

October 17, 2008

# Best Practices: Email Archiving

by Jo Maitland

for IT Infrastructure & Operations Professionals



October 17, 2008

## Best Practices: Email Archiving

by **Jo Maitland**

with Rachel A. Dines and Simon Yates

### EXECUTIVE SUMMARY

With so many vendors now offering products for email archiving, getting beyond the hype to what you really need is becoming more and more of a challenge. Many of these tools were created for financial services companies dealing with compliance regulations, but they are now being adopted by the rest of the marketplace. Much of the functionality is overkill, and many firms are getting tangled up in the complexity of these products. We talked with archiving users as well as leading vendors and professional services organizations to narrow down the essential best practices in this field. Smart companies use four key strategies to succeed with email archiving: They rightsize the infrastructure; walk through eDiscovery scenarios ahead of time; consider archiving other data types besides just email; and staff and manage the archive for the long term.

### TABLE OF CONTENTS

- 2 **The Trouble With Email: Best Practices From People In The Know**
- 3 **Best Practice No. 1: Rightsize Your Infrastructure**
  - What Infrastructure Pitfalls Should You Be Aware Of?
- 4 **Best Practice No. 2: Walk Through eDiscovery Scenarios Ahead Of Time**
  - What eDiscovery Pitfalls Should You Be Aware Of?
- 6 **Best Practice No. 3: Consider Archiving Other Data Besides Email**
- 7 **Best Practice No. 4: Staff And Manage The Archive For The Long Term**
  - Pitfalls To Avoid
- 8 **Forrester's Email Archiving Next Practices**
- 8 **Identifying Your Challenges**
- 10 **Case Study**
  - Media General Deploys Scalable, Affordable Email Archive
- 10 **Supplemental Material**

### NOTES & RESOURCES

Forrester interviewed 30 vendor and user companies, including Autonomy, Microsoft, Mimoso Systems, LiveOffice, and Symantec.

#### Related Research Documents

["Archiving: Finding Data In 2050"](#)

July 1, 2008

["The Forrester Wave™: Message Archiving Software, Q1 2008"](#)

February 14, 2008

["The Forrester Wave™: Message Archiving Hosted Services, Q1 2008"](#)

February 7, 2008

## THE TROUBLE WITH EMAIL: BEST PRACTICES FROM PEOPLE IN THE KNOW

If there is ever a smoking gun in an eDiscovery case, it's usually an email. The Federal Rules of Civil Procedure were amended in December 2006 and mandated that businesses be able to quickly produce all electronic content relating to a case. Email has been the most difficult content to pin down. Business users send and receive hundreds of emails a day internally and across the globe to customers, making these electronic gems extraordinarily tough to keep a grip on. Email archiving tools have emerged in recent years to better contain and preserve email and also to search for it in the event of litigation.

To understand how to successfully implement and manage an email archive, we conducted more than 30 interviews with IT operations managers in charge of archiving, as well as vendors, integrators, and academic experts. For outside opinion of which best practices really make a difference, we interviewed IT directors from Chevron Phillips Chemical, City of Chicago, Johns Hopkins University, Media General, and Mitel Networks. To get the most out of email archiving environments, four best practices are: 1) Rightsize your infrastructure; 2) walk through eDiscovery scenarios ahead of time; 3) consider archiving other applications besides just email; and 4) staff and manage the archive for the long term (see Figure 1).

