email?**

Most likely, this will cover internal and external email and probably Instant Messaging. Consider also email generated from Smartphones, users using webmail services such as gmail from within the work environment and other specialist messaging systems such as Bloomberg & Reuters.

## **2. Archive Retention Policies**

### **2.1. How does the solution allow for the configuration of Granular Archive Retention Policies?**

You will find that you will need to define retention policies for different classes of message, which will affect how long you store that class of archived messages. It may vary by context/user. For instance, email classified for HR, Executives, Legal, Tax & Sarbanes Oxley all may need different retention policies to be applied. Similarly, you may find a need for retention by subject, matter, client, by creation date or by the folder from which the email originated.

### **2.2. How does the solution support multiple-use archives?**

Will you archive into a single container or into Use-Specific containers? It is quite reasonable to deposit all archived email into a single repository. However, if email has different classifications for retention or there is a need to separate certain user groups into a different archive it may make more sense to set up different archive containers to support these different classifications. Similarly, you may wish to segment the Senior Management archives from the rest of the employee archives. You may also wish to have different archives stored at different points on your network or on different storage media.

### **2.3. How does the solution allow the purging of archived messages?**

This may be message age, message size, mailbox folder, user-initiated, mailbox size, message content or some other combination.

### **2.4. How does the solution support HSM requirements?**

Where will you store archived messages, should they be non-editable (WORM), should the archive store get cycled to less expensive storage as data gets older (eg: first 90 days in mail file, next 360 on SAN, remainder on tape etc.). You will always need to plan a suitable storage medium as any archiving system is going to generate a substantial storage demand.

## **3. Archive Retrieval Policies**

### **3.1. How does the solution allow users to search, view and retrieve their own archived messages?**

If driven by Regulatory requirements, you may wish to lock away the collected archives from access by users. Otherwise, there needs to be a self-service mechanism for users to access their own archived email without any special software or processes.

### **3.2. How does the solution allow a Supervisor to search, view and retrieve archived messages for their subordinates?**

You may have a need for a team supervisor to search through the email of their subordinates.

### **3.3. How does the solution allow a Compliance Manager or an Administrator to be able to search, view and retrieve from the whole archive?**

There will generally be a requirement for a Compliance Officer, Legal Counsel or an authorized IT Administrator to have the ability to search through the whole archive.

### **3.4. How does the solution support traveling users that need access to their own archive email while disconnected?**

If email has been moved from the mailbox to the archive, even if a stub file has been left behind, when a user is disconnected from the network they will not be able to see their messages in the archive. There may be a proportion of your users who frequently travel with a laptop and need to get access to all of their email, including those items that have been archived.

### **3.5. How does the solution allow users to accessing their archived email - from a standard mail client, via a web browser, through Smartphones?**

Consider which devices your users will use to retrieve items from their archived email.

### **3.6 How does the system search and retrieve email that was encrypted by the user?**

Consider the process for searching content and extracting from email messages that were encrypted.

## **4. Migration & Co-existence**

### **4.1 How does the solution help in planning a Migration Strategy?**

It is not too early to be thinking about migrating from the solution you choose, so you need to be careful not to get locked into an architecture you cannot get out of! Consider also a possible migration to a new email platform after which you would need to continue to capture messages and still get access to the archive for old messages. As you look at solutions make sure you understand the collection and storage architecture? Does it lock you into an architecture from which you cannot exit? Does it offer support for all the major email systems to which you may migrate or may inherit as part of an acquisition? Does it have a documented method for you to easily migrate to whatever will be the next generation of Archiving Solutions after, say, 5 years of collection?

### **4.2 Do I need to capture Personal archives into the final solution?**

PST files and Lotus Notes archive files can become a very burdensome issue – there will be potentially respondent data, but where is it, how much is there and how much do you need to keep?

### **4.3 But wait, doesn't Microsoft now offer email archiving within Exchange 2010.**

Yes, Microsoft has introduced a second mailbox that allows users to move email from their primary Exchange store and PST files. But, unlike external Email Archive solutions, it does not support journal capture for Regulatory needs, it requires PSTs to be imported by the user one-by-one and it provides no single-instancing nor compression.

## **5. Cost of Ownership**

### **5.1 How much storage will the archive consume and does it take advantage of single-instance storage and compression techniques?**

An email archiving solution will always require a significant storage medium as any archiving system is going to generate a substantial storage demand. Even with compression & single-instancing, this can add significant cost to the overall solution and should be sized according to your retention policies.

### **5.2 What software pre-requisites are required for the Email Archiving Solution and are these compatible with my environment?**

Is an SQL database needed or an information lifecycle management (ILM) tool, for example.

