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.

The explosive data growth coupled with federal mandates to retain data records for longer periods of time is fueling IT spending on ILM solutions. ILM is not a product or a solution but a strategy that combines processes, people and products to effectively manage the most critical asset of the company, its data. The data in an organization can be classified into three categories; unstructured data such as documents and images, semi-structured data such as email and structured data stored in databases. Most vendors today are focusing their ILM solutions on unstructured and semi-structured data. The more pressing issue for their customers, however, is how will they manage the lifecycle of

their most business-critical data - their structured data. Princeton Softech brings ease in dealing with ILM of structured data through database archiving.

Structured data resides in databases supporting enterprise applications (Oracle, Peoplesoft, JD Edwards) and drives every business. Loss or slow retrieval of structured data, which is growing 125% annually according to the Meta Group, can cost a company millions in revenue. Overall, this is the data that drives the heart of every business.

Princeton Softech, a pioneer of database archiving software for structured data and the market leader (with 56% market share according to Gartner), is helping customers such as Bank of New York implement a comprehensive ILM solution across the entire enterprise, encompassing all major ERP, CRM and business applications.

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

NetApp

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

Princeton Softech provides customers with strategy sessions to discuss ILM deployment. The first step is data and application classification, followed by data archive business specifications and data access requirements. Built into these sessions are an assessment of customers' data environments and implementation plans.

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.

In addition to the strategy sessions described in the previous question, Princeton Softech can discover and profile the current data usage through our software.

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.

Princeton Softech's Active Archive Solutions enable the customer to store the archive data in the most cost effective medium which can be online in an archive database, near-line on a file server, off-line to tape or disk based WORM devices. The appropriate storage media selected will be based on the business value of the data and the data access profile. Princeton Softech offers strategy workshops to assist customers in developing this profile.

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.

As part of the archive implementation, the data retention periods are reviewed to determine how often and when data is archived. Included in this process is a recommendation on when data should be deleted. Once this has been identified, the archived data may be automatically deleted.

Policy Articulation. Describe how policies are created and how they are applied to existing data.

Princeton Softech's Active Archive Solutions include application specific solutions which are packaged with all the business archive rules. These out-of-the-box solutions can be easily customized to meet the specific needs of the business using a graphical user interface. Archiving a custom in-

house application leverages the same user interface to define the custom archive business policies. All of these policies are used to perform archiving on a regular process, enabling the database to be maintained at a manageable size.

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?

Princeton Softech Active Archiving™ solutions support storage devices including SAN, NAS and CAS.

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.).

N/A

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?

Yes, Princeton Softech solutions support heterogeneous environments (databases and operating platforms) on both the production and storage side, increasing performance, availability and ROI for enterprises. For example, Princeton Softech's Active Archiving Software integrates well with EMC, IBM, HP and NetApp.

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?

Performance characteristics are always unique to each environment and application. What is critical to measure is the end-to-end performance of the data being archive and placed in your specified destination of choice. In addition, what are the resources consumed to achieve this process and what impact does the solution have on the production system.

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?

Princeton Active Archive solutions comply with standard network protocols and are compatible with these networks.

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).

Active Archive Solutions comply with standard storage network protocols and are compatible with these networks.

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.?

Princeton Softech support and identifies the subsets of data to archive based on user-defined specifications such as database table names, selection criteria, archive policies and archive index requirements that allow for fast retrieval. Specifications also include user-selected relationships to define the traversal path for archiving data.

Archive processing begins with the first table the user specifies in the traversal path. Data from related tables are processed next based on selected relationships and the logical keys of the data retrieved from the first table.

As part of the archiving process, the metadata, which includes database and table definitions, indexes and relationships, is extracted and stored by the archiving software with the archived data. This metadata provides the basis for maintaining the referential integrity of archived data and provides the foundation for accessing archived data at any time. The result is a self-contained, transportable archive file that ensures future access to archived data.

Cost. List the cost of your solution or provide some means for calculating cost for a specific environment.

N/A

Standards. List any relevant open standards upon which your product is built.

N/A

Other features. List other pertinent aspects of your solution.

1. Support for various storage devices to ensure cost efficient solution
2. Provides high performance archiving with minimal to no impact on the production environment.
3. Easily re-startable in situations when required.

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.

A comprehensive enterprise database archiving solution must provide the capabilities to archive related subsets of rarely accessed data from the various relational databases (Oracle, DB2, UDB, Sybase, SQL Server and Informix). This includes federated access, archiving from more than one database at the same time. It also includes complete interoperability to archive data from one database and restore to another database or even a newer version of the database. This ensures both enterprise support as well as long term protection for any future direction of the company. The ability to deploy a single archiving methodology across all applications will make archiving consistent, efficient and easy to maintain across the enterprise. Only the archive business policies should be changed from one application to the next.

Access to the data must be easy and on demand. The user must have the ability to access all the data or just a selective subset.