Document Retention and Information Technology Controls

Designing an E-Mail System in a Post Sarbanes-Oxley Environment

Disclaimer: The author of this document is not an attorney and is not qualified to give legal advice. All summaries of laws and regulations are based upon publicly available sources and are fully referenced. Readers who rely upon the legal conclusions contained herein do so at their own risk. This document was written as a guide for IT managers. Each company should review their specific legal requirements with their corporate counsel, prior to implementing any solutions recommended herein.

Executive Summary

By far, the most important issues facing Information Technology (IT) managers this year are compliance and business continuity/disaster recovery. Public companies are subject to a variety of compliance and regulatory issues, such as Sarbanes-Oxley (SOX). IT must develop processes and solutions to support their company's overall compliance strategy. E-Mail retention is a specific subset of the total compliance obligations a typical public company faces, but the risks involved with non-compliance are not well understood by most. CFOs worried about the financial controls requirements of SOX may fail to focus on the document retention requirements. IT managers have a duty to create and recommend solutions that fit within the framework of the company's strategy. This paper is meant as a summary of the major compliance regulations and as a guideline to designing an effective e-mail strategy.

Included herein is a summary of Sarbanes-Oxley, Graham-Leech-Bliley and other regulations. Every company has unique compliance requirements. Public companies are subject to different laws than private companies. Financial companies and health care organizations may have additional obligations that do not apply to manufacturing companies. This paper is best considered a starting point, rather than a complete solution. The designs herein will provide a solid baseline for evaluating the e-mail compliance strategy of a public company.

Historically, businesses have considered document retention to be a liability. The general rule of thumb was to keep documents the least amount of time possible. Current compliance obligations are requiring companies to re-think this approach. SOX section 802 imposes criminal liabilities on the improper destruction of business documents. The problem that IT faces in light of SOX is that the legislation does not specify precisely what information needs to be retained, nor does it clearly define retention durations, but it does impose criminal penalties if the wrong data is retained, or it is not retained long enough. Because of this paradox, the only course of action IT can reasonably take is to establish policies and technical solutions to retain any document that could fall under the regulation for the maximum foreseeable time period necessary. The penalties provided under SOX make the risks involved with retaining a document lower than the risks involved with destroying it.

This change in attitude presents a unique problem for IT with regard to e-mail. Other documents, such as paper documents and documents on file shares, can be controlled by policy. File shares can be backed up and archived. Delete rights for individual files can be controlled. E-mail is different. Users typically have complete autonomy over message retention. Some messages are deleted immediately, thus they bypass scheduled backups. Other messages are kept too long, going beyond the business need and the compliance requirement for retention. IT must design e-mail systems that enforce company archive and retention policies regardless of user actions.

Traditional tape backup archival solutions fail in this regard, because the user has the opportunity to delete messages prior to archiving occurring. Compliance concerns indicate that a more reliable approach to archiving be implemented. Any message passing through the Mail Server must be intercepted and processed by the archive prior to reaching the user's mailbox. Messages in the archive must be indexed to facilitate search and retrieval, checksumed to ensure message integrity, and written to Write-Once Read-Many (WORM) media. The Archive solution must also enforce document retention policies by marking the earliest possible destruction date for a message, and must allow for document destruction to be suspended in response to litigation or government action.

This paper contains a generic design of a compliance-focused e-mail system, supported by analysis of the users of the system and their compliance-related requirements. The conclusion of this paper includes a 30-point checklist that can be used to measure existing mail systems and evaluate new products. Compliance is a process, not an event. While this paper is a useful guide, it is a snapshot in time. IT managers must stay abreast of changing regulatory requirements, update the compliance checklist, and re-evaluate their systems on at least an annual basis.

Introduction

Corporate Compliance has become the "Year 2000 Problem" of the new century. Many IT vendors, law firms, and auditing companies are using the banner of compliance to sell complex and expensive solutions to problems that may or may not exist. Even so, recent government regulations such as Sarbanes-Oxley and SEC Rule 17a-4 do impose new restrictions and create new risks. In light of this, it is important for IT managers to review the major compliance obligations and develop reasonable strategies for obtaining and maintaining compliance.

One aspect of compliance that is troubling public companies is the requirement to retain e-mail messages as business documents. This paper examines the laws and regulations governing e-mail messages and attempts to translate these rules into technical requirements for IT managers. The goals of this review are:

- Gain a general understanding of the obligations and risks of compliance as they pertain to IT departments, specifically focusing on document retention and IT controls.
- Understand the specific use cases of users, executives, and auditors with regard to e-mail compliance solutions
- Analyze the flow of messages through a compliance-focused e-mail system
- Create a general design for a compliance focused e-mail system
- Design a compliance testing plan
- Give IT managers the necessary tools and resources to test current mail systems, perform a compliance GAP analysis, and review new tools for their usefulness in meeting compliance obligations

The paper draws upon a number of resources to compile its understanding of the current compliance landscape including advisory memos provided by various law firms, recognized industry experts such as the SANS Institute, and vendor seminars on the topic of compliance. Below is a summary of major compliance issues and their applicability to IT.

Sarbanes-Oxley Act of 2002 (SOX)

Passed in 2002 as a response to the Arthur Andersen/Enron scandals, the Sarbanes-Oxley Act (SOX) is a broad sweeping piece of legislation pertaining to the proper controls and storage of corporate information. Although the majority of SOX regulations pertain to controls on financial systems, some pieces of the legislation do pertain to the IT. Specifically, section 404 defines the needs for proper control systems, which has been translated to include IT controls, and sections 802 and 1102 define penalties for the alteration or destruction of documents and other records. SOX carries with it a rapidly approaching deadline. Public companies must be compliant by the beginning of their first fiscal year following December 31, 2004. This means most companies have less than 3 months to meet their SOX obligations (SOX 2002).

Document Retention

Under SOX, "It is now a criminal offense – with a maximum prison sentence of 20 years, substantial fines, or both – to destroy corporate documents 'in contemplation of' a federal 'investigation' or 'administration of any matter.' The investigation or proceeding does not have to be pending or imminent at the time, nor must it be criminal in nature or carried out by a law enforcement agency…Since no federal investigation has to be actually pending or imminent to render destruction of documents illegal, any ad hoc destruction of documents, even if done innocently to 'clean house' may be viewed by prosecutors with a heightened degree of suspicion. The best and perhaps only way for a public company to reduce the heightened risk associated with the ordinary destruction of old and outdated corporate records is to implement a reasonable Document Retention/Destruction Policy that forbids the ad hoc or inconsistent handling of company records, and to educate all of its employees about the policy (O'Melveny & Myers LLP, 2003)."

As pertains to the IT, SOX section 802 and 1102 requirements would apply to file servers, departmental servers, documents on corporate desktops, and potentially all e-mail and instant messaging conversations.

The SOX legislation and subsequent SEC rules are not entirely clear on how these regulations apply to e-mail and instant messaging. However, based upon the advice of O'Melveny & Myers and other industry experts, it appears prudent to include these forms of communication into document retention policies and strategies.

IT Controls

SOX also places a heavy burden on control systems, including IT process controls as these can affect the reliability of financial information. According to section 404 of the Act, a public company's annual report must "contain an internal control report, which shall 1) state the responsibility of management for establishing and maintaining an adequate internal control structure and procedures for financial reporting; and 2) contain an assessment, as of the end of the most recent fiscal year of the issuer, of the effectiveness of the internal control structure and procedures for the issuer for financial reporting (SOX 2002)." Most accounting and audit firms have interpreted this language broadly to include the need to audit IT controls that could affect financial information. This includes security and access controls, backup and recovery procedures, disaster recovery and business continuity strategies, and change control plans. Many audit firms now use the COBIT standards as a guideline for auditing IT control systems.

