Architecting Email Storage
For Regulatory Compliance

White Paper

Paul Mayer
Product Manager
June 2003

Table of Contents

Abstract

Essential communication services

Virtually every analyst, business person, doctor, lawyer, teacher, bus driver, grandparent, or teenager would agree that email has taken an increasingly important role in providing essential communication services. As recently as the mid 90s, businesses largely viewed email as an adjunct to their core communications infrastructure. It largely served a novel role as a means of peripheral, interoffice communication. Any attempts to communicate with external parties via email were considered rogue or unnatural.

Email and business processes

Then in the late 90s, something changed. Email communication moved into the mainstream of the global-corporate communication infrastructure. Organizations started incorporating email into fundamental business processes:

- Supply chain communications were redirected to email
- Orders were accepted and confirmed via email
- Contracts were exchanged and negotiated via email
- Marketing campaigns were launched using email
- Lists containing the email addresses of potential buyers became a hot commodity, while faxes were reduced to "hassle" status

Importance of email

Researchers attempt to quantify how meaningful email has become to organizations:

- According to the International Data Corporation (IDC), the number of emails sent each day will grow from 9.7 billion in 2000, to over 35 billion in 2005. In addition, the storage demand from email volume is projected to be 230 petabytes in 2003, for a compound annual growth rate of 300% since 2000. The cause of this growth is mainly due to increasing email volume and email size, including attachments.
- According to the Gartner Group, 60% of business-critical data is now contained within email systems.

Much of the evolving focus on email use around the globe was not fully planned by either organizations or technology providers. Consequently, email-system infrastructures lack the fundamental capabilities to deliver Service Level Agreements (SLAs) surrounding email systems and their use as an essential business tool.

Continued on next page

Abstract, *continued*

Email as a communication tool

What are the effects of this shift to email as a communications tool? This shift has an impact on other forms of communications, and presents challenges in shoring up an organization's infrastructure to manage its email operations. The total impact is challenging to measure, but for starters lets look at some interesting facts:

- According to the US Department of Treasury, the volume of first-class letters has dropped by 5 billion annually.
- The United States Postal Service cut 30,000 jobs between 2001 and 2002.
- According to the United States Postal Service, 1.7 million new email addresses are added each year.
- Data storage managers have found themselves unable to meet SLAs surrounding Recovery Point* Objectives (RPO) and Recovery Time* Objectives (RTO) of email applications and data, due to outdated technologies and procedures.
- Email communications have appeared on the radars of several highly visible regulatory agencies, prompting numerous rules with a far-reaching impact on how organizations capture, manage, audit, and store email communications.
- According to Creative Networks, Inc., (CNI), due to scalability challenges nearly 75% of end users were unable to recover an archived email without assistance from the systems administrator.
- According to CNI, only 29% of organizations would be able to locate an email message that was six months old.

Email systems introduce challenges

It has become evident to most organizations that email has provided a means of communications for both internal and external users, offering unprecedented efficiency with previously unimaginable advantages. Along with all of its tremendous benefits, however, email systems have also introduced various challenges related to storing and managing the data from this growing mission-critical resource.

Audience

This white paper is directed toward organizational leaders who are seeking knowledge about storing and managing the continuous growth in email data, and the ever-increasing demand to have the email system available for internal and external users nearly 24/7/365.

* See terms and definitions at the end of this document.

Introduction

Challenges surrounding email systems

Most organizations are quick to agree that they have various challenges related to managing their email system and its associated data. The bad news is that there is no single solution that would resolve all of these challenges. Email system challenges have common themes, but they are not the same for every organization. These challenges come in a variety of forms, influenced by a number of environmental factors. Addressing these challenges typically requires some combination of people, processes, and technologies.

Email system assessments

Because of the many different challenges surrounding email systems, it is imperative to perform a multi-faceted assessment to determine the full scope of challenges surrounding an email system. This assessment needs to apply a number of different lenses to the data storage and management practices associated with the email system. The assessment should review the current requirements and capabilities of the email system from the perspectives of the following:

- End-users, storage and system administrators
- Business units that have incorporated email into defined business processes
- Human resources and compliance officers *(for content auditing purposes)*

Gap analysis

The assessment should be used as a foundation for a "gap analysis,*" which identifies the shortcomings, and prioritizes the needs to be addressed concerning the email system and its data.

Continued on next page

* See terms and definitions at the end of this document.

Introduction, continued

Email system capabilities and legal compliances

The primary areas that should be addressed when holistically assessing an email system's storage and management capabilities, and legal compliance are shown in the pyramidal framework below, followed by a list of questions that should be thought through to facilitate constructing an effective solution.

- Fundamental recovery – do you have the ability to consistently make reliable backups*, and then restore those backups in cases of data corruption, data tampering, data loss, or an inadvertent deletion occurs?
- Advanced recovery – do you have the ability to perform backup and restore operations at adequate performance levels within defined backup and recovery windows?
- Availability – do you have the ability to meet the necessary uptime requirements to support business operations?
- Archive – do you have the ability to accommodate the data archival requirements of your organization? Are you managing and backing up excessively aging data on costly primary disk storage subsystems?
- Compliance – are you subject to regulatory agencies that have implemented rules governing how your email data must be indexed, stored, and audited? Do your internal records management policies and procedures cover email records?

Continued on next page

* See terms and definitions at the end of this document.

Introduction, continued

How to proceed

The abovementioned pyramidal framework and associated questions can serve as a high-level guide through the process of analyzing and addressing the organization's email data storage and management practices, and its associated infrastructure. However, there will be a natural tendency to jump to the highest level of the pyramidal framework and then assume that by fixing that level, the email system will become more manageable and have fewer problems. As with any construction, it is important to build a solid foundation before adding the upper levels, and the same applies to a well-designed email system—this means addressing each level fully before moving on to the next.

