

ILM and Tiered Storage

By: the SNIA's Data Management Forum

There has been much discussion and confusion around tiered storage from a business perspective, its usefulness for IT customers as a preferred cost reduction strategy, and vendor tendencies to call tiering "ILM" as though this is all there is to ILM. Not so! Tiering is an important component of an ILM practice, but not all.

Deleted: Dean Morash, Quantum Corporationfor

Deleted: about

Deleted: architecture

Tiered Storage, a Primer

Tiered storage means to establish a hierarchy of storage systems based on service requirements and cost. Why place inactive or expired data on your expensive primary enterprise arrays? What if there were a lower cost alternative? Well, there is. The cost advantage tiered storage offers to I.T. is immense and proven. Most enterprise implementations see a ROI in under 6 months with reductions in purchase and operating costs in the \$Millions per year.

Deleted: and also an attempt to demonstrate the need for amending the acronym to include an I for Intelligent Tiered Storage (ITS) so that it can join the other data migration schemes (ILM, DLM, HSM) that already have three letters. And finally, a more meaningful objective than to enter into the acronym fray is to establish the various methodologies of architecting a tiered storage strategy based on a divergent set of available vendor technologies and the deployment experience to date from the customer's perspective.¶

Adoption of tiered storage as a datacenter architecture is underway. This transformation is fueled by the explosion of low-cost and reliable Serial Advanced Technology Attachment (SATA) disk array systems. These capacity oriented systems have been successfully displacing more expensive SCSI and FC disk and tape in the data center over the last three years. The impact of SATA is born out by IDC's new forecast in which they predict that 40% of all storage capacity installed in 2008 will be capacity oriented technologies like SATA disk.

Deleted: Tiered storage is a hot topic amongst trade publications, analysts, and vendor storage professionals and will get even hotter in the near future. F

Formatted: Default Paragraph Font

Formatted: Font color: Auto

Deleted: based on Redundant Array of Independent Disks (RAID) architectures, t

Deleted: less expensive

Deleted: in

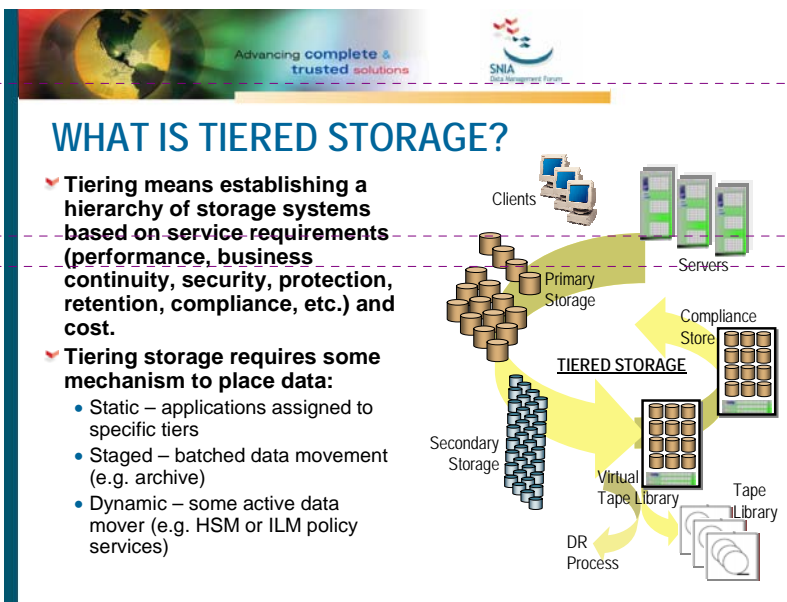
Deleted: According to

Deleted: With an additional tier of storage now available based on this technology trigger customers now have a plethora of additional targets on which they can store their data. These additional "targets" and the evolving data management software have begun to create a logical architecture "Tiered Storage" for many customers to adopt. According to IDC's Vice President of Storage Systems Richard L. Villars in a paper he authored in May of 2005, "IDC sees tiered storage as becoming the dominant organizational model for overall storage in the coming years".¶

Deleted: and migration

Deleted: created data

Now, add the intelligence of ILM-based practices. By applying retention, placement, protection, business continuity, archive, and compliance policies to information, based on its value, as it is created and as it ages the practices can be automated and operations costs reduced. Many new tools are available now to implement automated ILM polices and many others will come to market next year. This



is also the area where the SNIA is working on standards for ILM by extending the current SMIs (Storage Management Initiative Specification) standard. Once complete, SMIs for data, information, and security services will instrument the automation of ILM practices.

