

ILM advice: Collaboration is critical

Successful implementation of information lifecycle management starts with collaboration among various IT and non-IT groups within an organization.

By Michael Peterson



IF YOU ARE A CIO OR INFORMATION management professional, your job may be at risk.

This is one of the messages in a recent paper, *Collaboration: the New Standard of Excellence*, jointly written by ARMA International and the Storage Networking Industry Association (SNIA).

The message of “collaboration” should be taken holistically. Perhaps the hardest step in the process of implementing practices such as information lifecycle management (ILM) is to bring the information owning and administrating communities of an organization together to collaborate and align business objectives with the requirements for information. Although challenging, this is the most important place to start, because this is where the entire transformational process of moving to an “information-centric enterprise” begins.

The underlying collaboration message is the following: If you want to successfully solve the complexity and cost crisis in data-center operations, you have to change current practices and begin working together as an organization. If you want to successfully deal with the demands of regulatory compliance and legal and security risk, you have to focus on defining the business requirements for your information and not just focus on the network, security, or storage infrastructure.

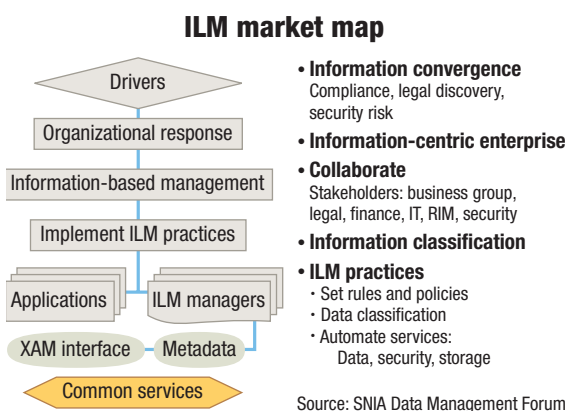
SNIA's contention in the collaboration paper is that you must make the shift to managing and operating the data center based on the value of, and requirements for, information to the organization. Step one in information-based management is to collaborate and define the

requirements for information so that it can be managed properly over its lifecycle. This is the key to ILM.

For those who have grown tired of the ILM story, it's important to understand that ILM is just coming of age. We are only three years into the market development cycle and already second-generation tools are in the market. Thousands of organizations worldwide are in various stages of ILM projects and their numbers are growing rapidly. ILM is no longer just something that vendors talk about as a way to promote their products; it is a proven practice.

The evolution of ILM is not much different from that of other storage technologies. For example, SANs were launched in 1996, work on Fibre Channel standards began in 1997, and market acceptance took off in 1999—three years after the initial launch. We're now 10 years into the market cycle and SANs are common. Similarly, iSCSI began around 2000, market acceptance followed a few years later, and we are now in year eight of a growing market. The process always follows a similar timeline, and ILM is no exception.

What is the connection between collaboration and ILM, and what does a complete ILM-based practice look like? Consider a market map, which is an illustration that aids in visualizing how all the pieces of a market fit together. For example, markets begin with user-centric drivers of demand, and solutions are then identified that address those needs. A market map for ILM-based practices is illustrated in the figure. Here is how the pieces fit together:



The leading drivers for using information as the basis for management (and to deal with overwhelming cost and complexity) are elements of what we call information convergence, which is the trend in which operations, practices, applications, and roles converge around

information and its value to the organization, creating the information-centric enterprise. The top information drivers are regulatory compliance, legal discovery risk, and security risk. Storage problems such as “out-of-control capacity growth” are not drivers; they are results of the drivers—the need to store more, retain it for longer periods, secure it, protect it, and retain more copies in multiple locations, etc.

Organizational response transforms the organization into an information-centric enterprise—an organization in which the value of, and requirements for, information are used as the central basis for management, administration, and security operations. This creates a collaborative relationship among information users, information owners, and information systems operators and administrators.

The best organizational response to an information convergence problem is collaboration among the communities that have a vested interest in the information: the business group, legal, finance, records and information managers, and IT and security professionals. The purpose of collaboration is to define requirements for information and information resources.

Collaboration is the first step in an information-based management approach, and the second step is information classification.

Information classification is the process of assigning value and scope to a business process. It is the process of aligning the value and requirements for information with the supporting infrastructure resources (e.g., compute, network, storage, and supporting data, security, and storage services).

The next step is to implement ILM-based practices by developing a comprehensive management practice around the requirements for information based on its classification. This is the foundation for executing ILM-based practices—to translate specific business and application policies and rules (service level objectives and agreements) into actionable practices. Tools are available to support these practices, with capabilities such as the following:

- Classifying data based on file metadata or content;
- Assigning service-level requirements to the data based on its classification;
- Classifying storage resources based on performance, availability, or other service level attributes; and
- Aligning the data with the most-effective storage resources.

CONTINUED ON PAGE 41

Global namespace is a term that's often associated with NAS virtualization. A global namespace essentially provides a single unified view of NAS resources. In user terms, it's the creation of a virtu-

al but unified directory structure where all resources are under a single parent directory, eliminating the need to have separate mount points for individual resources.

In technical terms, most vendors tout the use of a "virtual file system" that forms the foundation of the entire virtualized NAS environment. It is the virtual file system that allows objects from phys-

ically disparate resources to be merged and appear as one to clients or users.

Evaluating NAS virtualization

Here are a few tips to help evaluate NAS virtualization solutions:

Does your NAS appliance vendor offer virtualization?

Most major NAS vendors have some form of virtualization available in their product portfolio. While they may or may not be comparable to the stand-alone virtualization appliances, there are benefits to having a virtualization solution from the same vendor as your NAS filers, especially when it comes to integration, support, and manageability.

