

ZL UNIFIED ARCHIVE

Unified Email Archive

ZL TECHNOLOGIES | White Paper



ZL TECHNOLOGIES INC.

Evaluation and Selection of an Archival Solution

In December 2002, the SEC got serious, corralled some big banks, and collectively fined them \$8.25 million. Their offense? They didn't have an email archival system that the SEC deemed adequate. If you think you're exempt because you're not a bank, think again. Recent legislation such as the Patriot Act, HIPAA, Basel II, and the Sarbanes-Oxley Act reaches farther than ever to affect the great majority of corporations and make compliance a major issue.

Performance is another issue. Existing email servers were never designed for today's volumes. Attachments are a big part of the problem. As much as 85% of all email data is due to attachments, according to The Radicati Group, Inc. As the email servers get bogged down by sheer volume, adding more email servers is expensive and not very effective.

And then there's that email-content elephant in the room, ignored by all. Who wants to admit that the bulk of corporate intelligence lies buried in email and attachments, but no one has figured out a cost-effective way to store, access, and search them? Instead, we spend as much as 20% of our time searching through our email and files, according to a Gartner Group estimate.

So, how does email archival go about addressing these diverse challenges? As a first step, the word "archival" should be, well, archived. Email archival today is a lot more than just spooling data onto tape, throwing it into a closet and retrieving it when necessary. Users now want archival to address multiple challenges, principally:

- 1) Compliance with government regulations
- 2) Storage offloading from overworked mail servers
- 3) Legal Discovery
- 4) Email content management

Partial solutions exist today, usually in some limited, stand-alone format. However, they fall short of an integrated approach, namely, a unified archive.

ZL Unified Archive

The new archival model integrates the four most-needed functions onto one platform:

- 1) **Compliance.** New regulations continue to lengthen the period of retention, expand the scope covered, and shorten the required time of data access, search and retrieval. The archival solution should not only address your compliance requirements today, but also be scalable and flexible to change.
- 2) **Mailbox and Attachment Management.** To lighten the data load on email servers such as Exchange and Notes, data-heavy email with attachments can be separated from the main message and replaced with a link. This is known as “stubbing”. Stubbing can be done for entire emails, partial emails or just attachments. When the recipient opens the email, he clicks on the link which points to the actual data stored in the archive, which is then downloaded. The effect is to dramatically reduce the volume of data stored and handled by the email servers, resulting in faster performance, lower costs and greater content control.
- 3) **Legal Discovery.** Organizations are able to investigate message and content archives, and provide a workflow solution around the litigation support process. Efficient and high volume document production capabilities make finding, retrieving, reviewing, refining, managing and document production activities simple, immediate and user-friendly.
- 4) **Content Management.** Enterprises now find that as much as 60%-80% of its corporate memory and intelligence are buried in email and attachments. Unlocking the value of this vast resource requires archival software which can organize huge stores of unstructured content and enable fast, on-demand searches. Also, access, search and restore privileges should be available, not just for the administrator, but for all employees, with built-in safeguards restricting access per company policy.

The archive should unify all four functions in a single system, and you should be able to turn on the functionalities you need.

So, Which Archival Solution Meets My Needs Best?
Below is a list of the four common objectives, including a summary of those capabilities most