Definitions

Tiered Storage, according to SNIA-DMF: - *Tiering means establishing a hierarchy of storage systems based on service requirements (performance, business continuity, security, protection, retention, compliance, etc.) and cost. A company may implement a number of tiers such as primary disk, secondary disk, backup disk, archive disk, compliance disk, tape backup, tape archive, optical archive, etc.*

ILM, according to SNIA-DMF: - *ILM is described as 'information-based management, Or more expanded, a 'standards-based, business driven management practice'. ILM is not a specific service; rather it is about using the value of information as the basis for establishing policies and services for the information. It is about establishing requirements and operating to standard practices and service levels.*

The key differences between ILM and Tiered Storage

To begin with, tiering is not ILM. It is common to see the two confused. Tiering is how we arrange the repositories. This is analogous to saying backup is ILM or archive is ILM. ILM is not a specific service. It is not the architecture. In a tiering context, ILM is how we manage the repositories based on the value of information over time. ILM aligns the business requirements and business processes with the service level requirements in an automated fashion, from data creation to data deletion. Whereas, tiered storage should be viewed as one of many components in the implementation of an ILM-based practice.

Applications Support Tiered Storage

Many applications have the ability to operate to ILM-based practices. You can start small and easy. Begin with a practice like backup by implementing a backup array or virtual tape system connected to your tape automation and let the backup media manager and virtual tape system automate the tiers. Or, implement an integrated email archive system which integrates multiple tiers for primary storage and archive or compliance stores. Then, when you get more ambitious, free up some of that space on primary storage by deleting the expired data. Or allow the database to archive it to secondary storage. You'll find that the leading databases and some key add-on tools instrument and support tiering very nicely. Similarly, many ECM, Enterprise Content Mgmt, products do the same.

Want to do more? Start using SRM tools to profile your unstructured information. Find the active, inactive, reference, and expired data and then set up classes and operate to service requirements based on the value of the information. Move that inactive or expired information off primary storage.

Deleted: it is within reason to envision data being stored at varying tiers of storage media based on it's criticality, access frequency, duplication, retention and destruction requirements. ¶

Deleted: a definition according to searchstorage.com

Deleted: The assignment of different categories of data to different types of storage media in order to reduce total storage cost. Categories may be based on levels of protection needed, performance requirements, frequency of use, and other considerations.

Deleted: ¶

Deleted: a definition

Deleted: The policies, practices, services and tools used to align the business value of information with the most appropriate and cost effective infrastructure from the time information is created through its final disposition. Information is aligned with business requirements through management policies and service level associated with applications, metadata, and data

Deleted: more completely

Deleted: and manages

Deleted: an important set of building blocks

Deleted: strategy

Deleted: In a tiered storage deployment it may be appropriate to manage much of the data retention policies for different applications manually or in a semi-automated fashion and therefore it is more labor intensive.

Deleted: ¶

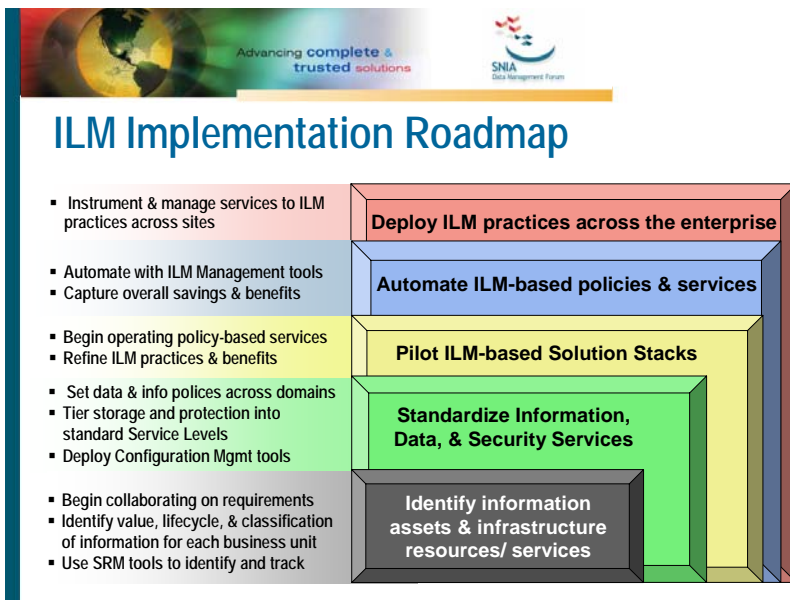
Yes, tiering storage requires some mechanism to place data and there are three basic methods:

- Static – applications assign information to specific tiers
- Staged – batched data movement (e.g. archive)
- Dynamic – some active data mover (e.g. HSM or ILM policy services)

The biggest challenge imposed by implementing ILM-based tiered storage is in getting agreement on the value of information and setting the classification requirements. This is I.T.'s holy grail. To have the business set the requirements and take responsibility for the cost. It takes collaboration with the business groups, the records managers, the security managers, and sometimes legal and finance to get this done, but it is worth the effort.

