

ZL UNIFIED ARCHIVE Top Differentiators

ZL TECHNOLOGIES | White Paper



ZL TECHNOLOGIES INC.

In the world of archiving and discovery, the following eight concepts are at the heart of the industry's greatest challenges. Most vendors do not call attention to these eight issues because they cannot solve the problems they raise. Alone, any one of these eight items could force an archive deployment to fail. Together, these eight items represent the dirty little secrets that other solutions do not want customers to know about. Addressing each challenge directly and effectively pushes ZL Unified Archive ahead of all other vendors.

1. Unmatched Scalability Delivers Industry's Fastest Search Speeds and Early Case Assessment

Amongst search technologies, some return results in seconds whereas others take days. When a simple query takes days to come back, do not expect to be prepared for initial disclosure any time soon. The responsiveness of ZL's search engine counted in seconds not only for general queries but even for queries made across ALL custodians and ANY date range. This is due to the power of the ZL GRID and next-generation design, which give ZL's customers the ability to perform instantaneous early case assessment across the entire corporate corpus.

Old generation solutions cannot support the volume of data produced by today's data systems. Volumes have increased by 20% or more each year for the past decade. Thus, when it's time to search, systems originally designed for archiving in the 90's are easily overwhelmed by the number of retrieved items, assuming they do not timeout first.

Consequently, in these older generation products, even a simple search often produces incomplete or inaccurate results. ZL, on the other hand, provides a next-generation platform designed to scale to volumes produced by hundreds of thousands of users, delivering complete, consistent results, search after search without any timeouts, regardless of result size or complexity.

2. Sophisticated Searching with an Intuitive Interface

The most well known, widely used legal tools provide advanced features like proximity search, enabling users to find relevant documents with keywords and phrases, associated by proximity. Proximity is a natural and intuitive way to associate words and phrases, producing highly accurate results. Combined with an easy-to-use interface, ZL advanced search features make searching extremely simple so that users quickly find what they need.

Other vendors have expensive, complex technologies that require users to "teach" the engine to associate certain "concepts," using the software's artificial

intelligence algorithms to do so. This “training” period is long and involved, often requiring costly “consulting” services which escalate costs further.

Though touted as being “more advanced,” the downside with this type of concept-based system is that the user is no longer in control of how results are gathered. Instead, results are generated by algorithms which users have no way of controlling, forcing the user to trust the accuracy of their results to a blackbox. No matter what claims these vendors make, trusting an “algorithmic” system that requires training ultimately limits the predictability of evidence obtained in this way.

With ZL, the user is always in control and will always understand why any result appears.

3. Incorruptible Index (Index Recovery)

Amongst the dirty secrets in E-Discovery, corrupted indices are the worst. Search indices are files that are always being written to, around the clock. As such, they cannot be easily backed-up. But like any file, indices corrupt. Without any back-up, every archive is an E-Discovery nightmare waiting to happen.

The only recourse for customers is a complete rescan of the raw data to recreate the index. Obviously, the time to perform a rescan is roughly equal to the first time the data was ingested. Or is it? In fact, time to rescan almost always takes longer due to the fact that only unutilized resources can be used to rescan while the production capture continues. Thus, an index for 6 weeks of raw data may require many months to scan. This means that no searches performed, no discovery requests responded to, and no regulatory audits answered during those months.

Even worse, most solutions do not have any mechanism to alert users to corrupted indices. The only time most customers learn that their index is corrupted is when they attempt a critical search and it fails.

ZL addresses these challenges with a unique patent-pending self-healing index design. First, ZL never requires customers to rescan data. ZL is able to recreate full indices in a matter of hours. No more waiting for reindexing. Secondly, ZL's automated monitoring system constantly checks indices for corruption or problems, so that the index is always ready when you need it. Snapshotting technologies further reduce downtime.

ZL is the only solution that combines an enterprise class search engine with self-healing properties to assure 100% search uptime and availability.

4. Digital Chain of Custody (DCC), Accounting for Every Mail

For enterprises, one missing e-mail can mean the difference between success and sanctions. Customers concerned about E-Discovery must be confident that their archive is complete. ZL's digital chain of custody (DCC), reconciliation processes, and audit trails assure 100% capture.

Digital chain-of-custody is an extremely difficult, challenging process. In fact, DCC is so difficult to do that unless a vendor has an absolute requirement to do it for their customers, it is usually not worth the time or testing required to engineer it properly.

