Clipper Notes™

Top 10 Things You Should Know
About Information Lifecycle Management

Analyst: Michael Fisch

Information lifecycle management (ILM) has exploded onto the storage scene with the sound and fury of the best of buzzwords. *Look, something new and wonderful! It's a must-have! A solution to your storage problems!* Vendors proclaim it; analysts analyze it; journalists write about it; and enterprises are reacting to all of the commotion. With such acclaim, the skeptic in us is tempted to dismiss it as the latest hype. However, there is substance behind the ILM concept, as well as a need to sort out what it really is, put it in perspective, and consider how to respond. In that spirit, we present the top ten things to know about ILM. Here is the list; each is discussed below.

1. *ILM is a concept for dynamically managing data over its lifecycle, from creation to deletion, as its value changes over time.*
2. *ILM is based on the premise that data has a lifecycle and a relative value that changes.*
3. *The business benefit of ILM is maximizing the value extracted from information while minimizing the total cost to store and manage it.*
4. *ILM requires viewing storage as a service, not just as a box.*
5. *ILM is an enterprise-wide strategy, not an individual product.*
6. *For the foreseeable future, implement ILM one step at a time.*
7. *The primary operational components of ILM are tiered storage, data classification/policy definition, and data movement.*
8. *Tactical triggers for considering an ILM solution include: meeting business and regulatory requirements for data retention and access; providing appropriate storage service levels; improving application and file system performance; lowering storage hardware costs; and speeding up data management operations like backup, restore, upgrades, or replication.*
9. *ILM is coming.*
10. *ILM is here.*

Read on for the details.

---

1. ILM is a concept for dynamically managing data over its lifecycle, from creation to deletion, as its value changes over time.

ILM seeks to apply the optimal storage service level to data at each point in its lifecycle. Service levels describe the particular performance, availability, recoverability, and cost characteristics applied to a data set. There can be multiple tiers of service – think in terms of "low, medium, and high" or "silver, gold, and platinum." The main purpose of ILM is to strike the optimal balance between meeting business requirements and minimizing storage costs - in other words, the right service level for the right data at the right time.

ILM navigates between two common extremes: (a) Assuming the value of all data is high

---

and spending too much to store, protect, and manage it or (b) applying inadequate service levels and suffering performance bottlenecks, downtime, data loss, or penalties for failing regulatory compliance. ILM offers a comprehensive, business-aware, and deliberate approach to storage and data management.

2. ILM is based on the premise that data has a lifecycle and a relative value that changes.

Similar to plants and animals, data follows a cradle-to-grave path of existence. A data object initially springs to life as the creation of an application or user. It could be a record of a financial transaction, an e-mail, a document, a video, or a thousand other types of digital information. Who or what created the data, the subject matter, the related business processes, and other factors determine its value and appropriate storage service level. Data often begins in an *active* state, where it is accessed and altered frequently, as in a document undergoing an editing process. After a period of time, data can become *fixed* or *inactive*, where an enterprise is much less likely to access or change it. Some data is even born fixed. In any case, fixed data still needs to be retained to meet business and regulatory requirements, but a lower storage service level and associated cost may be appropriate. An unexpected event like a financial audit or a long-dormant customer deciding to repurchase could revert data temporarily back to an active state. After a (possibly very long) period of time, the value usually falls to zero and data can be deleted. So ends the lifecycle. ILM exploits the natural rhythm and value differentials of data to the benefit of the enterprise.

3. The business benefit of ILM is maximizing the value extracted from information while minimizing the total cost to store and manage it.

This is a balancing act. On one hand, timely and reliable access to the right information at the right time is essential. Enterprises operate and compete on the basis of information as much or more than any other asset. Land, labor, and capital are no longer the exclusive economic inputs in the modern digital world. ILM ensures that data is stored in a service-level tier appropriate to its inherent business value, and that it is accessible when and where needed for enterprises to utilize effectively.

At the same time, storage is a sizable component of the IT budget. The cost of acquiring storage hardware and software and the overhead to manage them are significant. Controlling this cost is an ongoing challenge because the amount of data continues to rise and storage requirements tend to become more stringent with time. We have more and we expect more. Therefore, ILM also serves as a useful tool for coping with costs. The goods news is that the industry is poised to deliver a steady stream of innovations in the area of ILM to help enterprises remain informationally competitive at an acceptable cost.