This approach affords one the best possible knowledge of the email system's capabilities at each level, and exposes the data and storage management challenges associated with the email system—revealing the potential technologies and best practices that could be incorporated to build a well-designed and manageable email system.

About this white paper

This discussion is about the responsibilities and challenges surrounding email systems as to data storage and management, legal compliance, and the potential solutions that would resolve these challenges. The intent is to expose the needs that exist at each level of an email system using the pyramidal framework, and then discuss the appropriate technologies and processes that would address those needs.

Fundamental recovery

Demand for fundamental recovery

As logical and obvious as it might seem, organizations still find themselves scratching their heads about how to provide fundamental recovery of their email system's data. Although an organization perhaps has configured backup and restore functionality into the original system design, they now find that the capabilities originally specified for the email system do not meet the current demand for fundamental system recovery. While the problem is generally more complex than this, the general reasons for the inability of the email system to meet current demand are:

- The volume of data has grown by an enormous amount, and the backup infrastructure is severely taxed to meet this additional demand.
- The amount of time designated to perform backups is decreasing due to an increasing, around-the-clock reliance on email services.

Fundamental backup of email systems

Fundamental email system backup is done by the cold or hot backup method:

- A cold backup is performed with database system services turned off; the email system is therefore unavailable to end-users. By shutting down services, all of the pending processes are flushed from system memory, and a good copy of the data can be copied to tape, or to an alternative backup storage medium.

  This approach provides a consistent data copy that can be used to restore the entire system to a particular point in time. Note, however, that this process generally requires staging an entire separate email environment, complete with server, application software, and ample disk space for a complete email system restoration. Even if the goal of this exercise is to recover a single, inadvertently deleted message, this entire process must be undertaken.
- A hot backup is performed by putting the email system on notice that a backup is about to occur. This is done by using the backup software's Application Program Interface (API)* or multiple APIs. The goal of this process is to create a consistent point-in-time copy of the email system's data, to be backed up by the data protection software.

  This approach generally provides some ability to restore portions of the email system or its data more granularly, such as a mailbox or an email message, rather than restoring the entire email system. The primary downside to this approach is that while the email system is being backed up, there is a performance penalty associated with the additional load being placed on the server processors, as well as from the additional activity within the email databases.

  Additionally, email products such as Microsoft Exchange and Lotus Notes were not necessarily designed with 24x7-enterprise mission criticality in mind. Therefore, it is a challenge to achieve reliable backups that will pass the database integrity checks after an email system restoration.

* See terms and definitions at the end of this document.

Advanced recovery

Beyond fundamental recovery

Many organizations are beginning to incorporate email into a greater number of business processes; coupled with the declining costs of bandwidth, experts state that up to 60% of business-critical data is now contained within email systems. Consequently, organizations are increasingly finding that fundamental recovery is not enough to handle this growth. Email data is subject to the same generic pressures that are common to protecting other types of mission-critical data, such as:

- Shrinking backup windows
- RPOs
- RTOs

Beyond those three common challenges, email data presents specific protection challenges that often need to be addressed with special technologies and processes. These specific challenges are:

- Data redundancy
- Complex data structures
- No data-backup window due to 24/7 utilization
- Data consistency during live backups

Data redundancy

Due to the core capabilities of email systems, users regularly disseminate a common set of data files to broad audiences within a single email domain. For instance, it is common for a user to distribute a single document to every employee within a given group or organization. While an email system might have the inherent capability to establish and manage multiple reference pointers to a common file, backup products have traditionally been challenged to offer the same efficiency when backing up email data. This inefficiency leads to the following:

- Lengthy backup times
- Inefficient consumption of tape media and drive resources
- Unwieldy data-restoration procedures

Continued on next page

Advanced recovery, continued

Addressing data redundancy

To address the data redundancy challenge, some vendors have introduced products that have the optional capability of storing a single copy of a redundant file during a backup operation*, thereby decreasing backup time, and using tape resources more efficiently. Vendors have taken various approaches to solve this problem. Some approaches are more optimized for tape efficiency, while others focus more on backup performance enhancements.

Complex data structures

The architectures in which email data are stored vary dramatically from one email utility to another. Some products configure a discrete data file for each user mailbox while others embed an email hierarchy within the structure of a single file. In cases where a data file is shared, such as with Microsoft Exchange, there are particular challenges in performing backups to levels of granularity that are acceptable when you need to restore only a subset of data within the email system.

For instance, if an administrator has performed a backup of an email system at the domain level, and needs only to restore an individual mailbox or message, the restore process will be quite cumbersome and lengthy. In this case, the restoration process generally requires that the administrator configure a separate email server; the next step is to restore the entire set of email data and then sort through the data to find the desired message or mailbox. This is inevitably a time-consuming process.

Addressing complex data structures

Most major backup software vendors have added capabilities to back up leading email products without shutting down email services. By communicating with the email application via an API, these backup products generally acquire knowledge about the structure of the email data. They work in conjunction with the email application to set up the ability to restore subsets of the backed up data in an efficient and intuitive manner. This capability is generally referred to as a "mailbox-level restore," where the backup application has the ability to restore a mailbox for an individual user; or "message-level restore," where the software allows the restoration of an individual message within a user mailbox.

While these more granular approaches to email backup and recovery provide considerable time and resource savings during a restore, this capability does not come without a performance penalty on the front end.

