

Virtual Tape Libraries Are Very Real

By David G. Hill. Mesabi Group

The storage technologies trends spotlight shines on virtual tape libraries (VTLs) today. But why? What is the problem that VTLs solve? One answer is that a VTL delivers faster and more reliable backups and faster and more reliable restores. That is true — and useful — but not a sufficient justification in and of itself. Why not?

Although restoring from a VTL is typically faster than with a physical tape library, a VTL is not likely to qualify for high availability, recovery time objectives (RTO) that are typically only minutes a year. Other alternatives such as remote mirroring and continuous data protection can meet high availability objectives. Moreover, the concept of faster backups does not apply with those technologies and they are very reliable. So we have to look beyond what seems obvious about VTLs, but really isn't to find the real story about why more and more IT organizations are finding VTLs attractive.

Upon Further Review, VTLs Fill a Vital Role in Data Protection

So is there a need for VTLs? The answer is unequivocally yes. VTLs can play an essential and irreplaceable role inbreathing new life into traditional backup/restore environments, still the backbone of data protection in most enterprises.

Understand that standard backup is not the first line of defense for data protection. For physical data protection (say against disk failures), the first line of defense is RAID and remote mirroring to a designated disaster recovery site. For logical data protection (say against viruses or accidental file deletion), snapshots or newer approaches such as CDP offer a first line of defense.

However, any enterprise requires several layers of data protection even at the price of lower availability of data (which is a lot less than the price for no availability). Tape automation serves as the last line of defense with media at both remote and local sites in a media rotation strategy. Tape automation in conjunction with backup/restore software has long served as the backbone of the last lines of data protection.

Tape automation has been under increasing pressure due to ever-expanding amounts of storage and requirements for applications to be up on a 24 x 7 forever basis. One clear example is the "running out of night" problem where backups cannot be completed during the time available, but there are numerous other stresses and strains as well. The insertion of VTLs relieves those pressures and, in so doing, extends, cements, and reinvigorates existing backup/restore processes including tape automation.

Why rejuvenate instead of euthanize the old environment, since the use of a VTL implies that an organization plans to maintain its current backup/restore processes and that means keeping tape libraries? Although alternatives exist that may be appropriate in certain situations, any organization has to be cautious about taking a revolutionary approach where failure, i.e., if things don't work out as spelled out in the user manual, could be dangerous to critical business data and career-limiting for IT management.

Moreover, VTLs are an evolutionary step forward that offers a platform that can be smoothly assimilated within and also measurably improve IT data protection environments. VTLs can also serve as a centerpiece for data protection upon which some of the other data protection technologies can work in concert, such as CDP or remote replication to a disaster recovery site. The last lines of defense can work in conjunction or collaboration with the front lines of data protection and not in isolation. Therefore rejuvenation, not resuscitation is the operative word and goal.

This is important since a VTL is merely an additive to a data protection environment and not a replacement for existing investments. But even though VTLs add to the capital budget (CAPEX investment), an ROI case can be made for the investment.

And the Gold Rush Is On

The potential of VTLs has not escaped the notice of vendors (Table 1). Both the Goliaths and Davids of the industry are well-represented. However, not all VTLs are created equal. A key line of demarcation lies between VTL appliances and software-only VTL solutions. A VTL appliance is a turnkey solution that includes the software that drives the virtual tape library functionality, the server platform upon which the software sits, and the disk array that holds the backed up data. Some vendors loosely term the VTL software and the server upon which it sits as a VTL appliance, which fits one definition of an appliance as a dedicated-purpose server. However, we believe IT organizations interested in buying VTL appliances are best served by hardware-inclusive solutions that incorporate storage arrays. Perhaps the term VTL gateway (a la NAS gateway) would help differentiate between the two alternatives

The higher level of integration that comes with a full VTL appliance eases the burden of initial deployment as well as ongoing administration issues. On the other hand, a software-only (i.e., VTL gateway) solution lets the buyer use existing server and storage investments and does not dictate a proprietary solution, thus increasing flexibility and providing investment protection.

Other considerations for product adoption include both the usual suspects as well as some that strictly apply to VTLs:

- *Compatibility with existing environment* — Are the operating systems, server platforms, backup/restore software, network protocols, tape formats, and tape libraries copasetic?
- *RAS (Reliability Availability Serviceability)* — Does the VTL meet the standards of the IT organization?
- *Manageability* — How easy is to configure, tune, and administer the VTL?
- *Integration with physical tape libraries* — What is the data movement process between the VTL and a physical tape library and how does this affect how re-

stores are performed?

- *Space management* — How is compression handled both on the disk array and tape media? How does the system manage incremental backups? What data reduction techniques are used if electronic vaulting to a remote site is supported?
- *File restore granularity* — Can single files be restored easily (i.e., no volume restores necessary) and, if so, can these restores be done by a user (with the proper authentication)?
- *Value-added functionality* — What functionality beyond the basics does the VTL offer, such as electronic vaulting capability to remote sites or integration with other data protection technologies, such as CDP?

