

CYBERNETICS

Cybernetics Virtual Tape Libraries
Media Migration Manager Streamlines
Flow of D2D2T Backup

April 2009

Cybernetics has been in the business of data protection for over thirty years. Our data storage and backup solutions are designed to capitalize on proven technology for the most cost-efficient and data-protective solutions. Cybernetics recently announced the release of the newest feature set for the HSTC virtual tape library engine – media migration management.

This is now a standard feature in all Cybernetics VTL models:

- miSAN V Series
- HSTC RAID
- iSAN V Series

Who needs tape backup?

While disk-based backup is arguably the best solution for current, immediately accessible backup data, the advantage of tape for archives is much more than the longevity of decades for the tape media. With tape, the medium is not dependent on a single device for recovery. If a tape drive has a failure, the media can be read in another drive. Cybernetics makes it easy and affordable to create multiple backup tapes - if a tape media fails, another backup tape can be recovered. By contrast, recovery from a disk array is completely dependent on the health of the disks in that array. Even if there are 100 backup copies on a RAID array, not one of them can survive catastrophic failure on that RAID.

Tape is the best solution for:

- Legal Compliance for Archives
- Point in Time Archives
- Robust Recovery Options

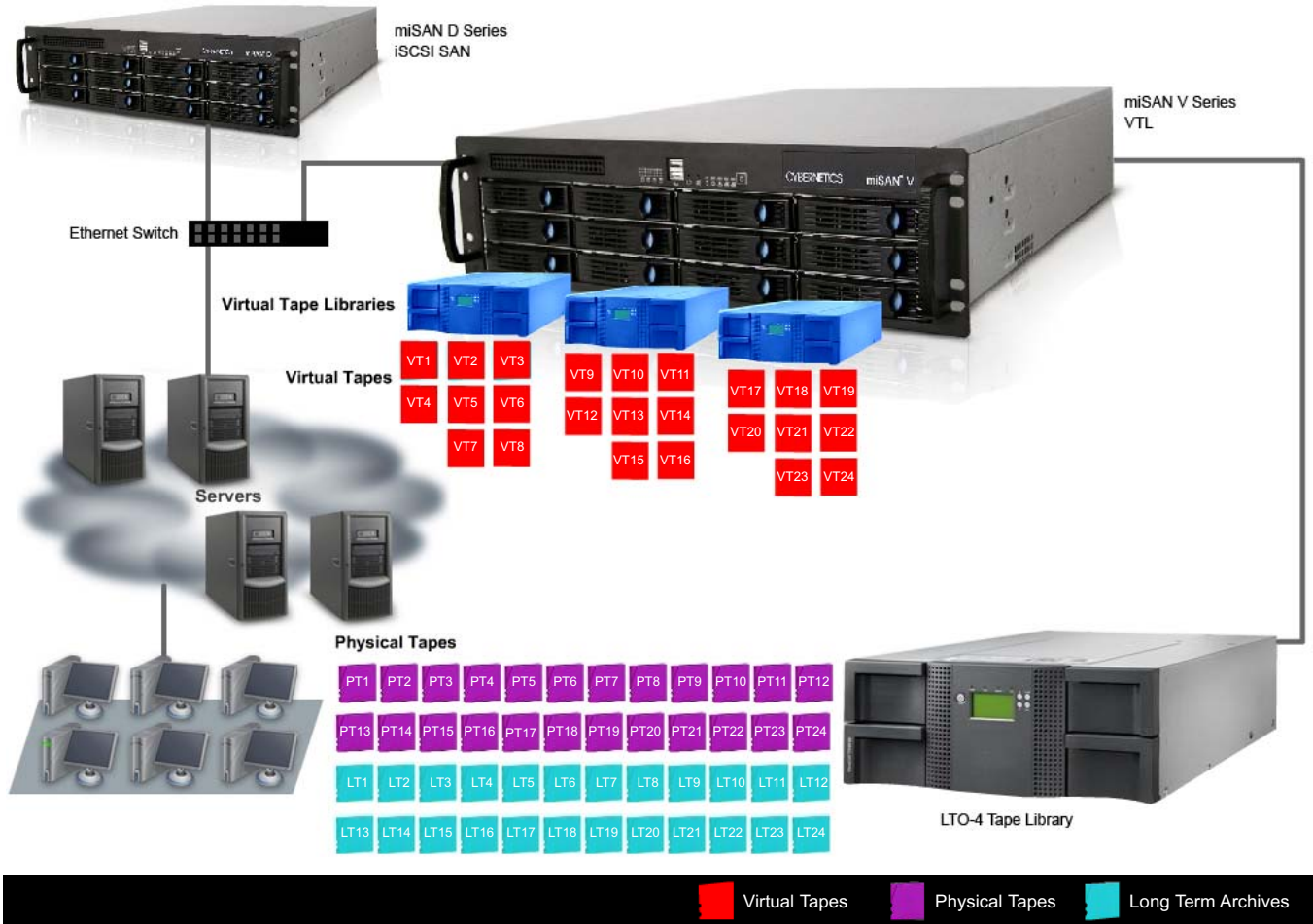
If tape is the safest and most economical medium for long-term archives, why does it seem to be losing market share?

