THE STRATEGIC IMPORTANCE OF EMAIL ARCHIVE MANAGEMENT

An email archive strategy is imperative as email grows exponentially

By Kevin Craine, MBA

Every day thousands of email messages are created as part of a company's business operations. Organizations find that email clogs computer networks and demands increasing amounts of their employee's time. Email is growing exponentially, and as a result, vast quantities of email reside on network servers, PC hard drives, backup tapes, and other mass storage devices located in corporate information systems. The mounting volume of email, which in many cases must be archived, is the number one corporate storage hog.

Email also constitutes a "corporate memory" of how an organization conducts its business – evidence of corporate decisions and behavior. Email messages are testimony to an organization's functions, activities and transactions. Recent headlines draw attention to the role of email in the Microsoft antitrust trial and the multi-billion dollar settlement in the "fen-phen" diet drug case. Ask any corporate counsel what one worry keeps him or her awake at night and the likely culprit is the multitude of email messages stored throughout their enterprise. Email represents a growing and elusive risk in litigation, so understanding exactly what your corporate email says and exactly where it is stored is critical.

A great deal of "corporate knowledge" takes form in electronic mail. Nearly all organizations depend on email to conduct business, yet very few regard email as an important corporate asset. Mining the value of email requires a well-thought archive strategy. Email records can easily be destroyed or lost due to hardware failures and routine clean up, leaving behind only fragmented traces of the corporate knowledge contained within. Archiving and restoring email requires costly administration that is often inadequate, and takes a toll on employee productivity. And when a virus strikes the losses can be staggering, as was the impact of the infamous "I Love You" virus in May 2000.

It is now axiomatic that the World Wide Web will expand the dimensions of human communication and that the Internet will shape new markets. The emerging issue, however, is email.

Despite the expansive aspects of email, electronic messaging systems do very little to help users manage their vast and dynamic stores of corporate email.

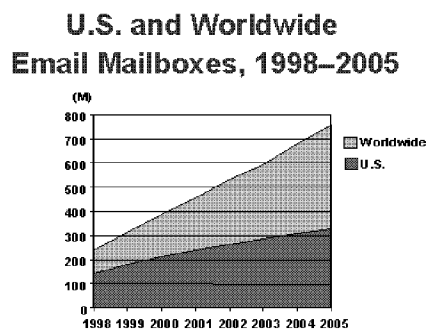
Despite the expansive aspects of email communication, electronic messaging systems do very little to help users manage their vast and dynamic stores of corporate email. The first part of this paper attempts to demonstrate the escalating growth of email, the risks associated with this corporate information, and the knowledge value contained within ever-growing corporate email stores. The second part of this paper provides a matrix of desirable features and explores some fundamental issues to consider when selecting an email archive system. In the end, it is the hope of this author that you will be better prepared to make informed decisions and form beneficial strategic plans with regard to your corporate email environment.

The Growth of Email

The growth of email use is phenomenal. Whatever watermark you use to measure the advance, this fact is clear: Email dominates the swelling tide of digital information – a tsunami of e-information that floods nearly every aspect of doing business and has begun to saturate corporate information storage systems.

Indeed, email has become so commonplace that it is hard to imagine getting along without it. This year, Americans will send and receive well over 6.8 trillion email messages – approximately 2.2 billion messages each day – compared with just 293 million pieces of first class mail. [1] Roughly two thirds of all American workers use email as part of their daily routine [2] and the average office worker will send and receive between 60 and 200 email messages each day. [3] Microsoft, for example, processes more than 3 million email messages every day. [4]

The convenience of email has an infectious appeal, and people are spending more and more time sending, receiving and storing email. Workers polled this year by the Gartner Group, a leading e-market research firm, spent an average of 49 minutes a day on email, 30 to 35 percent more time than they did a year ago. Ferris Research estimates that management level workers spend four hours a day working with email. Indeed, many of Intel's 88,000 workers spend up to 2 ½ hours a day on email. [5]



Email is not just an American phenomenon, however. Traffic is growing even more rapidly around the world. The number of email addresses worldwide increased 84 percent in 1999 alone, with the segment located outside the U.S. leading the way with triple-digit growth. [6] Today, over 225 million people worldwide can send and receive email, [7] and predictions are that an astounding 750 million email addresses will exist around the globe by 2005. [8]

The World Wide Web often grabs the headlines, but the fact is that there is 500 times more email created every year than the stock of Web pages (in terabytes). [9] Although a portion of this email is sent and received by individuals, most of the traffic travels via corporate email servers.