SOX Conclusions

The problem that IT faces in light of SOX is that the legislation does not specify precisely what information needs to be retained, nor does it clearly define retention durations, but it does impose criminal penalties if the wrong data is retained, or it is not retained long enough. Because of this paradox, the only course of action IT can reasonably take is to establish policies and technical solutions to retain any document that could fall under the regulation for the maximum foreseeable time period necessary. This is a fundamental change to the way companies view document retention and destruction. In the past, retention policies ensured the destruction of information as soon as possible, because it was viewed as potentially damaging. Under SOX, the penalties associated with premature destruction are deemed greater than the damage the information could cause. (Wright, 2003)

Gramm-Leach Bliley Act (GLBA)

The Gramm Leach Bliley Act of 1999 requires financial institutions to disclose to customers their policies and procedures for collecting and sharing information (Morimoto, 2004). "The law requires that financial institutions protect information collected about individuals; it does not apply to information collected in business or commercial activities. A company's obligations under the GLB Act depend on whether the company has consumers or customers who obtain its services. A consumer is an individual who obtains or has obtained a financial product or service from a financial institution for personal, family or household reasons. A customer is a consumer with a continuing relationship with a financial institution. Generally, if the relationship between the financial institution and the individual is significant and/or long-term, the individual is a customer of the institution. For example, a person who gets a mortgage from a lender or hires a broker to get a personal loan is considered a customer of the lender or the broker, while a person who uses a check-cashing service is a consumer of that service (FTC, 2004)."

The GLBA also defines a safeguard rule for the protection of customer data. "Under the Gramm-Leach-Bliley Act, the Safeguards Rule, enforced by the Federal Trade Commission, requires financial institutions to have a security plan to protect the confidentiality and integrity of personal consumer information (FTC, 2004)." Like SOX, the GLBA requires companies to have appropriate IT controls to protect data. This most often translates into requirements for security and access controls.

SEC Rule 17a-3 and 17a-4

In May, 2003, the SEC revised the regulations governing the activities of brokers and trading companies in response to the Enron and Arthur Anderson scandals. SEC Rules 17a-3 and 17a-4 govern the types of documents that must be retained and the retention durations. With regard to e-mail, rule 17a-4 (b) (4) requires the retention of "Originals of all communications received and copies of all communications sent by such member, broker or dealer (including inter-office memoranda and communications) relating to his business…" for a period of at least 3 years (the first 2 years in an easily accessible place). Rule 17a-4 (f) establishes additional requirements for the proper storage of electronic documents such as e-mail:

- □ Preserve the records exclusively in a non-rewriteable, non-erasable format
- □ Verify automatically the quality and accuracy of the storage media recording process
- □ Serialize the original and, if applicable, duplicate units of storage media, and time-date for the required period of retention the information placed on such electronic storage media
- □ Have the capacity to readily download indexes and records preserved on the electronic storage media to any medium acceptable by the Commission (SEC)
- □ Store separately from the original, a duplicate copy of the record stored
- □ Organize and index accurately all information maintained on both original and any duplicate storage media
- □ Each index must be duplicated and the duplicate copies must be stored separately from the original copy of each index
- □ Maintain an audit system providing for accountability regarding inputting of records to the archive (SEC 2003)

Although these rules only apply to brokerage companies governed by the SEC, the guidelines provide useful guidance to the design of e-mail archives.

Electronic Communications Privacy Act of 1986

Commonly referred to as the "Wiretap Act," the ECPA makes it illegal for organizations to monitor communications except when the individual has given consent or the monitoring is being performed to "protect the rights or property of the provider of the service (Morimoto, 2004)." This act could potentially apply to network monitoring, e-mail (SPAM) filtering, and monitoring of voice mail, even though the network is the property of the company. To protect against this, IT needs a monitoring policy and a logon banner that notifies the user that their communications may be monitored or recorded.

Health Insurance Portability Accountability Act (HIPAA)

Also known as the Kennedy-Kassebaum Act, HIPAA was passed in 1986 to standardize health insurance claims, protect privacy of patients, and prevent healthcare records from being published on the Internet (Morimoto 2004). HIPAA generally applies to health industry companies, such as hospitals and insurance companies. It is unclear what aspects of HIPAA may apply to general IT organizations. For instance, some data retained by HR departments regarding health insurance information may fall under HIPAA. The Centers for Medicare and Medicaid Services provides an online resource for determining if an organization is covered by HIPAA at http://www.cms.hhs.gov/hipaa/hipaa2/support/tools/decisionsupport/default.asp. Even if a company does not meet the general HIPAA obligations, it must take extra care to protect employee health information, such as insurance forms that detail social security information and pre-existing conditions.

USA PATRIOT Act - "Providing Appropriate Tools Required to Intercept and Obstruct Terrorism"

Passed in 2001 as a response to the terror attacks on the World Trade Center and the Pentagon, the PATRIOT Act gives the federal government broad powers to gather and share information. According to Wright (2004), "although sometimes these requests will be in the form of mandatory subpoenas, oftentimes the request will be issued informally without subpoenas, perhaps on an emergency basis. In advance of these requests, companies are wise to develop policies on how they will respond, with a view to avoiding the appearance of being unfair (discriminatory) to any employees or customers who might be implicated in the data disclosed."

It is difficult to foresee how the PATRIOT Act might apply to general IT departments. It is possible, for instance, that terrorists might compromise computer systems and use them as a platform for launching a denial of service (DOS) attack. As such, IT needs to establish a policy in advance to cover requests for information by law enforcement agencies.

Discovery Obligations

Although not a specific regulation, IT must also be aware of its obligations to produce information during discovery in the event of a lawsuit or investigation. A common example would be an employee who sues for sexual harassment and during discovery requests all e-mails from or to the employee during a specific period of time. IT would have a duty to produce the requested information and could suffer negative consequences for failing to produce.

In the fictitious sexual harassment example above, IT could be forced to review hundreds of desktop computers for pertinent information. In combination with obligations under other regulations, IT must develop policies and technologies to centrally store, archive and index data that may be subject to litigation to reduce the cost associate with discovery and to prevent the accidental or intentional destruction of data prematurely

Compliance Conclusions

Although none of the laws and regulations listed above specifically site e-mail, proper archival of messages is a requirement. Sarbanes-Oxley considers some e-mails as business documents. Requests for access to e-mail messages can be generated either by government subpoena to assist with investigations, such as federal action under the Patriot Act, or by parties to litigation as discovery requests. As such, these messages must be securely retained and archived.

Historically, businesses have considered document retention to be a liability. The general rule of thumb was to keep documents the least amount of time possible. Current compliance obligations are requiring companies to re-think this approach. The penalties provided under SOX make the risks involved with retaining a document lower than the risks involved with destroying it.

This change in attitude presents a unique problem for IT with regard to e-mail. Other documents, such as paper documents and documents on file shares, can be controlled by policy. File shares can be backed up and archived. Delete rights for individual files can be controlled. E-mail is different. Users typically have complete autonomy over message retention. Some messages are deleted immediately, thus they bypass scheduled backups. Other messages are kept too long, going beyond the business need and the compliance requirement for retention. IT must design e-mail systems that enforce company archive and retention policies regardless of user actions.

E-Mail System Requirements for Compliance

The challenge for IT is clear. Compliance obligations require IT to design an e-mail system that enforces document retention policies in an automated and consistent manner. The problem is that not all e-mail messages meet the definition of business documents that trigger the retention requirement. Many messages represent casual and/or private conversations of the users. These messages clearly do not need to be archived.

IT has two choices. The first is to design a mail system that attempts to distinguish between business messages and personal messages. This could theoretically be accomplished by indexing each message and comparing it to pre-defined key words, message characteristics and heuristics. Messages that exceed a set score would be flagged and archived. This approach is similar to the methods used by SPAM filters to evaluate messages.

Automated scoring is a risky approach for several reasons. First, scoring messages is an imperfect science. Most SPAM filters allow between 1% and 5% of messages through that should be marked as SPAM. They also block between 1% and 5% of messages that should not be blocked. There is no reason to believe that a scoring system to flag messages for retention would be any more accurate. Second, users who did not want their messages archived could intentionally attempt to trick the filter. Third, the penalties for not having messages that should be retained are severe. As such, implementing an automated scoring system is not recommended.

The second method IT can use to meet its obligations is to archive and retain all messages, for the maximum duration indicated under the law. As stated above, this approach assumes it is better to have a message longer than it is needed, than to not have a message that is needed. To enforce this type of strategy, the e-mail system must be designed to intercept all messages and archive them before the end-user can delete them. The design must also allow for messages to be retrieved from the archive in response to audits, compliance reviews and discovery requests.