### **5.3 How does the solution implement Single-instance storage**

There are a number of ways in which to implement degrees of single instance storage of messages. For example, given that attachments typically consume 80% of mailbox space, it may be acceptable to link the original message to a separate archive copy of the attachment.

### **5.4 How easy will it be for existing staff to include the maintenance of the archiving system in their day-to-day activities without needing to learn complex new technologies?**

An email archiving solution should not consume a lot of your staff time in administration or in learning and supporting new technologies.

### **5.5. How well does the solution scale to my future messaging traffic requirements?**

Make sure you have sized a solution that can support your message volumes now and into the future allowing for sensible growth..

### **5.6 What pre-requisite Professional Services & Training are required from the Vendor to make the implementation a success?**

Some solutions are sufficiently complex that you have to purchase Professional Services to implement their solution. If that is the case, make sure it is stated and sized upfront.

### **5.7 Is the solution based upon an Appliance, Cloud (SaaS) or Software solution model?**

Your preference will determine where the archive data will physically reside and how acceptable this is from a security perspective. Up until recently most cloud based solutions were acceptable for capturing journal-sourced email but not

practical for ageing data out of mailboxes. With the recent advent of solutions from organizations such as Microsoft, the cloud has much more potential for storing mailbox archives but there are operational and cultural issues that will dictate whether a Cloud solution is right for you. Internal appliance based solutions may have data volume limitations to be considered.

**5.8 How well does the solution fit into my existing investment?**

Make sure the technology is not forcing you into replacing other components of your existing investment.

**6. Email Creation Restriction Policies**

**6.1 How does the solution help the management of Email Creation Restriction Policies?**

A good archiving system is going to capture email you may prefer had never been sent! So you should consider implementing controls to restrict certain kinds of information from even being created in the first place.

**6.2 How does the solution put controls on email a user can send?**

This would most likely be based upon a combination of email size, word content, recipient address (such as competitor, press, analyst).

**6.3 What actions does the solution take when a user tries to send a non-compliant message?**

You may wish to warn a user before the final send that a message contains content that is non-compliant. You may otherwise wish to block the message depending on the severity.

## 7. What are the next steps?

### Back to that Google search.... What did you find?

As you can see from the Google search, there is no shortage of potential email archiving solutions available. But you should be wary of the hype. Many software and storage vendors seem to be emphasizing the importance of addressing the regulatory requirements of email archiving to capture every email and Instant Message, and how quickly you will be thrown in jail for not complying. So again, who do you trust? How do you choose a solution that is right for you?

**ReSoft International is here to help you.** We are an independent integrator and reseller of best-of-breed software technologies that address Policy Management, Email & Instant-Messaging Security, Regulatory Compliance and Storage Relief.

1. **Get Educated** – ReSoft has put together a comprehensive White Paper Library addressing many of the issues you are probably facing. Some of the areas the Library covered are highlighted below. You can download any of these white papers free of charge. <http://www.re-soft.com/wp>
2. **Keep up to Date** – review the ReSoft blog to keep up to date with latest issues at <http://www.re-soft.com/blog>.
3. **Contact ReSoft for further discussions on how we may be able to help you.** ReSoft has, for over 16 years, acted as a trusted advisor to hundreds of organizations, applying the relevant technologies and techniques from its broad set of Email & IM Security tools to reduce liability risks and save time. We offer services to assist in defining an appropriate Archiving solution.

This paper has set out to illustrate ReSoft's significant coal-face experience in providing professional services and solutions that have helped many organizations to define & implement policy-based archiving & discovery systems across industries and organizations of all sizes. We have defined what we believe are the key questions to be asked of a vendor before committing to a particular archiving solution. These questions are part of an Archive Planning Workshop we offer to assist organizations in determining whether they need to deploy email archiving, what components would be required, and how they would be justified.

Review the **ReSoft** website at [www.re-soft.com](http://www.re-soft.com) if you are seeking ways to:

- reduce Mailbox Sizes and manage Retention & Quotas in your Lotus Domino & Microsoft Exchange systems
- undertake Legal & Regulatory Discovery Searches of Email and IM archives
- reduce Usage Violations, Loss of Intellectual Property & SPAM in Email, Webmail and IM
- make Email & IM delivery more Secure to meet HIPAA, GLB, SOX etc.
- make Exchange OWA & Outlook Calendar more Usable and Secure
- measure email/web Bottlenecks, Availability & Cost-Recovery
- proactively defend against Spyware, P2P, Worms and other emerging threats

### For Further Discussions, please contact:

ReSoft International LLC  
PO Box 124  
New Canaan CT 06840  
<http://www.re-soft.com>  
203 972 8462  
[info@re-soft.com](mailto:info@re-soft.com)