



Managing Backup Data Life through Offline Migration

Over the last five years, several important changes have taken place in the area of data protection and preservation. The separation between disaster recovery and data archiving has largely disappeared under the pressure of two concurrent trends: the explosive growth in the amount of data that needs to be preserved over increasingly longer periods of time and the increase in power and sophistication of data backup technology.

In many cases, the same technology used for disaster recovery, (backup software storing data on tape media), is being used to 'preserve' data with retention periods of seven years or more. Many organizations have at least some data under infinite retention. This means that many IT departments are faced with conflicting requirements: disaster recovery technology is driven by performance and availability, while long term data preservation is driven by storage media capacity and 'longevity' of the underlying software and hardware technology.

From the disaster recovery perspective, there is a clear incentive in leveraging any technology advance as quickly as possible. Since the data retention periods average three to six months, by the time a company has made the switch to a new backup technology, the legacy archives will be out of retention and could be retired along with the legacy backup software.

A totally different scenario applies to backups used for long-term archiving of static data. There are various reasons for companies to move to a new backup platform:

- The new software product may offer important features and benefits.
- The current software's maintenance costs are too high.
- The current software is losing market share and may become obsolete in the near future.
- Consolidation and standardizing on a single vendor.

Typically, even more compelling reasons exist to change media technology:

- Increasing backup sizes and fixed backup windows require higher data-rate tape drive and media technology.
- The cost per Gbyte of the current media is prohibitive when compared to newer, higher density media.
- Data life is likely to exceed the life of media technology. Technology refreshing (e.g. replacing obsolete DLT media with LTO-3/4 media) is the answer.
- Evolving data access and management requirements: Data merging or segregation (e.g. partitioning of an archive into two or more separate archives based on specified criteria) may be required.
- Stricter data storage security and access control requires data encryption, both of new and of legacy data.



Whenever a company makes a change in its backup platform or media technology for one or more of the above reasons, it has to solve the problem of how to guarantee access to its legacy data.

Cloning Backup Sessions

Migration (also known as cloning) of backup sessions to new, higher density media is typically part of the functionality of existing backup applications. Such native cloning is, of course, only applicable when a change of backup platform is not required. But even in that case, there are several pitfalls associated with native cloning, especially when large quantities of vintage tapes are involved:

- Licensing or logistics problems may prohibit having tape libraries with legacy technology and with new technology simultaneously online.
- The cloning operation will consume massive bandwidth and will cause disruption or degradation of on-going backups and restores.
- Reading of large quantities of vintage media (read errors, dysfunctional tapes) will put a strain on data center staff, especially when resulting in drive failures.
- Enterprise backup software is not optimized for massive cloning operations.
- A session level catalog of the vintage tapes may be missing, which is a requirement for native cloning.

So, we have established that there exist many valid reasons to change backup platform and especially media technology, but that tools to accomplish large-scale backup tape migrations are either not optimal or, in the case of platform change, non-existent.

Off-line Backup Data Migration with MediaGenie Proteus

In response, Data Strategies has developed its proprietary MediaGenie Proteus software tool for large-scale, high-performance and cost effective migration of backup data across media technologies, hardware platforms, and software products. MediaGenie Proteus already supports the most popular backup software products, including NetBackup, BackupExec, Networker, ARCserve, TSM, HP Data Protector, NTBackup, Commvault Galaxy, etc.

MediaGenie Proteus is highly scalable and supports a broad spectrum of off-line Backup Data Migration services based on the same “non-invasive” technology and unique processes that characterize all our offerings.

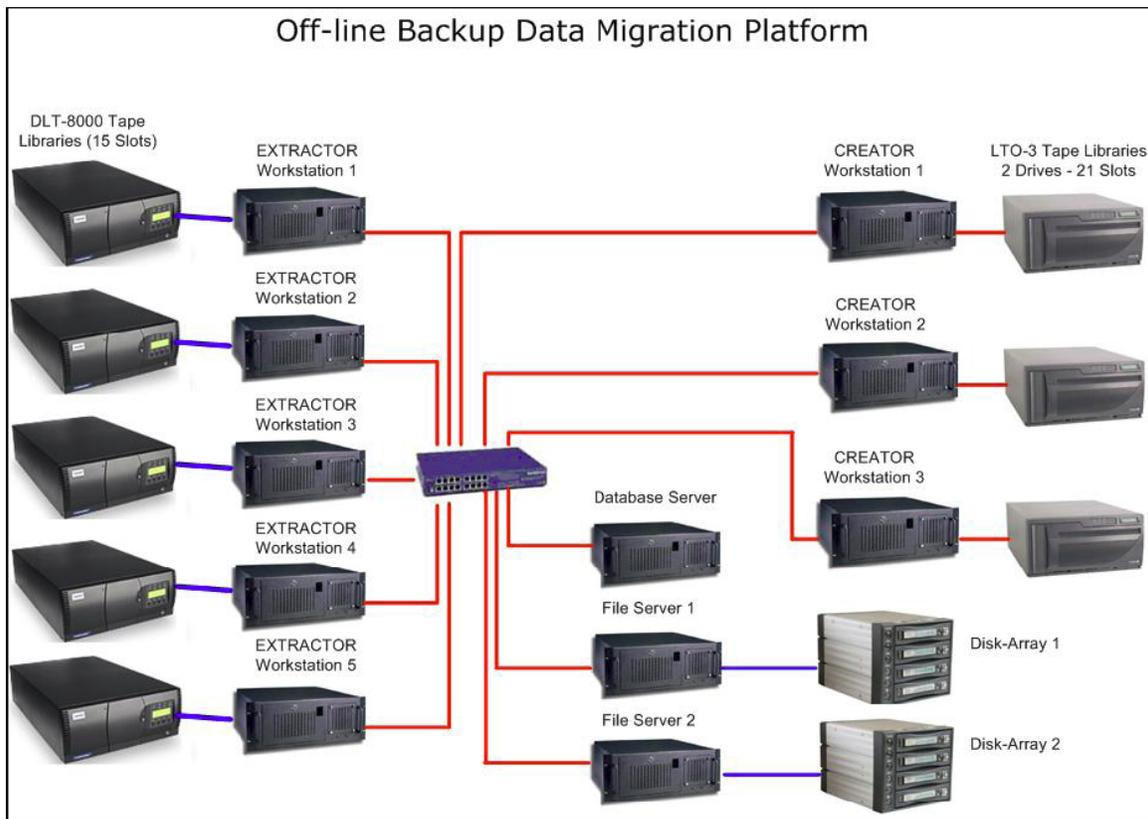
MediaGenie Proteus has the following functionality: Generation of session level and file level catalogs from media created by legacy applications; extracting data from legacy backup tapes; selective data migration (backup data splitting) based on customer specified criteria; session cloning to higher capacity media and data migration from a legacy backup formats to new backup formats (products).

