

You May Use Disk; You Will Use Tape
Disk-based solutions are vying to take over your backup processes. Tape may give an inch, but tape has plenty of inches where that came from.

Integrated Solutions, May 2003

Written by Tom von Gunden

Given the proliferation of disk-based backup solutions and cheap disks in general, end users may be wondering about the future of tape. But, despite the "tape is dead" hype, tape isn't going anywhere. Well, actually, it is going somewhere. It's going to at least one place where it has always gone - to off-site archival storage as a removable medium for disaster recovery reassurance and for long-term data protection. For those purposes alone, tape will remain a viable, cost-effective storage technology. But, off-site archiving isn't the only need that tape will continue to address. Tape is still and will remain a core technology for supporting local, centralized backup/restore operations. And, for some high I/O (input/output) real-time data collection operations (e.g. charting weather patterns from downloaded satellite information), it is increasingly being used for primary, first-line storage.

End users considering where and how tape might best target their overall data protection and storage management needs should consider, first and foremost, their current and anticipated data growth. Understanding how they acquire, use, and manage their data - and how much of it - will help users determine the level of tape solution required, from entry-level, single-drive devices to automated enterprise-class libraries attached to SANs (storage area networks).

Capacity And Complexity: Guess Now, Regret Later

Although the segmentation of tape solutions isn't neat and clean, most tape vendors identify three basic categories, which illustrate a range of scalability and complexity: entry-level, midrange, and enterprise-class. Entry-level tape solutions typically target the needs of individual desktop users, small workgroups, or small businesses. The most basic of entry-level solutions is a single tape drive directly attached to a desktop workstation or small server. With no automation functionality and very little software management capabilities, the drive supports only basic backup/restore operations and relies on manual removal and insertion of cartridges.

Slightly higher in scale, but still typically direct attached, are dual-drive entry-level units. At the highest end of the entry-level space, users might have single- or dual-drive autoloaders, which may be supporting several workstations on a small departmental or workgroup LAN. In general, entry-level solutions are characterized by the relatively low number of lower-end DDS (digital data storage) cartridges (fewer than 10) and the limited distances supported by their IDE (integrated drive electronics) or SCSI (small computer system interface) connections.

In the midrange, tape solutions tend to be handled by dual-drive autoloaders or by automated libraries with 2 to 10 drives. According to Paul Scheuer, midrange tape marketing manager for IBM (Armonk, NY), "If your capacity needs go beyond 10 cartridges in a small- to medium-sized business, you'll want to bring in at least the first level of automation." The deployment of faster, more functional libraries reflects the midrange user's need for tape solutions

that are now part of networked storage infrastructures. "At a minimum, a midrange tape solution is moving out of the department and onto the edge of the data center, if not right down into the data center," says Mike Lakowicz, VP, product strategy and business development for Seagate Removable Storage Solutions (Costa Mesa, CA). "So, in the midrange, you'll still see libraries attached to LANs, but it's becoming more likely users are attaching them to data center SANs. Those libraries need to support increasingly multiple access for backup jobs for increasingly higher amounts of transactional data."

Once companies move into the midrange tape solution space, they gain centralized and, perhaps, remote management capabilities, as well as enhanced availability. Says Gregg Ormsbee, product manager for tape library vendor Qualstar Corp. (Simi Valley, CA), "Midrange libraries provide a higher level of reliability - hot swap drives, redundant power supplies, and a modular design, which increases ease-of-service."

According to Lakowicz, enterprise-class tape solutions are "the midrange on a larger scale, with a greater multiple of 9s in terms of availability." Compared to the typical midrange library's 2 to 10 drives and 30 to 200 slots for cartridges, an enterprise-class tape library may have more than 10 drives and house from several hundred to several thousand cartridges. Those cartridges tend to be in the "super" tape class - with native capacities of 100 GB in latest generation LTO (linear tape open), AIT (advanced intelligence tape), or DLT (digital linear tape) formats, including S-AIT (super AIT) and SDLT (super DLT). "With today's super capacity drives and data compression, these tape systems can store from multiple terabytes up to several petabytes of data," says Scheuer.

In addition to highly scalable capacities, enterprise-class tape solutions can accommodate the storage needs of large, diverse, even global networks. "Enterprise tape libraries support multiple interfaces, such as SCSI, Fibre Channel, and iSCSI [Internet SCSI]," says Ormsbee. "They also offer software tools that support remote monitoring, provide automated alerts, and allow users to set policies for data movement and replication."

Shelf Life Gives New Life To Servers

In addition to considering how much size, complexity, and scalability they need in a tape solution, end users must also determine the role of tape within an overall data protection strategy. That strategy is likely to also include disk. The disk/tape relationship is particularly important now that hard disk and drive prices are dropping and vendors - even some traditional tape vendors - are touting disk-based, or D2D (disk-to-disk)-based, backup/restore and business continuance solutions. The second disk in the disk-to-disk scenario may be slated for work as a primary backup target or as an interim "backup cache" to back up data online before moving it to tape.

Tape vendors acknowledge the role of disk-based backup. According to Tom Yuhas, director, Sony Data Systems, Sony Electronics Inc. (Park Ridge, NJ), "At some level, it makes sense to move hot online files that become 'warm' to disk-based cache. But, it doesn't make sense to leave most data on disk arrays on your servers or storage devices for years and years. Even if there are compliance requirements to keep the data, most companies can't predict when it will be needed - or if it will ever be needed." Lakowicz agrees that disk can serve a useful, even cost-effective role in staging data for archival backup

but still sees tape as the actual archival repository. "In the midrange and enterprise spaces, there will be an increased use of low-cost disk at the front end of the data protection solution - perhaps even for copying data to a remote data center. But, as a last line of defense, companies will still want to put their data on removable media that can be securely vaulted."

Scheuer admits that recovering data from tape is a slower process than recovering it from disk. However, he sees the cost of maintaining backup copies in hard disk storage as cost-prohibitive, particularly as companies scale up their back/restore systems to match their data growth. "The advantage of copying to disk evaporates quickly when you are backing up several hundred GB," Scheuer says. "To keep satisfying your backup requirements on disk, you'd have to keep adding or expanding RAID [redundant array of independent disk] subsystems and keep those disks spinning. By contrast, having tape cartridges sitting on shelves doesn't expend any server power." And, those servers are too valuable to waste on aging data that, though it must be maintained, doesn't contribute to immediate business needs. As Yuhas puts it, "The data has value, but so does the server. As the value of data declines over time, if you leave it online on a server, the value of that server declines. So, save the server by saving the data to tape."