**Figure 1** Email Archiving Best And Next Practices

Best practice	"How to"	Pitfalls
Rightsize your infrastructure	<ul style="list-style-type: none"> <li>• Plan for fast growth.</li> <li>• Think about scalability.</li> <li>• Select an archiving product that has some awareness of backup built in.</li> <li>• Don't assume archive storage equals cheap storage.</li> </ul>	<ul style="list-style-type: none"> <li>• Turning off collection of email to the archive to expand your storage</li> <li>• I/O contention</li> <li>• Slow response times</li> </ul>
Walk through eDiscovery scenarios ahead of time	<ul style="list-style-type: none"> <li>• Avoid an email archiving product that's a hodgepodge of third-party tools cobbled together.</li> <li>• Define retention and deletion policies early.</li> <li>• Keep it simple.</li> <li>• Pinpoint essential discovery features.</li> <li>• Understand what email archiving doesn't do.</li> </ul>	<ul style="list-style-type: none"> <li>• Finding out the weakness of your archive during a discovery request</li> <li>• Keeping everything</li> <li>• Boiling the ocean on features and functions</li> </ul>
Consider archiving other data besides email	<ul style="list-style-type: none"> <li>• Be aware that SharePoint, IM, and voicemail systems are all discoverable.</li> <li>• Consider a common retention policy.</li> </ul>	<ul style="list-style-type: none"> <li>• Jumping to settle a case because eDiscovery seems too difficult</li> </ul>
Staff and manage the archive for the long term	<ul style="list-style-type: none"> <li>• Create a permanent, cross-functional team.</li> <li>• Have a strong DBA on staff.</li> <li>• Push the eDiscovery piece to the lawyers.</li> <li>• Plan for upgrades.</li> </ul>	<ul style="list-style-type: none"> <li>• Don't set it and forget it.</li> <li>• Don't breach your security policies.</li> </ul>
<b>Next practices</b>	<ul style="list-style-type: none"> <li>• This is a journey, not a project.</li> <li>• Maintain an audit trail.</li> </ul>	

## BEST PRACTICE NO. 1: RIGHTSIZING YOUR INFRASTRUCTURE

A common thread through many of our interviews was about legacy infrastructure and understanding whether it can scale to support the performance characteristics of the archive. One IT shop at a large manufacturer eventually figured out that its archive was slow not because there was anything wrong with the archiving software but because the database was running on an eight-year-old server! To ensure that you approach infrastructure needs correctly:

- **Plan for fast growth.** Buy more disk than you think you'll need and reclaim it elsewhere later if necessary. Terabytes of data can easily become a petabyte within a few years. Many IT professionals we spoke with were surprised at how fast their archiving storage needs grew. Moreover, if you plan to ingest PST files to the archive, or old Lotus Notes email for example, your storage requirements will increase fast.
- **Think about scalability.** Know your speeds and feeds. To help make sense of vendor claims around scalability, you will need to figure out the volume of email you get on a daily, weekly, and monthly basis and the average and peak email times. On this note, consider the kind of redundancy the archive should have. Should it be the same as the live email system? If your archive goes down and doesn't capture email for 24 hours, it's possible that during the outage users will delete email that will not be captured unless you have a failover system set up.
- **Select an archiving product that has some awareness of backup built in.** Archiving and backup processes must be aware of, but not interfere with, one another. In many instances, users cited examples where their backup application would capture a file but the archive would not; it would see some action happening with the file as it was being backed up and assume it was still in use. This prevented the file from being archived.
- **Don't assume archive storage equals cheap storage.** The concept of tiered storage has been whittled down to a single phrase: expensive disk for mission-critical applications, lower-cost disk for business-critical applications, and cheap disk for archiving. Unfortunately there's a lot more complexity to it than this, especially within the archiving tier. IT shops that have used one-size-fits-all cheap storage for archiving have run into performance issues, particularly when faced with eDiscovery requests. If you expect to turn around a quick response to a subpoena for information, your index to where everything is stored on the archive must be on high-performance storage.

### What Infrastructure Pitfalls Should You Be Aware Of?

Two of the most commonly cited mistakes were not dedicating enough capacity and power to the archive. Beyond those two, we found pitfalls around:

- **Turning off collection of email to the archive to expand your storage.** Running out of storage for the archive is a common trap we found among our interview base. Expanding on the fly puts pressure on the Exchange servers, as email now clogs up here and cannot be purged or deleted. It also burdens end users, as they can't delete email and will start calling the help desk.
- **I/O contention.** With free resources and spare virtual machines, it's tempting to run your Exchange and archive servers on a virtual system. But features in the archive like full text indexing are processor-hungry and will cause I/O contention, in many cases slowing down the responsiveness of Microsoft Outlook. We ran across several IT shops that started in a virtual environment and then broke out the archive onto dedicated resources.
- **Slow response times.** Don't assume an old server and a bunch of SATA disks will do. Your archive will perform only as well as the hardware you throw at it. Most companies use a mixture of SATA disks for the archival data and Fibre Channel disk for the index to ensure adequate responsiveness to search.