ZL Unified Archive

Storage	Discovery	Compliance	Security
<ul style="list-style-type: none"> ▪ Exchange Optimization ▪ Lotus Optimization ▪ File System Optimization 	<ul style="list-style-type: none"> ▪ Case Management and Litigation Support ▪ Enterprise Search 	<ul style="list-style-type: none"> ▪ Retention Management ▪ Pre-Review ▪ Post-Review 	<ul style="list-style-type: none"> ▪ Email Security ▪ Secure File Collaboration ▪ Secure Webmail
Search	Categorization	Policy Mgmt	Access Methods
<ul style="list-style-type: none"> ▪ Global Search ▪ Precision Search ▪ Incorruptible Indexes 	<ul style="list-style-type: none"> ▪ Automatic Categorization ▪ Lexicon and Taxonomy Management 	<ul style="list-style-type: none"> ▪ Archiving ▪ Stubbing ▪ Retention ▪ Privacy ▪ Restoration 	<ul style="list-style-type: none"> ▪ Web Browser ▪ Outlook/Notes Integration ▪ Web Services ▪ JAVA SDK
Scalability	Resource Mgmt	Data Capture	Open Standards
<ul style="list-style-type: none"> ▪ GRID Architecture ▪ Modular Indexing 	<ul style="list-style-type: none"> ▪ Single Instance Storage ▪ Vault Management 	<ul style="list-style-type: none"> ▪ MS Exchange ▪ Lotus Domino ▪ PST/NSF Files ▪ IM/Bloomberg ▪ File Systems 	<ul style="list-style-type: none"> ▪ Java ▪ MIME ▪ SNMP ▪ TLS

relevant to achieving them.

What I Need For Regulatory Compliance. Each regulated sector has highly specific archival requirements. Indeed, an archival solution should support the requirements of multiple sets of regulations, since organizations may fall into two or more compliance categories. The archival solution must be able to sit at the gateway (where all incoming and outgoing email traffic passes) in order to have a complete record of all incoming and outgoing email; otherwise, leakage occurs. A sophisticated rules engine is mandatory to accommodate complexity in compliance requirements. Flexibility is also critical since compliance parameters change often. The solution should enable easy retention and destruction policies. In addition, tools for random sampling and approval processes should be an integral part of the offering.

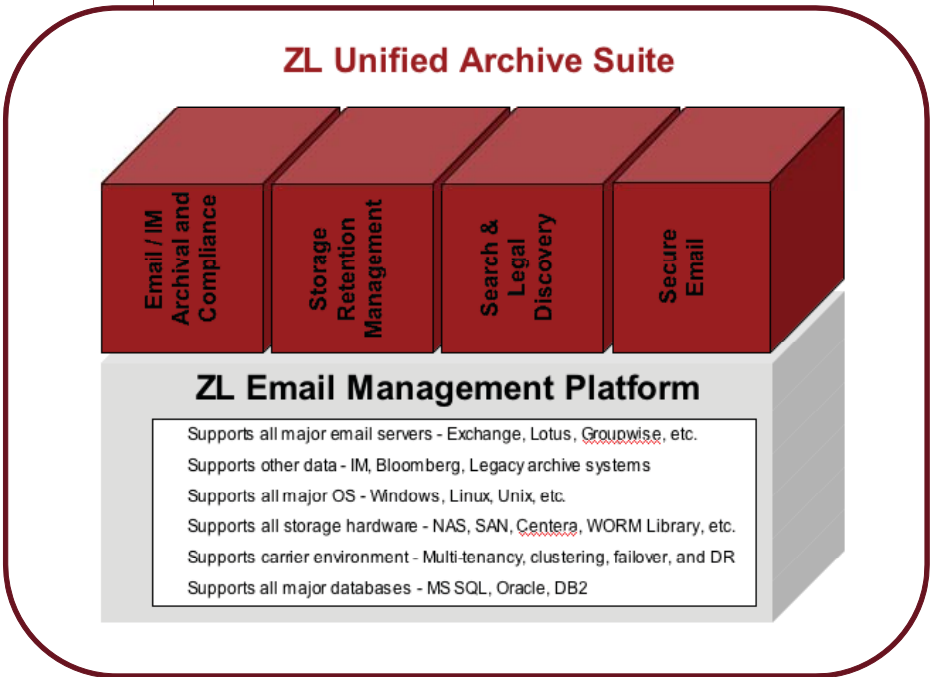
What I Need For Storage Offloading And Management. Storage offloading pays immediate dividends by relieving your email servers from handling large email attachments, resulting in faster performance and significant savings (you don't need to buy more email servers). You should look for the following: implementation at the gateway and server-level, since it's a lot easier and more complete than desktop implementations; single-copy approach, with no duplicates; Web-based attachment viewing by the end-user and administrator; "stubbing" or attachment separation and delivery for both internal and external recipients; and quick and comprehensive attachment search capabilities.