Tiering is a “Phase-II” activity in the course of implementing an ILM program. It is an important early step because so much benefit can be had which demonstrates the power of ILM as a management practice.

So, our recommendation is to add intelligence to you tiering efforts by integrating it into a broader ILM-based practice. You can get help with understanding this further from the SNIA’s End-User Council and the Data Management Forum at www.snia.org/dmf and at our upcoming conferences, Storage Networking World and Enterprise Information World.



By: Michael Peterson

Program Director, SNIA Data Management Forum
and Dean Morash, Quantum Corporation

October 2005

www.snia.org/dmf www.enterpriseinformationworld.com

Formatted: Indent: Left: 0.24",
Space Before: 0 pt

Deleted: What is most appropriate; a fully automated, semi-automated, or manual approach to policy management? This is somewhat the million dollar question and is often driven by the customer's IT environment and more importantly the preferences, experience of the administrators and comfort level with trusting automated provisioning software to manage the companies assets. With at least three different levels of data protection now available (see paragraph below "Correlation Between The Tiers and the Data Protection Spectrum" and the supporting diagram), by allocating applications based on business requirements IT administrators can save as much as 25 to 50% when compared to storing data on high-performance, high-availability FC RAID arrays. ¶

Deleted: setting and managing policies for a diversity of applications and media types. It is often a question of where the policies get managed as well as the extent of automation. Policies can be established and managed at three different layers within the storage network: 1. at the application itself (becoming more and more common) 2. the data migration software layer 3. integral to the disc storage array. ¶

Policy management at the application layer ¶
This appears to be nirvana, having applications that allow IT administrators to enable policy management at the application layer based on business process requirements and retention requirements imposed by the clients (i.e. finance manager, corporate attorneys, CTO and the Compliance Officer. This would enable complete access to all migrated and copied data from the application layer regardless of what tier of storage the data is stored at so long as the access to the data stubs is served back up to the applications. ¶

The importance of Data Migration (HSM) software to manage retention policies¶
Unfortunately, all applications are not this sophisticated and so the responsibility for policy management is often driven down to the data migration/HSM layer to assist in managing these retention policies. Numerous companies have created special purpose data migration software based appliances that provide a more universal approach in support of numerous structured and unstructured data coming from a diverse set of applications and operating systems. ¶

Correlation Between the Tiers a ... [1]

setting and managing policies for a diversity of applications and media types. It is often a question of where the policies get managed as well as the extent of automation. Policies can be established and managed at three different layers within the storage network: 1. at the application itself (becoming more and more common) 2. the data migration software layer 3. integral to the disc storage array.

Policy management at the application layer

This appears to be nirvana, having applications that allow IT administrators to enable policy management at the application layer based on business process requirements and retention requirements imposed by the clients (i.e. finance manager, corporate attorneys, CTO and the Compliance Officer). This would enable complete access to all migrated and copied data from the application layer regardless of what tier of storage the data is stored at so long as the access to the data stubs is served back up to the applications.

The importance of Data Migration (HSM) software to manage retention policies

Unfortunately, all applications are not this sophisticated and so the responsibility for policy management is often driven down to the data migration/HSM layer to assist in managing these retention policies. Numerous companies have created special purpose data migration software based appliances that provide a more universal approach in support of numerous structured and unstructured data coming from a diverse set of applications and operating systems.

Correlation Between the Tiers and the Data Protection Spectrum

There are a number of different versions of the data protection spectrum depicted below in a variety of storage periodicals. Arguably there are no wrong or right answers. In this diagram you will notice the addition of Continuous Data Protection (CDP) to the model. The principle purpose of this diagram is to associate the three tiers of storage with the underlying technologies. Hopefully this diagram provides some healthy dialogue regarding tradeoffs for customers and vendors alike regarding which technology to deploy to solve specific point problems. Some of these key tradeoffs are cost, availability, throughput, security, scalability, portability, etc.

>>>Please mark diagram Copyright Quantum Corporation