In the 1990's, as faster linear tape drives were released, an issue developed that hit straight to the heart of reliability for tape backup. Tape is primarily a linear technology, and tape drives are not designed for the variable rate data streams common for backup servers. The tape drives were simply too fast for normal network backup. When a streaming tape drive is starved for data due to variable transfer rates, it goes into a repositioning mode, best described as "shoe-shining". This back and forth thrashing is very hard on the tape drive and the media, resulting in tape backup failures. Cybernetics innovated the first virtual tape library – the HSTC (High Speed Tape Cache) - very specifically to make tape backup reliable and fast again. Cybernetics engineers virtualized the RAID array to present as tape, capable of accepting concurrent streams of backup data at varying rates of speed. RAID can accept the highest and the lowest rates of speed with absolutely no negative impact. After the servers have completed their backup, the HSTC streams data directly from high speed RAID over a dedicated SCSI bus to the tape drive to sustain the optimum, maximum speed to tape. It reduces overall dependency on tape, and more importantly, it dramatically improves the reliability of tape drives and media. Today, this HSTC engine is available in three hardware platforms to meet the needs of the smallest networks to the largest enterprises: miSAN V Series, HSTC RAID, and iSAN V Series.

What is the HSTC and how does it work?

The HSTC is an iSCSI SAN backup solution with a feature rich menu for virtualizing disk to appear as tape backup devices. A single unit can be configured to function as several independent tape backup devices, to consolidate backup across a network or enterprise. Virtual tapes are assigned to virtual tape drives and virtual tape libraries to support backup for a variety of servers, operating systems and backup software products. Each virtual tape device can operate independently of and concurrently with the others.