The following generic e-mail system design has been created to help companies understand the components necessary to archive and retain all e-mail messages with a focus on compliance. The Design is intentionally vendor neutral. The goal of this paper is to provide a guide for companies looking for compliance help. Obviously, some products will better implement the techniques described herein than others. Companies looking to implement this design should use the information provided to evaluate products for their effectiveness.

This paper uses Unified Modeling Language (UML) diagrams to describe system Use Cases. Use Cases describe the interaction of Actors with the System and with each other. Use Cases are often used by software developers to define and describe an application during the development process. This paper applies the principals of Use Case modeling to describe how compliance obligations affect the design of an e-mail system. The process of developing a Use Case model can be summarized in eight steps (Armour & Miller 2001):

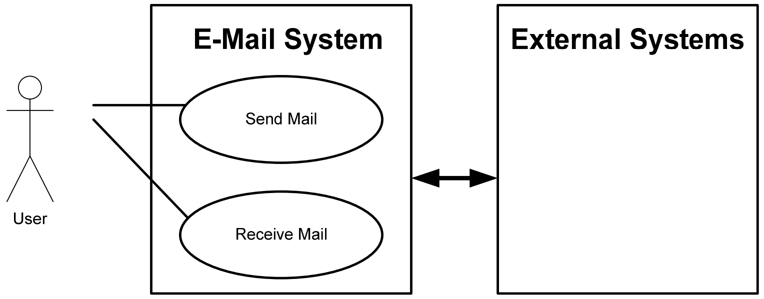
1. Define the system boundaries
2. Find the actors
3. Find the use cases
4. Describe each use case
5. Re-factor the use case model (optional)
6. Prioritize use cases (optional)
7. Add future requirements (optional)
8. Organize the use case model (optional)

These steps reflect the iterative nature of Use Cases during the software development process. This paper uses the first four steps of the Use Case modeling process to define the compliance-focused e-mail system. Implementing the designs below may require additional modeling, including re-factoring the Use Cases according to the unique needs of the company and prioritizing the Use Cases to determine which functions to implement first.

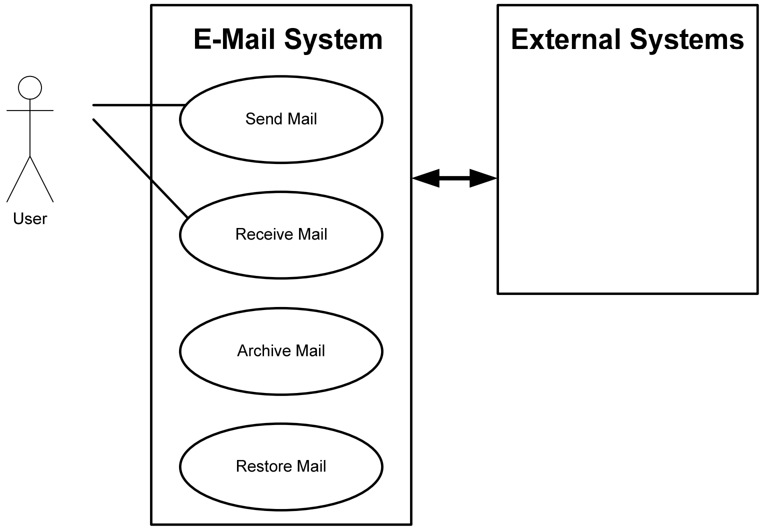
Compliance E-Mail System Boundary

Defining the system boundary is an often overlooked step in Use Case modeling. The system boundary can be thought of as the scope of the project. System boundaries include a name and description of the project, and may include a more thorough analysis of system stakeholders, intended users, and external systems that will interact with the system. On a Use Case diagram, the system boundary is represented by a rectangle surrounding the elements of the system. Users of the system and external systems are shown outside the rectangle (Armour & Miller 2001).

Traditionally, e-mail systems have been defined by simplistic system boundaries. Users send e-mail and receive email. Mail is sent internally to other users in the same system, and to external mail systems. The system boundary of a basic e-mail system can be diagramed as follows:



When defining a system to meet compliance obligations, additional features must be added to the system. Based upon the compliance analysis done above, the system boundary must be expanded to include archiving of e-mail messages and restoration of messages from the mail archive.



For the first step in defining a System Use Case model, this generalized definition and diagram of the system boundary is sufficient. As Actors and Use Cases are found, and described, the System Boundary definition will expand and become more specific.

Actor Specifications

Actors are entities that interact with the system to complete events. Actors can be users, devices, and other systems that interact with the system being described (Amour & Miller 2001). When modeling a System Use Case, the second step is to identify the likely actors and create specifications for those actors. From a compliance standpoint, e-mail systems have four user profiles that can be classified as actors: General Corporate Users, Internal financial controllers such as the CFO, Corporate Counsel, and external Financial Auditors. The CFO user and the Corporate Counsel user profiles are subsets of the General Corporate User profile, with unique requirements that extend beyond the average user.

Actor Specification	
Actor Name: Corporate User	**Abstract:** <No>
Description: Corporate users are every person who has an account on the e-mail server. Corporate users want to send and receive e-mail. They are not interested in compliance. The goal for compliance design is to avoid requiring corporate users to modify their behavior. Messages should be archived and retained independent of any user action. Compliance design will affect users in two ways. First, the ECPA and other laws and regulations regarding wiretapping make it advisable that companies notify users that their e-mail is being monitored and recorded. Logon banners should make it clear that users have no expectation of privacy when using corporate e-mail systems, regardless of message content. Second, users may derive ancillary benefit from investments in archive systems. A mail archive that allows users to search for, and retrieve lost messages can reduce support calls to IT help desks and avoid costly restorations from Backup. Any archive solution should include a self-service mechanism that allows users to retrieve messages from the archive.	

Actor Specification	
Actor Name: CFO	**Abstract:** <No>
Description: SOX requires the CFO, CEO and others to certify their financials on an annual basis. This certification includes assertions that IT controls are appropriate, and carries with it personal criminal liability. In most public companies, the CFO is leading the way towards SOX compliance since this is mostly a financial requirement. CFOs have two compliance needs of a mail system beyond their general mail use. First, the CFO must have confidence in the design and implementation of the system that allows them to certify under criminal penalties that they are compliant. Second, the CFO needs to be able to respond to audit requests. These requests will include compliance reports on the system and the retrieval of samples of messages as audit proof.	

Actor Specification	
Actor Name: Corporate Counsel	**Abstract:** <No>
Description: Corporate Counsel (internal or external) are responsible for responding to government inquires and discovery requests. Beyond their general use of the system, corporate lawyers need a message archival system that allows them to respond to these requests without IT intervention. The goal of the lawyers will be to fully respond to any request without providing any non-required messages. This means that the counsel will require a powerful and flexible search engine that lets them create queries that match the discovery order. Messages must be searchable based upon sender, recipient, date, and complex keywords. When creating responses, lawyers must prove that they have provided all requested information and have not modified any message. The process for creating a discovery response from the archive should include the creation of a single file that contains the query, the search results (messages), and checksums of every message. The lawyers can burn this file to a Write-Once Read-Many medium such as CD/DVD-ROM and provide it along with a diagram of the message archive system. In this scenario, system design documentation is critical. Lawyers must prove that the archive contains every message requested to avoid more costly discovery reviews, such as forensic analysis of user desktops.	

Actor Specification	
Actor Name: Financial Auditors	**Abstract:** <No>
Description: External financial auditors must also certify the financials of public companies. This certification includes a review of IT controls. During IT reviews, the auditors are looking for three things: IT policy, IT designs that enforce the policy, and evidence. Financial auditors will use the system in conjunction with the CFO to run compliance reports and extract sample messages as evidence.	

As the Use Cases are found and described, additional actors may be identified. These additional actors are likely to be external systems outside the e-mail system boundary.

Use Case Scenarios

Use Case Scenarios are used to describe the interactions between Actors and the System. After defining a System Boundary and identifying the Actors, the next step in the Use Case process is to identify and describe Scenarios. These Scenarios will become the guideline for procuring and implementing the system. Products will be evaluated based upon their ability to fulfill the Scenarios.

On a Use Case diagram, actors are shown outside the system boundary, and Scenarios are shown within the boundary. Relationships between Actors and Scenarios are shown as lines and arrows. Actors can be associated with each other, and with Scenarios, and Scenarios can be associated with Actors and other Scenarios. Include Relationships describe a relationship in which one use case uses the behavior of another. Extend Relationships describe a relationship in which one use case adds to the behavior of another (Armour & Miller 2001).