Email Size on the Rise

While email use is growing exponentially, the size of email (i.e., bytes) is increasing as well. In the early 1990's, most email messages were between 1Kb and 3Kb in size, but the size of a typical email today (without an attachment) is around 5Kb. A decade ago file attachments were rare, but today over 20 percent of email comes with an attachment. Graphics, spreadsheets and Word documents frequently accompany email messages. As a result, email with attachments average around 100Kb currently. [10] Analysts predict that the size of email and attachments will grow at a rate of 35 percent per year. [11]

Because of this byte-sized trend, email has spawned a boom in the digital archiving business. Analysts estimate that the demand for digital storage grew by more than 1,800 percent between 1998 and 2003. Customers will pay dearly for this digital storage space. Spending on managed storage services, a mere $11 million in 1999 is estimated to have grown to $4.8 billion in 2003. [12] Revenues at EMC, the leading provider of digital storage, have increased well beyond forecasts – by over 50 percent annually. [13] Increasing volumes of email are the cash cow in EMC storage farms.

To combat the ever-growing demand on server storage, system administrators often limit the allotted size of their users email store. Today, the median size of a message store allotted to a typical corporate user is 45Mb. [14] If we assume that a typical user sends and receives 70 messages per day (at 5Kb) and 20 percent of those messages have attachments (at 100Kb), then our user generates a total of 1.75Mb of email traffic per day. If all of this data were kept in his message store, our user would exceed his allotted storage space in just over 25 days.

Users get around this limitation by storing email in multiple places (often multiple times). Email takes up residence in a hodge-podge of hard drives, disk drives, tape drives, laser disks, compact disks and network servers of all sizes and types. A midrange estimate of the amount of data currently stored on magnetic tape is 2.5 exabytes (an exabytes is 1 million terabytes) with another 2.5 exabytes stored on computer hard drives. [15] Analysts estimate that 30 percent of stored email records and their attachments can be eliminated because they are redundant copies. [16]

Email Expense on the Rise

The quantity and size of email is growing beyond the ability of many organizations to effectively administer it. One study found that out of 926 organizations surveyed, 97 percent said that their email stores had grown faster than predicted. Nearly three-quarters had growth in excess of 50 percent, and a majority said that email backups consistently run into production periods. In 67 percent of the companies surveyed, restoring lost email for a single user takes over 24 hours to complete. [17]

The cost of email administration, therefore, is a growing concern. On average, an organization will spend nearly $200 per user each year to locate and retrieve information from backup tapes. When considering lost user productivity, revenue loss and administration and management costs, companies spend more on this problem than for all technical support and help desk combined. [18]

Another factor that contributes to the escalating expense of email archive and administration is found in the nature and evolution of storage technology over the past 30 years. In the 1970's data archives were almost entirely transactional and interactive in nature with large mainframe systems processing airline reservations or health insurance claims. The 1980's ushered in the era of desktop computing and a shift in the nature of data being stored began to occur. While traditional transaction-based data still accounted for the lion's share of information in data stores, a growing percentage became static in nature (e.g., stationary vs. dynamic) as corporate email led the "knowledge worker" movement. In the 1990's this shift became more dramatic and today approximately 70 percent of data stored is static rather than interactive in nature. [19]

Due to this evolution in data file structure many of the current approaches to email archive management are unduly expensive because they are based on an outdated model. This cost can be a barrier to entry for many organizations. The "classic" approach of the 1970's and 1980's worked well for the type of interactive data prevalent 30 years ago, but implementing "big iron" may not be the best approach to archive static email messages…especially as volumes grow beyond expectations. As a result, some organizations find that the costs of an archive solution outweigh the risks of not having one in place. "We'd rather take our chances and pay the fines," laments one anonymous senior executive (see "risks" below). Any email archive solution you consider must be weighed against the expense and risks involved.

