SPECIAL ADVERTISING SECTION

inside storage

BusinessWeek

Reprinted from a special advertising section that appeared in the August 20/27, 2001 issue of BusinessWeek

the power of Petabytes

As storage needs continue to soar into the Petabyte range and beyond, companies must learn to manage these oceans of data. More mobile devices and far-flung workers are further changing the landscape.

Early this summer, a senior IT manager at Owens & Minor was preparing for a difficult meeting with his CIO about the status of several major system rollouts. Don Stoller, director of information management at the $3.5-billion medical products distributor, had to tell his boss that there was a significant unexpected cost associated with one of the projects.

The project is a new system that would let customers and suppliers run a set of comprehensive business intelligence reports against O&M's data warehouse. This will be done via a browser using the company's extranet and will force O&M to dramatically increase its storage capacity.

Because the new system will include data from its customers' internal procurement systems about non-O&M purchases, Stoller is projecting that the company's 350-Gigabyte data warehouse could quickly balloon to as much as 20 Terabytes. That's an increase of 6,000 percent.

Although O&M's stupendous data storage surge may be extreme, Stoller is not the only IT manager with burgeoning data storage issues. Purchases of disk storage are projected to grow at double-digit rates, from $28.4 billion in 1999 to $53 billion in 2004, according to analyst firm International Data Corp. (IDC). In 2003 alone, IDC estimates organizations worldwide will purchase 1.6 million Terabytes of storage, which translates into 1,600 Petabytes and — get ready for the next storage term on IT vocabulary tests — 1.6 Exabytes.

All the new data — including e-commerce, business-to-business, audio, video, e-mail, images and Web content — being created is only the tip of the storage iceberg. Given the importance of applications and data warehouses, Stoller, for

Cover image: PhotoDisc

example, mirrors all O&M data twice. The first instance is for automatic failover; the second is as part of the company's backup and recovery strategy, in which mirrored data is subsequently copied to tape.

O&M's original 350-Gigabyte data warehouse actually consumes just over one Terabyte of disk storage when the mirrored copies are considered, immediately tripling the amount of stored data. And this Web application is just one in a portfolio of data warehouse applications that may eventually include product images, further exacerbating the storage challenge.

The Three Choices

Technology managers facing unrelenting demands for more storage have only three choices: Throw more online storage at the problem; start throwing data out; or try to more intelligently manage their data. The first option — adding more storage — is so costly that it can have a serious impact on the income statement. Despite per-megabyte costs amounting to mere pennies, consulting firm Meta Group estimates that storage spending will consume as much as 80 percent of the annual IT budget. Just throwing more disk at the storage demand — no matter how cheap the purchase price — is an increasingly untenable strategy. The cost of administering the storage in terms of people and tools can run seven to 10 times its acquisition cost.

"Management certainly is concerned about the size of the storage investment," says Bill Moran, data center operations manager at Incyte Genomics, a genomics technologies and products provider to the biotechnology and pharmaceutical industries. The 10-year-old company already has 13 Terabytes of primary storage (mirrored for a total of 26 Terabytes) plus many more Terabytes on Network-Attached Storage (NAS) appliances. Looking across all platforms and servers, Moran estimates the company manages 50 Terabytes of storage, and it continues to increase — rapidly.

Recognizing that storage was spiraling out of control, Incyte brought in consultants to help the company rein in its storage growth. The proposed ideas revolved around moving data from online storage to someplace else, tantamount to throwing out data. User resistance thwarted those plans. "The problem is that everybody wants all their data online," Moran says.

Like it or not, "companies are going to have to develop a storage resource management strategy," says Tony Prigmore, senior analyst at storage industry analyst firm The Enterprise Storage Group. Such an approach entails prioritizing storage requirements based on the applications and the need for performance and high availability. Companies then can implement the appropriate storage strategies, ranging from Storage Area Networks (SANs) and the creation of virtual storage pools to implementing storage archiving and Hierarchical Storage Management (HSM) systems.

The process is pretty straightforward, says Chris Gahagan, vice president and general manager for recovery and storage at storage vendor BMC Software. First, users need to decide what is important and calculate how much data they have and how long they need to keep it. They then need to determine the cost. Finally, "compare the cost to the value you receive," Gahagan says, while asking, "How much additional revenue will be generated by keeping all this data?" Much of the data is unnecessarily redundant or of negligible value and can be thrown out.