The actor specifications listed above indicate five primary Use Case Scenarios:

1. Send Mail – A user sends an e-mail message through the system.
2. Receive Mail – A user receives an e-mail message from the system.
3. Retrieve Message from Archive – A user receives an archived e-mail message.
4. Audit Archive – A user runs a report on the archive and checks the report against sample messages from the archive.
5. Respond to Discovery Request – A user receives messages from the archive to provide to a third party in response to a subpoena.

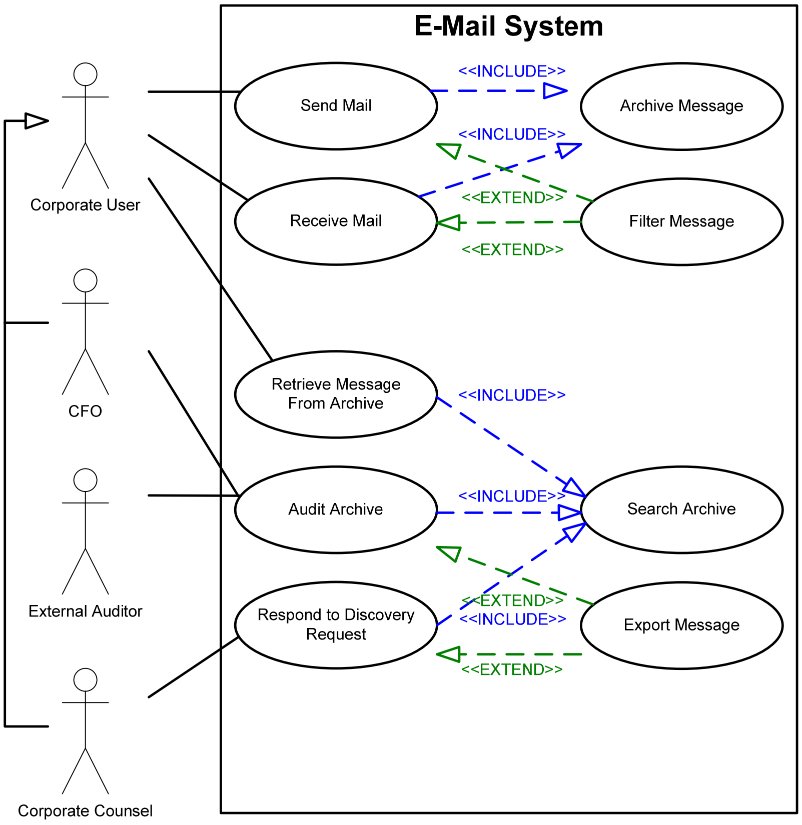
The primary Use Case Scenarios are extended by four secondary scenarios:

1. Archive Message – Send Mail and Receive Mail include automated archival of the message.
2. Filter Message – Send Mail and Receive Mail are extended by SPAM and Anti-Virus filtering of messages.
3. Search Archive – Retrieve Message, Audit Archive, and Respond to Discovery Request include a query of the archive index.
4. Export Message – Audit Archive and Respond to Discovery Request may be extended by exporting of messages and message checksums to a file.

The System Boundary, Actors, and Use Case Scenarios and their relationships are shown below on the Initial Use Case Diagram. Note from the diagram that the CFO and Corporate Counsel inherit the Send, Receive and Retrieve uses as Corporate Users, but have additional uses that are not applicable to other users.

Compliance E-Mail System

Initial Use Case Diagram

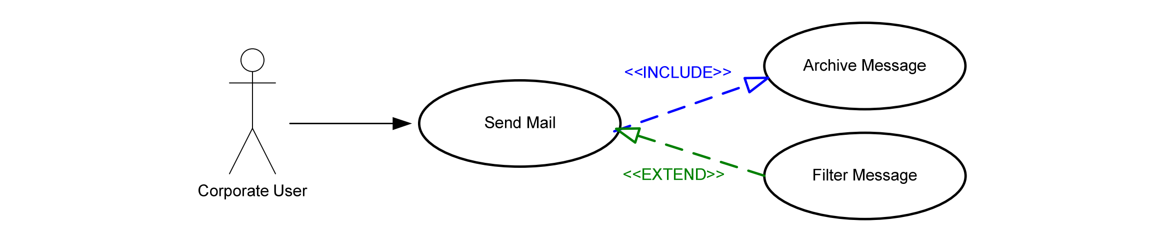


The Use Case Scenarios are further described below using Sequence Diagrams. Sequence Diagrams model the interaction between Actors and System Objects arranged in time sequence (Roff 2003). Time sequencing is an important factor when designing a compliance-focused e-mail system. The order in which events occur significantly impacts whether the system meets or exceeds compliance requirements. The most important time-sequenced event will be the flow of messages through the e-mail system to the archive. To ensure that the archive contains a complete record of e-mail messages, e-mail must be written to the archive before being presented to the user's mail box. Traditional tape backup archival solutions fail in this regard, because the user has the opportunity to delete messages prior to archival occurring.

Compliance concerns indicate that a more reliable approach to archiving be implemented. The Send Mail and Receive Mail Use Case Scenarios below introduce the concept of a Message Interceptor. Any message bound for the Mail Server must be intercepted and processed by the archive prior to reaching the user's mailbox. Although we do not yet know what form the Message Interceptor will take, certain functions of this subsystem can be inferred. First, the Message Interceptor must act as a gateway to the Mail Server in which all messages must pass. Second, like a firewall, the Message Interceptor should implement a "fail-closed" strategy. If the Message Interceptor stops working, all e-mail should be blocked. Third, the Message Interceptor should be transparent to the senders and receivers of e-mail. It should integrate seamlessly with the Mail Server.

The Scenarios described below will show the flow of messages and the function of the Message Interceptor. Once these processes are fully described and understood, the components of the compliance-focused e-mail system can be described.