Experts estimate that for every dollar spent in storage for email archiving, another five dollars are spent to manage the system. [20] To assume that the cost of email archiving is equal to the cost of the associated storage systems is not entirely accurate. You must certainly analyze your volume in terms of terabytes, but you should also consider your archive volume in terms of number of objects that make up those terabytes. When combined, the volume in terabytes and the volume in email objects can dramatically impact your expenses, as a large database and search engine infrastructure must be implemented to enable the accurate retrieval of such objects. Traditional systems are focused on reducing the expense of storage hardware, but do nothing to reduce the expense of infrastructure and administration.

In short, email messages are growing in quantity and contain increasing amounts of information. The time that employees spend using email is on the rise and more time and money is spent than ever before in attempts to manage the ever-growing tide of email. Traditional approaches to data archive may be outdated and cost-prohibitive to implement.

The Risk of Email

Email is more than just bits and bytes, however. Email is a cornerstone of corporate business and has quickly replaced written documents, faxes, and even the telephone as the primary communication conduit for most corporations. As computer technology has advanced and people have become accustomed to using email, the importance and complexity of email has evolved from a few simple text messages into a multitude of mission-critical documents. Negotiations, bids, proposals, contracts, legal agreements, regulatory forms, and a host of other vital correspondence now find form in email.

But with this progression comes increased corporate risk. Email is essentially a "corporate memory" of how an organization has conducted itself. A potent and powerful factor during litigation, courts in the U.S., as well as many other countries, have ruled that email relating to potential litigation must be preserved as evidence and is subject to disclosure. According to Richard Lazar of Fios Inc, an electronic discovery services company, "email can often show - with a high degree of reliability - who said and knew what." Lazar, a former litigator and now President and CEO of Fios, points out that "email often shows who else was involved in activity. It is a litigator's primary weapon in the search to uncover a needle in a haystack." [21]

A potent and powerful factor during litigation, courts have ruled that email relating to a particular criminal case must be preserved as evidence and is subject to disclosure.

As email becomes the default method of business communication the ability to archive, search and understand vast stores of email with effectiveness and ease is imperative in order to mitigate corporate risk during litigation. While email is often used for formal correspondence, the informal nature of email can bring additional risk. Email reaches beyond traditional business communication and bonds people in a web of informal banter that is potentially dangerous. Employees tend to be less formal when using email even when discussing formal business and often unknowingly treat email as an unrecorded telephone conversation.

The essential questions are: What does your email say about your company and the conduct of your employees? Would you be comfortable having your email exposed in a court of law, tantamount to a retroactive wiretap on your communications?

Email is the "Smoking Gun"

For trial lawyers email often represents the "smoking gun" that will win their case or force a settlement. Attorneys aggressively lobby courts for access to their opponents email stores. This in itself can be a significant corporate risk. One executive bemoans: "We had to produce all the email pertaining to the matter from the past 10 years. This was a tremendous expense. Some of the information was on tape backup, but was very difficult and time consuming to select, filter and reproduce. We also had to go to every user to get local email stores from laptops, desktops and even some employees' home systems. It was a nightmare and cost the company millions." [22] Judges have noted the costs and effort associated with retrieving email records, but have characterized this burden as "one of the risks taken on by companies which have made the decision to avail themselves of computer technology."

