

E-mail Archiving

Table of Contents

Executive Summary	Page 1
E-mail Archiving Defined	Page 2
Policy Development	Page 3
Technology Options	Page 4-5
Best Practices for E-mail Archiving	Page 6-7

Executive Summary

From an improved ability to meet compliance and e-discovery requests to better storage management and opportunities for knowledge management, most organizations can benefit from deploying an e-mail archiving solution. This can be especially true for state, county or federal governments who process Freedom of Information Act requests or must comply with open records or sunshine laws in their respective states, as well as educational institutions.

After the Federal Rules of Civil Procedure codified the expectations for e-discovery, the need for the effective archiving of electronically stored information, such as e-mail and file attachments, have increased dramatically. The amended rules require any organization that might be sued in federal court to have systems for retrieving electronic data — which could include e-mail, network activity logs, digital recordings, voice mail, spreadsheets and more — if the information could be considered evidence in litigation.

At their core, most e-mail archiving solutions sit somewhere in the messaging stream, indexing the contents of messages coming and going and copying the messages themselves to another locale. When it comes to searching or retrieving archived data, web interfaces and browser plug-ins are often incorporated to make retrieval easier and virtually transparent to users working within traditional inboxes.

System Overload

- 97 billion e-mails will be sent daily worldwide in 2007.
- 40 billion of those e-mails will be spam.
- Annual volume of business e-mails sent worldwide will approach five exabytes in 2007.

SOURCE: "Worldwide E-mail Usage 2007-2011" by IDC, Framingham, Mass.

What Is E-mail Archiving?

The Federal Rules of Civil Procedure, coupled with long-standing requirements to maintain a regular records-retention schedule at numerous organizations, has made implementing e-mail archiving technologies and creating acceptable usage policies necessary. The question for most organizations is how to start the process of evaluating e-mail archiving tools and to ensure that they're following recognized best practices.

Because no organization can do everything at once, experts suggest first getting a handle on an IT area that is often the source of most discovery activity: the e-mail systems in the organization.

"E-mail is still the killer application. With the growth of the web, it's become even more so," says Jason Baron, director of litigation for the U.S. National Archives and Records Administration.

A good starting point is to first establish a policy-based approach to e-mail archiving complete with a central repository. That policy should cover acceptable use and meeting public records requests with the Federal Rules of Civil Procedure (FRCP).

As regulations covering e-mail retention, public requests for information and the demands of the e-discovery process all multiply at once, archiving is becoming a more complex undertaking. To ensure that an e-mail archiving system meets the needs of various departments, some organizations are opting for user-directed archiving. Departments have long done their own e-mail retrieval through the archiving system, and now employees are being trained to set varying retention periods for messages that meet their specific requirements. Individual departments can also create folders in the archive that cross-reference related issues and speed retrieval.

Though there's a growing clamor for tools that implement archiving policies across all electronic formats and media, fully integrated systems are several years away, says David Ferris of Ferris Research, an analyst firm specializing in messaging technologies.

"Government has special obligations for open access to information, so there's greater demand

to archive everything with tools that will let you get at anything," Ferris says. "In principle, they'll all blend because they're all types of electronic documents, but e-mail is different. From a technical perspective, you're more interested in the structure of e-mail. E-mail archiving may stay essentially a separate technology for five years."

For now, the most prominent trends in e-mail archiving are improvements in the search capabilities of the tools available, Ferris says.

Your institution needs records management policies for electronically stored information. The policies should identify the types of records (e-mail, electronic spreadsheets, etc.) that must be kept, where and how they will be stored, when, if ever, they should be deleted and who has responsibility for deleting them. Train employees so they understand what "acceptable usage" means.

Tips for creating an acceptable usage policy

1. Convene a cross-functional team of department leaders to set expectations for how e-mail, Internet and computer resources are to be used relative to business goals.
2. Create a detailed document that defines what you consider to be appropriate — and inappropriate — behavior. Don't assume that your users know what you consider to be the difference between personal and professional messages.
3. Explain the consequences for violating the policy within the document. For instance, if misuse of the e-mail system will result in termination, say so.
4. Notify employees that you will be performing random, periodic audits of their mailboxes. Experts say this alone can be a great deterrent for misuse.
5. Present the document to all employees at several times during their tenure, including hiring, reviews, and staff meetings. Allow time for them to ask questions and provide clear answers. Have them sign a written statement confirming they have read and understand the policy.

E-Discovery and Electronically Stored Information

Being right isn't enough. Sometimes, you need to prove it. And that's why electronic documents are increasingly subpoenaed in civil cases.

But it's not because of a lack of documentation; most organizations have too much. It's because they've captured the wrong data or simply can't find the right data. That's why organizations need to consider e-discovery before they deploy any electronic data archiving and management system.

