

The advantages of near-line storage

Chris Green looks at the future of localised storage in the face of improved reliability and cost savings from centralisation.

Chris Green, Computing 30 Jan 2003

Talk about storage, and those in the industry will extol the virtues of iSCSI, Fibre Channel, network attached storage, storage area networks and all manner of big, powerful, cost-effective, networked and centrally manageable storage platforms.

Chances are they won't mention near-line storage. Despite a consistent level of spending on new, often centralised storage among businesses, near-line devices continue to hang in there at a desktop and workgroup level.

But just what is a near-line device? It's a CD writer, a DVD recorder, a Zip drive, a SyQuest drive, a desktop tape back-up drive or small workgroup autoloader.

Its large-format storage devices provide simple, cheap storage and archiving capabilities, but with the advantage that data can be stored on one of many pieces of removable media.

According to researcher IDC's 2003 forecasts, companies will invest about \$34bn (£20bn) in storage hardware and software this year alone, making it one of the few areas of IT to retain a significant slice of today's modest budgets.

The roles of online and near-line

IDC believes that budgets will be spent in three main areas:

- 77 per cent on disk storage (high-performance raid arrays, external HDDs, Zip and other proprietary magnetic formats)
- 13 per cent on fibre channel (storage area network technology)
- 10 per cent on tape storage (local and centralised)

In large organisations, shuttling data between localised near-line storage and larger repositories can create a more effective workflow process.

It is an idea that appealed to Yahoo which is deploying a kind of 'jukebox' approach to its data needs and moving it from near-line to online as needed.

The divide seems clear-cut: online storage works best where different users have to share or work on the same data across the company; or where huge amounts of data need to be stored and then accessed again at a moment's notice; or where data can come from any corner of the corporate map.

Near-line works best when the workflow process requires that key data is retained and/or archived locally for quick local recall, but where access by users other than the individual is not so time-sensitive.

Justin Sutton-Parker, of storage software developer ITP, explained that improvements to software management tools have kept pace with hardware demand, meaning that many near-line solutions can be almost as centrally manageable as consolidated arrays.

"As near-line storage has developed, so too has the software," he said. "The technology now allows data stored within the near-line zone to be accessed without IT administrator intervention and with little degradation to access speed.

"The ability to offer storage that is accessible in an almost online fashion, with the advantage of a removable format, is proving extremely valuable in medicine and design.

"Often the information has to be portable, whether it's for off-site transfer or interdepartmental information sharing between, for instance, private and NHS healthcare operating from the same building but on a separate network."

John Williamson, divisional director of Sony's shared storage solutions department, highlighted the data integrity issues.

"While online storage provides users with instant access to data, it is also susceptible to viruses and user error," he said. "Offline storage, on the other hand, provides an extra layer of security by allowing data to be copied onto a separate tape. However, users are not able to access the data so readily.

"The benefits of both are encapsulated in near-line storage, simply because it allows users to access data quickly and provides the option of allowing them to decide which data they keep online and offline."

The problem for some companies is that, while online may be more suitable, high demand and low capacity often forces data to spill off onto near-line devices as a short-term fix, which can soon become a long-term legacy.

"Business data storage today is characterised by greater volumes, higher capacity and faster access," explained Andrew Wyllie, commercial manager at NEC's business equipment division.

"Users are constantly demanding the ability to move and handle larger files, often containing large high-resolution images.

"Documents that were previously sent as hard copies to users are now sent automatically by email. The total quantity of data handled by businesses is doubling every three years, and people want to keep it to hand, storing more and more on their hard drives."

But the primary function for most near-line storage systems is for back-up and archiving. This is particularly the case for most tape and CD systems, which are used for complete system snapshots or for smaller fractional back-ups, such as a single issue of a magazine.

The falling price of DVD-RW and +RW, and the fact that some discs can hold more than 4GB of data, has made the platform particularly popular for archiving, even though drive speeds are still fairly slow compared with writeable CD.

At the local level, the available back-up window is shortening, as more time is being spent during the working day actually working.

So a supervised back-up needs to take the shortest amount of time possible, hence the popularity of cheap external disks over even cheaper, but slower, tape and CD.

At the server level, for workgroups and the back office, the window is narrower, thanks mainly to 24-hour trading and the internal hosting of web resources that also need 24/7 availability or which are privy to dynamically changing real-time data (stock quotes, web forums, news feeds and so on).

"The other key factor keeping near-line storage alive is the deterioration of available back-up windows," said Sutton-Parker.

"Information stored on removable media such as DVD can generally be omitted from a backup as it is already committed to a stable medium, and may even be stored on tape already as a tertiary procedure.

"This means that less data is included in the daily back-up and ensures that systems can be backed up within an allotted time frame or even in a live environment."

Tape vs disk vs CD

Where a genuine need for localised, direct-attached or otherwise isolated storage exists, the next question is which format? With near-line, you are looking at three platforms: tape, disk and CD.

"It is undoubtedly true that the cost of disk is coming down and the growth in the use of cheap secondary disks is increasing. The result is that the implementation of near-line storage is changing," explained Steve Mackey, sales director at tape array maker ADIC.