Consider the case of a female employee who brought a lawsuit against software giant Oracle. The woman had a romantic relationship with CEO Larry Ellison, but was eventually fired for an unrelated cause. In retribution, the woman sent an email message from the account of another Oracle executive

saying, "I have terminated [the woman] per your request." Based on the email, the employee filed a wrongful discharge and sexual harassment suit against Ellison. But the district attorney found the smoking gun in the Oracle email audit logs, and eventually proved that the woman had falsely created the email message. She was prosecuted for perjury and falsifying evidence. [23]

Perhaps the most popular example of damning email in litigation is the U.S. Department of Justice's antitrust trial against Microsoft. The content of email traffic between Bill Gates and other Microsoft executives was startling and captured attention in many corporate boardrooms. This message rang clear: If a damaging email exists, someone will find it, and it will most certainly resurface to haunt you.

To mitigate the potential risks associated with discovery, some companies have a policy to intentionally delete email messages (after 30 days, for example). This type of approach can backfire, however, as was the case in the "fen-phen" diet drug suit. One of the manufacturers was severely sanctioned for overwriting tape backups for several months after litigation had begun. The court instructed the jury to infer that the deleted messages contained evidence that was potentially damaging, and that the company had either negligently or intentionally destroyed the evidence.

Another risk that is often overlooked is the fact that a great deal of personal email is transmitted via corporate messaging systems. This can leave a company open to both litigation and theft. The free flow of email means the backdoor is potentially always open. One example is a recent case that involved a contract employee at Gillette Company who was caught using email to sell stolen plans for the company's new Mach-3 razor.

68% of companies surveyed characterized email abuse as widespread, with losses estimated at $3.7 million per company per year.

- Lee Bruno, "How Safe Are Your Business Secrets?"

"30% of the email going through our servers was not work-related."

- Jeff LePage, director of MIS at American Fast Freight Inc.

In short, the risks associated with email are significant in terms of litigation, exposure of corporate and employee behavior, and theft of proprietary information. In order to mitigate these risks, companies must understand where all of their email records are located, what they say and who is saying it.

The Knowledge Value of Email

A great deal of "corporate knowledge" takes residence in the form of email. About 35 percent of the business information used in day-to-day business is stored in email messaging systems. [24]
Firms are beginning to recognize email as a corporate asset, but for the most part they have been powerless to gain control or do anything constructive with this intellectual property.

Mining and managing the information contained in email is a challenge that is more demanding than simply moving bits and bytes to tape. The key to finding the "knowledge value" of corporate email lies in the ability to access, understand and use these archived messages in beneficial and collaborative ways. Email records need to be quickly and easily captured and archived so that they become part of the available corporate information repository.

Hidden Knowledge Assets

Important email information is often hidden from the rest of the organization. For the most part, individual users decide which email messages will be retained or deleted. This is especially true with older email

messages. For example, when an employee leaves an organization or is otherwise unavailable, the knowledge assets contained in their email store are generally unattainable or, in the worst case, lost for good.

Even when employees are on the job, lost or unavailable email presents a problem. As mentioned previously, if a typical user reaches the capacity of their allotted email store in less than a month, what are the alternatives at that point? He or she can either delete messages arbitrarily, or save email to network file servers and/or a local hard drive. Both of these ad hoc alternatives ultimately inhibit an organization's ability to maximize and benefit from the value of email.

Ultimately, turning email into an asset requires the ability to facilitate an exchange of knowledge between people. A foundation of email knowledge exchange is the ability to tap into the vast store of information captured in email repositories, yet the existence and location of archived email is oftentimes unknown to the workers throughout an organization. Employees need to be able to access from their desktop relevant information created anywhere in the company.

Access to Email Archives is Problematic

Most companies have backup procedures for email, many backup on a daily basis. Yet 83 percent of users cannot access the backup without getting an administrator involved. And the time it takes to actually retrieve a message is so long (an average of about five hours) that many users simply recreate the message (55 percent). Research indicates that it takes email administrators an average of 1.8 hours to recover a message if it is recent (within a month). The time stretches to 11.2 hours if it is over a year old. By that time many users have given up. [25] The result is that valuable information is lost and time is wasted looking for or recreating messages. The retrieval process in many companies is simply too bothersome for most users, or it lacks the timeliness that they require.