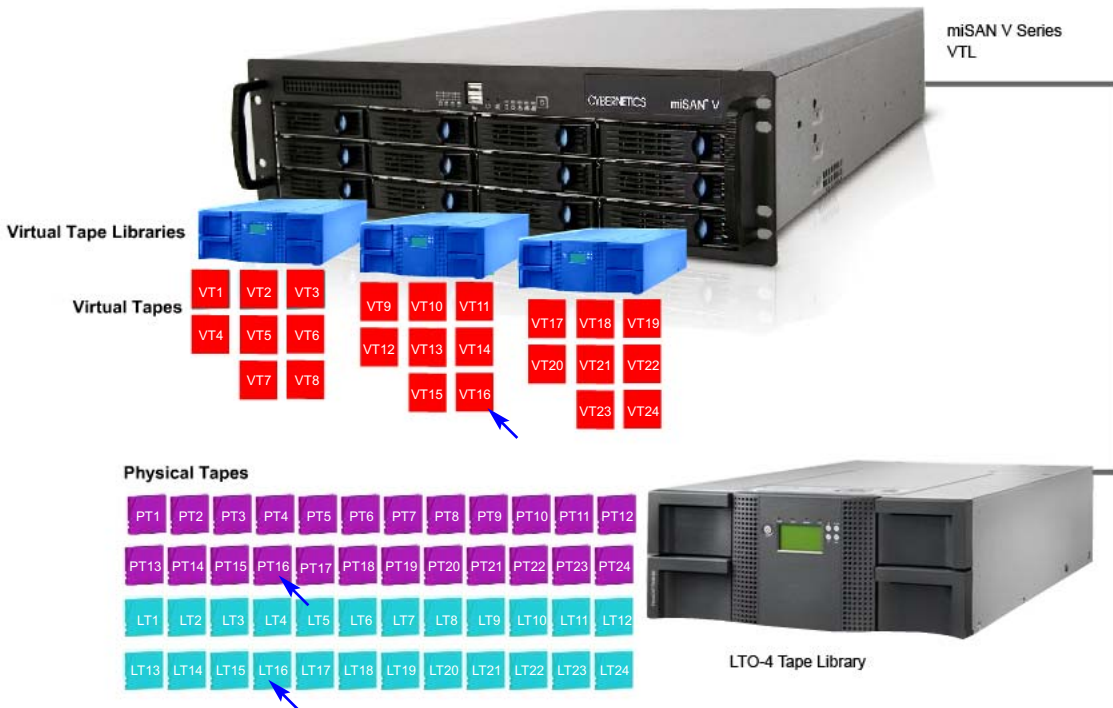
The HSTC moves selected virtual tapes to removable tape on demand, or automatically according to user-defined parameters, and the migration is performed in a completely serverless mode.



How is the HSTC different than other VTLs?

The new Media Migration Manager feature within HSTC manages each phase in the life of backup data as it flows from near-line disk to removable tape media in an archival tape library. Ingenious and elegantly simple, the HSTC is designed to maintain the most current backup files on high-speed disk for fast restore, while older backup files are pushed out to removable tape media in the archive tape library. This migration feature is serverless, and completely automated.

The Cybernetics VTL is configured to present itself to the backup software as an exact image of the archive tape library – even if there is a disparity between the number of virtual and removable tapes. Each virtual tape is logically associated with a removable tape, and all virtual tapes are migrated to a removable tape automatically. Every virtual tape is identified according to its eligibility for use in a new backup job, and as near-line virtual tapes are overwritten, the removable media tape copy remains in the long-term tape archive library. The Cybernetics Media Migration Manager synchronizes the backup catalogs for both virtual and removable tape libraries to seamlessly traverse virtual and removable tapes, with the intelligence to determine the most expedient path for backup or restore. If the data resides on virtual tape, the restore commences from the disk array. If the data has been migrated out of the virtual tape library, the backup software will load the tape for restore from the tape drive. Most importantly, the process is completely automated by the HSTC Media Migration Manager, which works with any backup software product that supports tape.



All 24 of the Virtual Tapes reside on Physical Tapes

The Media Migration Manager also facilitates the use of a small removable tape library as an export port for a larger virtual tape library. As virtual backup tapes are archived to a small tape library, the virtual tape is migrated out of the daily rotation stack of virtual tapes. As the removable tape copy is migrated out of the removable tape library for off-site storage, a copy can be migrated into the VTL archive tape space. The archive backup data can be recovered directly from the VTL until they are overwritten, eliminating the hassle of locating and transferring the actual offsite tape media for the tape restore process.

All Cybernetics VTLs running HSTC are available with options for hardware offload engines for data compression and data encryption. The hardware data compression feature can improve backup capacity by as much as two times. If the data is compressible, it is easy to double accessible backup capacity without doubling cost. Hardware data encryption secures access from potential interlopers on disk and/or on tape media. Both of these powerful features are implemented at the VTL hardware level, offloading the heavy processing burden from network servers. All data is compressed before encryption, and Cybernetics hardware data compression can actually improve throughput performance at the same ratio as the capacity benefit.

Who can use the HSTC?