## BEST PRACTICE NO. 2: WALK THROUGH eDISCOVERY SCENARIOS AHEAD OF TIME

Our case study of Media General, a nationwide publishing company, shows how one company successfully installed email archiving for eDiscovery purposes. A key takeaway from Media General's experience was tapping HR and legal for their requirements before IT bought and implemented the archive.<sup>1</sup> To make the most of eDiscovery:

- **Avoid an email archiving product that's a hodgepodge of third-party tools.** Supporting a complex mix of archiving tools when something goes wrong will result in finger-pointing among the vendors. Most of the products out there today differentiate in just a few areas, and eDiscovery is one of them. It's important to understand what is native in the product versus sourced from partners, as this could impact your ability to recover email in a timely fashion.
- **Define retention and deletion policies early.** Establishing retention and deletion policies early will prevent your storage growth from becoming unmanageable. Under litigation hold circumstances, you will not be able to delete data, but otherwise a policy that states why you have a retention and deletion strategy and how it is followed is acceptable in court. It should say that you keep email in Exchange for a certain number of days and on the archive for a certain number of weeks or months before finally deleting it. A policy like this is defensible in court as long as you can prove that you adhere to it. If you don't have this policy you will be stuck with keeping everything forever. And don't even consider arbitrarily deleting data during the legal process. The courts will find out, assume you're up to no good, and fine you heavily. The SEC fined Morgan Stanley \$15 million for its blatant inability to produce documents.

- **Keep it simple.** A standard deletion policy among the companies we interviewed was 90 to 120 days. Many are classifying email, so that accounting email has a certain retention period, payroll email another retention period, contract email another retention period and so on, deleting it according to a pre-defined policy. For now this is a fairly manual process, with IT creating the classifications and users expected to move email to the correct folders, but eventually it will be automated.
- **Pinpoint essential discovery features.** Many of the email archiving tools on the market today have hundreds of features, most of them around electronic discovery. Understanding the key features and how you will use them will help get beyond the hype to what you really need. Our interviewees helped us identify the most commonly used features for eDiscovery (see Figure 2).
- **Understand what email archiving doesn't do.** Set expectations correctly with your management that email archiving is not an eDiscovery tool in the same league as CT Summation, Guidance Software, or Ringtail. These kinds of products perform very sophisticated tasks around processing, review, and analysis of data for discovery purposes. None of the email archiving tools out there can match this level of data analytics and were not designed to.

**Figure 2** The Most Commonly Used Features For eDiscovery

• Allow legal to search across a designated set of mailboxes for a case
• Let legal search message properties, content, and attachments quickly
• Let legal mark messages with custom notes
• Let legal mark messages for legal hold
• Save results of search
• Get notified when new items hit an existing search
• Export results of search
• Allow IT/legal to import legacy email (e.g., PST files, back-up tapes) into the archive
• Configure policy for email retention
• Delete messages at the end of period
• Role-based access only/authorization
• Support for other content types (Lotus Notes, IM, Database, SharePoint, file-storage)

### What eDiscovery Pitfalls Should You Be Aware Of?

eDiscovery is a stressful process during which your IT infrastructure decisions will be put to the test. Many companies are unprepared for this challenge. You want to avoid:

- **Finding out the weakness of your archive during a discovery request.** The subpoena has been passed along the management chain and is now in your hands as director of IT operations. Try explaining to your CEO that you need to hire a consulting firm to help produce the required data, as your archiving infrastructure was not designed with search and electronic discovery in mind.
- **Keeping everything.** Your storage costs and ability to manage the environment will quickly get out of control if you attempt to keep everything.
- **Boiling the ocean on features and functions.** Anecdotally, Forrester has found that most of the functionality of the top email archiving products rarely gets used. IT organizations often get caught up in the hype of all the latest features without thinking about how these features will actually help or fit in with their existing process.