4. ILM requires viewing storage as a service, not just as a box.

"Storage as a service" is foundational to ILM. The service level is the common ground where business and IT meet and agree how data is to be managed. This view is a step up from looking at storage as a box of spinning disks on the data center floor. It is like the difference between a telecommunications network and a long-distance plan, or a power plant and the electrical service delivered to your home. As consumers, we appreciate and experience the service delivered to our doorstep. The infrastructure that provides it, including all of the points of manufacturing and distribution, are behind the scenes and taken for granted. In a similar vein, storage service levels are what the application and user actually experience. They represent the consumer- or demand-side view. Users decide the storage quality that they need based on application requirements and what they can afford[1], and the IT department delivers to the specification. It is important for enterprises to make this mental and operational shift and treat storage as a service. Not only will this help enterprises to better deliver and responsibly consume storage resources, but it is also an essential step in the ILM journey. Matching data to service levels is the foundation of ILM.

5. ILM is an enterprise-wide strategy, not an individual product.

You can think of it as a lifestyle – a holistic way of approaching information management from cradle to grave. While a product might provide an ILM solution within a defined scope (e.g., database archiving), ILM, in its fullest sense, is an enterprise-wide strategy that encom-

---

[1] See *The Accounting Pendulum Swings at Storage (or, Why The Taxman Cometh?)* in the issue of **Clipper Notes** dated March 10, 2007, at http://www.clipper.com/research/TCG2007038.pdf.

passes a variety of systems and procedures. The concept takes into account data in all of its forms, capacities, and uses. ILM involves technology, people, and processes.

6. For the foreseeable future, implement ILM one step at a time.

The ILM vision may be holistic and enterprise-wide, but in the near term, enterprises are likely to deploy point solutions in response to specific needs or pains, such as complying with an industry regulation or improving storage utilization. An enterprise-wide deployment in one fell swoop is not yet feasible. For one, the technology has not evolved enough for such a broad solution. There would need to be a common metadata repository and integrated software across a variety of platforms. While some vendors are working in this direction, they have their development work cut out for them. In the long run, and after ILM has demonstrated a track record of success, enterprises may take a comprehensive approach to it. But, in the meantime, it makes the most sense to approach ILM in one incremental step at a time, solving current problems and demonstrating tangible progress through a series of successful projects.

7. The primary operational components of ILM are tiered storage, data classification/policy definition, and data movement.

Tiered storage is the baseline for delivering differentiated service levels. It means there is more than one physical type of storage. A tier can be a particular type of media, such as disk, tape, or optical.[2] A disk tier, for instance, can be differentiated by drive type (Fibre Channel, SATA, SAS, etc.), RAID level, platform architecture (high-end global cache versus midrange dual controller), and replication (point-in-time copy, synchronous, and asynchronous mirroring). Network bandwidth and connection redundancy also play a role. Tape libraries and drives can also be differentiated by performance and capacity characteristics. Furthermore, modern tiered-storage environments are likely to include multiple disk tiers, including SATA for secondary storage purposes. While service levels are ultimately defined in terms of specific metrics[3],

tiered storage is the foundation for delivering multiple, differentiated service levels.

Data classification and *policy definition* is the process of sorting data into meaningful groups and applying lifecycle management policies. This requires metadata – or relevant information about data that allow it to be described and qualified – and smart software that can leverage and apply it. Policies need to be shaped by the business. No single group has a full sense of what can or ought to be done. Functional personnel have the best sense for their particular business process and application requirements. IT personnel have the best understanding of technical issues and ramifications. Where regulatory and legal requirements are concerned, executives and even legal counsel will want to provide input and direction. How data is handled is critically important to an enterprise, so communication and a consensus-oriented approach are recommended to ensure data classification and lifecycle management policies reflect the needs of the whole business.

Finally, *data movement* shifts data between tiers to accommodate changing service level requirements. This is key for dynamic lifecycle management. Data movement is ideally an automatic and non-disruptive process handled by intelligent software – meaning that human intervention would not be required once policies are programmed, and that data remains accessible to applications during and after data movement. This would deliver all the benefits of ILM while minimizing management costs and maintaining application availability. Other related ILM components include centralized management of the storage infrastructure, resource virtualization/logical abstraction, data search and retrieval, and flexibility in applying levels of data protection and recovery.

8. Tactical triggers for considering an ILM solution include:

- ***Meeting business and regulatory requirements for data retention and access*** – The recent trumpeting of compliance concerns has pushed the issue of data retention into the spotlight, but it has always been an important business issue. Best practice suggests that all enterprises should define and enforce data retention policies, if only to meet the legal, tax, and operational require-

---

[2] For details, see *Tiered Storage Classes Save Money – Getting The Most Out Of Your Storage Infrastructure* in **The Clipper Group Explorer** dated August 29, 2002, at http://www.clipper.com/research/TCG2002030.pdf.