Email administrators require an average of 1.8 hours to recover a message if it is recent and 11.2 hours if it is over a year old.

Information Agility

The value of information is directly correlated to how quickly you can access and retrieve it. Simply archiving an email message is only part of the equation. Information is only useful if you can find it when you need it. Companies must have *information agility* in order to effectively react to dynamic changes in their marketplace. Traditionally, change in the marketplace was somewhat predictable; business increased or decreased in a reasonably linear pattern and competitors entered or exited the market in a relatively logical and predictable fashion. Today, however, the economic, technological and societal factors that influence change are moving simultaneously and unpredictably. If firms are not competitive and agile in using the information they have within their enterprise, they will be less able to face these competitive pressures. An email archive strategy ensures that an organization can find, use and keep email information with agility and effectiveness.

While access and agility are vitally important, retrieving the *right* information is essential. No matter how quickly an email can be retrieved from archive, the information is of little value if it is not the information you actually need. Searching corporate email messages by sender, subject or date alone is not enough. For example, suppose you need to retrieve a specific email sent by Martin Mason, sometime in April, which refers to shipments of South American coffee beans. The trouble is that Mr. Mason sent and received nearly 3,000 email messages during that period, and none of the subject lines have anything to do with coffee. The ability search the full content of email messages is essential in order to access and retrieve the right information quickly.

Statutory Retention Regulations

Effective management of email archives is essential in order to comply with statutory regulations and other compulsory policies. Non-compliance of these retention requirements exposes an organization to the risk of fines and penalties, shareholder mistrust and diminished customer confidence.

Laws relating to document retention vary depending on your industry and your location. Some legal retention periods stretch from seven to 15, and even up to 26 years. For example, many manufacturing organizations need to archive their email due to an Office of Fair Trading ruling (many make it common practice to retain email for seven years). In the United States, organizations are scrambling to comply with a variety of new legislation (the Gramm-Leach-Bliley Act and the Health Insurance Portability and Privacy Act are two good examples).

One study of 926 companies shows that while 79 percent of the organizations surveyed were aware of legislative requirements to retain business documents, only 57 percent had a formal policy regarding email retention. Even more surprising, 100 percent of the companies interviewed left it up to the end user to determine which messages are archived and which are deleted. [26]

In short, email represents a significant corporate asset. Unfortunately, the bulk of this information is generally hidden from or unavailable to the whole of the company. Recovering lost email is time consuming for administrators and frustrating for users, and existing email archive strategies often prove inadequate to comply with statutory and regulatory requirements.

Requirements of an Email Archive System

Backup, archive and retrieval of email messages is a more significant problem than is commonly known. Organizations generally have a lack of control over email, and without an effective email archive strategy firms are exposed to these liabilities:

- Overtaxed server storage.
- Rising cost and effort associated with email administration.
- Loss of critical email records.
- Impact to employee productivity.
- Increased risk during litigation.
- Inability to comply with statutory and regulatory retention policies.
- Inaccessibility to corporate knowledge throughout the enterprise.

Given these liabilities, the overriding questions are: What are the requirements to consider when purchasing an email archive solution? What are the benefits that you should expect?

Below is a matrix of features to contemplate when evaluating and selecting an email archive solution for Microsoft Exchange. My aim is to provide you with food for thought in the most important areas to consider. Use this matrix to evaluate software vendors and for internal discussion during your selection process. Focus on the features that are the most important for your organization and look for aspects that make the most sense for your particular situation. Each feature is accompanied by a brief commentary on its importance and what to look for during your evaluation process. Use the matrix to question potential vendors, establish decision criteria and evaluate potential solutions for the best fit.

Requirements Matrix

Requirements	Description

COMPABILITY

Exchange production platform:	MS Exchange 5.5, 2000 or mixed.
Operating systems:	WIN 95 / NT / 2000
Email Clients:	Exchange/ Outlook 97 – 2000 / Outlook Express / Outlook Web Access.
Relational Database Management System:	Oracle, SQL, Sybase, dB2, Ingres, etc.