What I Need for Legal Discovery. Key features of a legal discovery solution are: searches instantaneously across large volumes of email and files; automated special handling of client-attorney privileged email; easy access to data in storage media; the ability for legal counsel to directly search and retrieve email; a solution around legal support processes; automatic classification of records; specialized numbering conventions; avoidance of spoliation and secure collaboration among internal discovery users and external legal counsel.

What I Need For Email Content Management. Capabilities essential for email content management include: quick searches of the email and attachment content; hierarchical access privileges; detailed audit trails; and timely and scalable indexing of new content.

Getting It All Together. The four main archival applications, namely, compliance, management of email attachments, legal discovery, and content, should ideally be integrated to run on a single system. The benefits of unified archival include

considerable cost savings, a single point of control, integrated reporting and dramatically lower administration overhead. Done right, unified archival provides you comprehensive email content control well into the future.



Features of the ZL Unified Email Archive Suite

a. Integration of Compliance, Mailbox and Content Management on a Single System.

Running all three applications on one system can save a lot of money. For example, if you run three separate systems, you need three servers, or six, for redundancy. The same applies to the storage system. You also need multiple administrators; separate monitoring and response systems; and multiple backup-and-restore devices and processes. And after all this, there’s still no integrated reporting, nor a single point of control. By contrast, if these solutions were on one unified system, the costs would be much lower since you need only one production system and one backup system. Administration would be easier with a single interface, single point-of-control and integrated reporting.

b. Ability To Sit At The Gateway, Ability To Block Email.

The archival system should be able to sit at the “gateway,” i.e., the edge of the network where the internal corporate network meets the external Internet. All incoming and outgoing email passes through the gateway. Very few archival systems have the horsepower to handle this torrent of data. Sitting at the gateway is important because it catches all incoming and outgoing traffic and, therefore, provides a complete record. More importantly, blocking of email at the gateway is best done by the archival solution because:

- (a) The archive can route the problem email through compliance officers who can decide whether to hold or release it; and
- (b) The archive can keep an accurate record of blocked email. Archival solutions that cannot sit at the gateway provide incomplete or even inaccurate records.

Such solutions depend wholly on the internal email server to provide them copies of email. The email server, unfortunately, gives an incomplete picture. For example, a separate content filter may block email at the gateway, but the email server inside is oblivious as to which were blocked. Another source of data loss is the blind copy feature (“bcc:”), whereby the email servers are effectively “blind” to the bcc: messages. And still other data leaks come from messaging systems, which bypass the main email servers, e.g., automated CRM email, statement delivery systems, etc. All these leaks can be caught at the gateway. If you must maintain complete records, the archival system must operate at both the gateway and the email server level.

c. Scalable and Flexible Engine (MTA).

Email archival involves mind-boggling volumes of data. In a single year, a 12,000-strong company generates more email data than all the printed content to date in the U.S. Library of Congress. The software engine must have the horsepower to handle the combined burden of attachment, compliance and content management. You should ascertain in detail the scalability and flexibility of the base architecture and the core engine or MTA (Message Transfer Agent).

d. Industrial Database, plus Metadata for Scalability and Flexibility.

The solution should incorporate any well-known database, or preferably, offer a choice of databases such as Oracle, DB2, MS SQL, and so on. Beware of proprietary databases. They are unlikely to be robust and you may face “vendor lock-in,” along with difficulties in integrating, trouble-shooting, and modifying software. Archive solutions that use no database may be seriously handicapped. For larger installations, a “meta-data” layer improves scalability and flexibility. The meta-data acts like a library index card, if you will, pointing to the actual books. As you can imagine, it’s a lot easier to move and sort index cards than heavy books. Also, it enables the building and changing of multiple boxes of index cards at will, thus enabling greater flexibility. (Without metadata, we’d be stuck with the original book classification permanently.)

e. Highly Available Architecture, with Fail-over and Load-balancing.

