The building blocks of information life-cycle management

Opinion by Jehoshua Bruck, Rainfinity

A lot of people in the storage community are talking about information life-cycle management (ILM). Here's the short version: Lots of vendors are selling into the ILM space but none has a comprehensive ILM product. ILM is a good idea, but it's not easy. *Still, it's worth it.*

It's worth it because IT professionals are struggling to keep up with storage and provisioning, to make users happy and to observe regulatory requirements. The Sarbanes-Oxley Act and the Health Insurance Portability and Accountability Act, are just two examples of legislation that are redefining what information must be kept, for how long, and who gets to see it. ILM provides a means for storage administrators to keep up with these requirements while also improving their ability to provision storage and maintain high levels of data accessibility for end users. Why should you look closely at ILM? Well, do any of the statements below look familiar? If so, you're not alone:

The polite statement	The IT translation
"Storage management challenges are escalating."	"I've got a 200-port SAN, a big NAS and a slew of direct-attached storage, and I'm trying to manage it with a spreadsheet!"
"Users want instant and uninterrupted access."	"Users are big crybabies, but I've got to make sure they're happy."
"Senior executives need to make sure that corporate data is protected, accessible and restorable."	"If government attorneys come calling, don't blame me."

ILM helps companies to protect, manage and use their data by matching data with the storage media it needs most at given points in the life cycle. Life cycle

means data creation, modification, retirement and (sometimes) deletion, while matching changing data value to appropriate storage targets.

Making it work

There's no "giant maker" ILM application. ILM is made up of software that enables four basic building blocks: transparent data movement, global name space, policy engines, and discovery and classification.

Transparent data movement: ILM needs transparent, automated data movement that happens without annoying users or interrupting applications. Serverless backups, transparent migrations and hierarchical archiving can all transparently move data to different storage targets, and users never know the difference.

Global name space: ILM also needs a global name space, which fools users into thinking that the file name they're seeing is actually where the file is physically residing. The abstraction layer lets administrators change the location of files and directories without having to update the end user's access or notify them of the new location.

Policy engine: Once data movement is completely transparent, administrators can set policies to automate it across storage devices. Several storage management applications come armed with policies that should operate on a heterogeneous level. (Upcoming Common Information Model and Storage Management Initiative Specification standards will make this job easier.) Well-crafted policies lift a large manual burden off administrators, allowing them to get to the interesting parts of their jobs.

Discovery and classification: This helps administrators know what files are where, who owns them and how they can best group and manage them. Administrators use software tools at the file level that capture information about the files, their locations, their applications, their subjects and how people are using them. By grouping files by classification and mapping them to applications and users, ILM can perform tasks such as assigning bandwidth or multiple computing paths to critical data.

In addition to forming the foundation for ILM, these building blocks can serve as stand-alone solutions that provide significant short-term benefits and results. A company can gain advantages right away by starting to implement the building blocks that are available, such as transparent data movement. The first step is to analyze your environment and determine your priorities.

Analyze your environment

Questions to ask when analyzing your environment:

What are your storage requirements and needs? Consider:

- User needs
- Regulatory compliance
- Round-the-clock access to applications
- Data intensity, such as graphics and image-production programs
- Records management

Match your environment to an ILM strategy

- Match requirements to ILM building blocks.

Develop priorities

- Prioritize projects by asking which deployments would have the greatest positive impact.

Observe budget restrictions

- You might already own some of the ILM tools. ILM is a modular approach that avoids forklift replacements.

How to avoid ILM minefields

More than one company has started an ILM project and then dropped it because it "just didn't work." So what's the problem? Here are some possible answers:

- **Companies think they can find one vendor with the whole solution.**

  There is no one vendor that can do it all, no matter what the pricey storage behemoth told you. You need tools for transparent data movement, global name-space capability, policy engines and classification tools. And you need them across multiple systems, multivendor environments and applications.

- **Companies overemphasize backup and recovery.**

  Backup and restore are vital components of ILM, but they're not the whole story. ILM doesn't just look to protect data, it looks to manage it. For example, what and where is the most current copy, how many layers of protection should it have, what is its priority in the corporate/regulatory universe, and how important is it at a given stage in its life cycle?

- **Companies think they can buy more storage to solve storage management problems.**

You can buy storage until the cows come home, but if you can't manage what you've got, you haven't solved the problem. Spend your money on tools to help you manage increased complexity, not just on storage capacity.

Building an ILM environment

ILM environments operate with two important basics: round-the-clock data access and heterogeneous support.

- **24/7 data access:** Downtime is bad business. ILM targets data movement based on how valuable a given piece of data is at a given point in time. An ILM strategy that keeps interrupting service to end users or applications will fail, and might severely harm the business to boot. Transparent data locations are also important to end users, who don't care where their data is stored as long as they can easily find it.
- **Heterogeneous storage support:** ILM needs heterogeneous storage environment support. Even if an organization is a single-vendor shop, setting up ILM with multivendor capabilities is insurance for future open systems. For example, when Serial ATA disk came along, many companies happily adopted it and saved quite a bit of money, whether their primary vendor offered it or not. If they had been tied to a proprietary environment, they would have had to cool their heels until their vendors decided to add ATA to their product lines.

Remember, ILM is a process. ILM manages and streamlines complete data life cycles by matching data's point-in-time value to prioritized storage resources. By treating ILM as a series of building blocks, companies can improve their existing storage environment right away while working toward full and successful ILM