[3] Storage service level metrics include performance (MB/s, IOPS), availability (percentage uptime), recoverability

(recovery point and recovery time objectives), and cost ($/MB).

ments of the business. In reality, though, it may be the requirement to comply (and fear of penalties for noncompliance) with regulations like Sarbanes-Oxley, HIPAA, SEC 17a-4, etc., that cause many enterprises to take action. Data retention means not only keeping data but also having the ability to readily search and access it.

- *Providing appropriate storage service levels* – As noted above, an enterprise may be providing inadequate data protection and access performance or spending too much on storing data of minimal value. It is in the spirit of ILM to recognize and correct service-level imbalances.

- *Improving application and file system performance* – Databases and file systems are like cargo ships in that they can slow and become less stable when overloaded (i.e., too much data). User response times increase, which can affect worker productivity. Instability and longer rebuild times interfere with availability. Pruning inactive records or files and moving them to a less-expensive storage tier offloads the extra "cargo", streamlines the system, and make it faster and more nimble. It saves on the higher cost of primary storage and helps manage and maximize platform utilization through smart data migration. The time-consuming and costly alternative is continuously to tune the database, manually, migrate files, and/or periodically throw more hardware at it in the form of server upgrades, additional NAS appliances, and top-tier storage capacity.

- *Lowering storage hardware costs* – Archiving inactive data saves the cost differential between the primary and secondary tiers of storage. This is especially attractive if the amount of data is large. Storing data on ATA disk and/or tape can be a fraction of the price of high-end storage arrays with Fibre Channel or SCSI drives. The savings are magnified if an enterprise uses multiple replicas of data for purposes like disaster recovery, testing, and development – as many do. In that case, the total savings would be a multiple of the number of replicas. Smart ILM policies also help promote high utilization of storage resources, especially of more costly primary storage.

- *Speeding up data management operations like backup, restore, upgrades, or replication* –

The larger a database or file system is, the longer it takes to backup, restore, replicate, or perform a software upgrade. If any of these activities take an inordinate amount of time, it can cause unacceptable downtime, interfere with production operations, and/or delay application testing and development cycles. There may be more than one way to solve this problem, but using ILM to streamline the primary repository may be the most cost-effective and durable solution.

9. ILM is coming.

As a long-term trend, ILM will blossom because it meets a broad and deep need for more effective cradle-to-grave management of enterprise information. This need will only increase as data grows and service level requirements increase over time. For its part, ILM helps address everything from storage costs to regulatory compliance and improved service levels. Furthermore, there is a strong focus on delivering on the ILM promise in the vendor community. Vendors are pouring major resources into development, acquisition, and integration of ILM solutions. Although some vendors proclaim the goodness of ILM more loudly than others, no single vendor owns it. It is an industry-wide initiative and long-term trend. So, there is demand and there is supply, and the ILM "transaction" is bound to happen.

10. ILM is here.

In the near term, there are a variety of ILM solutions available. They typically address specific applications, and, in total, they encompass databases, e-mail applications, file systems, and unstructured content. More generally, these represent the categories of structured, semi-structured, and unstructured data. If your enterprise has a requirement that ILM can address, you don't have to wait – there are good solutions on the market now.

Conclusion

If any of these ten considerations aren't on your ILM agenda, they should be. Better get moving!
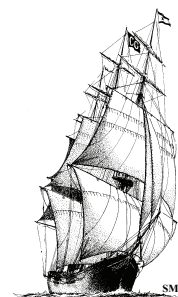
About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

➢ *The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.*

About the Author

Michael Fisch is Director of Storage and Networking for The Clipper Group. He brings over ten years of experience in the computer industry working in sales, market analysis and positioning, and engineering. Mr. Fisch worked at EMC Corporation as a marketing program manager focused on service providers and as a competitive market analyst. Before that, he worked in international channel development, manufacturing, and technical support at Extended Systems, Inc. Mr. Fisch earned an MBA from Babson College and a Bachelor's degree in electrical engineering from the University of Idaho.

➢ *Reach Michael Fisch via e-mail at mike.fisch@clipper.com or at 781-235-0085 Ext. 211. (Please dial "211" when you hear the automated attendant.)*

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** *Captain's Log*, **The Clipper Group Voyager,** **Clipper Notes**, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.