The process of capturing the index information to enable these restoration capabilities can add considerable performance degradation to the backup process. So much so that many organizations having this capability within their email system choose not to use it because of the processing overhead which eats into their already-strained backup window. Therefore, these capabilities may be best coupled with off-host backup capabilities described later in this document.

Continued on next page

* See terms and definitions at the end of this document.

Advanced recovery, continued

Shrinking backup windows

Increasing 24x7 use of email systems eliminates the backup window. Backup used to be relatively easy. The company spent 8 to 10 hours creating data, and the system administrator would have full run of the system until the next morning, to do backup, upgrades, migration tasks, etc. This stretch of time was known as the backup window. Today's production email systems are placing previously unimagined demands on the email infrastructure, leaving little time for backup or restoration of this now mission-critical resource. This pressure exacerbates other challenges surrounding data protection, as it makes it difficult to spend the necessary time to accommodate other requirements.

Addressing shrinking backup windows

Several technologies have been introduced to address the shrinking backup window challenge. These technologies from several vendors exist in both hardware and software forms, spanning many different approaches. The common theme of these technologies is to isolate a point-in-time copy of the email data for recovery purposes:

- To use the data as a backup copy, the administrator would simply roll back to this online copy of the data. This can generally be accomplished in minutes, compared to hours or days for a full system restore of a production email system.
- To use the data as a source to transfer an archived copy of the email data to tape.

Creating point-in-time copies of data

There are two general approaches to creating point-in-time copies of data, and either approach is available in both hardware and software technologies *(although the software capabilities are not available for all operating systems, and the hardware solutions are not available for all storage platforms)*.

1. The first approach is called a split mirror, in which a full duplicate set of data, which was managed as a live RAID mirror, is separated from the production copy and maintained as a static copy.
   - The hardware approach to split mirrors offers host-independent processing of RAID management, which can save a small amount of CPU overhead compared to software RAID systems.
   - The software approach to split mirrors offers an improved ability to flush the write buffers from memory before affecting the mirror separation, thereby minimizing the chance for data corruption on the split mirror.

Continued on next page

Advanced recovery, continued

Creating point-in-time copies of data (continued)

2. The second approach to create a point-in-time copy of data is to generate a pointer-based snapshot of the data. While the split mirror method requires complete duplication of the storage space for each mirror, the pointer-based snapshot approach locks down the blocks in use by the production data at the time of the snapshot, and uses cache to manage any updates to the live data. When the administrator chooses to roll back the data to a previous point-in-time, using a snapshot, the administrator simply activates the snapshot so that the blocks referenced within the snapshot can now be presented to the host.
   - As a hardware-based solution, this capability is generally part of the RAID management tool set provided by the vendor.
   - As a software-based solution, this capability is generally available as part of a journaling file system, or as a feature of virtualization software products.

Data integrity during live *(hot)* backups

The challenge in achieving data integrity during a live *(hot)* backup is due to the active nature of email systems, given the relative immaturity of their internal backup utilities. This means the system administrator must interact with the email system's log files during the backup and recovery process. There is the possibility that some transactions could remain resident in system memory during a backup operation, and therefore, not be written to the backup media.

The restoration process for an email system generally requires the application of log files to bring the database to a current and consistent state. Depending upon what processes were used to create the backup copy, the email system could be challenged to return the data to a consistent state, and if the problems cannot be remedied, the data could be worthless to the email application.

Addressing data integrity

While data integrity continues to be an industry-wide challenge, vendors have introduced technologies to address this problem. By tightly integrating with an email system's API, the data recovery technology can more reliably establish a clean breakpoint, so that the email system can be rolled back safely to some point in time, without experiencing data corruption. This process is currently left almost entirely up to data-recovery solution vendors, but in the future, some of this responsibility will more than likely be assumed by the email-application vendors.

Database vendors, such as Oracle, who faced the same challenges in creating backup copies of live relational database files, offer a successful solution model for this problem. These vendors rose to the challenge by incorporating robust utilities into the database management system, which allowed for sophisticated manipulation of a live database for backup purposes. The database is put into hot-backup logging mode, and the backup operation is then able to produce a consistent and reliable copy of the data consistently.

Email system availability

Availability

As many organizations achieve a solid data recovery infrastructure for their email data, they begin to focus their attention on shoring up the availability of the email application. By definition, availability means that the email system is resilient to interruptions, and offers uninterrupted service in the face of impediments that would otherwise cause system downtime. Just as there is a need for less intrusive backup and faster recovery of email systems, there is also a need for higher levels of availability, which stems from increasing utilization, leading to users relying more on these systems.

A focus on ratcheting up the availability of an email system helps to minimize the impact of potentially disruptive events that could cause planned or unplanned downtime, which would cause the temporary suspension of email services.

Planned and unplanned downtime

There are many events in an email system that could lead to downtime, or application performance degradation, which can be minimized by using high-availability technologies. The following table shows common events that lead to planned and unplanned email system downtime:

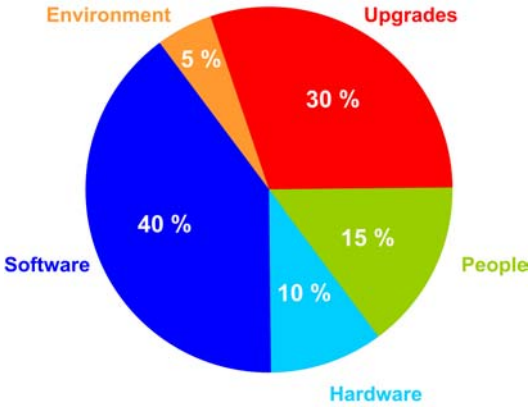
Availability Solutions	Planned Email system Downtime Causes	Unplanned Email System Downtime Causes
Fundamental Availability	Storage subsystem component upgrade	• Disk drive failures • Storage controller failures
Basic High Availability	Network infrastructure component upgrades	• Network infrastructure component failures • Unsustainable spike in network traffic
Advanced High Availability	• Server hardware upgrades • Operating system upgrades • Application version upgrades • General application maintenance • Growth of disk volume for expanding log files or general mailbox storage	• Server hardware failures • Operating system failures • Email application failures • Data corruption • Administrator errors • Viruses • Hacker attacks
Remote High Availability	Off-host/off-site backup	Site disaster

Continued on next page

Email system availability, continued

Causes of system downtime by percentage

According to IEEE computing, causes of system downtime in percentages appear in the pie chart below.



Availability subcategories

Just as with the broad topic of email solutions, the subcategories of availability have their own myriad of technologies that can be applied hierarchically to meet the needs of the organization. The subcategories include the following:

- Fundamental availability: At the foundational level of availability technologies, we find RAID technologies focused on raising the level of fault tolerance within the primary storage infrastructure for the email system. RAID technologies offer a proven, cost-effective means of increasing the reliability of a storage subsystem by reducing single points of failure, and by providing varying levels of protection against hard drive failures. Modern RAID architectures allow tremendous design flexibility to meet the performance, availability, and redundancy requirements of any data center operation.
- Basic high availability: At the basic level of high availability, introductory technologies are designed to create a more robust storage network infrastructure. In a Storage Area Network (SAN) environment, high-availability fabrics result from deploying redundant switches, Host Bus Adapters (HBAs), and cabling. This approach can be taken to eliminate single points of failure, and to allow for continued high throughput in times of high system utilization. In a LAN infrastructure, dedicated high-speed LANs can be configured to isolate potential bottlenecks from backups or Network Attached Storage (NAS) access.
- Advanced high availability: After establishing a solid availability infrastructure, more sophisticated availability technologies can be implemented to deliver greater email system uptime through direct support of email operations. At this layer, technologies such as server clustering and virtual-volume management are available.

Continued on next page

Email system availability, continued

Server clustering technologies

Server clustering technologies allow multiple physical servers to assume the shared identity of a single logical application server. Generally, clustering configurations allow for two types of failover: active/active and active/passive.

- In an active/active configuration, all nodes in the cluster are engaged in active workload processing, and in the event of a downed server *(either planned or unplanned)*, the other servers in the cluster can be configured to assume the tasks assigned to the downed server.
- In an active/passive configuration, one or more servers in the cluster are running idly, waiting for an event that would trigger a transition of processing tasks to one or more of the idle servers.

Server-cluster failover

There are three categories of events that could trigger a failover in a server cluster:

1. Primary server fails to respond: Any number of events can occur that could unexpectedly cause disruption to an organization's email services, including
   - Failed server hardware
   - Crashed operating system
   - Application failure
   - Memory leak
   - Hacker attacks, etc.

   Most clustering software products can be configured to monitor a server by sending repeated "heartbeat checks" over a TCP/IP * connection to see if a server is responsive. Generally, when a server fails to respond to a configurable number of attempts, the tasks that are assigned to that host server are reassigned to an available server in the cluster. Then the failed server can be remedied offline without the pressures of halted production, which could compromise the quality of the fix.

Continued on next page

* See terms and definitions at the end of this document.

Email system availability, continued

Server-cluster failover (continued)

2. Sensors indicate application under strain: Some clustering software products can be configured to audit thresholds within the email application to determine when the application enters the early stages of a problem that may eventually take the application down. By triggering a proactive failover, the email application can be redirected to another server resource, and the problem can be diagnosed and resolved in the downed server without affecting production.

   In some cases, where the problem consistently appears, the corrective action may be automated in such a way that minimal human administrative action is required. This can be used as a way to prop up an email application in environments that stress the application beyond its normal performance capabilities.

3. Planned downtime: As with many applications, email systems are notorious for requiring administrative maintenance, which could require downtime. For example:

   - Email-application upgrades that require a server reboot
   - Server-hardware upgrades
   - Operating-system upgrades
   - Backup-software upgrades on the email server, etc.

These all can cause interruption to email services within an organization. By using clustering technology, many of these disruptions can be averted. Email processing responsibilities can be redirected to another server at a convenient time, and administrators can actually have the luxury of performing system maintenance tasks during regular business hours.

Volume management software

Running out of storage space can have a disastrous impact on an email system. It can cause abrupt termination of the application and complicate the recovery process tremendously. Volume management software virtualizes the storage infrastructure so that administrators can interact with a logical and friendlier representation of the storage rather than the sometimes-daunting physical storage subsystem itself. Feature sets vary from product to product, but the leading value proposition of volume management technologies to email system administrators is the ability to optimize the storage for the email application, and to grow volumes dynamically without interruptions to processing.

Continued on next page

Email system availability, continued

RAID 1

Additional value is realized when the volume management technology allows RAID 1 mirrors to be configured and split from the primary storage for mounting by another server. The data can then be used for off-host backup *(described earlier)*, or for testing or development purposes, where an up-to-date copy of data is desirable. This approach is particularly effective when inexpensive ATA-based or repurposed, older disk storage subsystems are used to store the auxiliary mirror, which will reduce the cost of protection, and minimize the potential for becoming locked into a single storage hardware platform. This capability also leads to the ability to migrate data to new disk storage subsystems as a seamless background task, which helps to ensure data integrity and a smooth transition to the new disk storage subsystems.

