Network Computing's storagepipeline

Questionnaire

ENTER
CATEGORY
Comparative Review

Dear Data/Information Lifecycle Management Solution Provider,

The following is a questionnaire for our use in understanding and evaluating the numerous solutions that are collectively called data or information lifecycle management solutions by vendors. Our working definition of these types of solutions is “a software and/or hardware solution for migrating data through the storage infrastructure using automated policies that match data content and/or access characteristics to storage platform capabilities and costs.” If you would like to suggest amendments or revisions to this definition, please make your suggestions as part of this questionnaire.

Please respond as per the cover letter attached to this questionnaire. You can edit this document directly to provide your information.

Usage Scenarios

Since the capabilities of these solutions may vary, we are going to assess their effectiveness in the context of different usage scenarios.

Capacity Utilization Efficiency. These users mostly use a D/ILM solution to place data on platforms or media that are least costly, most capable and best suited to their access characteristics and “inherited” requirements in terms of retention, security, and criticality. Automated data migration may also be seen as a strategy for reducing TCO by reducing or capping administrative staff requirements.

Capacity Allocation Efficiency. These users are primarily interested in D/ILM as a mechanism for sharing capacity in the most efficient way and for eliminating unnecessary replication or junk data – often to defer additional hardware purchases. Automated capacity allocation may also be seen as a strategy for reducing TCO by reducing or capping administrative staff requirements.

Disaster Recovery and Business Continuity. These users seek to use a D/ILM solution to segregate data by its restoration priority in the wake of a disaster and to route data to the appropriate protection process (mirroring or continuous data protection, snapshot or disk-to-disk replication processes or tape backup). These users may also seek to leverage D/ILM to provide input to change management processes designed to keep plans up to date with burgeoning data.

Questions

Development Status and Objectives. Describe your vision of D/ILM and identify what features or functions your platform currently provides and what features or functions are still in development.
ADIC differentiates between DLM (Data Lifecycle Management) and ILM (Information Lifecycle Mangement). DLM, which ADIC supplies through its StorNext Management Suite data management software, provides an automated system for managing files on specific storage resources. DLM matches data access, protection, and retention policies to business requirements, and it provides a system for changing the way that data is treated over time as the business requirements change. DLM provides data management value for a number of uses, one of which is providing a storage management foundation for ILM.

ILM uses applications that can look inside files, and use the content of them to help an end-user determine their value and different use requirements. ILM applications are very specific, dealing with a particular application area (SEC regulations for email retention, for example), and end users are likely to use several different ones. ILM can help users understand different requirements for files, but to actually manage map files to storage resources—to actually move data, protect it, access---ILM needs to use DLM.

Development Partners. List the vendors with whom you are working to deliver your D/ILM solution functionality.

ADIC is working with several partners in different areas (some will provide ILM applications, other specific resources): EMC, LSI, StorageTek, HP, Cray, CGG, among others.

Data Naming Scheme. What is your method for identifying data storage requirements or characteristics and for using those requirements or characteristics in building policies?

Many of the specific requirements will come from ILM applications as they mature. The ADIC DLM solution understands the characteristics of different storage resources and the file management policies that provide different levels of protection, access, retention, levels of service, and cost. Users can set policies from standard access and age patterns within the DLM environment, but ILM will add additional application-specific capabilities.

Access Frequency. Identify the mechanism by which your solution identifies how frequently specific files or datasets are being accessed and whether and how access frequency is used in migration policies.
The foundation of ADIC's DLM system is a unified storage environment that links different platforms and resources using a file system. Since we own the file system, all the metadata levels of information (access, age, owner, etc.) are directly available to use for data management policies. The resources in the unified environment include different classes of storage (performance disk, value disk, tape) from different manufacturers, and we can present that environment to applications in different ways (as a standard disk resource or as a virtual tape library, for example).

Storage Platform Characterization. Explain the mechanism that your solution provides for characterizing the performance capabilities and costs of specific hardware platforms for data storage so that this information can be used to target the appropriate storage platforms as destinations for automatically-migrated data.

ADIC's SNMS provide a system for users to identify and name different storage pools made up of resources that users can associate with different cost, performance, access, and resiliency characteristics. The data management policies in the SNMS suite move data between these pools, creates copies, and provides for access across them.

End of Useful Life. Explain how your platform facilitates the automated removal and clean-up of data that has outlived its useful life and restoration of freed capacity for use by applications.

The key to re-use of resources once files are deleted is solved by the SNMS's use of its file system to provide a single name space that is available to all hosts. Since the file system owns all the storage, it can easily and directly re-allocate blocks to new uses once data in them has expired.

