

Arkivio - December 2003

ARKIVIO's auto-stor version 3

A Start-Up Ahead of the Pack

It seems that almost every vendor in the storage market has jumped on the Information/Data Lifecycle Management (ILM/DLM) bandwagon. ILM processes address an organizations' need to treat data differently, according to the value of that data to the business, throughout the lifecycle of that data. This is not a new concept, a key aspect of ILM, HSM (Hierarchical Storage Management), has been practiced in mainframe environments for many years. However, ILM for heterogeneous, open system storage environments in general has required services organizations to tie multiple vendor solutions together to provide data protection or archiving solutions.

A large number of storage vendors have realized that ILM/DLM for open systems is not just a nice to have, pie in the sky dream that end users share, but rather a mission critical business process that users are now demanding. Employing effective ILM/DLM processes can help enterprise organizations reduce management and capital expenditure costs, ensure regulatory compliance and avoid costly litigation.

Arkivio is a start-up that recognized this need well ahead of the majority of vendors now touting ILM/DLM solutions. ARKIVIO has just released the third revision of the ARKIVIO auto-stor solution, adding a number of features that enhance an organizations ability to automatically manage and protect data throughout its lifecycle. Arkivio's auto-stor solution enables organizations to automatically move/migrate data to designated resources based on user defined policies. The Arkivio software allows for multiple levels of "data valuation", where various criteria are weighted before any action is taken. This granularity elevates the software well beyond many of the simple data movers/migrators that are available today. ARKIVIO's auto-stor is an intelligent Automated Data Migration (ADM) solution that can help users realize their ILM goals of reducing costs and more effectively protecting data throughout its lifecycle.

Real Results

In ESG terms, ILM is not a product, it is a business process. Organizations must first assess how their storage resources are being utilized, and determine what data should reside on what resources. Organizations

must also determine retention periods and deletion criteria. These decisions are based on the perceived value of the data; determined by the relevance of that data to the business, availability requirements and compliance or corporate governance regulations. Once these decisions are made, organizations can then set policies and invoke ILM processes in order to move, migrate and protect the data according to these policies.

Enterprise organizations have realized significant cost savings and streamlined management operations by implementing ARKIVIO's auto-stor solution. One of ARKIVIO's customers was struggling to keep up with storage growth (100% each year), while adhering to corporate data retention policies. In addition, the IT design group was under a corporate mandate to reduce costs and find efficiencies with existing resources. The customer first looked to better manage the data on 14 file servers in their corporate data center.

The customer is utilizing the ARKIVIO auto-stor solution to migrate data from the primary storage attached to the file servers to secondary ATA arrays. Policies have been set up to move data when it has reached a defined age, and has not been accessed in a defined period. With these simple policies, this customer estimates they will be able to move 1.5TB of data off of primary disk, and reduce primary storage expenditures by over 33% in the first year. Their estimated cost differentiation between primary and secondary disk is over 3x, so being able to continuously move data to secondary disk will result in significant cost savings over time. In addition, now that the data has been moved off of primary disk, the customer has reduced their backup window for these servers by 40%.

This is just one example of how organizations can benefit from running the ARKIVIO solution. ARKIVIO's auto-stor has a number of features that this particular company has yet to take advantage of such as replication and data valuation. Regardless, the customer expects that they will net what they paid for the solution within a year (this is beyond ROI; this is savings in addition to what they paid for their current implementation of auto-stor). Although a number of vendors are now pitching their ILM solutions, ARKIVIO is truly delivering on the promise of cost savings and ease of management.

The Product

ARKIVIO's auto-stor is what ESG refers to as an Automated Data Migration (ADM) solution. The distinction is important, as so many vendors have different definitions of ILM solutions; it is easy to confuse what each solution does. An ADM solution provides the ability to assess and monitor data (SRM like capabilities), set data migration policies, and enforce those policies by moving/migrating the data. These features are the core of an ILM process, however the distinction is that end-to-end ILM processes (being touted by many large storage vendors) will also integrate in a number of other products, including storage management, provisioning, and data management applications such as snapshot and backup. In addition, a few vendors are offering content management and application specific solutions to their ILM portfolio. While ARKIVIO may not have the same portfolio of offerings as HP, IBM or EMC, they do provide a core ADM solution (available today), which addresses key aspects of ILM- classification, policy based migration, movement and replication of data. ARKIVIO has integrated their software with backup applications to ensure that migration prior to a file being backed up does not occur. This level of integration ensures that the ARKIVIO software can be an integral piece of the end-to-end ILM process. Further integration with front end content management solutions will enable users to utilize the ARKIVIO software in ILM processes.

ESG has outlined a five step process to implement ILM. ARKIVIO's auto-stor enables users to go through these steps of assessing (discovery), organizing, classifying data, then creating and implementing these policies. A key strength to ARKIVIO's solution lies in its granular classification and policy management features.