Volume management and email applications

When volume management software is integrated with the email application, a powerful synergy can be realized. By communicating with the email system via API, the volume management software can request the I/O buffers for the email system. This level of communication also provides the ability to optimize the storage configuration based upon application configuration parameters, so that the volumes can be ideally configured for the specific requirements for log file and mailbox access, and growth within a given email environment.

Email system resilience

Remote high availability: The combination of higher reliance on email systems as mission critical, and an increased focus on business continuity and Disaster Recovery (DR) has encouraged many organizations to consider investing in leading technologies to provide uninterrupted access to email systems and their data in the event of a disaster. In the course of architecting a DR infrastructure for an email system, an organization must measure three separate variables with regard to its email system:

- The first variable is the RTO for the email application itself. In other words, how much time can you afford to have users not able to communicate via email?
- The second variable is the RTO for the email data; if it is deemed critical that users have as little application downtime as possible, is it equally important that users have access to their previous messages within the same timeframe?
- The third variable is the RPO for the email data. In other words, how many messages can the organization afford to lose in the event of a disaster: one day's worth, one hour's worth, or none?

Continued on next page

Email system availability, continued

Resilience technologies

Resilience technologies that would help design an email environment for resilience in the event of a site disaster include the following:

- Wide Area Network (WAN) or campus server clustering: Based upon local server clustering, these technologies allow one to configure a remote server as a passive, tertiary member of a local server cluster, used only in the event of a site disaster. This approach would allow for fast resumption of email services to users, so that this core communication infrastructure can be preserved, and vital business processes can continue with minimal interruption.

  In campus clustering, server processes are redirected over the production email Local Area Network (LAN) or Metropolitan Area Network (MAN) infrastructure where live data duplication occurs, using data mirroring or data replication technology.

- Remote replication: Data replication technologies provide the ability to create a secondary copy of data at an off-site facility, primarily for disaster recovery purposes. This capability is highly complementary to remote clustering technologies, enabling full relocation of both the email system and its data, with minimal interruption to production. Software-based and hardware-based replication technologies each have strengths and weaknesses:
  - By using hardware-based technology for replication, the data transfer occurs without affecting the application or file servers.
  - By using software-based technology for replication, better integration with the file system and applications occurs, permitting better data consistency, which is particularly critical for database applications. In addition, on modern server architectures, the additional host cycles for software-based replication are generally considered nominal.

Continued on next page

Email system availability, *continued*

Data replication technologies

Data replication technologies are diverse technically, and at a high level fall into two categories: scorecard or write duplication.

1. In scorecard replication, a baseline file system or disk-block bitmap is created and periodically monitored for changes. These changes are then applied to the replica data.
2. In write duplication, disk writes are intercepted by the replication software and applied at the replica site in the same order as they are written to the primary site. Write duplication happens synchronously, asynchronously or in some cases, semi-synchronously, meaning:
   - Synchronous data replication applies each write to the replica system while the primary host waits for verification. In environments with bandwidth constraints, this can have performance implications on production storage.
   - Asynchronous data replication places each write into a queue, so that primary storage performance is not affected by WAN latency.
   - Semi-synchronous data replication allows the system to operate in synchronous mode until certain performance thresholds are reached, and then automatically switches to asynchronous mode.

Figure 1 shows an example of both server and storage replication. In server replication, the replication software is placed on the involved servers. In storage-to-storage replication, the replication software is placed on the involved storage subsystems.

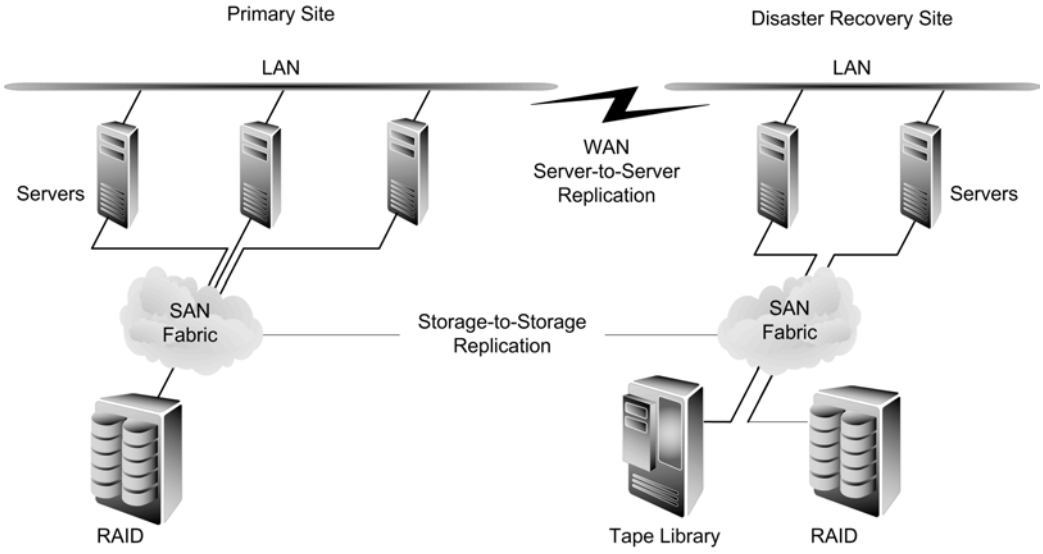


Figure 1: Server and Storage Replication

Email archives

Email growth rates can have an adverse impact on IT organizations

Organizations are increasingly finding that they are not efficiently able to keep pace with the tremendous data growth rates characteristic of email systems. These growth rates have a rippling effect throughout an IT organization, for two major reasons:

1. The initial cost of disk storage to host the data, which is a decreasing part of the Total Cost of Ownership (TCO) due to the decline in disk subsystem prices; however, the management of the storage does not inherently carry the same cost reduction benefit over time. In fact, management of data stored on a primary disk can actually rise on a per megabyte basis as the data ages, given the rising cost of maintaining aging storage hardware and the conversion costs of manually migrating data from a retired storage subsystem to a newer one.
2. The performance of the email system can suffer dramatically when individual mailboxes become bulky with email attachments.