Experts acknowledge that the process of "getting there" may seem unending, but maintain that agencies can accomplish much groundwork in the first nine months, simply by creating an interdisciplinary team of legal personnel, records managers and IT folks who meet regularly to hammer-out policy.

The process of evaluating your organization's e-mail archiving system and process is an opportunity to meet key legal and regulatory requirements for preserving electronic data and preserving it for an investigation. In 2006, amendments to the FRCP specified that the discovery process applied to electronic documents and provided guidance as to how those documents should be handled. The amendments codified what had been a reality in the legal system for several years, according to David Goldstone, a partner at law firm Goodwin Procter LLP in Boston.

"Courts have been treating electronic documents in the same way they treat paper

documents for many years. It's just that the number of electronic documents is multiplying exponentially. For example, some people send 50, 60, even 100 e-mails a day," says Goldstone.

The issues raised by the FRCP and e-discovery for the private sector are exacerbated for government and education, which gather massive amounts of electronic information and often have limited resources to deal with the data, says Goldstone. Compared with private businesses, government entities with their agencies and department mandates for public transparency must also deal with a wider array of retention requirements, he says.

Most organizations have long-standing requirements to maintain a regular records-retention schedule for the mountains of paperwork they produce. As such, you might think these organizations had a leg up on the rest of the world when it comes to handling legal discovery requests or a Freedom of Information Act inquiry.

Sadly, judges in a few U.S. government cases found otherwise. Fast-forward to the present world of e-discovery and the need to now comply with various amendments to the Federal Rules of Civil Procedure — it's enough to make even the most stalwart CIO tremble as thoughts turn to IT preparedness for legal holds, preservation requests and the ongoing production and protection of a wealth of electronically stored information (ESI).

In civil court, you only have to be 51 percent right," says Trent Livingston, a principal in the e-discovery practice group of expert services firm LECG of Emeryville, Calif. Even then, organizations have a tough time proving that "they did the right thing."

Important Ruling: Phoenix Four Inc. v. Strategic Resources Corp.

The U.S. District Court in New York ruled in 2006 that the defendant, investment adviser Strategic Resources of New York, had overlooked the equivalent of 2,500 boxes of documents during the e-discovery portion of its trial with investment firm Phoenix Four of the Bahamas. The reason: The IT staff never spotted a partitioned hard-drive section containing data. The judge found that the "duty in such cases is not to retrieve information from a difficult-to-access source, such as the server here, but rather to ascertain whether any information is stored there."

Technology Options for E-mail Archiving

The new electronic discovery laws mean IT departments have to retain e-mail and other digital documents in case they are needed as evidence in lawsuits. Most IT leaders have probably told staffers to archive their own e-mail or to print the messages that need to be retained and keep them. Unfortunately, this isn't good enough any longer.

Now these documents and others, such as network activity logs, digital recordings and voice mail, must be retained as long as the organization's policy states. Luckily, technology is able to help; products designed to handle e-discovery typically allow the administrator to archive, index, classify and search content.

Symantec Enterprise Vault

Symantec Enterprise Vault is one of the better-known products in the e-discovery toolset. The flexibility and thoroughness of Enterprise Vault is really what sets this product apart. E-mail archiving can be configured for Microsoft Exchange, Lotus Domino and SMTP servers. Enterprise Vault can also integrate file systems, content management systems, SharePoint sites and instant messaging content. A typical deployment of Enterprise Vault would consist of two servers: One would act as the indexing and archiving server; the second would be a Microsoft SQL Server. The SQL server is required for storing the configuration data and item archiving. Users can search archived data through a simple web interface.

Enterprise Vault also offers an extension to their product called Discovery Accelerator, providing an automated, defensible and efficient means to extract data for further legal review

This provides several benefits including:

- **Integrated workflow** — E-discovery administrators can take advantage of an integrated solution to extract files from their Enterprise Vault archive system and deliver them into third-party tools.
- **Minimal labor for production** — Without the need to manually transfer potential evidence, this lowers the total cost of ownership.
- **Efficient marking and review** — Items marked in third-party tools (for example, as attorney — client privileged) during the review process can be returned to the Discovery Accelerator database to reduce the cost of further review.
- **Tracking and management of internal or external productions** — Inside counsel maintains centralized visibility into what has been produced and where it has been sent with reporting about work completed by outside counsel or internal investigators.

GFI MailArchiver

GFI MailArchiver is a software solution that can reside on your Microsoft Exchange Server or on a separate server. MailArchiver uses the journaling feature of Exchange to archive all copies of messages into either a Microsoft SQL database, its own SQL database engine or a file system. MailArchiver will also allow employees to search through their own e-mail for lost or deleted messages that no longer show up in Outlook. This feature can reduce the workload on IT departments.

Administrators (and users) can search e-mail through several criteria, as well as search the contents of any attached files. The interface is web-based; there's no need to load software on client stations. Administrators can import PST files, ending the dependence on archiving all messages in the cumbersome PST file format. Pricing is based on the number of active mailboxes on the Exchange Server.

