

NetworkWorldFusion

Disk saves the day (and night)

With back-up and retrieval speeds that run 100 times faster than they do for tape, disk is quickly winning favor as the preferred medium for backing up primary storage.

By [Deni Connor](#)
Network World, 03/24/03

"Back-up to disk, archive to tape."

That, says Michael Peterson, a senior analyst with Strategic Research, has become the mantra for storage backup.

Disk-based options quickly are replacing tape as the primary means of backing up data, analysts and customers say. Increasing cost parity between tape and disk is one reason, but back-up and retrieval speed is the biggie. Businesses that run 24-7 and rely more on electronic data simply can't afford to suspend operations for the time it takes to deal with tape backups.

At one time, performing nightly backups presented no problem. IT managers could take databases out of operation or make them inactive (putting them into a "quiesced" state) when their companies closed for the night. But as data stores grew, incremental backups became necessary at many companies — time did not permit entire backups nightly. And with the rise of the nonstop business operation over the past few years, no time is right for taking down an application for back-up purposes.

Backing up to disk provides huge relief because it is about 100 times faster than backing up to tape, analysts say.

"The average seek time of an Ultra ATA drive is 8.5 milliseconds, transferring data at 100 megabytes per second," says Pete Gerr, an analyst with Enterprise Storage Group. "For a Linear Tape Open tape, you are going to get about 30 megabytes-per-second transfer, and the average seek time is anywhere from 10 to 85 seconds."

Gerr uses a 2G-byte data set as an example. Presuming local tape drives, backing up data would take five hours while restoring the data would take seven hours. It would take all of 30 minutes to back up the data to disk and only 20 minutes to restore it.

With disk, IT managers also can back up more than one server at a time. With tape, only one server could write at a time.

Although the advantages of tape backup over disk are decreasing, users say they aren't throwing out their tape drives and libraries or forcing them into disuse. Rather, they are using tape to double protect and archive their data. Some users are even replicating their data to remote sites as a sort of triple whammy of data protection.

Here's a look at how three corporate storage managers are approaching backup today.

Always available

In March 2002, Scott Mastre, server engineer at Wells Fargo Trust Operations in Minneapolis, was in a predicament. He was running out of time — not to mention getting short on patience — because he needed to back up transaction logs from as many as 40 SQL Server and Oracle databases on Windows 2000 servers every day.

Because he was using tape to store each of the transaction logs, only one server could do its backup at a time. And in many cases, he needed to take down databases so he could back up the transaction logs.

"We could only back up nightly when we shut down the databases," Mastre says. It took Mastre about two hours to back up transaction log by transaction log for those 40 databases. Mastre figured switching to disk would not only save him time, but also keep the databases constantly available.

He chose Xiotech's Magnitude storage arrays for local and remote data storage, CommVault Systems' Galaxy back-up software and St. Bernard Software's Open File Manager data-protection software. Open File Manager lets the Galaxy software capture files that are open and in use during the back-up process. This means data is backed up completely and uncorrupted even though the databases remain in use.

Disk back-up options

A variety of vendors offer hardware storage devices that use inexpensive ATA drives, allowing users to adopt disk-based backup in lieu of slower tape options.

Vendor	Product	Maximum capacity (in terabytes)
ATTO Technology	Diamond Series RAID Storage Array	3.8
EMC	Centera	160
Network Appliance	NearStor R100	96
Quantum	DX30	3
StorageTek	BladeStore Disk Subsystem	160

"Those backups are all 'hot' in that the databases are still online while the backup takes place," Mastre says, noting that the main targets for open file management were the Windows file servers. "There we have a combination of user files and data from older legacy applications. And typically we have problems getting users to log off at night, so we weren't usually in a state where we could back up the files."

Now, at any given hour, Mastre says, as many as 40 jobs are streaming transaction logs to disk without interrupting operations. This can be compared with the two hours nightly he had to quiesce databases during the tape back-up process.

