

Managing the Lifecycle of Your Data

A tiered storage strategy can help IT departments better manage their data and cut storage costs.

Leaders in the storage market are touting a strategy that will help IT managers better organize their data, which will make the backup and archiving process much faster and result in substantial cost savings.

Information lifecycle management (ILM) is the combination of business processes, policies, and disk storage hardware and software that an IT organization uses to most cost-effectively and strategically handle data, from the moment it's created to its disposal.

An element of ILM is tiered storage, in which the newest or regularly accessed data is stored in higher-end disk storage devices and unused data is archived in a second tier using lower-cost disk storage devices. Storage management software helps IT organizations manage the flow between the different storage tiers.

"Data is growing so fast that you can't afford to treat all the information the same way," says George Symons, chief technology officer of information management at EMC. "You can't just keep the data on the most expensive hardware. You have to manage it properly."

As companies grow, IT managers today are faced with an unending avalanche of data. In addition, new compliance regulations, such as Sarbanes-Oxley, and the threat of lawsuits require companies to keep better records of e-mail and other corporate data and have the ability to quickly retrieve that information.

However, most companies have not begun to reconfigure their IT environments with different tiers of storage, says Tom Valiante, vice president of Americas Reseller Channel, Hitachi Data Systems (HDS). According to technology analyst firm IDC, less than 20 percent of corporate data requires storage in high-performance disk storage systems.

"A lot of customers are using standalone storage, so there's a huge opportunity for them to tier their data, make data management easier and reduce their costs," Valiante says.

Benefits of Tiered Storage and ILM

By moving dormant, nonessential data from a primary storage system to another storage tier, IT organizations can reduce the amount of data that they need to

regularly back up. That, in turn, speeds nightly and weekly backups of a primary storage system and diminishes the chance for backup errors, according to a report by the Enterprise Strategy Group (ESG), an analyst firm in Milford, Mass. Off-loading data from primary disk storage will also boost its performance. Because it will hold less data, users can find and access data more quickly.

Having a separate tier for archived data will also make it easier and faster for IT departments to find older, unused data when it is needed, according to ESG. Cost savings are another bonus. Companies will spend less for primary disk storage and spend less IT time on administering storage systems. Storage hardware made specifically for archiving data is generally easy to use and requires less IT management, the ESG report found.

Strategies like ILM and tiered storage are still in its infancy, but will become more important in the coming years as IT organizations learn about its benefits, says Marc Farley, analyst at Building Storage Inc. and author of the book, *Building Storage Networks*.

"Today, there's a lot of hype around ILM. We're still in the early part of the adoption stage, but there's no question that data management is huge," he says.

Farley predicts that the technology industry will focus the next 10 to 15 years on safety and security issues, and that includes proper data management. Archiving data will help companies comply with regulatory requirements and protect them from litigation, he adds.

"It's a part of managing risk," Farley says of good data management. "If you can't produce the electronic documents that are required in litigation, it may cost you a lot of money."

In fact, 42 percent of 485 companies surveyed by ESG in January said they had to search and retrieve electronic records because of a lawsuit or regulatory discovery process. The study found that 54 percent of respondents failed to produce requested records.

Hitachi's Storage Products

While EMC uses the term "ILM," HDS prefers to call it "tiered storage." Regardless of the name, the goal is the same: deliver hardware and software solutions for IT »



ILM is a strategy for classifying different types of data, and migrating those data categories to the right storage medium at the right time.

departments to better manage the growing reams of data in their enterprise.

To build different tiers of storage, companies should store their most recent, critical and frequently accessed data on high-end Fibre Channel disk storage products. For archival data, companies should move them to lower-cost Serial ATA (SATA) disk storage drives, according to Hubert Yoshida, chief technology officer at HDS. "Using lower-cost media will reduce costs," he says.

To help companies simplify data management, HDS takes an approach called "storage virtualization" that allows IT departments to centrally manage multiple storage devices from a single Hitachi system, including storage arrays from competing vendors.

In this scenario, a smart device called a controller sits in front of external storage systems. Software embedded into the controller, combined with Hitachi's HiCommand suite of storage management software, gives IT organizations one central view and one central location to manage all their data, Yoshida says.

In the past, if companies wanted to upgrade their storage systems and unite their data, IT managers had to migrate all the information into the new storage devices, a time-consuming task that also resulted in system downtime, he says. The HiCommand software suite allows IT managers to unite new and old storage systems, create different tiers of storage and move data between those tiers without disrupting operations. "No one else has this capability. We can do it dynamically," Yoshida says.