The virtual tape devices in the Cybernetics HSTC support almost any operating system and backup software package. All hardware models with the HSTC VTL engine include both SCSI and iSCSI interfaces. The SCSI ports can be configured for archival tape devices or legacy host connections. In addition, the optional Cybernetics iClient module will connect a SCSI tape controller in any legacy computer system to the iSCSI backup SAN. A single miSAN V Series running the HSTC VTL engine can provide backup for an IBM iSeries, Windows, Linux, Unix, OS X, Netware, and more.

The HSTC can overcome host server limitations as well. For example, the IOPlless tape controllers in the IBM iSeries cannot support functionality of a tape changer, or tape library. The HSTC can auto-change a pool of assigned tapes to a virtual tape drive to provide automated tape changer functionality.

All archival to removable tape is performed by the HSTC itself. Since removable tape archives are streamed in a serverless mode directly to the tape drive, this data never hits the network after the initial backup to disk. The archived removable tapes are universally compatible. There is nothing proprietary about the tapes, so they can be recovered directly from any tape drive of the same format, any time and anywhere. Cybernetics HSTC can even overcome tape media format incompatibilities. Every major tape media format can be converted to new tape formats. For example, a legacy DLT tape can be read in as a virtual tape, and written out to a removable LTO tape.

Does the HSTC de-duplicate data?

Data de-duplication is a buzz-word used to describe a number of technologies designed to reduce capacity requirements for backup data. The most commonly advertised de-duplication techniques are based on hashing algorithms. With the objective of eliminating redundancy to save space, backup data is scanned for sequences that are identical to previously stored data sequences. This is accomplished by reducing data sequences to hash codes – values calculated based on data bytes. If a sequence of data returns a hash code that exists in a table, it is discarded and replaced with a pointer to the lone copy of that sequence. With all hashing algorithms, it is possible for different data sequences to result in the same hash code, but in absolute terms, the random statistical probability seems very low. Unfortunately, when this probability is applied in the real world to the incredibly large number of non-random data bytes within the gigabytes, terabytes and petabytes of enterprise data, corruption can become quite possible. Over a large data volume, a hashing algorithm could discard a unique data sequence based on a hash code that appears to represent a duplicate sequence. More problematic, a link could be updated, mis-populating thousands of pointer-dependent files. Since the very nature of backup is retention of redundant data so that any lost data can always be recovered, hashing based de-duplication is in general, pragmatically flawed. Cybernetics recommends a safer form of saving space, and the HSTC VTL makes it easy.

The term de-duplication was not yet coined when the oldest forms of reducing backup redundancy were implemented, but de-duplication is descriptive of incremental and differential backup methods. Backup software writes out data that has changed since the last backup - unchanged data is not duplicated in the backup file. If these differential backup files are written direct to tape media, restore can be tedious and risky. With sets of differential backup tapes, a restore involves recovering each backup set sequentially, and usually involves locating and loading multiple tapes. If any single tape is bad, the entire restore is a failure. By contrast, virtual tapes are stored on a RAID array protected by redundancy, parity and a hot

spare. The virtual tapes are accessed and restored at high performance RAID disk speed. With the Cybernetics VTL, de-duplication through incremental or differential backup is a very compelling solution. While the ratio of change level data to the full volume varies for every application, a good rule of thumb is about 10%. Assuming daily differential backup, this is a 10:1 capacity reduction. Implementation of the hardware compression feature of the HSTC VTL can further reduce backup capacity by about 50%, or 2:1. Combined, the daily incremental backup and data compression can reduce backup capacity by a ratio of 20:1 – and these methodologies are completely lossless, so there is no risk of discarding unique data. Admittedly, this is not trendy, but it is highly productive and extremely safe.

Cybernetics protects data through every phase.

Cybernetics solutions cover the full spectrum of data storage and data management. The miSAN and iSAN D Series models are based on our SANDR iSCSI SAN engine to serve live, primary disk storage across the network for virtual and standalone servers. These D Series disk solutions also incorporate layers of front-line data protection for maximum day-to-day uptime with snapshot, high availability, and replication/synchronization features. The V Series models with HSTC manage near-line and long-term archival backup with a tiered strategy for archiving backup data as it ages. Paired, these technologies provide the highest level of data accessibility and protection for today and tomorrow.