Users as email archivists

To work around the aforementioned problems, email systems offer internal functionality that would allow users the ability to create their own personal archive folders and to set up rules to move older messages to this folder after a certain time period. Many organizations exploit this capability, and administrators often enforce mailbox quotas designed to force users to either delete unneeded messages or move them to a personal archive folder. While this approach provides short-term relief for one problem, it essentially allows users to become the administrator of their own data, which raises several red flags:

- Users are notoriously sporadic at managing their own data. When industry studies indicate that up to 60% of an organization's intellectual property is contained within its email system, a prudent organization would centralize the management of this critical resource.
- The personal archive-folder approach generally allows for a configurable storage location for the archived data. This means that administrators have little control over the location of this data, and as a result, little control over its backup.
- The decentralized approach to data archiving provides a poor foundation for adding auditing capabilities to meet regulatory compliance *(described later),* or simply to adhere to best practices for internal auditing purposes.
- A recent industry study by CNI indicated that as much as 81% of business end users are not able to retrieve email data from personal archives without assistance from help-desk personnel or systems administrators.
- A Gartner report indicated that enterprise email users spend about 90 minutes per day managing their mailboxes.

Continued on next page

Email archives, continued

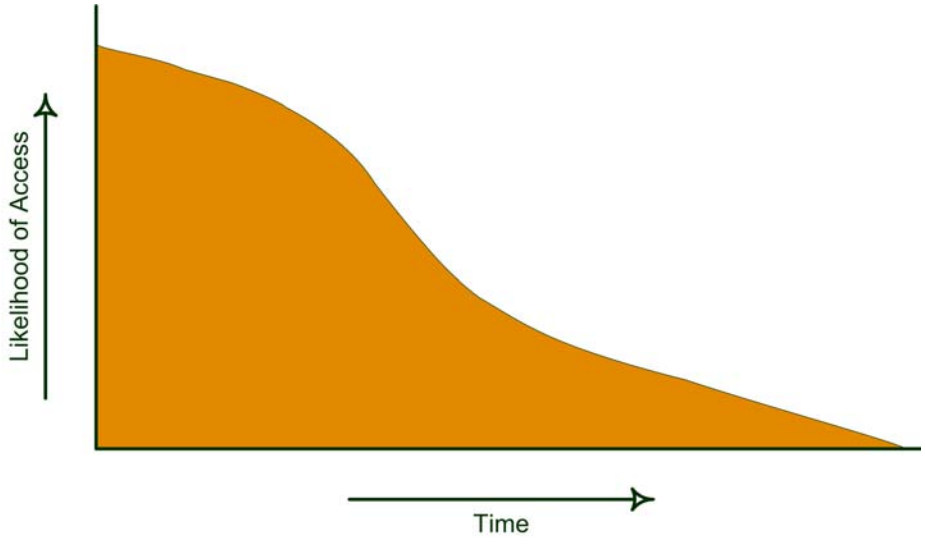
System Administrators as email archivists

An alternative to this approach is to push the manual archiving responsibilities to the system administrator, which addresses the data centralization and protection concerns, but fails to address the scalability and efficiency challenges adequately. Current email applications do not normally provide adequate tools for management. According to CNI, organizations that choose this approach without deploying proper tools to automate this task, find that their system administrators spend an average of five to six hours per week retrieving archived messages for users. This overly burdensome process does not account for the time spent archiving the data originally, and then managing the archive media.

Automated email archiving

To address these challenges, technologies have been introduced to automate the email archive and retrieval process. The technologies are generally built on existing Hierarchical Storage Management (HSM) foundations, which traditionally provide for migration of data on a standard file system such as NTFS* or UFS.*

Traditional HSM technologies provide the ability to manage the life cycle of file system data by automatically migrating it from expensive primary disk storage subsystems to lower cost and performance storage mediums, such as magneto optical, CD, DVD, tape, or Advanced Technology Attachment (ATA) disk subsystems. See the information life-cycle graph below.



Continued on next page

*See terms and definitions at the end of this document.

Email archives, continued

Technologies to automate email archiving

Email archive systems apply the same concepts as file system HSM technologies, with the added benefit of integration with the email system via an API to offer the ability to extract, migrate, and restore data within proprietary email repository architectures.

These technologies can be applied to reduce the amount of time spent significantly by both system administrators and end users on activities related to archiving, managing, and recovering email data—particularly attachments, which represent the bulk of email data. By removing this considerable portion of data *(commonly 70% to 90% of overall data)* from the primary storage environment, administrators can perform backup operations in a fraction of the time previously spent performing this task. Furthermore, this archiving ability provides a solid foundation for most regulatory compliance requirements to which the organization could be subject.

