Information Lifecycle Management Roadmap

October 12, 2004

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Information Lifecycle Management Roadmap

The *Vision for Information Lifecycle Management*, ILM, developed by the Data Management Forum defines ILM as a new management practice for the datacenter. ILM is not a specific product nor is it just about storage and data movement to low cost disk. It is a standards-based approach to automating datacenter operations by using business requirements, business processes, and the value of information to set policies and service level objectives for how the supporting storage, compute, and network infrastructure operates.
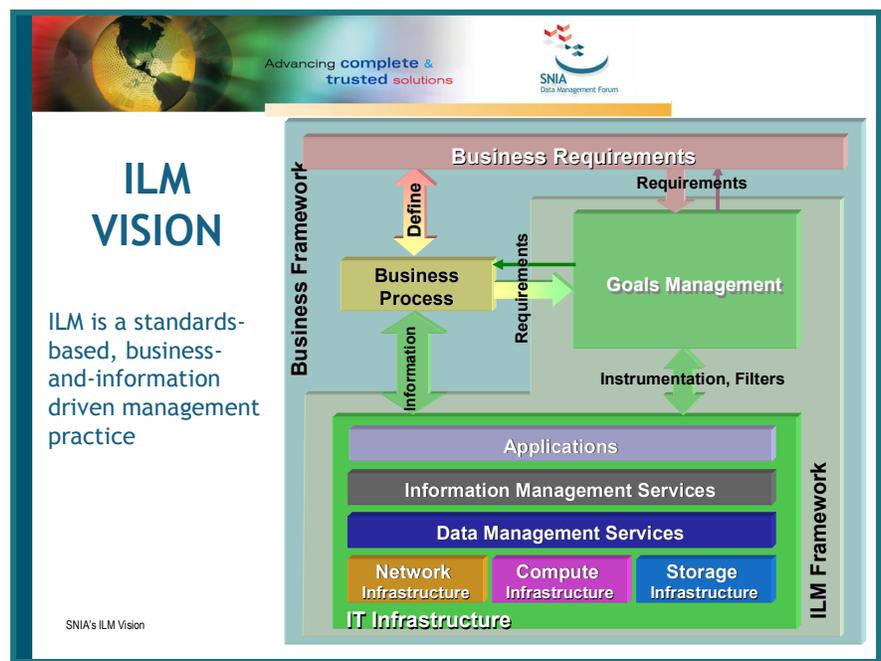
The key question that flows from this 'vision' of ILM is "How do we get there?" since these capabilities don't fully exist today. This is the work of SNIA and the Data Management Forum: to unify the industry towards a common goal, to develop the relevant standards, to facilitate interoperability, and to conduct market education around ILM. It begins by shifting our view from storage management to data and information management – the lifeblood of today's *business processes*. It expands by leveraging and cooperating with other industry groups to relate the management of infrastructure services to the same data and information. This process requires a roadmap to guide the way and enable us to measure progress.

Definition for ILM

To begin, consider the power of the DMF's definition and vision for ILM.

> *Information Lifecycle Management is comprised of the policies, processes, practices, and tools used to align the business value of information with the most appropriate and cost effective IT infrastructure from the time information is conceived through its final disposition. Information is aligned with business requirements through management policies and service levels associated with applications, metadata, and data.*

The ILM framework illustrates the definition. Goals Management abstracts capabilities from the IT Infrastructure in terms of policies and practices, and rationalizes those against business requirements and information value. This allows for automation of the IT infrastructure and its associated services. A critical goal of this ILM vision: a new set of data-centric standards based on SNIA's *Storage Management Interface, SMI*. The resulting automation sounds simple, but will require years of hard work to achieve heterogeneous interoperability.