For example, generic, consumer-based search technologies that enter into the enterprise search space learn very quickly the challenges that differentiate searching within enterprises versus for consumers (e.g, a typical Google query). Simply put, one missing result for a consumer may only mean one less choice of restaurant for lunch, but one missing result for an enterprise may be the smoking gun that means millions of dollars in damages in a legal battle. In reality, even the majority of enterprise search solutions lack this level of rigor.

The challenge for this level of digital accountability is the same as when providing a physical chain of custody: every step in the chain must be known and under control of a single authority in order for the audit trail to be complete and the chain of custody to be unbroken. The difficulty in doing this when archiving lies in the fact that most vendors rely heavily on multiple OEM technologies to provide their functionality. Each of these OEM technologies is like a dark, locked room. Data enters the room, but no one knows who touches that data or how it is modified, only that the resulting data exits out of the room. If any modification, corruption, or alteration of the data occurs while in that room, the archiving vendor has no way of knowing. Consequently, digital chain of custody is broken each time it is passed to an OEM component.

ZL's architecture addresses this by having:

- Complete Code Control – all major components of ZL are developed, designed, controlled by ZL.
- StraightThrough Processing – all actions conducted against data in ZL are performed in a single straight through process where the data is opened and read only once, after which all actions are applied and the data is then closed.

- Fundamental Customer Requirement – in ZL’s target market, digital chain of custody is a requirement, not an option. Because of this, ZL has been designed from the ground up to deliver DCC throughout the application, as opposed to adding DCC controls as a snap-in or afterthought.

5. True GRID Advantage

As the use of e-mail grows exponentially, systems used to manage or archive e-mail are some of the most heavily stressed applications within the enterprise. Moreover, many of these systems are responsible for legal holds and capture, compliance, or records management around the clock, all tasks of critical importance. The combination of extremely high volume and critically high importance forces customers to demand the most performance, reliability, and scalability from its archive vendors. However, with very few exceptions, vendors have fallen short in delivering this.

Most are architected as monolithic, siloed applications that require dedicated servers along with separate buffer and back-up servers for each function. Others use an archaic design of parent-child architectures, which are still vulnerable with single points of failure and manual processes for failover. Still others make the noble claim of being first generation grid architectures, yet fail to deliver True GRID due to the use of OEM components for critical tasks (e.g. search engines). Since such components are OEMed, it means they are not built using the grid architecture the vendors claim, and as such, no benefits are derived.

ZL, on the other hand, has repeatedly differentiated itself against competitors using its True GRID architecture, which combines self-healing and distributed processing to deliver nearly infinite scalability, no single point of failure, reliability, and disaster recovery. Customers have consistently remarked upon how much easier it is to rely upon and increase performance on ZL deployments using fewer servers, typically 75% less server hardware than the closest competitors.

Not only is ZL able to deliver all of the above advantages, but when coupled with the use of virtual machines, customers running ZL with True GRID can further reduce the number of physical servers used as well as ongoing Total Cost of Ownership (TCO) from power consumption, space, and administrative overhead traditionally seen when using more physical devices.

6. True Single Instance Storage

Of the many secrets that archiving vendors keep from customers, the most hidden is the value of True Single Instance Storage (SIS). All archiving vendors claim to perform single instance storage, the magic of storing a single copy of mail or file when multiple copies of the same mail occur in the archive due to multiple sends, copying, cc'ing, etc. Vendors never hesitate to tout the benefits and subsequent return on investment this feature delivers. This selling point is often a key driver for CIOs trying to meet storage budgets in their project implementations.

The reality however, is a stark contrast from what these vendors claim. In fact, if one examines the best practices guidance for the largest vendor in the space, one finds that when sizing storage, it is recommended that any potential SIS benefits should not be considered, that in fact, the amount of data stored within the archive system will remain the same as in the mail system. Wait, what happened to the magic of SIS?

Unlike its competitors, ZL's approach to SIS, True Single Instance Storage, differs dramatically, the most significant outcome being the reduction in size of the e-mail data when it is collected and stored within the ZL archive. Not merely a groundless claim, what once existed as 1 TB in Exchange or Lotus actually becomes a mere 250GB or 200GB in ZL, a 75-80% reduction. When sizing, ZL recommends customers use a conservative expectation of 50%. No other vendor delivers that level of reduction.

ZL accomplishes True SIS through its unified vault architecture. Most vendors waste much more storage because they cannot single instance when:

- The same mail is sent to two or more groups.
- The same mail is captured by two or more methods.
- The same mail is used by two or more applications.
- The same mail is stored in two or more vaults.
- The same files are attached in two or more e-mails.