Send Mail Scenario

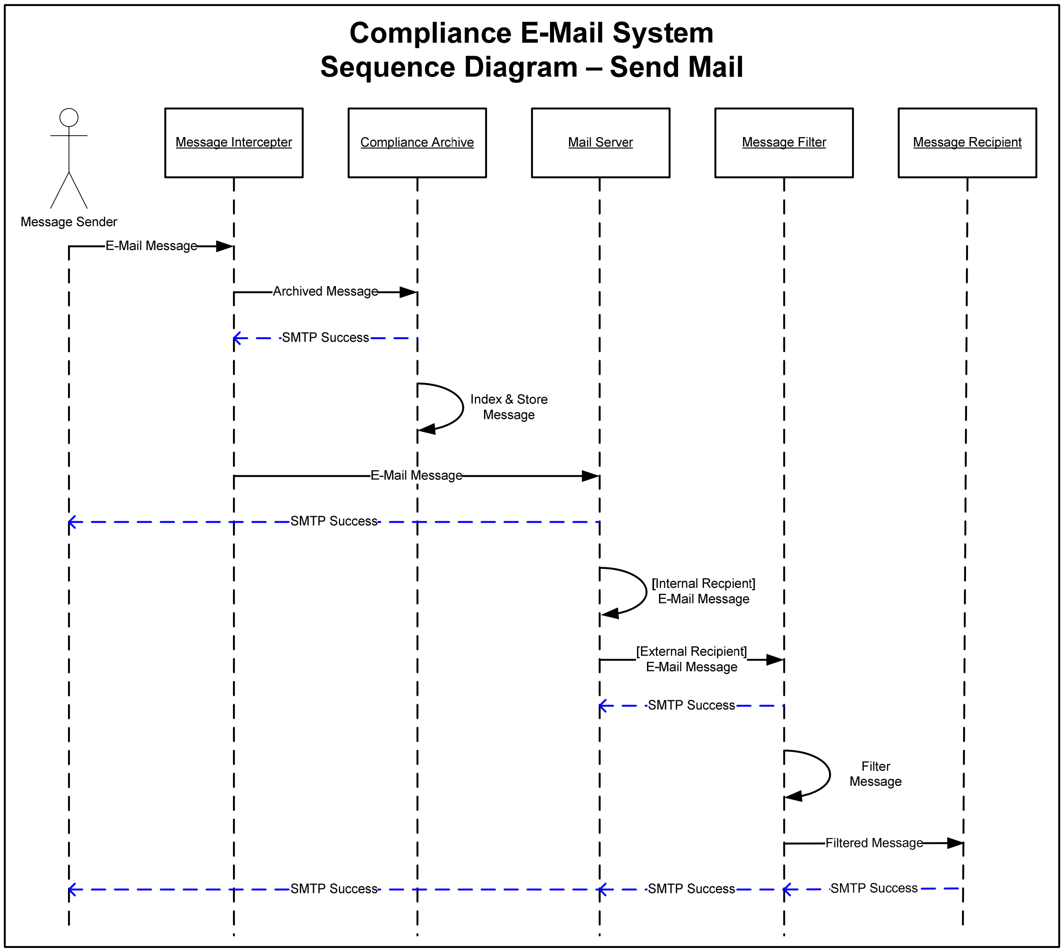


Use Case: Send an e-mail to through the system

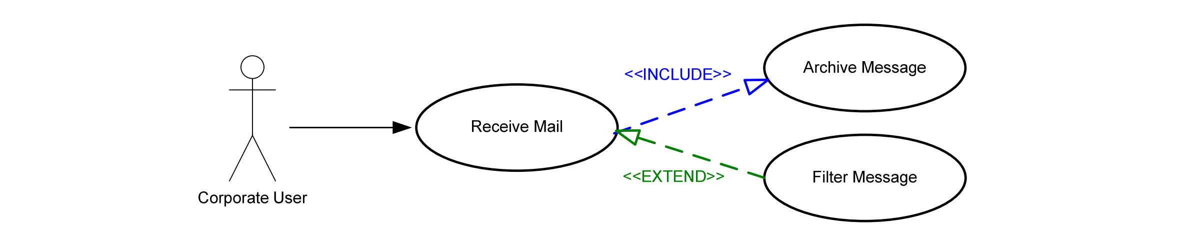
Actor: Corporate User

Description: The user sends an e-mail using their local client. The e-mail client opens an SMTP session to the mail system. The message interceptor receives the message and forwards it to the compliance archive and to the mail server. The compliance archive indexes and stores the message. If the message is bound for an internal recipient, the mail server moves the message to the internal mailbox. If the message is bound for an external recipient, the mail server sends the message to the message filter on the outbound mail relay. The message filter checks the message to determine if it should be blocked or forwarded. If the message passes the message filter, the mail relay forwards it to the message recipient.

The Send Mail Use Case can be diagramed using the following sequence diagram:



Receive Mail Scenario

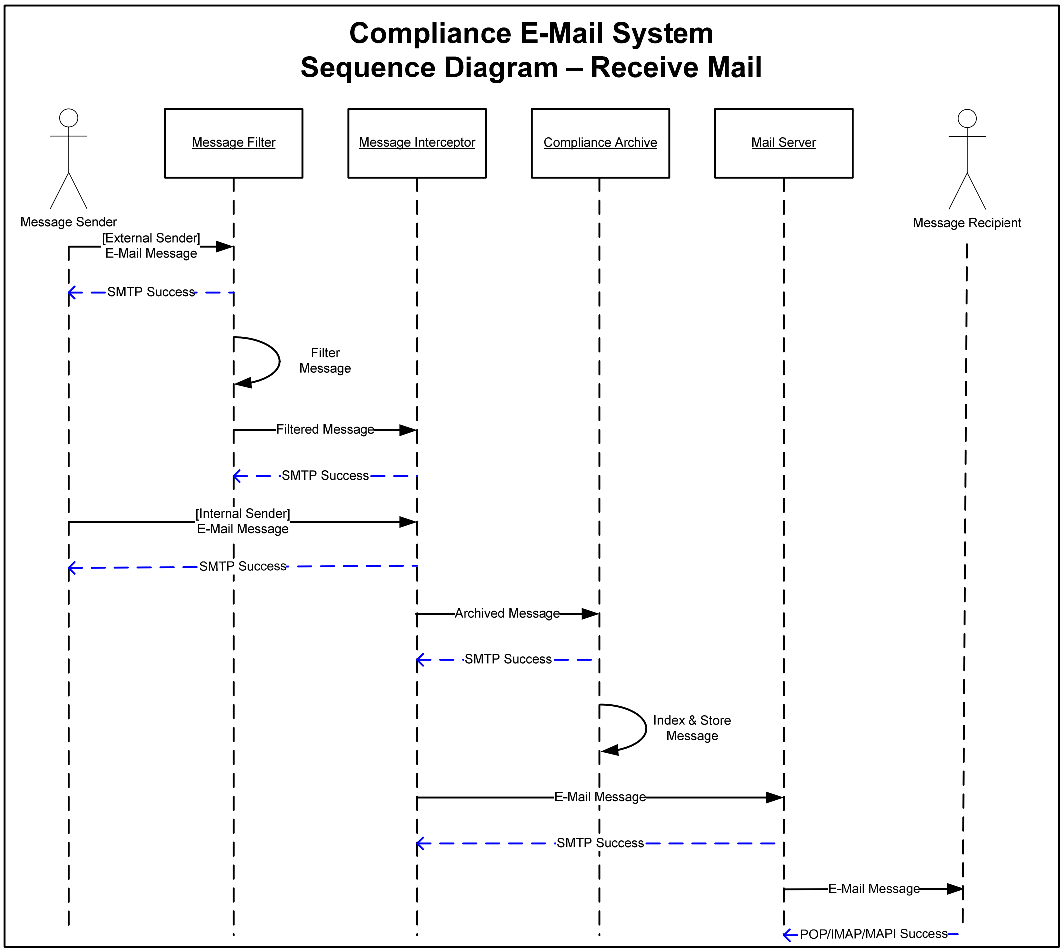


Use Case: Receive an e-mail from the system

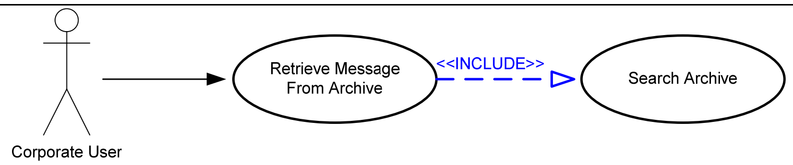
Actor: Corporate User

Description: An e-mail message is sent into the mail system via SMTP. If the message sender is an external sender, the message is received by the message filter process on the in-bound mail relay. The message filter checks the message to determine if it should be blocked or forwarded. If the message passes the message filter, the mail relay forwards it to the message interceptor. The message interceptor receives the message and forwards it to the compliance archive and to the mail server. The compliance archive indexes and stores the message. The mail server stores the message and awaits message retrieval by the recipient. The user connects to the mail server via POP3, IMAP, or MAPI and retrieves the message from the mail server.

The Receive Mail Use Case can be diagramed using the following sequence diagram:



Retrieve Message from Archive Scenario

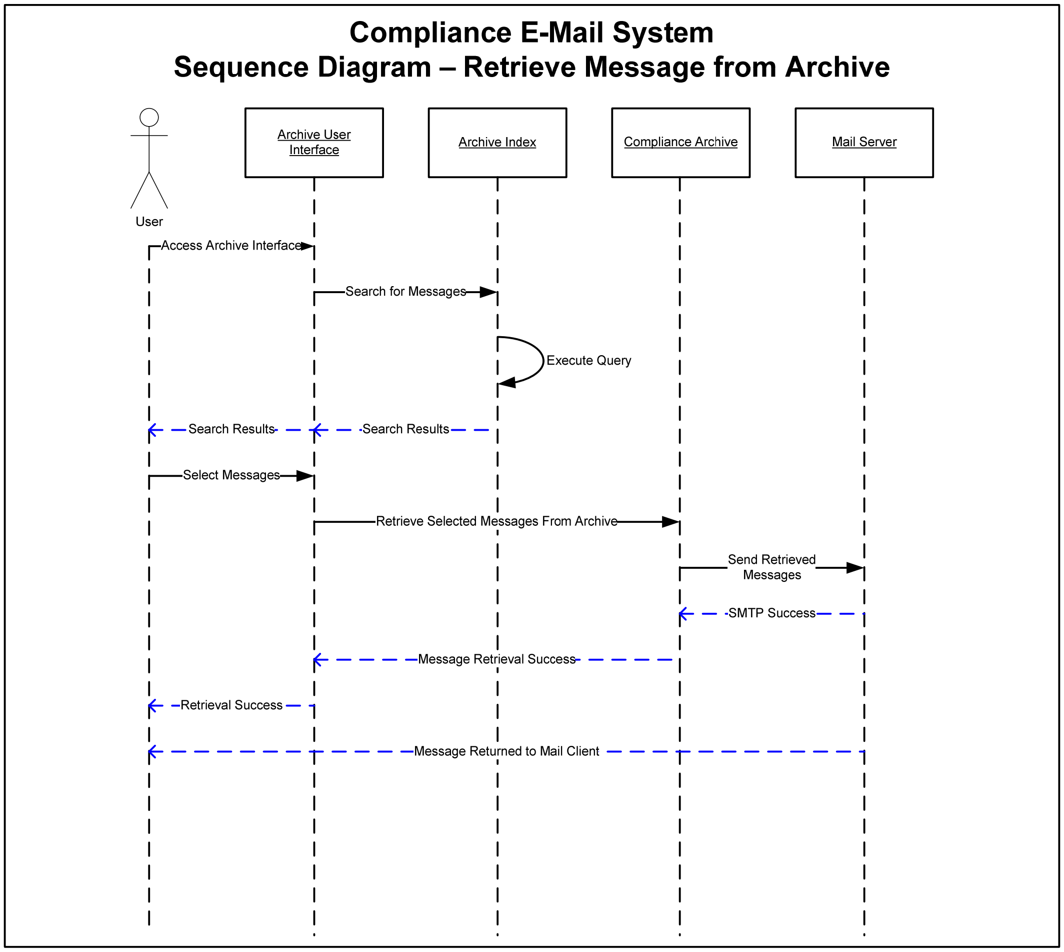


Use Case: Retrieve a Message from the Compliance Archive

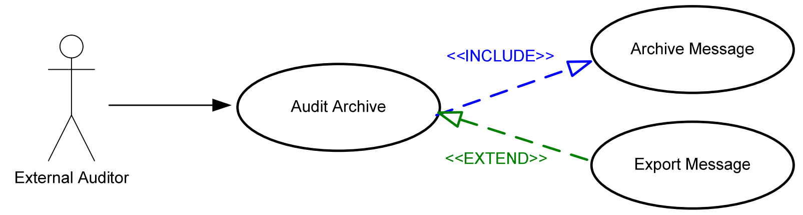
Actor: Corporate User

Description: A user accesses the archive user interface and uses the interface to search for a message or messages from the archive. The archive executes the query against the archive index and returns a list of matching messages to the user interface. The user selects the desired messages from the list to be retrieved. The user interface instructs the archive to forward the messages back to the mail server. The messages are then retrieved by the user's mail client from the mail server.

The Retrieve Message from Archive Use Case can be diagramed using the following sequence diagram:



Audit Archive Scenario

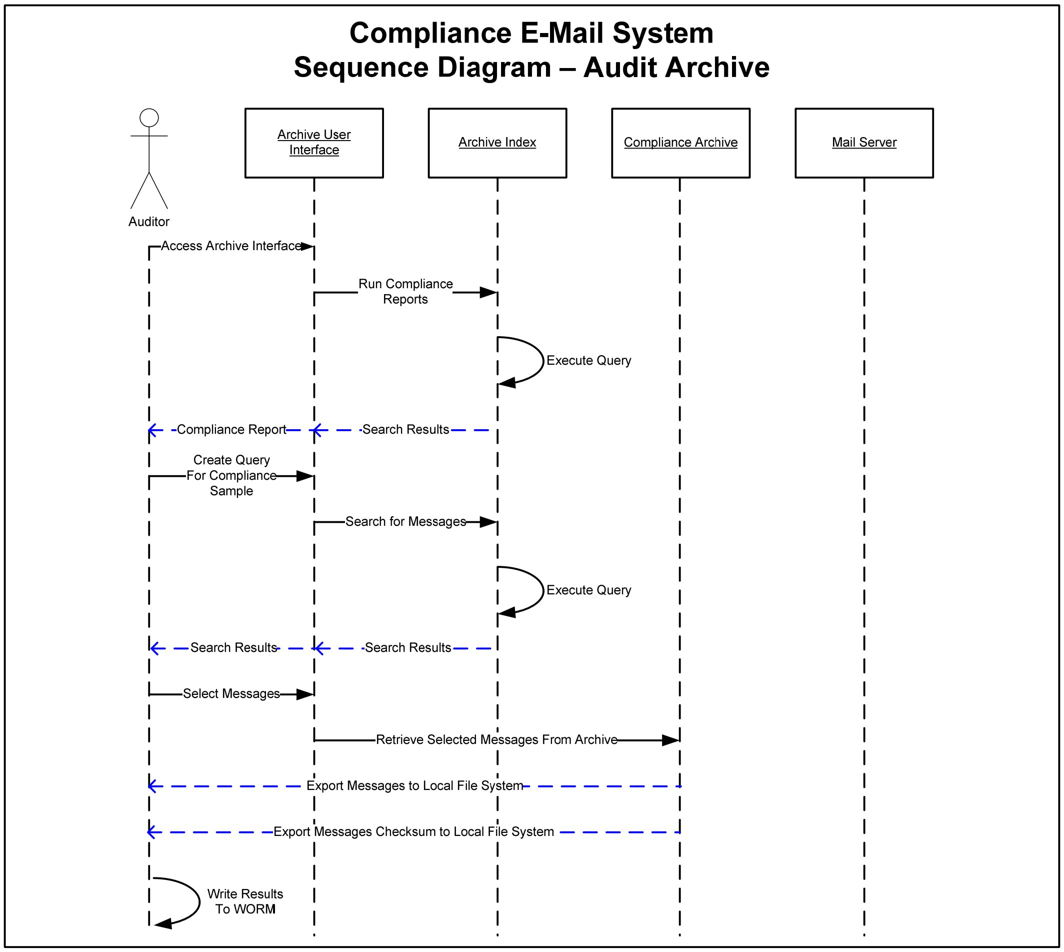


Use Case: Audit Archive for Compliance

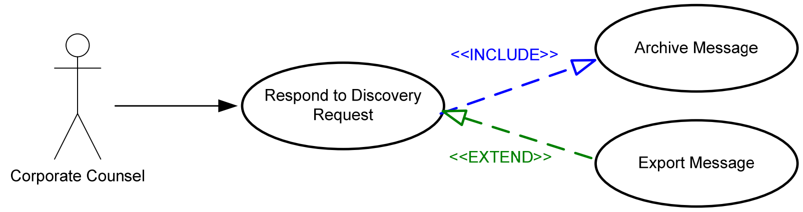
Actor: External Auditor or CFO

Description: The auditor accesses the archive user interface and runs the desired compliance reports. The archive executes the query against the index and returns the report to the user interface. The auditor creates a query for a sample of messages within the archive that will be used as evidence. The archive executes the query against the archive index and returns a list of matching messages to the user interface. The Auditor selects the desired messages from the list to be retrieved. The archive exports a file to the Auditor's file system that includes the query, the resultant messages, and the checksums of the resultant messages. The Auditor records the exported file to Write-Once Read-Many (WORM) media.