ILM VISION

ILM is a standards-based, business-and-information driven management practice

SNIA's ILM Vision

Business Requirements

Requirements

Define

Business Process

Requirements

Goals Management

Instrumentation, Filters

Information

Applications

Information Management Services

Data Management Services

Network Infrastructure

Compute Infrastructure

Storage Infrastructure

IT Infrastructure

Business Framework

ILM Framework

Roadmap for ILM

The DMF has defined a roadmap to guide the deployment of ILM along a logical path that starts with best practices using today's solutions and leads towards ILM enabled practices using solutions developed in the future. The roadmap begins with the need to establish a strong foundation and culminates with fully automated, multi-site, heterogeneous "Enterprise ILM". The roadmap is divided into 5 Phases, each describing a distinct step towards an ILM-based operating environment and useable as a means to describe where in the process a particular datacenter's implementation stands. These phases are described below.
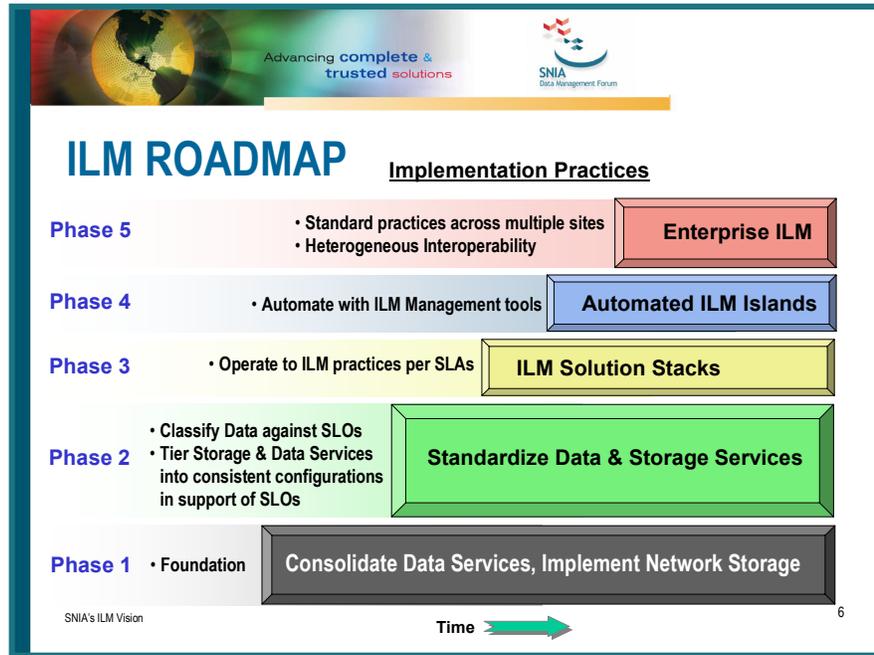
Phase 1: An ILM-based platform requires a strong foundation. The recommendation is to first consolidate your storage and data services around network storage. Reduce the variety and begin standardizing the configurations you are using. In other words, get lined up to move to Phase 2 by preparing the foundational architecture upon which you can build ILM-capable services. Define your data center architecture, its requirements, and what it is capable of doing.

Phase 2: The objective of Phase 2 is to standardize your data and storage services so that they are consistent, repeatable, and efficient processes. It begins with a data classification effort. Whether it is conducted for a single line of business or for the entire company, it is necessary to understand the *requirements* of existing and planned data first. This can, and should, include a ranking of applications – and their data, relative to one another, based on the value of that information to the business. This value can include intrinsic value based on how much money is lost due to unavailability of the data or the application, and it can include extrinsic values such as risk and company image associated with loss of the data. The value of the data then leads to the definition of Service Level Objectives (SLOs) – acceptable levels of service for performance, availability, operational recovery, disaster recovery, security, and more.

By associating the applications and data with standardized configurations that achieve the desired SLOs, you have a better opportunity to associate cost-performance appropriate configurations to your applications. Associating the value of information with the appropriate infrastructure and services is the guiding principle of ILM. And with consistent configurations, you have a better chance of achieving consistent, predictable, behavior at lower costs. Now your process begins to scale.

Today, many companies conduct classification engagements with professional services consultants who can help you through this potentially complex process. If you begin this yourself, then keep it simple. Understand and rank the value of applications and data to your company and match that against an inventory of existing and needed services. The most crucial step here is to get buy-in from across the company. Seek cooperation from legal, compliance, I.T., and the business units.

Tomorrow, some of this will be greatly simplified as SNIA defines standard sets of data classification profiles across the different dimensions of IT services that may be provided today. These profiles will provide a blueprint for how to organize data classification SLOs and will align your organization with how ILM will evolve in the future.