What's more, the leading vendor uses a proprietary format that is three times the size of the original e-mail, making the atomic size of each archive record three times that of ZL even before SIS is considered. SIS? ROI? It is unclear how these vendors arrive at their claims.

7. Database Count

When archiving was initially unveiled to businesses, it was primarily for the purposes of storage management and the reduction of e-mail storage of regular enterprises. This was back in the mid 1990s, when most businesses saw barely 10 e-mails per individual per day and each e-mail averaged less than 5KB in size. Today, employees average over 100 e-mails per day sent and received, with each e-mail averaging over 100KB. Early designs used databases much in the way that current systems do, but with one major difference: whereas old-generation designs can accommodate 5 to 10 million records per database, next-generation solutions like ZL can store roughly 1 billion records per database.

The operational and financial consequences of this difference are staggering. An average 1,000-employee company using standard e-mail usage metrics with a retention policy of five years will retain nearly 200 million mails in that time. An average 5,000-employee company with the same metrics and retention policy will retain nearly 1 billion mails in that time.

For the leading vendor to accommodate the 1,000-employee company using best practice guidelines from their manuals will require roughly 26 separate database instances. To accommodate the 5,000-employee company, the leading vendor requires over 120 separate databases. Given that general database administration practices recommend roughly one database administrator for every 20 databases, one can quickly see the cost of ownership for the leading vendor becomes untenable very quickly.

The difference with ZL is that its architecture was designed at a time when e-mail had already shown its exponential pattern of growth. More importantly, ZL was designed not to scale for standard businesses, but for carrier-class accounts with over 500,000 mailboxes to support.

ZL architecture was designed targeting these metrics for performance and scalability. As such, each ZL database must scale to much higher record counts than its competitors and consequently have been engineered to do so successfully.

In the above comparison, for a 5,000-employee business, ZL requires a single database versus the leading competitor who requires 120+. This translates to a single part-time DBA vs. 5+ full time DBAs. At roughly \$150,000 per year, DB cost of ownership for ZL? \$30K. DB cost of ownership for the leading competitor? \$750,000.

If one includes the headache and infrastructure needed for backing-up and administrating that many databases regularly, the cost of ownership goes even higher. This assumes that one can even back up that many databases, a well-known issue with the competing solution.

8. Consolidated Mail Systems

Big businesses become big by either growing organically over time or acquiring or merging with other companies. One of the greatest challenges of any large enterprise that actively grows through mergers and acquisitions is the consolidation of the IT departments and, in this case, specifically the e-mail systems. Although the logical notion would be to merge the mail systems, often times, firms cannot decide which direction to go and end up putting the decision off until some later date, which may be years away. Meanwhile, legal discovery, storage management, compliance, and records management requirements must be met.

Different organizations choose different enterprise mail solutions all the time, such as MS Exchange, Lotus Notes, or Sun JMS. The great majority of leading archiving solutions only do a single mail system well, since they generally focus on one during their development.

Moreover, even if the solution considered does support multiple mail systems, they must deploy completely separate, standalone archiving infrastructures (back-up, disaster recovery, availability, storage, servers, etc.) to accommodate these mixes. Furthermore, customers unfortunate enough to have to manage duplicated archiving solutions are at a significant disadvantage even at a day-to-day level. For instance, for any search or discovery, admins must perform identical searches in duplicate locations, set policies in duplicate systems, and manually manage users moving between departments in the two systems, hoping the entire time that the different coded applications they are running will behave and perform identically. This does not include all of the physical duplication of resources and associated costs.

Needless to say, from a compliance, regulatory and discovery standpoint, there are a lot of moving parts and manual processes that administrators must consistently perform and do so accurately in order to avoid any mistakes.

Here again, the ZL Unified Archive's centralizing architecture has dramatic advantages over less flexible solutions. Because ZL is a unified archive based on

open standards, the same installation supports any mail system. Proven in real world deployments, ZL can capture, index, archive, and store MS Exchange users, Sun JMS users, and Lotus Notes users within a single archive and infrastructure. No duplication is necessary, meaning no extra deployment and no difference in code-base or application. In fact, when it comes time to consolidate the entire organization onto a single messaging platform, ZL's Unified Archive dramatically reduces the time and amount of data that needs to be migrated from the retired system, often reducing it by over 95%, saving time and space.

General counsel are able to confidently search and perform E-Discovery via one interface, knowing that they will capture all information across the company. Records management and retention policies can be set in a single location and are applied properly across the enterprise. Legal holds, preservation orders, system administration, supervisory review, and storage all occur under a common unified archive regardless of the original mail system, saving money and reducing complexity.