STORAGE

Open storage architecture:	It is important to be able to write to any type of storage media that is accessible through Windows NT/2000. Confirm that near-line and off-line storage can be written and retrieved successfully.
The type of Media supported:	NAS, magnetic, RAID, optical disk, tape, CD, DVD.
Records/Document Management Application:	You should be able to integrate via your existing RDMA.
Compression to near-line devices:	Compression should be least 50% or higher.
Minimize impact on network resources:	Messages should travel across your network in compressed format (in all directions).
Single Instance Storage of redundant email:	Single instance storage of archived email is important due to the potential redundancy build up in the central archive should your organization be dealing with multiple Exchange servers. It is critical that only ONE instance of your email is archived. Look for this feature from multiple databases, within all multiple data stores, across all corporate servers. Expect zero redundancy in near-line storage repositories.
One object storage:	The entire email message (including the message profile, text, and all attachments) should be treated as one single object. This is important in order for the record to be considered a legal document and admissible as evidence in court proceedings.

OFF-LINE STORAGE

Migration to off-line storage devices:	It is critical to have the ability to move archived data from near-line to off-line storage devices to satisfy retention periods.
Users access to files migrated to off-line:	For search identification, it is important that the system keep a partial message portion on-line.
Storage in open, readable format:	The system should always decompress data when moving to off-line storage. Avoid dependency on proprietary 3rd party software to access archived messages.

RESTORE

Original Folder Structure formats:	Look for the ability to restore email and attachments using the same mailbox folder hierarchy found in the archive.
Processing time:	An important performance benchmark is the ability to restore mail at faster rates than via straight backup tapes.

ARCHIVE CONDITIONS

Management of email from cradle to grave:	The ability to implement comprehensive email retention and disposition policies that can be applied to users and/or groups based on business requirements is critical. Non-adaptation to such strict retention guidelines could expose your organization to the risk of fines and/or penalties under industry or government regulatory compliance.
Custom archival rules and attributes:	Look for the ability to build custom rule formulas from policy-based templates and to select specific email to archive (based on size, age, subject, source, etc.).
Schedule:	Look for the ability to establish single and multiple schedules for specific users and/or groups, as well as the ability to accommodate production archive schedules executing automatically.
Security:	It is imperative that files moved to near- and off-line storage are archived in read-only formats to prevent modification or tempering. This is required also for compliance with SEC, NARA, NA Canada, FOIA and DOD regulations, as well as ARMA guidelines.
Run archive process during business hours:	Another important feature is the ability to run archiving processes during regular production business hours and against accounts while users are currently active on your system.
One way permissions:	You should be able to accommodate management access to employee's email, or for compliance personnel that are required to poll the entire mail of the company.
Permissions automatically set:	It is important to apply the same permissions to the archived email stores that were in place prior to the email being archived.
Synchronize with Exchange address lists:	If you are using MS Exchange, you should be able to synchronize the archive address list with the Exchange address list. This will assist in the management of your users archive environment and in setting up archive groups where user and group settings can interact on a variety of different levels.
Archival fault tolerance:	Look for a fail-safe method to verify that messages have been archived successfully before they are removed from production email servers.
Public Folders:	You should be able to archive information from the Public Information Store, Calendars and Tasks.
Override of archiving rules:	Users should have the ability to override the archive rules.
Server-based Journal:	The system should support the Journalling feature in MS Exchange.
Purging messages:	You should have the ability to purge or migrate data from the near-line archive. This will allow the elimination of data that exceeds retention schedules and give you the capability to remove data that is unwanted.