Email data management and regulatory compliance

SEC Rule 17a-4 and the Sarbanes-Oxley Act

It seems as though you cannot read a newspaper these days without reading a story about an organization that failed to manage its email data properly and therefore, find that they face very serious consequences. Two legal compliances are note worthy due to their significance to organizations using email as a business tool: SEC Rule 17a-4* and the Sarbanes-Oxley Act.*

- SEC Rule 17a-4 requires that companies under its jurisdiction maintain accessible, secure business records and have the ability to produce records meeting stipulated audit criteria, quickly. As several organizations have recently found out, the penalties can be severe for non-compliance.
- The Sarbanes-Oxley Act of 2002 takes this accountability to an individual level, and places personal requirements on corporate executives to endorse their organization's financial statements, and to ensure that the organization has the ability to audit its business records, including electronic communications. By failing to meet the requirements of this act, an executive's personal liability could potentially include financial penalties, exclusion from managing public companies, and even imprisonment.

Who sets up the guidelines and laws?

Governing bodies such as the Securities and Exchange Commission (SEC), the National Association of Securities Dealers (NASD), the Department of Defense (DOD), the Federal Food and Drug Administration (FDA), and the Internal Revenue Service (IRS), as well as the federal and state legal systems have established precedents and guidelines, and are enforcing far-reaching laws associated with managing business data. Non-compliance carries serious consequences for a wide-range of public and private organizations.

Continued on next page

* See terms and definitions at the end of this document.

Email data management and regulatory compliance, continued

Guidelines and laws

While these emerging guidelines and laws are precisely defined by each respective governing body, and certainly vary from source to source, several common themes emerge across the rules:

- Electronic business records including email communications are being placed under the same scrutiny as their paper-based counterpart. Failure to manage these records properly and failure to reproduce certain email data upon request could be deemed an obstruction of justice.
- As part of the discovery process in a legal proceeding, a company could be required to produce email data as far back as several years, with limited time to accommodate the request. The email audit requirements could include extensive search and retrieval capabilities, including message text, header information, attachment content, and links.
- Requirements placed on organizations to archive email data systematically as a standard course of business, and in some cases, on a non-alterable storage medium to bolster the evidentiary value of this data.
- Organizations are being increasingly called upon to document data management practices and requirements related to email and to communicate this information to employees so that all stakeholders are on notice as to the liabilities associated with email messages.
- Metadata or descriptive information about the email data itself is increasingly viewed as a key consideration when evaluating the evidentiary value of archived email messages. Practically speaking, email messages need to remain associated with pertinent information, such as date and time, sender, receiver, etc.

Understanding the laws

It is imperative that every organization understands the specific legal requirements that they may be subject to by any of the aforementioned governing bodies. At the same time, they must also examine their internal needs for email data storage and management.

Continued on next page

Email data management and regulatory compliance, continued

Adequate coverage and compliance

Much of the time, many organizations seek to beef up email data management to improve organizational effectiveness rather than to meet compliance requirements. This management effort could be in place to support productivity-enhancement initiatives, or be used to investigate potential employee misconduct. Regardless of the motivation, when an organization feels that it has adequate coverage as to legal compliance because it has archive copies of its email data on backup tapes, that organization could find itself legally exposed. This approach would work until it became necessary to restore a specific archived message by a particular date.

Consider the following hypothetical inquiry: An organization is required to produce every email from the last four years that was sent by anyone in the firm with the words "stock," "guarantee," and "rich" in the text of the message—and that organization is given one week to comply. How many person hours will it take to fulfill this request? Considering costs for that week, how many millions of dollars would have to be spent to accommodate that request? Unfortunately, many organizations are learning the answers to these painful questions in real time.

New technologies and email data management

The imperative for better email data management has driven the development of new technologies to support this mission. These emerging technologies are specifically designed for the dynamic nature of email systems that contain a combination of structured and unstructured data.

These email management technologies can be components of a broader set of systems geared toward management of all types of business data, including scanned images, paper, microfilm, microfiche, and host-generated ASCII reports, etc., or they could be dedicated strictly to the management of email data. In any case, the functionality offered by these new technologies is generally focused on meeting the applicable requirements of one or more of the aforementioned governing bodies. In addition, enforcement of best practices often drives added operating efficiencies.

Continued on next page

Email data management and regulatory compliance, *continued*

How these new technologies function

Whether targeted for regulatory compliance, judicial discovery, or best practices, the functionality offered by these technologies generally falls into the following areas:

- Metadata extraction – the ability to automatically index messages as they enter the system, using such key attributes as sender, date, recipients, etc. In addition, some technologies allow for full text indexing of the message text, as well as any files attached to the message.
- Search capabilities – the ability to query a records-management database to locate messages containing specific attributes. This query is generally performed outside the email system, due to the limited capabilities of email systems to perform queries simultaneously on active data along with archive data, or across multiple mailboxes in a single search.
- Migration – the ability to migrate email data, attachments, or attachments along with their associated message to a secondary or tertiary storage medium. This is used as a foundation for archive capabilities, as well as message life-cycle management.
- Message life-cycle management – this capability allows users to establish a set of policies that determine the data management rules that apply to each category of email data. As a possibility, email messages pertaining to anything related to taxes can be placed on a seven-year retention schedule, with requirements for fast access for a period of six months. After six months, the likelihood of retrieval declines considerably, and after seven years, the records are systematically purged.

  Another possibility would be that email deemed to be SPAM is not placed under the system's management capability, but left to the discretion of each employee to delete.

  To determine which life-cycle management profile applies to a given message entering the system, a set of criteria is established for each profile that categorizes the message based upon a set of key words, or a lexicon. Once a message is identified as business data, it is managed according to its life-cycle requirement policies, including storage migration, archival, and removal of data from the system as well.
- Security – Prudent email data management includes secured access to sensitive data. Not only should a system offer the ability to reserve access to certain data for those with necessary security clearance, it should also offer the ability to track which users have attempted to access documents to which they have not been granted access. A preferred security model allows a configurable hierarchy to be applied so that users can be placed into groups for access to some categories of data, while other data may be associated with specific PIN numbers due to its sensitivity.