The Audit Archive Use Case can be diagramed using the following sequence diagram:


Compliance E-Mail System
Sequence Diagram – Audit Archive

Respond to Discovery Request Scenario

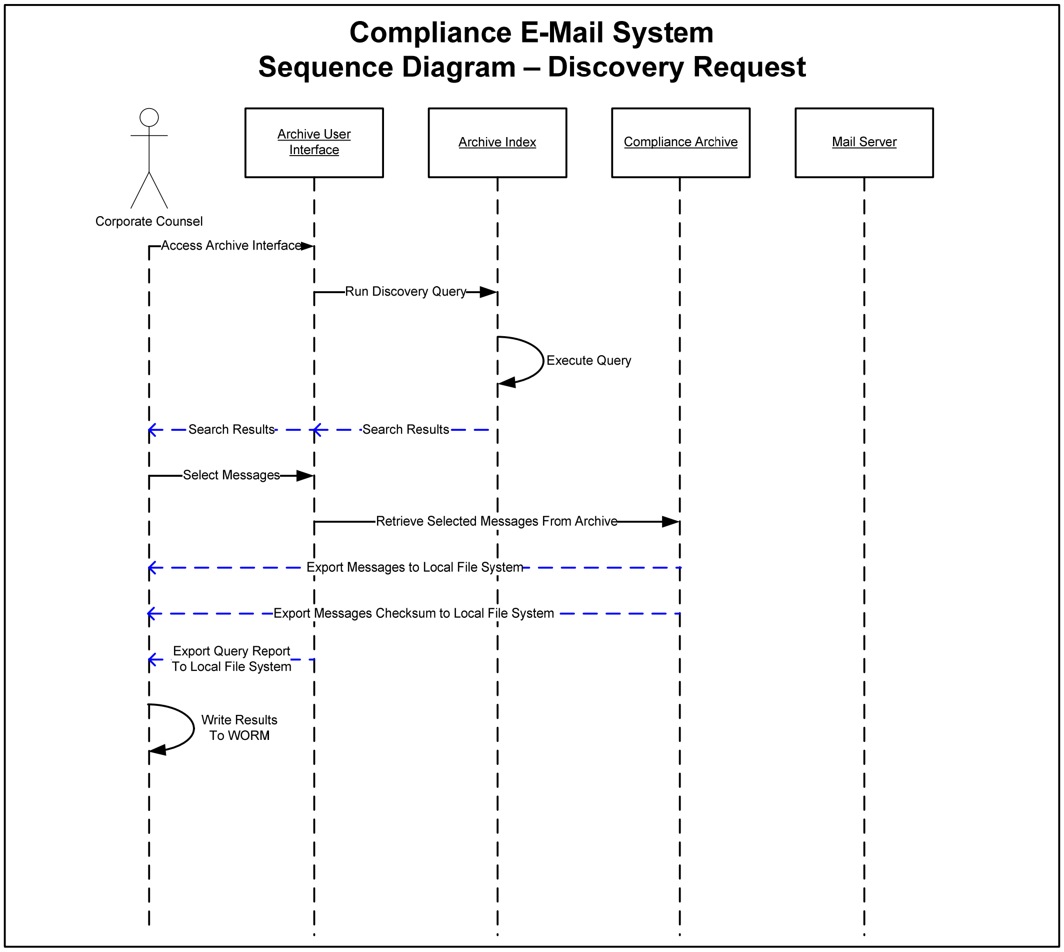


Use Case: Respond to Discovery Request

Actor: Corporate Counsel

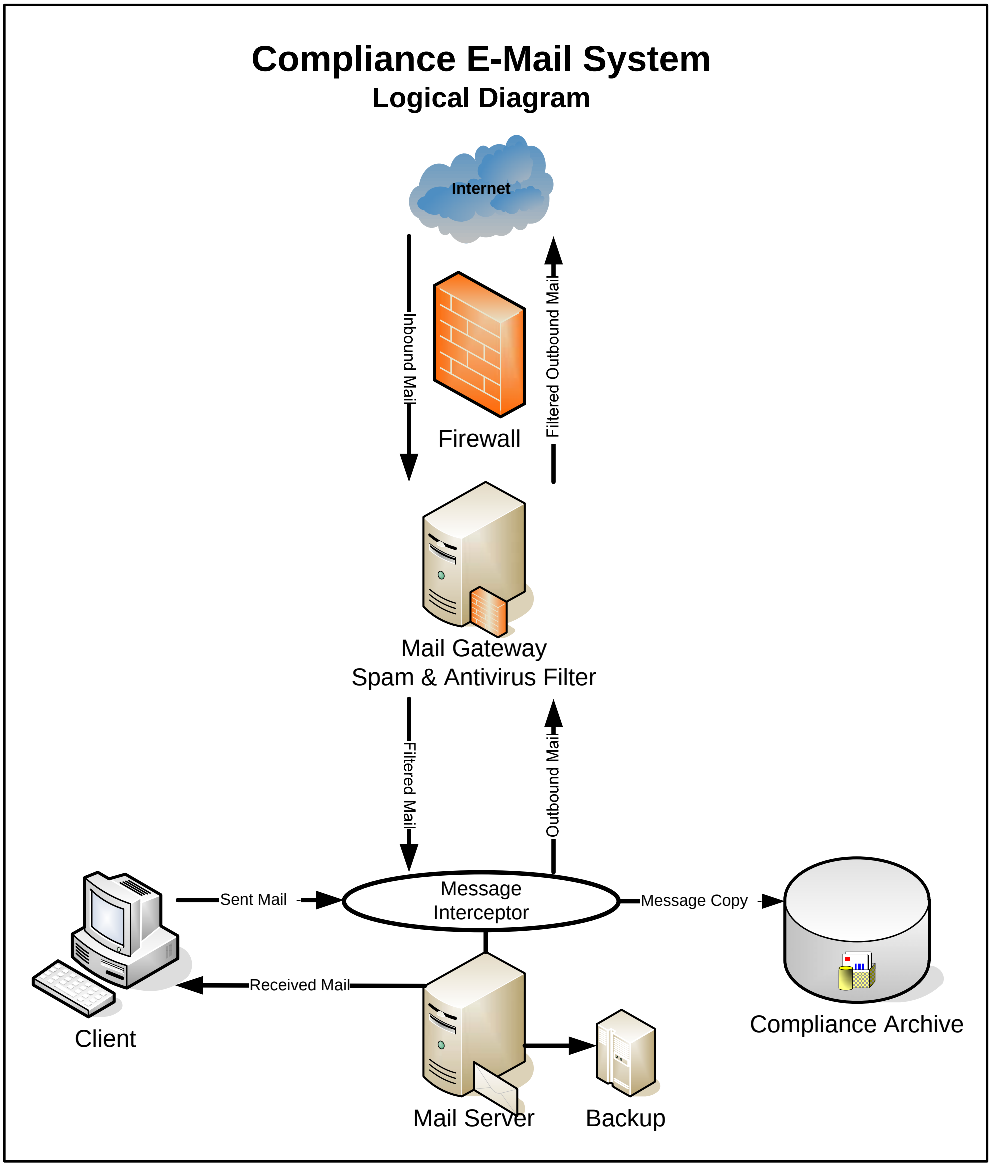
Description: The Corporate Counsel receives a discovery request involving e-mail. The Counsel accesses the archive user interface and builds a query to match the request. The archive executes the query against the index and returns the message list to the user interface. The counsel selects the desired messages from the list to be retrieved. The archive exports a file to the Counsel's file system that includes the query, the resultant messages, and the checksums of the resultant messages. The Counsel records the exported file to Write-Once Read-Many (WORM) media, and provides it to the requestor, along with a certification that the discovery request has been fully answered.

The Respond to Discovery Request Use Case can be diagramed using the following sequence diagram:



System Components

Once the Actors have been identified and the Use Case Scenarios have been explored, it is possible to identify the components necessary to implement a system that matches the System Use Cases. The next step in the design process is to describe the System Components and diagram the logical flow of messages through the system. At this point, the system design begins to resemble network diagrams and system specifications traditionally used by IT. The System Components will form the basis of the compliance checklist, which will be the last step of the design process.



Based upon the above Use Cases, the following system components have been identified:

- □ Boundary Firewall – Any system directly attached to the Internet is vulnerable to attack. Firewalls should be used to protect the mail system from external access.
- □ Message Filter – SPAM and Anti-Virus filters eliminate unwanted message prior to reaching internal mail servers. This is important from a compliance standpoint because blocked messages do not need to be archived. This reduces the size of the archive. Implementing mail filters also helps protect companies from human resources claims with regard to sexual harassment and hostile work environments. The COBIT standards that most auditors are using to evaluate IT controls require SPAM and Anti-Virus protection mechanisms.
- □ Message Interceptor – All messages that flow through the system must be sent to the archive prior to allowing user access. Some mail servers provide a journal mailbox feature that will perform message interception. A copy of every message that traverses the server is written to the journal. If this function is not built into the mail server directly, a third-party solution is needed.
- □ Compliance Archive – Archive solutions must checksum and index every message, prior to storing the message on a WORM media. Checksums protect the archive by providing proof that the message has not been modified. Indexes allow the archive to be queried for specific message attributes. The archive should also manage retention policies by marking media with a DO NOT DESTROY UNTIL flag. When extracting messages from the archive to respond to discovery and compliance requests, the archive solution must guarantee message integrity by writing the query used, the messages, and the message checksums to a single file.
- □ Mail Server – Mail servers should support the archiving process by preventing users from bypassing the message interceptor. Ideally, the mail server should also integrate its search mechanism with the archive index.
- □ Backup & Recovery – Proper Backup and Recovery mechanisms are important to protect the mail system. The archive can be used to recover small groups of messages, but it cannot restore the entire mail database.