RETRIEVAL

Retrieval ability:	It is essential to provide end-users with the capability to retrieve archived messages and attachments (with their original attributes) through your existing email client.
Web Browser Client Access:	User access via a Web browser is a beneficial feature.
Outlook Web Access (OWA):	Look for secure retrieval to archive items.
Familiar archive client format:	Users will require less training and assistance with a familiar mailbox structure (i.e., Outlook).
Seamless geographic retrieval:	The system should have seamless and secure access regardless of archive storage location.
Notebook remote retrieval:	If your users travel or use a “virtual office,” your system should provide remote access to email archives from notebooks and laptop computers.
Transparent access by end-user:	It is important that the system provide an image of the archived message to end-users and not require that you restore the actual message.
Re-using retrieved archived messages:	You should be able to "reply to" or "forward" messages restored from the archive that will mean the new message becomes a new email object.

SCALABILITY

Enterprise deployment:	The system should transparently support multiple email servers regardless of their geographic location in the organization.
Dynamic deployment:	The system should give you the ability to add more archive servers into an environment dynamically. This is critical for organizations with exponentially increasing email traffic and where demand for storage space is high.

SUPPORTABILITY

Enterprise end-user mail access:	It is imperative to accommodate the temporary relocation of users in a geographically dispersed server network without the need to migrate their mailboxes.
Internationally disperse network:	Similarly, it is critical to accommodate user access and retrieval to centrally located archives while users are traveling abroad.

CLIENT INSTALLATION

Perform centrally:	Installation should be simple.

CAPACITY PLANNING

Analysis tools:	The system should provide analysis tools that help determine email volume distribution.
Volumetric tools:	The system should provide tools to evaluate storage capacities and utilize “single instance” storage.
Reporting tools:	Look for the ability to generate “what if” reports for decision-making and archive strategy development.

SEARCHING

Within the email client:	It is important for users to have the ability to utilize the native search capabilities on the archive pointers (i.e., via Outlook).
Search archived message and attachments:	End-users must be able to render full text expressions against all archived mail messages, attachments, public folders and mailboxes.
Multiple language support:	You should be able to support multiple language character sets for full text searching of data fields.
Third party search engines:	Another beneficial feature is the ability to interface with Index Servers and Search Engines.

ACCESS POINTERS

Message archive pointers:	Archive pointers are helpful because they leave a portion of the archived mail message on your email server. This way users have more information before electing to retrieve an archived message.
Configuration:	It is important for the access pointer to be configurable in size. This will provide users with the ability to view the message portion of metadata selected or the entire message body left to view.
Single point of access:	The system should be convenient for end-users. Look for the ability to access archives directly from MS Outlook 97-2000, or Outlook Express, Messaging, etc. No additional interfaces or passwords should be required.
Management life cycle:	Must be able to set retention on archive pointers.

PST FILES

Migration tools:	It is critical for the system to provide automated tools to import PST files from desktops and/or shared drives to the archive repositories.
Message archive pointers:	Look for the ability to create message archive pointers for each PST file migrated (archived) to near-line archive stores.

REPORTING

Distribution lists:	Look for the ability to record all the recipients belonging to each distribution list at a specific point in time.

ACCESS CONTROL

Online Synchronization:	It is imperative that the system automatically synchronizes the access rights to archive message files.
Transparence:	Access to the archived mail stores should not require any additional end-user applications. Archived messages should be accessed within existing email clients (i.e., MS Outlook, Outlook Express or messaging, etc.).

SECURITY

No passwords:	The archival process should be convenient and not require additional passwords.
Firewall Protection:	To protect against intruders, it is essential that the archive repositories be blocked from client machines. End-users should not have direct access to the archive.
Virus Protection:	Archive stores should be protected from any virus. Users should be able to purge corrupt messages from the stores.
Archive client:	The system should provide full text search capabilities and use the same security authentication as your existing email system (i.e., MS Exchange).

PERFORMANCE

Multi-processor machines:	The system should operate as a multi-threaded application and capitalize on multi-processor server configurations.
Load balancing:	Look for the ability to configure multiple archive processes.
Access speed to scan a message:	You should be able to benchmark and measure the average number of email messages that can be scanned per second.
Access speed to archive a message:	Establish benchmarks and measure the average number of email messages that can be archived per second.