What features do you need the most?

Your purchase should be based on evaluation of features that matter most to you, not the frills that you may never need. After all, the virtualization appliance is going to be in the center of your environment, so if the appliance fails your entire NAS environment is down. "Un-virtualizing" is painful, and you do not want to do it often. So if your goal is to consolidate multi-vendor, multi-platform NAS islands into

a single unified space, your evaluation should be based on the vendor's interoperability matrix and not, say, a single vendor solution.

Another important consideration is scalability. How many physical resources, total amount of data, and number of files can the appliance support? Keep in mind that file-sharing environments tend to serve lots of small files—sometimes millions. Find out if the appliances you are evaluating have limits.

What about backup and recovery?

Integrating the virtualization solution with your existing backup environment is critical. The virtualization appliance may have its own data-protection features; however, you don't want to create a whole new backup environment just to support those features. Integration with your existing backup environment is particularly important if your NAS shares are used to store data that needs to be protected for legal and compliance reasons.

The other critical question to consider is

how to recover from a failure of the appliance. For example, how is the appliance metadata and virtual file system backed up?

What are the security and authentication issues?

If you have a single authentication infrastructure across your environment, whether it is LDAP or Active Directory, you need to ensure the virtualization appliance not only supports, but fits in with, this authentication mechanism. You'll need to be able to migrate to your virtualized NAS environment with user profiles and permissions intact.

Encryption may be another requirement. If security is paramount, check to see if the virtualization appliance supports encryption of data at the source and in flight, and verify that the appliance addresses your security requirements.

When implemented for the right reasons, NAS virtualization will prove to be a valuable asset in your environment. □

Multi-protocol NAS virtualization

One aspect of NAS virtualization that is fairly new is the ability to abstract iSCSI resources, as well as NAS resources. Many NAS vendors offer iSCSI with their NAS appliances, with the iSCSI LUNs residing on the file system as special files. For example, a 100GB iSCSI LUN resides on the file system in the form of a 100GB file with special attributes. Imagine a situation where the virtualization appliance simply treats this LUN as a special file and presents it via its own network ports as an iSCSI LUN. At that point all of the features of virtualization (e.g., data mobility, migration, and security) can be "mapped" in the same fashion as file virtualization. □

Ashish Nadkarni is principal consultant with GlassHouse Technologies (www.glasshouse.com).

VENDORS MENTIONED

Acopia, BlueArc, EMC, NeoPath, Network Appliance, NuView

SNIA on Storage continued from p. 39

Applications and ILM management tools implement and automate the policies for their respective domains using available services. These data, security, and storage services, as illustrated in the figure, right, provide the tools and technologies by which IT delivers the infrastructure to meet service-level requirements that can flexibly change over the data's lifecycle.

The next stages of market maturation for ILM-based practices have two dimensions:

- Introduction of better automation and management tools, with centralized management platforms and practices integrated with enterprise applications; and
- Release of the initial specifications for two standards being developed by SNIA, the ILM-related portions of SMI-S and XAM, and their integration into products.

The first of these standards is SNIA's Storage Management Initiative-Specification (SMI-S) for ILM services. Just

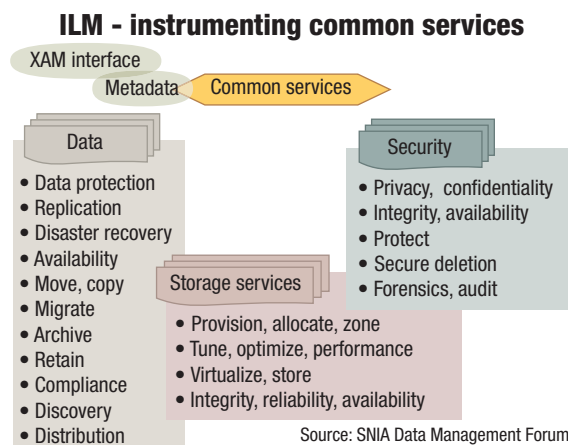
as SMI-S has provided standard management interfaces for managed devices and network elements, SMI-S for ILM services will allow heterogeneous services to be instrumented and driven by central ILM management tools and applications. SMI-S-based instrumentation of services is essential for automation of ILM-based practices. The first such specifications will appear in SMI-S 1.2.0, with more management interfaces to follow in subsequent releases.

The second standard is the eXtensible Access Method. XAM, which is currently being developed by SNIA, provides applications with a standard interface to storage, beginning with object storage systems such as content-aware storage (CAS), with the capability to write metadata relevant to ILM practices. As XAM becomes incorporated in a broad spectrum of applications, it will become strategic to ILM-based automation.

Although your ability to instrument and automate your infrastructure and

services will be limited until these standards are completed and adopted, there are many practices you can implement today. For example, you can build and automate solution stacks for ILM-based practices around applications such as e-mail or database archiving. You can implement tiered storage, data protection, compliance, and archiving solutions, and automate them using either ILM management tools or utilities integrated into virtualization platforms.

The most-effective way to get started with ILM is to have the CFO or CIO engage with a professional services organization for an ILM assessment of the effectiveness of data-center processes. This engagement also allows the professional services organization to engage with the data center, lines of business, and key enterprise stakeholders such as



legal, security, and records and information managers to help your organization take the first steps of collaboration and information classification.

For more information, visit the SNIA Data Management Forum's Website at www.snia-dmf.org. □

Michael Peterson is the program director of the SNIA Data Management Forum and president of Strategic Research Corp. (www.sresearch.com).