System Compliance Checklist

The last step in the design process is to create a System Compliance Checklist. The checklist will be used to evaluate existing systems and new products. The checklist should also be combined with logical, physical, and Use Case diagrams to document the final mail system. This system documentation will be used by auditors and in-house counsel to describe the system as implemented, to show the effectiveness, accuracy, and completeness of the mail archive.

The Compliance Checklist is a list of control measures that test whether the mail system matches the above design. These control measures are general in nature. Companies that face compliance regulations beyond SOX, such as stock brokers (SEC 17a-4), health care providers (HIPAA) and financial institutions (GLBA) may need to add control measures to the list, prior to review. Each measure is listed as a statement of fact. Systems being evaluated are given either a pass or fail for the measure. The failure of any single measure may constitute a failure of the system to match the above design. Any failure should be fully described, including any mitigating factors, using the Compliance Failure Analysis form provided below.

Once the Compliance Checklist has been completed, a system that passes all measures can be described as "compliance-focused." Note that compliance is a process, not an event, so a system can not be judged as complaint or non-compliant. Laws and regulations change frequently, so the compliance checklist should be updated on an annual basis and systems should be reevaluated.

If the system fails the compliance checklist, a GAP analysis should be done to determine the steps necessary to correct the deficiency. The cost to correct can be measured against the risk of non-compliance via a Cost/Benefit analysis. In most cases, the risks involved with non-compliance under SOX will outweigh the costs to correct the failure.

E-Mail System Compliance Checklist

Compliance Control Measure	Pass	Fail
1. Public Internet connections are protected by a boundary firewall.		
2. Public Internet connections to e-mail system are limited to TCP Port 25 SMTP		
3. Outbound connections to TCP Port 25 SMTP are limited to mail servers (Prevents users from using rouge mail relays to bypass archive)		
4. External E-mail is routed to e-mail system through a mail gateway		
5. SMTP relay on mail gateway is limited to routing from external sources to internal domains only (no open mail relays exist on public Internet)		
6. Messages are scanned for Viruses and SPAM on mail gateways before being forwarded to mail server		
7. Messages are scanned for Viruses and SPAM on internal mail servers before being routed (provides protection from internal message sources)		
8. Messages that are flagged as containing SPAM or viruses are cleansed or blocked		
9. Blocked messages are temporarily quarantined and then automatically discarded (allows administrators to recover messages that trigger a false positive)		
10. Virus definitions and SPAM definitions are updated automatically when released by the vendor		
11. All messages that pass through internal mail server are intercepted, journaled, and stored in archive prior to being routed, regardless of source or destination		
12. A checksum of every message is stored by archive		
13. Every message (including attachments) is indexed by archive for easy retrieval		
14. Archive stores messages on Write-Once Read-Many (WORM) media		
15. Archive de-duplicates messages sent to more than one recipient (Message and attachments are stored once to reduce archive size)		
16. Messages are marked with a “do not destroy until” date based upon retention policies		
17. Retention periods on a message, once set, can be extended but never be reduced (reductions in retention periods only apply to new messages)		
18. Message destruction can be suspended to comply with litigation or government action		
19. Archive searches are limited by user-based and role-based security		
20. General users are allowed to search for their messages only, and may restore messages to their mailbox without IT intervention		
21. Managers of users requiring managerial message review (such as stock brokers) can search for messages of their employees without IT intervention		
22. Privileged users such as Auditors and Corporate Counsel can run reports, perform queries, export messages, and suspend destruction policies without intervention from IT (allows for independent review of IT administrators)		
23. Auditors can run compliance reports on performance of archive		
24. Auditors can export archive messages for compliance review		
25. Corporate Counsel can build complex queries against the archive that match discovery requests		
26. Corporate Counsel can export archive messages to respond to discovery requests		
27. Messages are compared to message checksum prior to being exported to ensure message integrity		
28. Exported messages are written to a file, along with a description of the query, and the message checksums		
29. Mail Servers and Archive are backed-up nightly		
30. Backups are stored off-site		

Compliance Failure Analysis

Failed Control Measure	Description of Failure	Mitigating Factors

Conclusions

Although this paper was written using a vendor neutral approach, the capabilities described above are based upon currently shipping products from major vendors. Microsoft Exchange and Lotus Notes/Domino both support message journaling. Veritas Enterprise Vault (formerly KVS) and Hewlett-Packard Reference Information Storage System (RISS) support indexing and archiving messages in WORM media. Anti-Virus and Anti-Spam solutions from Trend Microsystems, Symantec, and Surf Control provide multiple layers of filtering. As a result, all recommendations listed above are technically feasible.

Compliance is an expensive proposition, however. Designing, building and supporting a compliance-focused mail system requires large investments in time, money, and technical resources. Companies struggling with SOX deadlines are focusing most of their efforts on the financial and IT controls requirements. The document retention requirements that are largely being ignored will not go away. IT managers have a duty to explore these regulations with their CFOs and Corporate Counsels.

By developing a long-term strategy and following the recommendations above, IT can begin the process. Steps can be taken on the road to compliance. Even without a full archive solution, message journaling can be enabled on most mail servers. Turning on message journals, and backing the journals up to tape is one of several actions that will improve the position of a company, even if it does not meet the full requirements identified on the checklist.

Compliance is expensive, but the risks of non-compliance are far more costly, and time is running out. IT must identify its obligations, perform a GAP analysis, and begin correcting deficiencies immediately.

References

Allen, J.H. (2001). The CERT Guide to System and Network Security Practices. Boston, MA: Addison-Wesley.

Amour, F.; Miller, G. (2001). Advanced Use Case Modeling. Upper Saddle River, NJ: Addison-Wesley.

Barman, S. (2002). Writing Information Security Policies, (1st ed.). Indianapolis, IL: New Riders Publishing.

Booch, G.; Rumbaugh, J.; Jacobson, I. (1999). The Unified Modeling Language User Guide. Indianapolis, IN: Addison-Wesley.

Brink, T.; Gergle, D.; Wood, S. (2002). Designing Web Sites That Work. Usability For The Web. San Diego, CA: Academic Press.

EPIC (2004). The Graham-Leach-Bliley Act. Retrieved December 18, 2004 from http://www.epic.org/privacy/glba/

Morimoto, R., Ph. D. (2004). Seminar: Applying Government Laws and Regulations in a Windows Environment. San Francisco, CA: Convergent Computing.

O'Melveny & Meyers LLP (2003). Document Retention/Destruction Policies in Response to Sections 802 and 1102 of the Sarbanes-Oxley Act of 2002. Los Angeles, CA: O'Melveny & Meyers LLP.

Pfleeger, S.L. (2001). Software Engineering Theory and Pratice. Upper Saddle River, NJ: Prentice Hall.

Roff, J.T. (2003). UML A Beginner's Guide. Berkeley, CA: McGraw-Hill.

SANS (2004). The SANS Security Policy Project. Retrieved December 5, 2004 from http://www.sans.org/resources/policies/.

Savarino, W.F. (2003) E-Mail Management: An Issue Businesses Cannot Afford to Ignore. Washington, DC: Cohen Mohr LLP.

SEC (2003). Rule 17a-4 -- Records to Be Preserved by Certain Exchange Members, Brokers and Dealers [Effective until May 2, 2003.]. Retrieved March 8, 2005 from http://www.law.uc.edu/CCL/34ActRls/rule17a-4.html

SOX (2002). Sarbanes-Oxley Act of 2002. Retrieved December 10, 2004 from http://news.findlaw.com/hdocs/docs/gwbush/sarbanesoxley072302.pdf.

Wiegers, K.E. (2003). Software Requirements, 2nd Edition. Redmond, WA: Microsoft Press.

Wright, B. (2003). Business law and Computer Security. Achieving Enterprise Objectives through Data Control. Bethesda, MD: SANS Institute.