Policy Articulation. Describe how policies are created and how they are applied to existing data.
Policies for treating data are established through a "style sheet"-like interface, available in both a GUI and a CL interface. Policies are applied today to defined groups of files—specific

filesets or specific directories. Additional options based on file meta-data can also be supported.

Device Support. What storage devices does your product support? Are their any proprietary devices (controllers, arrays, HBAs, switches, SAN topology, virtualization products, etc.) that are required for your solution to work?

The ADIC SNMS system is a software-only solution that is vendor and technology independent—it does not require any proprietary hardware platforms or storage. It supports UNIX, Linux, and Windows platforms and operating systems; different storage classes and interconnects (FC, iSCSI); different brands of disk (EMC, LSI, HP, IBM, and many other disk vendors for FC or SATA); different fabrics (can mix switches); and different tape types (LTO, AIT, 9x40, DLT from StorageTek or ADIC).

Resource Consumption. How much bandwidth and server CPU "overhead" is introduced by your solution (e.g., to support polling processes, migration processes, agent processing, access frequency counting, etc.). SNMS uses an out-of-band metadata server so that access and transfer operations can take place at wire speed with the least possible added overhead when speed is required. The system does require a dedicated resource operating on any one of several UNIX, Linux or Windows platforms.

Flexibility. Does your solution have multi-vendor support? Can data, once integrated into your D/ILM scheme be migrated readily between your solution and other solutions in this space? Is you solution interoperable with other solutions in this space?

The SNMS unified storage environment is file system based, supports multiple platforms (UNIX, Linux, Windows), provides a full API, and can support virtually any applications that can write to a file system.

Speeds and Feeds. How should a prospective customer compare the relative performance of competitive solutions in this space? What are the appropriate performance measures or metrics to use in evaluating competitive solutions?

For systems of any size and any rate of growth, a key measure will be providing a range of access times based on quality of service needs across different platforms and operating systems. For the fastest requirements, access needs to be at raw disk I/O rates and should not require translation between different operating systems. We think that means ownership of the storage by a file system like the StorNext FS. And to get raw disk I/O, we think that the system should manage out of band—not introduce a gateway or other in-band virtualization appliance.

Interconnect Support. List the network or fabric interconnects that you support. Is there an optimal interconnect for data movement for your solution? What about for management?

We assume a storage area network interconnect; SNMS servers also communicate via Ethernet.

Protocol Support. List the storage networking technologies supported, including FCP, Ethernet, iSCSI, FCIP, iFCP, Parallel SCSI, SAS, and network file system protocols (NFS, CIFS, HTTP, DAFS). Our mechanism for creating the unified environment is a native SAN file system that can be exported via CIFS or NFS. We provide shared storage and automated file system policies via FC, iSCSI, FCIP, iFCP or other interconnect. We do support different storage devices (FC and SATA as well as different tape libraries and drives).

Data Type Support. List the data types supported by your solution, including file types and file systems, databases, and hybrids such as email. Comment on the granularity of your

D/ILM solution: will it support the migration of database components or subsets, subsets of email files, etc.?

We support both structured and unstructured data. The approach of DLM in general we believe should operate primarily at the file level while providing block level access,

Cost. List the cost of your solution or provide some means for calculating cost for a specific environment.
Good starting price would be $21,000 for a Windows system with 4 servers and 1 TB of managed storage.

Standards. List any relevant open standards upon which your product is built. The SN File System provides interface for POSiX and Windows.

Other features. List other pertinent aspects of your solution.
Key elements are heterogeneity—specifically the ability of the StorNext unified storage environment to allow storage to be shared transparently between different platforms, operating systems, and applications to keep performance high and users' product development cycles short.

Policies implemented allow data to be moved between resources, automatically protected, and made available transparently to different hosts. The value to the end user is that they save time and money by having a DLM environment that puts data in the right place at the right time—and that improves overall resource utilization.

The heterogeneity and integrated management of SNMS makes it easy to transparently add new resources over the life of the data and system—whether it's adding new classes of storage, new platforms, or newer versions of technology.

Miscellany: Please note any additional information that you think would be worthwhile for prospective customers to consider about your solution or other solutions in this space.
ADIC's DLM technology is in active use today, providing unified storage, automated protection, and dynamic management for industries with very large, fast growing, and mission critical datasets (science and engineering, digital media, oil and gas, broadcast and entertainment, financial services companies). Data growth, access requirements, new levels of storage resources (like ATA disk), and the emergence of ILM applications—which need DLM to actually map requirements to resources—are now showing how DLM can provide value to a much broader range of IT organizations.