SANs Popular

For what is left, SANs are emerging as the preferred approach to gaining control over storage. The SAN consolidates Terabytes of storage attached to systems throughout the enterprise onto a separate high-

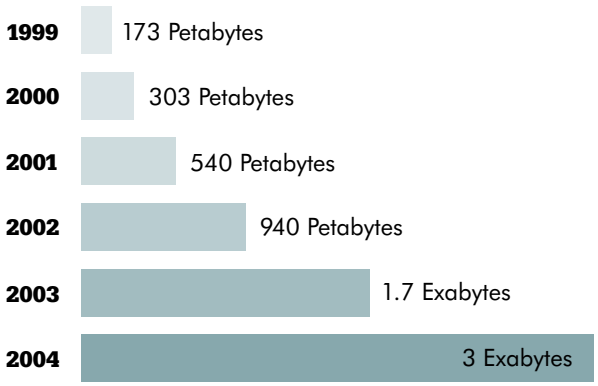
TIPS

Are SANs Sans Security?

Beware: The networking protocols that create Storage Area Networks do not address security. If a server attached to the SAN is compromised, the stored data is exposed completely. Fixes are currently underway.

Storage Shipments To Soar

Corporate storage is moving beyond Petabytes into the uncharted Exabyte neighborhood.



Source: IDC

speed network. It removes storage traffic from the corporate network while enabling any network server to reach any data stored on the SAN. Without a SAN, applications typically access only data physically connected to the server.

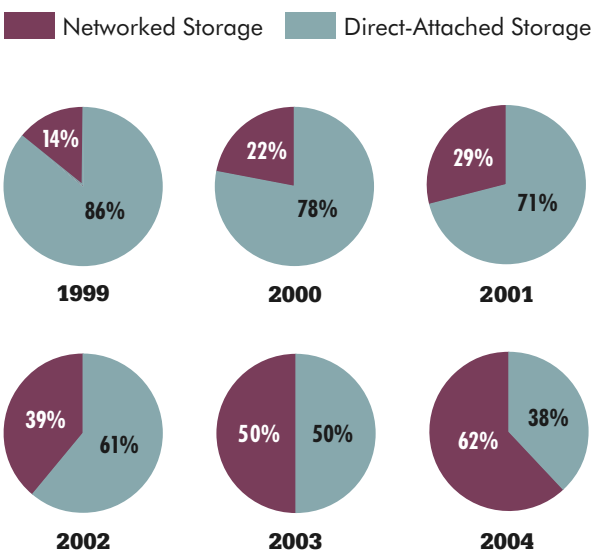
In effect, the SAN centralizes storage, allowing for more efficient administration and better storage utilization through virtual pooling, the ability to treat all the storage capacity as a single pool that can be dynamically allocated among different applications and uses. Without the SAN, one application can be desperate for storage while extra storage attached to a different server languishes unused and unreachable by the needy application.

Having seen rapid growth in both its primary mainframe storage and its server-based storage, discount shoe retailer Payless ShoeSource is turning to a SAN. The initial implementation is small — about 800 Gigabytes — to handle some server-based database applications and IBM's Lotus Notes. The benefit: "I can allocate storage by going to the console and redirecting it. If Notes needs more storage, I just click some buttons," says Dick Gorman, senior enterprise storage administrator. The SAN also greatly simplifies backup, he adds.

But SANs are no panacea. The technology is immature, interoperability among components from different vendors is limited and security — a critical concern whenever valuable information assets are stored — presents a challenge.

Storage Will Soon Get A Lot Less Direct

Companies will push more data onto servers that can be managed directly by IT (networked storage) and move away from storage drives directly connected to computers (direct attached).



Source: IDC

NAS, HSM Options

Network-Attached Storage also helps address the storage problem, but NAS alone does not provide a complete answer. It acts similar to a SAN by eliminating the need to directly attach storage to a particular server. It also allows any server on the network to access data stored on a NAS device. In general, NAS handles files associated with applications while SANs handle blocks of data associated with large databases. As a result, NAS is increasingly regarded as complementary to SAN and part of a total storage strategy.

Among other storage options are data archiving and HSM, which offer some relief by moving data from online disk to lower cost media, such as tape. HSM systems combine software that monitors data usage and moves infrequently used data to an automated tape library with the ability to quickly bring back the data when it is requested.

But HSM requires overcoming user resistance and enforcing rules about data retention, and it still might not solve the problem.

Even companies with the best data archiving intentions are stymied by the new types of data, to which conventional archiving approaches don't apply. "We have HSM for data archiving, but a lot of the data types now are so different, such as streaming media, that we really need a different kind of HSM system," complains a senior IT manager at a major consumer food products firm.