The first phase of the process, discovery, can be done in a number of ways depending on the way an administrator wants to implement the product. The ARKIVIO auto-stor architecture is very scaleable, utilizing either agent less or agent based methods of data collection and management. The Central Server software resides on a Windows 2000 server, and can simply use standard CIFS and NFS protocols to discover and evaluate file level information on each managed host. Server Agents can be installed on a host to perform data collection and invoke policies set up by the Central Server. Administrators may want to use the Server Agents if they want to reduce management traffic on the network, or reduce possible security issues that may occur if commands are constantly being sent via an open network.

In order to scale in large enterprise environments, ARKIVIO utilizes Remote Server Agents, which collect data, monitor, and report and enforce policies on behalf of the Central Server. The Remote Agents are in constant communication with the Central Server, which acts as the master in large scale environments.

Once discovery is complete, administrators can group files and volumes in ways that allow them to organize data according to business needs. Files across the network are automatically grouped by type; administrators can also set up custom groups based on criteria such as owner, size, and file path. This way, administrators can set policies on certain groups of files across the network, which aids in ensuring compliance and lowering management costs. The ARKIVIO software also automatically groups volumes by host and device. Host based volume grouping shows all volumes associated with a host, where as device based volume grouping shows all volumes on a single device (such as an array on a SAN) or on multiple devices that are allocated across a network. Administrators can also set up custom volume groups, which may reflect department or application allocation. By being able to see storage allocation on a volume level, administrators can better understand usage patterns and determine target destinations for moved/migrated data.

A unique aspect of ARKIVIO's auto-stor solution is the Policy Automation Engine (PAE), which enables data valuations according to a number of criteria. Unlike other data movers that can only move data based on a single criterion such as age or last time accessed, ARKIVIO's PAE allows users to set policies that weigh a number of criteria simultaneously. This is important for a number of reasons, administrators do not want to migrate data to secondary disk or move to archive if it is still deemed important by business users. A file may age to a point that meets a pre-defined policy requiring that it be archived, however if users have been accessing that file, it should not be moved to a place where it is more difficult to access. Because ARKIVIO can weigh the policies, their solution can determine if the file should be moved based on multiple criteria, each with a different weighting value. In the case outlined above, if an administrator sets the policy to reflect that user access outweighs age, the data will not be moved (or may be placed on secondary storage instead of archived).

ARKIVIO refers to this weighing system as a Data Value Score (DVS). The auto-stor software also provides a Storage Value Score (SVS) for volume groups. SVS is determined by volume attribute, such as cost or utilization.

Storage Knowledge - Business Results

DVS and SVS scores are internal to the PAE and determine when data should be moved, migrated or replicated. This level of flexibility helps meet the goal of making sure important business data is always on the right resource and available to end users.

Finally, once all of the policies are in place, the ARKIVIO software will automate actions based on those policies and scores (DVS and SVS). Decisions are made to migrate, relocate, replicate or delete data. After any of these actions occur, the solution will provide reports which show the results of the actions.

The ARKIVIO software performs two types of migration activities, tag migration and link migration (the type of migration is determined up-front when policies are set). Tag migration will leave a stub when a file is moved to another volume. If a user needs to access that file, the file is automatically moved back to the original volume. This is good for financial data that may become stale until year end when it is needed again for year end reports. Link migration will relocate data to another volume and leave behind a symbolic link that provides a direct-read of the files from the original location. When a user needs to access this file they will “see” the same path, however the data will stay on the new volume.

The ARKIVIO software can also relocate data, permanently moving the data to a new location. This method of data movement is used for archiving and consolidations purposes. Auto-stor can also replicate files and volumes based on pre-defined policies. Replication may be performed for on-line backup or DR purposes. By

targeting only specific data for replication, administrators can be assured that their most valuable data is always available. Finally, the Arkivio software can delete files based on criteria set up by the administrator to “clean-up” unwanted data such as temporary files, log files, or data that has passed its retention limit.

One really nice feature of the auto-stor software is the ability to simulate all of the policies and report on the results before any actual action takes place. This way, administrators can be assured that they have not set up policies that will inadvertently overload a volumes capacity or delete valuable files. They can also determine the ROI impact of migration policies by simulating the capacity that will be freed on primary storage devices.

The Bottom Line

ARKIVIO is offering a very strong software solution to help users reduce costs and more effectively manage their data. ILM may mean many things to many different users, but at its core, ILM is about intelligently managing data according to user defined policies. ARKIVIO’s auto-stor solution enables intelligent data migration, movement, and replication and archiving of files; all these capabilities can help users more effectively utilize resources, comply with regulations or corporate governance and reduce management costs. Bottom line- ARKIVIO can help users realize the promise of ILM.

Author: Nancy Marrone-Hurley

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Storage Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Storage Group, Inc. and is intended only for use by Subscribers or by persons who have purchased it directly from ESG. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Storage Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482.0188.