High availability is a critical requirement for regulatory compliance and attachment management. The solution should “fail-over” or transfer the workload of a failed server to a working one. It should also enable “load-balancing,” continuously distributing the work-load across a number of servers.

f. Platform Independence.

The archival solution should run on commonly used systems, including O/S’s, databases and application servers. Even if it happens to run on your current platform, you still want it to accommodate others, because change is a constant factor. For example, a merger may change your platform requirements overnight or your company may start a Linux strategy.

g. Software Technology Platform.

The archival system should be built on either “Net,” championed by Microsoft, or Java/Java 2 Enterprise Edition (“J2EE”), championed by IBM, Sun and others. Either standard generates benefits, including easier modifications, lower maintenance and faster integration to other applications. By contrast, legacy technologies such as C/C++ are more expensive to maintain, modify and integrate.

h. Compatibility with All Popular Mail Servers.

You should prefer an archival solution which works with multiple email platforms because of mixed environments and changing circumstances.

i. Email and Attachment “Stubbing” or Storage Offloading.

To improve email server performance, the archival solution should provide a complete storage offloading capability by replacing all data-heavy email with large attachments with a link in the original message. Desired features include: the ability to download the actual data through the link for both internal and external recipients; no client software to install; single-copy, multiple views; ability to sit at the gateway; and lifecycle management of stubbed data.

j. Granular Access, Search and Restore by End-user, with Restrictions and Privileges.

The archival solution should extend access, search and restore functions not only to administrators but also to every employee, subject to company policies and restrictions. To do this, the archival solution must enable hierarchical privileges, e.g., an employee may search only his folders, the supervisor across his department, and the CEO across the corporation. This is key to secure oversight of the archive contents. It also offloads housekeeping tasks, especially search and restore requests, from the administrator. This lowers the cost of administration significantly, maintains security, and increases the value of the archived data by giving access to people who can use it best. An additional useful feature could include restore services to the client for off-line reading, preferably without installing special client software.

k. Rules Engine.

Enterprises need to define and enforce archival rules in a flexible and granular fashion, including rules based on parameters such as sender, recipient, subject, date, size, domain, body text, and so on. The Advanced Search with Proximity and Context Sensitivity Matching feature enables greater granularity and refinement of searches according to use within sentences, paragraphs, and whole documents. Lexicon-based Mail Auto-Categorization allows for revolutionary simplified automated categorization of emails according to topic, genre or category, example "Attorney-Client Privilege", "Partners", "Customers" etc.

l. Attorney-Client Privilege Flagging and Privilege Logs.

These ensure that sensitive email exchanges between employees and legal counsel will not be turned over to external auditors.

m. One-click Search and One-click Export.

The ability to search and export with a click of a button greatly simplifies and

reduces the required process times for adherence to regulatory bodies' time limits, which can range from 1 week to mere hours depending on the severity of the review or subpoena. Disclaimer-Auto Exclusion easily prevents language used in disclaimers from being flagged should they happen to contain potentially sensitive wording such as "financial" or "secure." ZL also features Quick Summary for Lexical Analysis and Compliance Analysis for Reviewers, for easy summarization of compliance review logs, frequency and occurrences.

n. Compression. Storage costs are a significant component of archival costs.

Compression of content is the most direct way of minimizing costs, freeing network bandwidth and shortening the time for data transport.

o. Fast, Indexed Search of Message Header, Body and Attachments.

The ability to quickly search, find and retrieve data drives the value of an archival solution. Ideally, the software should support near real-time indexation of header data, body text and attachments. The ability to do Simultaneous Cross-Alias Searching (Email address, IM userID, Bloomberg) allows compliance officers and counsel to prevent double, triple or quadruple searches of the same user in the corporate directory, saving time and resources in the event of an audit.

p. Detailed Audit Trails and Integrated Reporting.

It is important to maintain a record of activities at the message level, the user level, and administrator level. Such audit trails are useful for performance profiling, security and accountability.

q. Administration Tools with Web-based UI.