Using the CommVault software, Mastre backs up the databases to the disk-based storage array for speed and concurrency. Once the data is on disk, he generates copies from the storage array to tape for off-site storage. "We perform hourly transaction log backups and then daily full backups," he says.

When users inadvertently erase files, Mastre can get them back faster from disk than he could from tape. "Previously, because we didn't keep tapes onsite, it could take a day to get the data back, and more time to restore it," he says. "Now with disk, retrieval is immediate."

The edge of the platter

For Rich Banta, senior enterprise systems engineer at St. Vincent Hospital & Health Services in Indianapolis, the day was 12 hours too short. Backing up the 6 terabytes of data on the 926-bed medical center's 350 servers tape took 36 hours, making nightly backups impossible. Tape backups simply wouldn't work any longer, especially because Banta expects to add another 100 servers this year as St. Vincent's grows through acquisition and normal business operations.

"We did some research, and I've seen some successful disk-to-disk-to-tape implementations on a smaller scale, so we thought the technology and products had finally matured," Banta says.

Plus with ATA disks, those commonly used with desktop computers, now costing less than a penny per megabyte, Banta says "we were finally able to make the sale to management." Comparatively, SCSI costs 3 to 5 cents per megabyte and Fibre Channel 7 to 15 cents per megabyte, according to Giga Information Group.

Most of the medical center servers run Windows NT and NetWare, but Banta also has some on Macintosh and Unix operating systems. He is backing up data generated from McKesson DICOM and ALI UltraPACS digital imaging systems. PACS is the term for picture archiving and communications systems used in diagnostic imaging in healthcare.

Banta supplemented his StorageTek L700 ACSLS tape library with a disk-based system that, according to vendor claims, is faster than tape. To that he added software from Evault called InfoStage, which performs block-level incremental backups of all 350-host computers to a StorageTek BladeStore Disk Subsystem. BladeStore uses ATA drives.

"What really makes this work, more than cheap ATA disk, is the block-level nature of Evault's InfoStage, which works with NetWare, Windows NT and Macintoshes, as well as Solaris and AIX," he says.

Once data is on disk, Banta can back it up to tape using Legato Systems' NetWorker software at his leisure and without affecting server operations. Then, using StorageTek's Application Storage Manager software, he makes asynchronous copies of the backups that he sends over an OC-3 line to a BladeStore subsystem 14 miles away. Banta eventually will upgrade the OC-3 to fiber to allow off-site mirroring.

Disk has saved his staff not just time, but money, Banta says. "We could have added another 100 servers into our back-up world without spending a ton of money. But we saved \$1.1 million [going to disk instead]. That covers our disaster-recovery costs," he says.

"We were in terrible pain before - and pain's a great motivator," Banta says.

Another triple whammy

Jeff Polsgrove, CIO for Scottrade, a financial firm in St. Louis, has about 8 terabytes of historical trading data he restores from tape to disk so he can get it faster. He moves the data to an EMC Centera "content-addressable" storage system, which uses ATA drives, Windows 2000 servers and EMC Symmetrix arrays. Content-addressable storage is used for data that doesn't change over time.

The need to comply with U.S. Securities and Exchange Commission regulations on keeping electronic financial records "readily available" spurs Polsgrove's tape-to-disk transfer. Accessing data stored in tape drives was too cumbersome, he says.

"When we wanted to reconstruct a transaction going back a number of years from tape, we had to search tape by tape for an account number and then restore the data we needed," he says. "And, because there was no index attached to a [tape-based] record, retrieval wasn't something the customer could do themselves. IS had to do it.

"With disk, we can get to data hundreds of times faster than having to sort through tapes, dump the data somewhere, mount another tape and go through the process again," Polsgrove says.

"And, the new way gives the data back to the customer," he adds.

Each night, Polsgrove backs up data from one disk to another, then dumps the data from the second storage array to tape. "We are using disk as a staging area because speed is of the essence. For disaster-recovery purposes, we need to send the data to tape and then off-site," he says.

"Back up to disk, archive to tape." The mantra grows ever stronger.