\$3,500: The cost to produce deleted e-mails from a single backup tape, according to an estimate from Kroll Ontrack.

Barracuda Message Archiver

Like other offerings from Barracuda Networks, its Message Archiver is an appliance-based solution. The appliance will instantly archive and index all e-mail, allowing immediate retrieval by authorized users. The device can integrate with Microsoft Exchange, taking advantage of the journaling feature to retrieve the messages. Most other e-mail servers can be configured to send copies of all e-mail to a specific address, which is then retrievable by the Message Archiver system.

Once installed, the system requires minimal maintenance. There are three models of Archiver, depending on the size of your organization or how many messages you need to retain. In addition to archiving, the system will allow the administrator to set up alerts to notify administrators when policies are violated. Barracuda Message Archiver has no per-user license fees.

The manufacturer also offers Barracuda Spam Firewall 300. Mountains of unwanted solicitations, ranging from absurdly spelled pharmaceuticals to stock quote recommendations to utterly meaningless blather, inundate your inbox and harm productivity. But picking spam messages off one by one at the e-mail client level simply will not work, not to mention the extra taxation on your e-mail archiving resources that occur by letting the spam get that far. The Barracuda Spam Firewall 300 eliminates the need for e-mail archiving systems to waste resources on completely irrelevant e-mail.

Barracuda allows IT to configure black and white lists to include specific e-mail addresses or a range of addresses that you know you must receive e-mail from or block completely. As part of annual maintenance, it also provides access to popular online blacklists from organizations that specialize in maintaining lists of known offenders.

Barracuda's Intent Analysis checks e-mail for certain key characteristics often found in spam,

and then it assigns a score. If the score adds up to a certain number, it will perform a task that you have assigned. For example, you can have it tag the subject with a standard phrase if it scores six out of 10, but block the e-mail completely if it scores a 10. IT can also create dictionary lists of keywords that the server will search for in each message, including wild cards to catch differently spelled words; for example, "Vilagria." Barracuda also has optical character recognition software built in to check image spam, and it blocks attachments with certain file extensions, such as .BAT, .EXE and .VBS.

Sony Intradyn ComplianceVault

Intradyn and Sony have partnered on the ComplianceVault appliance. This system is based on Sony's AIT tape drive with WORM (write-once, read-many) technology. The appliance installs in a manner consistent with other e-mail archiving appliances and can be set up in 10 minutes or less. The ComplianceVault supports virtually all POP3- or IMAP-compliant mail servers, as well as Microsoft Exchange and digital faxes. Conversations in instant messaging systems from Akonix or IMLogic can be archived as well. The system continuously archives all e-mail from the mail server and can keyword search 1 million e-mail messages per second.

All e-mail is stored to the 1 terabyte of internal drive storage and copied to WORM tapes to ensure the archives cannot be tampered with. The ComplianceVault does not require any per-user licenses and is offered in several storage sizes.

There are several ways to satisfy your e-mail archiving needs, and this list is not complete. When choosing your solution, consider what capabilities you need, as well as your deployment concerns. All of these products are designed to reduce search time. These systems can pay for themselves quickly if your organization is involved in any lawsuits or public records requests.

Purchasing Pointers

As you evaluate e-mail archiving wares to determine which best meets your organization's need, consider these questions regarding performance and scalability:

- How scalable is the solution in terms of number of users supported, number of servers supported, message throughput per hour, number of records supported and number of reviewers for pre- and post-send management?
- Can the vendor provide the results of performance and stress tests in real-world settings?
- To what extent does search and retrieval performance suffer as the number of records in the archive increases?
- What load-balancing functions are built into the product?

SOURCE: Osterman Research

Ferris Research estimates

the number of users on e-mail archiving systems will grow 55 percent between 2008 and 2010, to 32.3 million.

Best Practices for E-mail Archiving & E-Discovery

So, which organizations need to be aware of e-discovery? All organizations that have a computer, because most documents today begin their life in an electronic form. Although it's a bigger concern for larger organizations, e-discovery still applies to even the smallest institutions. The rules are intentionally imprecise, leaving room for the state courts to interpret them. Consider these tips to ensure that your institution is protected.

ONE: Transition from paper to electronic record-keeping.

Surprisingly, many organizations have grown accustomed to keeping their records in paper form. For these institutions, their first hurdle may be making the move to digital record-keeping. Many organizations have records schedules that are still based on paper. For these organizations, fielding a request — e-discovery or otherwise — is much more time-consuming.

TWO: Create a clear and consistent retention policy.

All organizations need to create and follow strict records-management policies, says Diane Barry, senior managing consultant in the e-discovery practice group of LECG, an expert-services firm in Emeryville, Calif.