Administrative tools should offer a Web-based user interface, enabling control from anywhere, via any device, including wireless. The solution should also offer hierarchical administration, where housekeeping tasks can be shifted from the administrator to end-users. Complete End-user Webmail View and UI. The archival solution should offer Web-based access, either through the browser or popular clients such as Outlook. The user's folder structure format should also be preserved. Compliance Tools and Process Applications. The archival solution should provide tools to make the compliance process easier for auditors and compliance officers. It should enable random sampling and automatic distribution of the sample to inspectors for examination, with tracking of each batch until approved.

r. The Complete Archive, including Internal-Internal Email.

To be a complete email content management solution, the archival solution must capture and manage not only internal-external email but also internal-internal email. In fact, the latter volume is often many times larger than the former, and you should verify the solution's scalability.

s. Load Balancing, Internally and Externally.

Most systems offer rudimentary load-balancing, which does a simplistic allocation of inbound traffic to the appropriate server for processing. Superior systems, however, also offer internal load-balancing, where the servers allocate tasks amongst themselves internally to achieve maximum throughput. The result is greater scalability, faster performance, more efficiency and higher availability from any given system configuration.

t. Snapshot Management.

Snap-shot management enables tracking of images of the archive at specific points in time, for retention or restoration. Some snapshots go further by enabling traffic management and finely-tuned archival, such as setting snapshot intervals to any arbitrary period. Snapshots must be taken with negligible impact on network or system performance, via control of data transfer rates. Snapshots can enable robust recovery, comprehensive journaling, and flexible data management, such as storing different periods in different locations or media. They also help users access older data as the archive software itself goes through different versions.

u. Security, Encryption, Authentication, Authorization.

The sensitive information often contained in archives, combined with the need to access it via the Internet, requires comprehensive end-to-end security, management of privileges, authentication, authorization, digital signatures, audit trails and, of course, tamper-proofing.

v. Multiple Modes of Access to the Archive.

As the Web redefined the concept of data access, the archive solution must provide multiple modes of access via open standards, including the Web user interface, IMAP4, WebDAV, MAPI, POP3, SMTP, HTTP/s, Web services, etc. This enhances accessibility and adaptability.

w. Single-copy, Multiple-views of the Archive.

Many archival solutions tolerate duplicate copies and display archive data in a limited way. You should demand single-copy capabilities, while providing multiple virtual views of the same content. For example, a user can view his archive in the same folder format as his primary email.

x. Flexible Presentation, Including Unicode Support, Wireless Delivery.

The email archive should offer flexible presentation including Unicode support for double-byte languages, and wireless access capability via any device (PC, PDA, handset) and any wireless protocol (WAP, i-mode, etc.).

y. Lifecycle Management and Granularity of Retention Policies.

The solution must enable lifecycle management of archived material, from initial storage to reclassification to destruction. Granularity of control should be fine-tuned down to the message-level. The solution should offer hierarchical storage management which automatically migrates data to different media, including disk, tape, Write-Once, Read-Many (WORM) drives, etc. Lifecycle management requires a sophisticated rules engine.

z. Near Real-Time Indexing vs. Batch Updates.

Most archival systems are unable to do near real-time indexing and updates. Instead, they do it in batches, with time between updates varying widely. The archive data may not be up to date. Ideally, the archival solution should support near real-time incremental indexing and updates, and be capable of parallel indexing across multiple machines to ensure scalability.

About ZL Technologies

Established in 1999, ZL Technologies, Inc. (ZL) provides cutting-edge enterprise software solutions for e-mail archiving, regulatory compliance, litigation support, corporate governance, content management, and secure email. ZL's flagship product, the Unified Archive, offers comprehensive email archiving and management for companies using Lotus Notes/Domino, Microsoft Exchange, Bloomberg, and others. The suite provides a highly flexible framework that is fully scalable, enabling organizations of all sizes to meet legal discovery, compliance, and storage management requirements. With a proven track record and an extensive list of clients, including Walgreens, Bank of New York Mellon, Pacific

Life, and Morgan Keegan, among other top global institutions, ZL has emerged as the premier provider of email archiving and compliance solutions. For more information, please visit www.ZLTI.com