Table 1: Virtual Tape Library Product Sampler

Vendor	Product	Product Focus	Technology Foundation
ADIC	Pathlight VX family	Targets both midrange and enterprise disk-based backup where disk backups can export to physical tapes without impacting backup application or server performance	Integrates not only software and a disk array, but also an ADIC tape library into a VTL appliance solution
COPAN Systems	Revolution 200T	The write once/read occasionally-optimized disk used gives cost efficiencies that encourages longer storage of backups online than might otherwise be possible	Uses its MAID (Massive Array of Idle Disks) technology as the hardware platform upon which the VTL software runs.
Diligent Technologies	VTF Open	A VTF Open server appears to backup servers as one or more open enterprise-class tape libraries	VTF Open runs on a dedicated standard Linux server that can be attached to any Fibre Channel-attached disk array
EMC	CLARiiON Disk Library Series	CLARiiON-based disk emulation of a tape automation system with an emphasis on ease of deployment and ease of ongoing management	The disk library takes advantage of the CLARiiON modular, pay-as-you-grow architecture for performance, capacity, and reliability
FalconStor Software	VirtualTape Library (VTL)	A VirtualTape Library solution stresses extra features such as an active/active failover option for reliability and electronic vaulting for remote disaster recovery	The VTL solution is powered by FalconStor's well-proven IPStor server software and connects to industry-standard disk arrays to FC or IP/SCSI connected backup servers
HP	StorageWorks 6000 Virtual Library System Family	Virtual Library System Family appliances emphasize reducing the complexity of shared storage along with maintaining the manageability of a single system	A VLS6000 family appliance uses the HP StorageWorks Modular Smart Array 20 Enclosure for its storage to emulate StorageWorks ESL, E-Series, and MSL tape libraries

IBM	TS7510 Virtualization Engine	Virtualization — emulation — of tape libraries, tape drives, and tape media for open systems servers connecting over Fibre Channel physical connection	Combines IBM server technology, disk technology, and tape technology into an integrated solution
Neartek	Virtual Storage Engine (VSE)	VTL software that sits between hosts and storage peripherals. The storage can be local as well as remote through electronic vaulting for disaster recovery protection	The VSE software runs on standard, off the shelf PC servers. Two logical components are an out of band Controller for coordination and communication and stackable Transfer Units for in band data movement and tape drive emulation
NetApp	NearStore VTL600 and VTL1200	Although a NearStore VTL appliance is the target of backup applications, source data can reside on any open systems storage environment, including EMC, HDS, HP, and Sun	A NetApp VTL appliance features capabilities, such as self-tuning performance (dynamically load balancing backup streams across disks) and tape smart sizing (efficient compression management to reduce the number of physical tapes required)
Quantum	DX-Series	Quantum has a DX-Series VTL appliance for any size organization from small to mid-size businesses to larger enterprises	A VTL appliance with features such as Optyon In-Line Data Compression, which is stated as giving 2:1 compression without performance degradation for either the backup host or the DX-Series processor
SEPATON	S2100-ES2 Virtual Tape System	Turnkey VTL appliance that emphasizes performance and scalability as well as efficiency for restores for both small and large data volumes	Architecture features a Dynamic File System (DFS) for volume restoration, SEPATON I/O Subsystem (SiS) for disk space usage, and a Scalable Replication Engine (SRE) for performance scaling
Sun	StorageTek Virtual Tape Library Series	Virtual Tape Library appliances offer a single point of administration that manages data from multiple backup servers and multiple backup applications	The Virtual Tape Library appliance combines a server, disk storage, and software in a single rack.

Source: Mesabi Group, February 2006

Mission Accomplished?

VTLs are obviously very real, but there is still some unfinished business. The biggest issue relates to with moving backup copies from disk to a physical tape library. There are basically two approaches to doing that function:

- The backup server controls the movement of data from the virtual tape library to the physical tape library. Although the physical tape library is visible to the VTL, the backup server controls the process. The advantage of this model is that if the backup that the backup server needs is not in the VTL, the backup server has the correct catalog information to be able to restore from the physical tape library. The disadvantage is that the backup server has to use resources to make a copy that the VTL could have just as easily made in the background

- The VTL simply performs a copy and export function to the physical tape library. Technically, this is the preferred solution since the backup server does not have to get involved in the duplication function. Practically speaking, however, having two backup copies that are identical (i.e., have the same bar code) can create consistency problems for the backup server. For example, if the backup server was not given the information that a tape copy returned from offsite (as in the case of a VTL failure) was the correct copy, the backup server would assume that the correct copy was in the VTL and reject the physical copy. Workarounds exist, but workarounds require compromises.

The Virtual Tape Library Special Interest Group (VTL SIG) within the Data Protection Initiative (DPI) of the Storage Networking Industry Association (SNIA) is working on a general solution to this issue and others related to the use of virtual tape libraries. That work constitutes a positive step forward and should not inhibit IT organizations from examining whether or not VTLs are appropriate and implementing them as necessary.

However, IT organizations should just not rush to deploy VTLs because the stresses and strains of trying to management current backup/restore processes are exacting an unacceptable toll. Rather they should take their time to determine which solution is most likely to treat the immediate symptoms and also serve as a foundation building healthier, more effective data protection infrastructures.

David G. Hill is principal of the Mesabi Group (www.mesabigroup.com). The Mesabi Group focuses on the revolutions in Storage Networking and Storage Management, and helps clients make the best and most efficient use of information for business value.

© 2006 Mesabi Group. All rights reserved.