The policies need to explain what kinds of records are kept (financial records, e-mail, IMs and blogs); how they will be kept (for example, whether copies of all files are automatically archived); in what format (such as tape backup or paper); when and if they are to be destroyed; and who destroys them. Organizations also need to decide how long they will keep IT records — for example, logs noting who is accessing which servers.

THREE: Purge files consistent with your set policy.

What records do you need to keep? "There's no law that says you simply have to keep everything," says Barry. "There are two reasons to keep files: if you have [an organizational] need for it, or if you have a legal need for it," she says (for example, financial records for tax purposes, or compliance records for regulated industries). Everything else should get tossed during routine purges of electronic files, she says.

Hoarding too many files, whether electronic or paper, carries financial and legal risks. Additionally, data storage per se is inexpensive. But sifting through electronic data is not.

That's one reason legal experts keep as little as possible. But, if litigation should ensue, you're obligated to halt the destruction of files.

Get rid of documents that aren't needed for regulatory or organizational purposes, when it's possible under the law. Hoarding files can result in unnecessary legal and financial risks and can make it more difficult to search for electronic documents in response to an e-discovery request.

FOUR: Understand legal holds.

Even if your policy sets an automatic deletion date, your organization is still obligated to preserve records as soon as the organization is in litigation or the subject of regulatory investigation reasonably anticipates such an action, for paper records as well as electronically stored information. Your organization can be sanctioned if any relevant information is lost when litigation is anticipated. The legal term for this is "spoliation of evidence," and penalties can include fines or criminal liability.

FIVE: Ensure that archived records are searchable.

The problem is, retaining records for an appropriate amount of time is just one aspect of the challenge. The other and trickier part involves searching those records to find the smoking-gun e-mail that could prove or disprove a case.

"Many [organizations] have well-managed systems for organization continuity," explains Jim Barrick, CEO of Control Discovery, a San Francisco firm that specializes in e-discovery services. "Unfortunately, that train comes off the track when they have to retrieve that information. While the task is storage, the goal is retrieval."

Making sure e-mails and other electronic documents are archived correctly is important, but the goal is efficient retrieval, and that depends on effective search algorithms.

SIX: Build a topographical map of your electronically stored information.

When it comes to navigating the realm of systems, databases, applications and e-mail messages, another e-discovery directive emerges: Before you can respond successfully to any legal request, you need to first get your own IT house in order. Increasingly, initial "meet and confer" discussions between opposing and defense counsels now rely on the availability of a content-rich (and context-sensitive) "data map" that describes not just where certain systems are, but also the type of data they contain, how often the data is backed up and the policies usually in place to automatically archive or delete data.

Think twice before you rely on a traditional IT architectural map or network topology diagram for the task, says Jonathan Redgrave, chairman of law firm Redgrave Daley Ragen & Wagner and editor-in-chief of *The Sedona Principles*, one of the Sedona Conference's industry-leading works on e-discovery and the FRCP.

"You need to be able to pull together some type of mapping of applications, databases and systems most likely to be called upon or looked to in either FOIA requests or

discovery proceedings," he says. "Instead of having an IT architectural map, however, you need a description of each of the data sources so that a nontechnical person can understand what and where the data is, and if the data is subject to any auto-deletion."

SEVEN: Archive e-mail and centralize data.

Prepare a server/storage system for centralized file management. Create one shared file folder for departments or groups. Create another shared file folder for shared data and add one subfolder per division, such as accounting or the IT department. These subfolders should more or less mirror your security model, so that accounting is the only group who needs access to the accounting folders, for example. Some of these subfolders will simply be logical ways to organize data.

EIGHT: Don't confuse backup with archiving.

Besides e-mail, risks lie in the disposition of backup tapes as well. "If you had to triage your problems, risks and things that get agencies into trouble, it's e-mail, it's backup tapes," says Baron. Distinguishing between backup processes and those used for archiving is key. "Backup tapes shouldn't be viewed as record-keeping systems. They should just be for disaster recovery."

Redgrave shares this view, which is also discussed in *Sedona Principle #8*. "You really need to have a good handle on what is being done, both in the archiving of information for medium-to-long-term storage as well as what's being done in the area of backup," he says. "Data should be kept only as long as necessary for backup, and then those tapes and media should be truly destroyed or rewritten — unless there is a legal hold. A lot of times, people use backup tapes for archive and preservation." In the area of information management programs and policies, the *Electronic Discovery Reference Model* shares criteria you can use to apply to data used for backup versus archiving. ♦

Before an organization knows what to throw

out, they need to know what they have. Most don't. E-discovery case law is filled with examples of organizations who sabotaged their own defense by not knowing where their data was. Indeed, 30 percent of small institutions surveyed by law firm Fulbright & Jaworski in 2007 said so-called "pre-production" efforts accounted for a fifth or more of overall litigation costs.