### BEST PRACTICE NO. 3: CONSIDER ARCHIVING OTHER DATA BESIDES EMAIL

A subpoena for information never requests just email, or just files; it will ask for “all electronic content” that relates to the case. To be prepared:

- **Be aware that SharePoint, IM, and voicemail systems are all discoverable.** Records from all three of these systems have been requested in eDiscovery cases recently. Some email archiving tools support file storage and SharePoint, but none are able to cover the breadth of applications you will need to search in a discovery request. Look for tools that will include at least email and file storage, and check vendors’ road maps for when they will be supporting other kinds of applications such as SharePoint.
- **Consider a common retention policy.** It will make life simpler during a discovery request if you have a single retention policy across your email archive and enterprise content management (ECM) applications. If the judge asks for “all electronic content” from December 2002 to September 2008 and you have multiple different policies across multiple different data repositories, it’s easy to see how things get complicated fast.

Scrambling to search across many different repositories with inconsistent retention schedules will be a thankless task. Keep in mind a typical subpoena gives you 10 days for electronic discovery. More often than not cases are being settled out of court because the discovery process appears too burdensome on the company. In cases that have won, IT leaders have been able to negotiate what is discoverable and what’s not. You are the expert here so make sure you share your knowledge of the environment with your management team.

## BEST PRACTICE NO. 4: STAFF AND MANAGE THE ARCHIVE FOR THE LONG TERM

A consistent message from our interviewees was to treat this as a strategic project that will benefit the company for years to come if done correctly. In that spirit, we recommend:

- **Creating a permanent, cross-functional team.** Choose representatives from legal, IT, and business that will form a permanent team responsible for the preservation of important corporate information.
- **Having a strong DBA on staff.** Most email archiving products use SQL or Oracle to store the index, and there will be hiccups and inconsistencies in the database. The majority of customers we talked to said that DBA skills above administrative level were important for patching and upgrades. The email archiving index is going to grow fast, and it's not a system to practice on.
- **Push the eDiscovery piece to the lawyers.** Searching email takes time. The lawyers know exactly what they're looking for while IT doesn't and will often return with "everything we could find." Sometimes by the time IT gets a chance to work on the search, the legal department has changed the criteria. Legal needs to be a partner in the email archiving implementation from the beginning, and since it will have a direct impact on their jobs, should be trained on the system. Finally, the user interface must be simple, or else IT will spend more time going over work the legal team should be doing.
- **Plan for upgrades.** Keep budget in mind for technology refreshes and ongoing management of the archive. Email growth is not slowing down any time soon.

### Pitfalls To Avoid

A lack of consistency in staff to manage the archive, and IT trying to do it all were common problems from our interview base. Here's what not to do:

- **Don't set it and forget it.** This is not a plug-and-play system you can walk away from. IT needs to monitor it regularly to make sure that the archive jobs are performing according to policy and that the system is not about to run out of storage capacity. In addition, policies need to be adjusted to reflect changing regulatory requirements or organizational rules.
- **Don't breach your security policies.** If IT staff perform the search, they will be reading not only the requested emails, but also many others along the way to finding the right ones. This has implications for privacy violations. Clarify who should have access to the archive, when, and for how long, and monitor this process.



## FORRESTER'S EMAIL ARCHIVING NEXT PRACTICES

While our research uncovered a number of email archiving best practices, here are some next practices that IT infrastructure and operations professionals should focus on once they've mastered the basics:

- **This is a journey, not a project.** Good compliance programs will enable the company to go beyond just archiving emails. You should be able to leverage the data for business intelligence, better decision-making, and risk reduction.
- **Maintain an audit trail.** Identifying who accessed which email records, when, and for what reason, should be tracked by your email archiving software. You should be able to produce detailed and summary reports for audit trails, especially for patient or financial records. This provides your company with assurance that there have been no unauthorized changes to the data, and that privacy and security measures are in place, which is especially critical when dealing with HIPAA regulations.

## IDENTIFYING YOUR CHALLENGES

Where should you start? Use this diagnostic tool to assess your current capabilities — and opportunities for improvement — and see how you stack up against your peers (see Figure 3). Scores will be calculated automatically for online readers. All scores are anonymous.