The cost of hard disk per gigabyte is now as low as £1, putting hard disks on a par with branded CD-RW disc media, and undercutting the cost of Zip which, even in its 750MB variant, works out at about £12.50 per gigabyte. Tape remains more cost-effective, but disk is now cheap enough to mean that the premium for the convenience is justifiable in a number of environments.

"It is critical is to compare the cost of the complete sub-system remembering that, with disk, it is essential to factor in the cost of backing up to ensure that the disk is protected," said Mackey.

There are also logistical elements to consider. Tape can be time-consuming, making it unsuitable in environments where back-ups need to be done quickly or incrementally.

"Traditional back-up tools, predominantly tape drives, have to deal with a constantly shrinking back-up window," explained Joseph Saroukhanian, senior product manager at Iomega.

This is particularly true in a server environment, where online trading and communications make it harder to co-ordinate and agree on a period to isolate a server and take a snapshot of it. The period of downtime can easily add up to a significant financial or productivity loss.

"The optical arena is developing quite nicely, while growth rates for CD-RW are slowing down," said Saroukhanian.

The emerging standards such as DVD+RW are beginning to steal the thunder of CDs.

According to IDC, the installed user base of CD-RW drives this year will be about 60 million. While the growth rate is slowing, it still represents massive sales, and compares favourably with Zip, which has around 45 million active users worldwide.

Disk also has the advantage of usually being based on standardised interfaces: SCSI and IDE on the inside; parallel, USB, SCSI and FireWire on the outside for externally connecting drives.

But Rupert Beeby, head of technical operations at storage solutions provider Sagitta, warned that many near-line storage systems, particularly non-hard disk magnetic technologies, have the downside of creating a lock-in environment. Once deployed, it's simply not cost-effective to break away from them.

"However appealing the concept, it is important not to forget that near-line storage is often based on closed proprietary architecture," he said. "Organisations are also squandering golden opportunities for open storage management offered by a digital assessment management system."

The very nature of near-line storage is such that bad planning or simply losing control of media purchasing and distribution can lead to much wider problems.

"Most organisations have their data monsters," said Beeby. "Simply confronting them can be a heroic task in itself.

"Even bearing these factors in mind, near-line forces organisations to confront their data fears. Throwing money into the storage pit simply feeds the data monster and it will grow. The real challenge is taming it. Organisations that successfully manage their data will get the most value from it."

So the near-line debate remains unresolved. Even with the changing dynamics of the storage business, there appears to be a role for near-line storage among a significant portion of the user base.

Others clearly benefit from banishing localised external mass-storage in favour of a more centralised repository that can enjoy economies of scale, easier management and the application of a single set of policies and digital document initiatives.

FOR AND AGAINST NEAR-LINE

For

- Localised devices can be more convenient
- Individual users can use near-line storage for rapid data transfer between users and devices
- Platforms are familiar to end users and are trusted
- Many removable platforms offer high-capacity and low hardware and media costs.

Against

- Centralised storage is easier to manage
- Multiple near-line devices can be expensive to support
- Disk/tape acquisition and reuse is hard to enforce and control
- Removable storage can be open to misuse.

REINVENTING THE ZIP DRIVE

The humble Zip drive deserves a special mention. While end-user tape drives have been around since IBM launched the original PC, they have generally been big, clunky and terrifying for all but the savviest user.

Backing up at the end-user level remained a fraught process, while restoring was cumbersome and slow, and early desktop tape drives simply were not sophisticated enough to allow the user to get at a single file. It was complete restoration or nothing.

Enter, in the early 1990s, the Zip drive: a proprietary disk format developed by little-known tape drive maker Iomega.

It was a hit. £12 bought you a 100MB disk, helping to drive another nail into the coffin of the floppy disk, and making small back-ups and hard drive clean-ups quick and non-threatening because the Zip worked just like a very high-capacity floppy drive.

Until recently, *vnunet.com's* sister title *Computing* used exactly this method for archiving the QuarkXpress files that make up each issue.

But the Zip has begun to fall foul of another near-line platform: the CD. The rapid collapse in prices of both CD-R and RW media, along with fast CD writer drives, has led many to cut back on their use of 100MB rewritable Zip disks, and the 250MB Zip platform introduced two years ago.

"It all depends on the application," explained Saroukhanian.

Naturally, he does not accept that Zip has been overtaken in the cost and convenience stakes.

"The trend among smaller businesses or in big companies on a workgroup level, is to use a great deal of removable near-line media, and the Zip drive continues to play a big part in that," he claimed.

To fight off CD, Iomega has revamped the platform yet again, taking into account the latest platter technology and producing a 750MB disk. It is larger in capacity than a CD, but still the same size as previous Zip disks, and with drives that are backwards-compatible.

"The capacity of CD was quite clearly superior to Zip, so we have now addressed that. Zips and, more recently, portable hard drives, continue to be among the most

convenient media to take the data with you any time you are moving to another location or need a duplicate copy of a piece of data," said Saroukhanian.

But look at the figures: a 750MB Zip disk costs the best part of £10, and drives cost about £129. Compare this with a 48x CD Rewriter for £49.99 from PC World and a 100-pack of CD-R write-once media for an average of £20, and it is easy to see why CD has had such a profound effect.

This article was printed from vnunet.com.
VNU Business Publications
© 2004 All rights reserved