HiCommand gives a visual map of a company's storage devices and includes a "policy manager" that automates the movement of data between different storage tiers, he says. For example, a company can create a policy to move data from the first tier to a second tier after data is 30 days old and then from the second tier to a third tier after 60 days, he explains. The software also features a "tuning manager" that monitors storage capacity and alerts IT managers of potential problems, such as a storage device that's about to reach capacity.

HDS's storage virtualization architecture solves the problem many companies have of having different islands of storage systems that are underutilized, Yoshida says. The firm's architecture allows companies to take advantage of underutilized storage resources before spending money on more storage systems, he says.

Hitachi offers a family of virtual storage controllers, the TagmaStore entry-level USP 100, midrange 600 and the high-end 1100 models, and the midrange TagmaStore NSC55, which is aimed at small- to mid-sized companies. Controllers are also available with built-in storage, so if companies want to create storage tiers, the controllers can serve as primary storage, Yoshida says. The NSC55, for example, has a capacity of between 365GB to 72TB of data.

The company's family of storage devices, which include the TagmaStore AMS (Adaptable Modular Storage) 200 and AMS 500, can serve as secondary- and third-storage tiers. Both products support Fibre Channel and SATA, as well as offer RAID (Redundant Array of Independent Disks)-6 protection. The AMS 200 features a capacity of 30TB on Fibre Channel and up to 40TB on SATA, while the AMS 500 can support up to 67.5TB on Fibre Channel and up to 88.5TB on SATA.

Alternatively, smaller companies that are just starting out, don't have a lot of data and don't need the controller technology, can build a tiered-storage architecture by using the AMS 200 and AMS 500 as standalone products, and using the HiCommand software to manage the data.

EMC's Offerings

From EMC's point of view, IT organizations can deploy ILM in three steps: First, set up a tiered infrastructure. Second, use software to manage and move the data to the different tiers. And last, classify the data, which means IT departments define each piece of data by placing it in different categories. That will allow IT managers to use software to automate the movement of the data into different tiers, depending on the type of information, its importance and how frequently it needs to be accessed, says Symons, EMC's chief technology officer of information management.

To build a tiered infrastructure, a variety of EMC disk storage products are available for storage and backup, including EMC's CLARiON family of networked storage systems, from the lower-end CLARiON AX150 to the higher-end CX3 series.

EMC's Centera archive storage technology, which the industry calls "content addressed storage," stores a wide range of content, including e-mail archives, computer-aided designs, electronic documents, images, voice archives and X-rays. It's a disk-based WORM (write once, read many) device that has the intelligence to seek out common files, such as five copies of the same Microsoft PowerPoint presentation, and save it only once, Symons says.

Centera, which uses SATA drives and can store between 7.7TB to 41TB of data per cabinet, executes policies, such as how long certain data must be retained before it's deleted. Even though the data is moved from primary storage, users have a search capability on Centera that allows them to search for files and quickly access the data, he says. For disaster recovery planning, IT organizations can replicate a Centera system to a backup Centera system. "What you want to do is put the information on the appropriate storage tier, so you can exploit the value of each of the tiers," Symons says.

After creating a tiered-storage environment, EMC offers software to manage the content and move it between tiers. For example, EMC DiskXtender allows IT staff to set policies to automatically move unused files from primary storage to an EMC

Short on time and staff?
Ask about CDW technology services to bolster your IT effort.

Centera device. Similarly, EMC EmailXtender is an e-mail and instant messaging archiving software that stores inbound and outbound e-mail and messages into Centera. If litigation occurs, requiring the company to access archived e-mail, IT staff or management can search e-mail by date or specific keywords.

The third step to ILM is classifying data, Symons says. EMC currently relies on its own consultancy practice and partners to help customers classify their data, but later this year, the company will ship software that automates the process.

With the software, a company can create rules for different types of data, such as financial records, e-mail or encrypted information. For data that's considered classified, for example, a company can create a rule that requires the information be stored for three years. Based on the classifications found on a document's

metadata, the software will automate the process of moving the document between tiers of storage, Symons says. "It will move, retain or delete files. The purpose of this step is to automate the process and make it less people intensive."

EMC and Hitachi began talking about ILM and tiered storage several years ago, but only now are IT organizations beginning to understand their importance in helping them streamline and simplify data management.

"Historically, companies have just thrown more disks at the problem, but with data growing somewhere over 80 to 100 percent a year, organizations are forced to look at a tiered infrastructure for information," Symons says. "Now, enterprises see step one — tiered storage — as critical and step two — the software that manages the data between tiers — as something they're looking at." ♦