**Figure 3** Email Archiving Self-Diagnostic Tool

<b>Part 1: Rightsize your infrastructure</b>	<b>Yes</b>	<b>No</b>
Do you know the availability requirements for each tier of storage and the cost implications?		
Have you projected the capacity growth of the archive?		
Are you ingesting PST files or old Notes email?		
Have you figured out the volume of email you get on a daily, weekly, and monthly basis and the average and peak email times?		
Have you considered what kind of redundancy the archive should have?		
Do you know how the archive will affect your back-up process?		
<b>Total</b>		

  

<b>Part 2: Walk through eDiscovery scenarios ahead of time</b>	<b>Yes</b>	<b>No</b>
Have you figured out what's native in the product versus sourced from third-party vendors?		
Have you defined a retention and deletion policy early?		
Is the policy usable and being adhered to?		
Do you know what eDiscovery features are essential?		
Does your management know the archive is not an end-to-end eDiscovery tool?		
<b>Total</b>		

  

<b>Part 3: Consider archiving other data besides email</b>	<b>Yes</b>	<b>No</b>
Have you incorporated other data into the archive such as file storage or SharePoint data?		
Have you created a common retention policy across archival repositories?		
<b>Total</b>		

  

<b>Part 4: Staff and manage the archive for the long term</b>	<b>Yes</b>	<b>No</b>
Have you created a permanent, cross-functional team?		
Does your IT staff have the right skills to manage the archive?		
Have you set aside budget for ongoing tech refresh and management of the archive?		
<b>Total</b>		

## CASE STUDY

### Media General Deploys Scalable, Affordable Email Archive

There are many reasons why organizations are implementing email archiving — to maintain compliance, relieve the burden on email servers, and provide legal protection. Whatever the case, it helps to look at how other companies have already deployed email archiving and learn from their successes and mistakes. Despite a tight budget, Media General deployed a successful email archive by following the best practices of rightsizing its infrastructure, nailing eDiscovery requirements upfront, and building a cross-functional vendor selection team.<sup>2</sup>

## SUPPLEMENTAL MATERIAL

### Online Resource

The online version of Figure 3 is an interactive self-diagnostic tool that helps clients assess how their current practices stack up against those of their peers.

## ENDNOTES

- <sup>1</sup> Before selecting a vendor, Media General asked legal and HR what their requirements were for an archiving solution and what they hoped to use it for. Some of these requirements included administrator granularity, the inclusion of prefabricated reports, and the ability to allocate permissions by role. The vendor selection team then used these as strict criteria for vendor selection. The result is an archiving system that does what they expected and hoped it would do. See the October 17, 2008, “[Case Study: Media General Deploys Scalable, Affordable Email Archive](#)” report.
- <sup>2</sup> Forrester published a detailed case study outlining Media General’s use of email archiving best practices. See the October 17, 2008, “[Case Study: Media General Deploys Scalable, Affordable Email Archive](#)” report.

# FORRESTER®

Making Leaders Successful Every Day

## Headquarters

Forrester Research, Inc.  
400 Technology Square  
Cambridge, MA 02139 USA  
Tel: +1 617.613.6000  
Fax: +1 617.613.5000  
Email: [forrester@forrester.com](mailto:forrester@forrester.com)  
Nasdaq symbol: FORR  
[www.forrester.com](http://www.forrester.com)

## Research and Sales Offices

Australia	Israel
Brazil	Japan
Canada	Korea
Denmark	The Netherlands
France	Switzerland
Germany	United Kingdom
Hong Kong	United States
India	

*For a complete list of worldwide locations, visit [www.forrester.com/about](http://www.forrester.com/about).*

For information on hard-copy or electronic reprints, please contact the Client Resource Center at +1 866.367.7378, +1 617.617.5730, or [resourcecenter@forrester.com](mailto:resourcecenter@forrester.com). We offer quantity discounts and special pricing for academic and nonprofit institutions.

Forrester Research, Inc. (Nasdaq: FORR) is an independent technology and market research company that provides pragmatic and forward-thinking advice to global leaders in business and technology. For more than 24 years, Forrester has been making leaders successful every day through its proprietary research, consulting, events, and peer-to-peer executive programs. For more information, visit [www.forrester.com](http://www.forrester.com).