MediaGenie Proteus implements a start-to-end, comprehensive approach to off-line backup data migration. This starts with the ability to absorb (pre-process) backup catalog

data and media management data from backup servers or, alternatively, to recreate catalog and media data from the backup media itself. It continues with the ability to map, extract, split, compose, convert, and edit backup data at various levels (volume, session, fragment, etc.), independently of the source or target backup software. It ends with the ability to create not only the new target media, but also the 'auxiliary' data needed to create or update the target backup server catalog and media management database.

Typical Backup Data Migration Platform

The figure below shows a typical hardware platform to migrate backup data using Data Strategies' MediaGenie Proteus software tool.



Case histories

NetBackup to NetBackup migration

A leading entertainment company was maintaining a large archive of some 3,000 DLT-IV, SDLT, and Sony DTF2 cartridges with infinite retention under Veritas NetBackup 5.1. The scope of this project was to migrate the backup sessions to LTO-3 media, maintaining the Veritas NetBackup format. Customer had formed two business groups and established rules for the separation of the backup data between the two groups.



Using these rules, the migration process filtered the data, creating two distinct sets of target tapes for each business group's NetBackup target server. To insure against possible future media loss, a second copy of the migrated data was generated concurrently.

The customer was able to provide a catalog backup on tape, which was utilized to load the "source" tables of the Proteus migration database. After the completion of the migration to LTO3 media, new NetBackup catalogs were generated for each of the target servers, using the information in the "destination" tables of the migration database.

Networker to NetBackup migration

A leading software company was maintaining a large archive of some 13,000 DLT-IV, cartridges with sensitive enterprise data under Legato Networker 6.2. Because of the company's internal security policy, this data could no longer be stored offsite without encryption. The scope of this project was to migrate unexpired backup sessions to LTO-3 media in NetBackup 5.1 format, using Decru Datafort for encryption. All information on the target media and in the corresponding catalog needed to match the information on the source backup tapes, regarding client name, backup type (full/incr), file path, backup timestamp and expiration date.

The migration process consisted of four sequential steps, each under control of MediaGenie Proteus: Extraction of fragment files from DLT tapes, restoring of savesets, (re)backing up to NetBackup disk storage units and writing to encrypted LTO3 tape. When writing to tape, client names, file paths, and backup timestamps are added and/or edited on the fly.

The customer was able to provide catalogs of all nine legacy Networker servers, which were utilized to load the "source" tables of the Proteus migration database. After the completion of the migration to LTO3 media, a NetBackup catalog (including volume database and media database) was generated for the target server, using the information in the "destination" tables of the migration database.

Networker to Networker migration

An electronics retail company was maintaining a mid-size archive of some 500 DLT, SDLT, LTO1 and LTO3 cartridges under Legato Networker 7.3, partly containing backup sessions with sensitive consumer data. Because of a new security policy, this data could no longer be maintained on backup tapes. The scope of this project was to clone or, if necessary, edit backup sessions on the file level, removing all files that contained sensitive data (as defined by certain file naming criteria). The target LTO-3 media should maintain the Networker format.

A catalog of the source tapes was not available, and MediaGenie Proteus was utilized to catalog all tapes on the session level. Based on the client name, a subset of sessions (both Windows and Netware) were cataloged on the file level.

Backup sessions that did not require editing (including agent based SQL and Exchange backups) were cloned to the new output media. Sessions that required file-level editing



were restored to disk, filtering out unwanted files, and re-backed up to the new output media.

After the completion of the migration to LTO3 media, a report detailing the relation between legacy sessions and migrated sessions with corresponding output volumes was provided to the customer.

Contact Information

Dr. Johan C. de Haas
johan@go-dsi.com
713-782-7183 (office)
713-301-6066 (cell)

Dr. Giancarlo Gaggero
ggdsi@go-dsi.com
713-782-7183 (office)
832-216-8347 (cell)

Data Strategies Int'l, Inc.
8562 Katy Fwy, Suite 154
Houston, Texas 77024