Questions to Consider

I hope that this dissertation has provided you with a useful overview of the liabilities and risks associated with corporate email as well as a pragmatic look at some of the key features and requirements of an email archive solution. My aim has not been to provide a recommendation for a particular product or solution, but rather to present the salient questions to consider and understand when evaluating a potential product. In closing, consider these questions when determining your next steps.

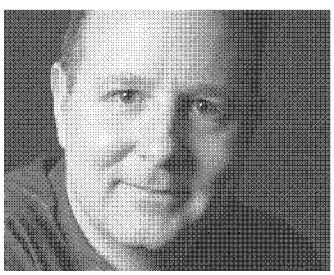
1. Is email growth expanding faster than you have anticipated?
2. Is your company experiencing poor system performance as a result of email overload?
3. Is the backup of your email stores difficult to accomplish?
4. Do backups exceed the available time limit?
5. Are your email users forced to delete messages so that space can be recovered?
6. How many times (instances) is a particular email (also attachments) stored within your organization?
7. How much time do users spend selecting which emails to delete?
8. How much time do system administrators spend recovering lost email messages?
9. Do you have difficulty locating old email?
10. How much space does email utilize on local PCs and other shared network drives?
11. What are you doing for backup or disaster recovery of email?
12. Are you incorporating email as part of your records management policy?
13. Are you aware of the retention requirements associated with your email?
14. Are you aware of the legal implications of email?
15. Is email strategic to your organization? Do you believe that email contains knowledge that is valuable and critical to your company's operations? What happens if a virus brings down your email system?

Conclusion

Electronic messaging systems were designed to provide fast, efficient communications. Email gives us the ability to easily communicate with instant and reliable delivery and universal access. The trouble is that none of the many redeeming facets of electronic messaging tackle the demands associated with the management, archive and retrieval of email messages. As email grows in both volume and complexity, this gap in design presents increased liability and risk to most corporations. As a result, it is crucial that organizations manage the life cycle of corporate email from "cradle to grave." Archive and retrieval of email must be viewed as a strategic imperative and managers must develop a holistic view email of communication.

With the ever-increasing volume and size of email traffic, system administrators must have a better way to manage message stores, mitigate the increasing demand for storage, and muster efficient performance from their servers and networks. Corporate counsel and compliance officers must have a better way of monitoring the content of email messages to both prepare for and avoid litigation. And mining the vast "knowledge value" of email requires the ability to merge formal and informal email communication in a way that facilitates a collaborative knowledge exchange between workers, management and clients.

About the Author:



Kevin Craine, MBA is the author of the book "*Designing a Document Strategy.*" With over 20 years in the information-processing field, working in the Insurance, Aerospace and Pharmaceutical industries, Mr. Craine is regarded as an authority on document strategy design and business process improvement. Mr. Craine received his BA in Organizational Communications and his MBA in the Management of Science and Technology. For more information visit www.document-strategy.com.

Notes:

1. *U.S. News & World Report.*
2. Approximately 89 million, according to Messaging Online. www.messagingonline.com.
3. *Business Week.*
4. *Windows Magazine* .
5. Messaging Online.
6. *Newsweek.*
7. International Data Corporation
8. University of Berkley, School of Information Management and Systems, www.sims.berkeley.edu/how-much-info.
9. Creative Networks, Inc. www.cnilive.com
10. *Inside Gartner Group.*
11. International Data Corporation. *InternetWeek.*
12. See: www.emc.com.
13. Creative Networks.
14. University of Berkley.
15. kVault Software. www.k-vault.com.
16. EducomTS Inc. www.educomts.com.
17. Creative Networks.
18. Persist Technologies Inc. www.persistcorp.com.
19. Ibid.
20. See: www.fiosinc.com.
21. Creative Networks.
22. *Wall Street Journal.*
23. Creative Networks.
24. Ibid.
25. EducomTS Inc.