Continued on next page

Email data management and regulatory compliance, continued

Technology and compliance

It is important to note that while these email-system management technologies offer tremendous capabilities and can serve as a platform on which to build compliant email-data management practices; however, in and of themselves, they do not produce compliance to any legal requirement or best practice. These technologies only serve to augment and to an extent, automate the formal email-data management policies and procedures that an organization chooses to implement.

Summary

Significance of email

As mentioned earlier, IDC predicts that email volume will grow to 230 petabytes in 2003, for a compound annual growth rate of 300% since 2000, which means that businesses are increasingly realizing and relying on the many business benefits of email systems. This growth can be attributed to increasing email volume and email message size, including attachments. It is clear that email has become an essential tool for internal and external business processes and communications at all levels of the organization; yet most email systems are not designed for the increasingly rigorous demands being placed on them by today's organizations.

Multi-faceted assessments

From requirements mandating that email systems be fully available 24/7/365, to the expectation of being able to track and reproduce any message, including attachments sent, at any time in history, to or from any person in or outside the organization is clearly asking much of today's email systems. Because of the many different challenges surrounding email systems, it is imperative to perform a multi-faceted assessment to determine the full scope of challenges surrounding the email system, and to identify the technologies and processes necessary to resolve these challenges.

Assessment outcomes

While evaluating options for improving efficiency or adding strategic value to the email infrastructure, it is prudent to keep in mind a hierarchy of needs that would apply to most organizations. By assessing the organization's needs and current capabilities at each level of the pyramidal framework discussed earlier, an organization can gain insight into its actual business needs, along with the required SLAs to set expectations surrounding the performance of their email system in the following areas:

- RPOs
- RTOs
- Backup windows
- Uptime needs
- Records management requirements
- Legal compliances

Using this approach, the process of optimizing the resilience and manageability of an email system should become less formidable.

Continued on next page

Summary, *continued*

Partnership with Datalink

When time and perhaps internal skill sets are the concerns when addressing challenges such as those related to email systems, a partnership with Datalink would be a wise choice.

As a leading independent information storage architect, Datalink is sensitive to the objectives of those having the responsibility to pursue storage solutions that optimize performance and functionality, and offer cost efficiency as well. Datalink's engineering practices, services, and relationships with key vendor partners keep us abreast of the rapidly changing trends in the storage industry.

We work objectively with our customers to plan and implement customized storage infrastructures that store, protect, and provide continuous access to data. Given that, Datalink is positioned to facilitate the decision-making processes of those IT professionals tasked with the responsibility to manage their growing data assets, as well as their storage infrastructures. From assessment to implementation and support, Datalink has the technical knowledge and resources to ensure that your existing storage infrastructure, or a brand new one, can meet your current business needs and those of tomorrow. A partnership with Datalink affords our customers expertise in all of those areas.

Datalink Corporation
8170 Upland Circle
Chanhassen, MN 55317
1.800.448.6314
www.datalink.com

Terms and definitions

Term	Definition
API	API, an application program *interface (API - and on occasion spelled application programming interface)* is the specific method prescribed by a computer operating system or by an application program by which a programmer writing an application program can make requests of the operating system or another application.
Backup	A secondary copy of data that is generally used for restoration in the event of damage to the primary copy of data.
Backup Operation	The process of preparing and copying selected data from primary storage to secondary storage.
Email	Email, *(Electronic Mail)* is the exchange of computer-stored messages by telecommunication. Email messages are generally encoded in ASCII text. However, you can also send non-text files, such as graphic images and sound files, as attachments sent in binary streams. Email was one of the first uses of the Internet and is still the most popular use. A large percentage of the total traffic over the Internet is email. Email can also be exchanged between online service-provider users and in networks other than the Internet, both public and private.
Gap Analysis	In information technology, a gap analysis is the study of the differences between two different information systems or applications, often for determining how to get from one state to a new state. A gap is sometimes spoken of as "the space between where we are and where we want to be." Gap analysis is undertaken as a means of bridging that space.
NTFS	NTFS *(NT File System; occasionally called New Technology File System)* is the file system that the Windows NT operating system uses for storing and retrieving files on a hard disk.
Recovery Point	The point in time at which data is restored; for example, if a backup is performed at midnight, and that data is used to restore a system at noon the next day, the recovery point is midnight. This is a measurement of how much data is lost between the last backup and the time of data loss or corruption.
Recovery Time	The time that it takes to restore a system using backup data, when the primary data has been lost or becomes corrupt
Sarbanes-Oxley Act	Sarbanes-Oxley Act, established in 2002, to protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws and for other purposes.

Continued on next page

Terms and definitions, *continued*

Term	Definition
SEC Rule 17a-4	This rule promulgated under the Securities Exchange Commission Act of 1934, regulates the methods by which subjected organizations manage and retain their business records.
TCP/IP	TCP/*IP (Transmission Control Protocol/Internet Protocol)* is the basic communication language or protocol of the Internet. It can also be used as a communications protocol in a private network *(either an intranet or an extranet)*. When you are set up with direct access to the Internet, your computer is provided with a copy of the TCP/IP program just as every other computer that you may send messages to or get information from, also has a copy of TCP/IP.
UFS	UFS, *(Unix File System)*, is the file system that the UNIX operating system uses for storing and retrieving files on a hard disk. Every item in a UNIX file system can be defined as belonging to one of four possible types: • Ordinary files • Directories • Special files • Links