

ARTICLE POSTED March 3rd, 2003

## **Disk and tape coexist harmoniously in the backup market**

By Robert L. Scheier

Tape lives.

True, vendors are announcing new, lower priced disk backup systems that can restore data more quickly than tape while almost matching tape's traditionally low price. But according to customers, analysts and vendors, the demand for tape will increase, not diminish, because it's still the only high-capacity media portable and inexpensive enough for long-term, off-site archiving.

"We don't see it as disk vs. tape at all. It's disk and tape," says Jeff Laughlin, director of strategy for the Automated Tape Solutions Group at StorageTek, a vendor of both disk- and tape-based products. Laughlin goes on to say that disk-based backup will be used to recover from local system outages, while tape will get the nod for long-term storage of data to meet legal requirements, or to recover from disasters that destroy an entire data center.

"Nobody we've talked to believes that disk is going to replace tape," says Mark Silverman, president and CEO of Bocada, Inc., whose software tracks the efficiency and reliability of backups to either tape or disk.

"Using inexpensive disks to recover from non-disasters makes sense," says Steve Duplessie, senior analyst at Enterprise Storage Group in Milford, MA. "But companies will still put a tape backup in a bunker in Provo, Utah, just in case a bomb goes off."

While falling prices will shrink the worldwide tape drive market from \$2 billion in 2002 to \$1.9 billion by 2006, unit shipments will grow from 2.2 million to 2.4 million, according to IDC's Worldwide Tape Drive Forecast and Analysis. For every megabyte of disk backup a customer buys, says Laughlin, they also buy between four and 10 times as much tape storage.

### **Disk prices dropping**

Exact cost comparisons are difficult, but Duplessie says low-cost disk systems now cost between 1 cent and 2.5 cents per megabyte, compared with about half a cent per megabyte for tape. That's a far cry, he says, from the previous gap of 10 cents per megabyte for disk vs. half a cent for tape.

Where organizations really need fast recovery from a disaster, most agree that disk is faster. "Depending on the speed of the library, completely recovering a multi-terabyte tape system might take more than a day," says Amar Rao, senior vice president of marketing and business development at Kashya, Inc. Kashya is developing an appliance that will use compression technologies to allow long-distance data replication across heterogeneous disk systems. Tape is slower, Rao says, because tapes can only be read sequentially (from beginning to end) rather than via random access provided by disk, and they must be physically loaded in and out of tape drives.

Another vendor aiming to exploit the advantages of disk-based backup is Avamar Technologies. Avamar claims its Axion appliance cuts the amount of storage space required for backup up to 98% by eliminating redundant sequences in files, databases and objects.

The argument that disk could replace tape is made most strongly by Diamond Lauffin, executive vice president of Nexsan Technologies Ltd. Nexsan's InfiniSAN ATAbay series uses RAID arrays built around serial ATA disk drives at prices starting at about \$10,000 for 500 gigabytes of storage. Nexsan backs up files in their native file format, eliminating the need to convert restored data from the proprietary file format used by some tape vendors, and allowing customers to automatically switch from a failed system to its disk-based backup in under two minutes.

However, adds Lauffin, reliability — not speed — is the major advantage of disk-based backup. Unlike tape, where corruption of a single cartridge can ruin a backup, Nexsan's RAID arrays guarantee a single

disk can fail without any loss of data. Tape's slowness, he says, also makes it unfeasible to do the data integrity audits recommended by tape vendors.

But other factors will keep tape viable. They include the complexity of adding new elements to already complex storage systems, the need to make the most of existing tape hardware and the trend toward tight corporate budgets.

Truck manufacturer Paccar, Inc., for example, chose to stick with tape when it upgraded its backup and restore capabilities two years ago. Before the upgrade, support staff had been making nightly backups of data on each of its 125 Windows and Unix servers, a process that was both labor intensive and unreliable, says Barbara Newell, a technology consultant at Paccar. Even when backups did work, Newell says it took as much as two hours to retrieve the tapes from an offsite facility and recover the data.

Paccar considered a disk-based backup system but was hit by sticker shock when it realized disk would (at that time) cost two to three times as much as tape-based backup. With about \$1 million to spend on the upgrade, Paccar instead took advantage of unused capacity in its three StorageTek PowderHorn tape libraries.

By installing 20 9840 tape drives and using the NetBackup software from Veritas Software, Paccar is able to do a full system restore every night of 4 terabytes of mainframe data and 1.5 to 2 terabytes of data from its distributed systems. It has also cut its restore time from two hours to two minutes.

With the drops in price for disk-based backup over the last two years, Newell says, the company will probably take another look at disk-based backup as part of a new storage and server consolidation project. "But right now, we're happy with tape," she says.

#### **The Cleveland Clinic approach**

Tape libraries may actually cost more upfront than disk systems, but they can be scaled at a much lower cost for each incremental amount of storage, says Robert Cecil, director of networks for radiology and cardiology at The Cleveland Clinic Foundation. That's why the health care provider stores 100 terabytes of images such as X-rays, ultrasounds and other patient images on a StorageTek PowderHorn 9310 tape library.

The clinic doesn't use the system for backup, he says, but rather for rapid access to historical imaging data for purposes such as comparing current images of a patient with previous images. Both the radiology and cardiology units have their own servers, which each use about 3 terabytes of direct-attached disk to store the most recent and frequently accessed information. These servers also handle meta-data describing where to find historic data in the tape library.

While the tape library was more expensive upfront than buying disk drives or a CD-ROM library, says Cecil, each successive increment of storage is less than one-tenth the cost of a CD-ROM or disk. When the PowderHorn tape library is full, Cecil can upgrade to faster, higher capacity tape drives while maintaining his investment in the jukebox and the software to manage it. Another important benefit, he says, is that such upgrades can be transparent to the servers that send and retrieve images from the tape library.

According to Silverman, some customers are also leery of adding any new elements to their already complex storage environments. Adding anything that makes a storage system more complicated makes it harder to manage, less reliable and less efficient. To ease the need for administrators to learn another management interface, vendors such as Quantum Corp. have announced disk-based storage systems that appear to software as a tape library.

With tape and disk serving similar needs and coming closer in cost, it's not surprising that some vendors are combining the two. StorageTek's Virtual Storage Manager for the mainframe combines tape and disk technology in the same system, using virtualization to maximize the available storage and using disk arrays as buffers for data to be stored on tape. In the first half of 2004, StorageTek plans combined disk/tape libraries that offer many of the same capabilities for Unix and Windows applications, says Laughlin.

This virtualization not only makes the most efficient use of the available storage space, but it also masks failures in the entire storage subsystem from the applications that use them. "In the long term," says Laughlin, "we see these libraries as data protection storage repositories where users won't know or care whether their data is on disk or tape."

**About the author**

Robert L. Scheier is a freelance writer and editorial consultant in Boylston, MA. He can be reached at [rscheier@charter.net](mailto:rscheier@charter.net).