Phase 3: This phase introduces the concept of *"solution stacks"*. A solution stack conveys the idea of a set of homogenous data and storage services that have been instrumented to work together in an ILM context in support of an application and its valuable information. A solution stack is a complete (or nearly complete) ILM environment around that application. For example, based on the SLOs for a specific application and its information, define where active primary storage goes, where long-term online retention is held, the protection, operational recovery, and disaster recovery processes, and the regulatory and archive repositories. Begin operating this solution stack to ILM principles. This is the first step towards automating ILM, around an application and its information's requirements.

Phase 4: ILM automation will require the introduction of a new class of management tool that will automate I.T. infrastructure practices as homogeneous islands. Why homogeneous? Industry experience is that it is far easier to first instrument a single vendor solution stack. Phase 4 implementations require that the supporting data and information services plus the network, server, and storage infrastructure be instrumented so that they can be operated centrally. This requires standards and is one of the goals of the DMF's ILM vision, and it must leverage work currently underway in utility computing, grid computing, web services, and service level management by industry groups such as DMTF, OASIS, IETF, GGF, EGA, ITSMF, ITIL, and more.

Phase 5: Interoperability is achieved. Interoperability means IT products define their capabilities and provide instrumentation in terms of services provided – and one product can be replaced with another in an IT infrastructure even if it provides those services in very different ways. Interoperability is peaceful co-existence so that one data management product does not interfere with another product's services. And, yes, interoperability means defining interfaces so one product can hand off services to another where it makes sense. Phase 5 scales from the largest of data centers down to branch offices so that a single set of processes, practices, and policies to be set in place and managed throughout the enterprise.

The value proposition of this vision for ILM is in its ability to solve the overwhelming complexity problem in the data center, achieving substantial reductions in the TCO. By being based on business requirements and standards, ILM is the perfect framework around which we instrument and automate the data center. SNIA is engaged in the standards development efforts and needs your help to complete this work. Join us and participate in our online operations portal at www.snia.org/dmf.

****

Information Lifecycle Management Roadmap

By: Michael Peterson, President, Strategic Research Corp. - and Program Director, SNIA Data Management Forum

Edgar St. Pierre, Consulting Engineer, EMC - and Co-Chair DMF ILMI Technical Liaison Group and Co-Chair SNIA ILM Technical Work Group

About the Author(s)

Michael Peterson: is President and Senior Analyst of Strategic Research Corporation based in Santa Barbara CA. In this capacity, Mr. Peterson is a prominent network storage and system I/O industry spokesperson and leader, often called the "catalyst of the industry". Mr. Peterson's engineering, business, and marketing background achieved through 32 years of experience in the storage, software, and systems industries along with his years of research into the MIS community give him unique business, systems, and market expertise. He is widely known as the leading market analyst in the "datacenter infrastructure industry (Severs, System I/O, Datacenter Networking, Storage, and Management) and is the founder and past President of the Storage Networking Industry Association and is currently the founder and Program Director for SNIA's Data Management Forum.

Edgar St. Pierre: has over 26 years of experience in software engineering, including 8 years in storage software development and the previous 18 years in the communications industry. His experience in the communications industry was focused on network management product development at Prime Computer and Motorola/Codex, applying protocols and data models derived from participation in OSI Network Management Forum activities. At EMC, he has been responsible for requirements analysis and software architecture for several of EMC's portfolio of information management products. Edgar received his BS in Computer Science from Roger Williams University. He is currently co-chair of the SNIA DMF's ILM Initiative Technical Liaison Group, and co-chair of the SNIA ILM Technical Workgroup.

About the Data Management Forum

The Data Management Forum is an initiative of the SNIA focused on building a community of I.T. professionals, integrators and vendors for the purposes of being the leading authority and resource on data management infrastructure and information lifecycle management and to facilitate and promote interoperability among ILM solutions and data services. The online operations portal for the DMF is at www.snia.org/dmf. In addition, the DMF runs several online resource sites, www.ilm-info.org, www.dp-info.org, www.cas-info.org, and its conference site www.ilmsolutions.org.

About the SNIA

The Storage Networking Industry Association is a not-for-profit organization made up of more than 300 companies and individuals worldwide spanning virtually the entire storage industry. SNIA members share a common goal: to set the pace of the industry by ensuring that storage networks become efficient, complete and trusted solutions across the IT community. To this end, the SNIA is uniquely committed to delivering standards, education and services that will propel open storage networking solutions into the broader market.