Companies need to get a handle on storage now; the problem is not going away. As the Internet Age unleashes oceans of never-before-available critical customer data, somebody is going to have to manage it, before it drowns them. ■

It's 11 p.m. Do You Know Where Your Data Is?

More data sits away from traditional offices and with remote employees on laptop and handheld computers. The challenge is to keep that data safe and useful — no matter where it is.

Mobile computing is great — until a senior executive loses a laptop in an airport, along with all the data stored on it.

That's what happened to Concert Communications last May, and it was enough to spur the global telecommunications giant to deploy software that automatically backs up critical data and settings from remote PCs, says Mark O'Malley, a Concert senior manager for projects and implementation.

Although he declined to identify the data on the lost laptop, O'Malley says Concert employees routinely travel with customer proposals and project plans that take months to produce. With more than five thousand laptop computers deployed, "our data is no longer centrally contained," and employees usually don't think to back it up, he says.

It's the dark side of mobile computing: Every salesperson polishing a presentation on an airplane or finishing a proposal at home is producing data that is stored only on a mobile device. Since that data often is never backed up on a corporate server, it can be lost forever if the device is misplaced or stolen, or if it just stops working.

Prepare Now

Keeping mobile data safe is only part of the mobile computing challenge. To stay on top of market conditions, headquarters may need to update mobile users with, among other data, new price lists, customer leads and revised forecasts as salespeople complete their calls. Right now, neither the software nor the wireless access is robust enough to do this easily. But now also is the time to prepare, before laptops become even more ubiquitous and handhelds, such as Palm's namesake units and Compaq Computer's iPAQ, gain added data storage capabilities.

Ninety percent of corporate IT departments fail to do regular backups of individual computers, says Chris Gahagan, vice president and general manager for recovery and storage at storage vendor BMC Software. "Our clients, to date, have done a lousy job of making sure laptops are properly backed up," adds Ken Dulaney, an analyst with Gartner, a research and consulting firm. "They've spent all this money on disaster recovery for the mainframe, but some of the most sensitive information they have is sitting on laptops completely unbacked-up." That is a scary prospect, since Gartner estimates businesses bought 18 million laptop computers last year, each with a five-Gigabyte hard drive waiting to be filled with critical information.

Some corporations run software that automatically backs up systems linked to fast corporate networks. That same software, however, can cause intolerable delays over a pokey dial-up connection in, for example, a hotel room. Backup can't be left to busy individual users, "because they won't do it or won't do it properly, or they won't do it often enough," says Jennifer DiMarzio, an analyst with market researcher Summit Strategies.

Backup To Go

While laptop computers are the greatest concern now, handhelds will pose a greater threat within a few years as such devices gain the ability to store more data and run more sophisticated applications, Gartner's Dulaney says. Many companies, adds BMC's Gahagan, don't even know how many handhelds employees buy on their own or what data from the corporate network they are downloading to those devices. Adding to the confusion are the many employees paying for handhelds out of their own pockets and the vendors still scrambling for dominance of

TIPS

Forbid The Kitchen Synch

Allowing everyone to freely synch data from desktop computers to laptops and handhelds can pose security risks. Consider using server-based tools and programs that centralize backup implementation rules.

the protocols governing wireless data transmission.

But the number of smart phones and handheld PCs capable of storing data remotely will only grow as more workers spend time on the road and as wireless broadband connections become faster and more common. Market researcher International Data Corp. predicts the number of mobile and remote workers in the United States will rise from last year's 39 million to 55 million in 2004. And while many current handheld devices store little data, Gartner's Dulaney predicts that "every handheld sold next year will have an [expansion] slot, and people will purchase at least 32 Megabytes of additional storage" for them.

While "virtually every one of our enterprise customers has some portion of their workforce using PDAs," says BMC's Gahagan, few back up or synchronize any data on them except for e-mail.

Smart Backup

Corporate information managers also must be able to tweak the backup process so specific users can receive updates on — or back ups to — different information as users' jobs change.

"As robust as these mobile devices are getting to be, they're not as robust as the central database servers," says Bill Jones, vice president of product management for wireless software developer Synchrologic.

Indeed, many remote devices have substantial communication challenges. Today's handhelds, for example, have much slower connections, dramatically smaller memory and even decidedly less practical screens. Security protocols that require changes in data access privileges — such as when an employee's job or location changes — also require changes to data synchronization procedures. This is especially vexing when both the devices and the employees are remote.

TIPS

May I Audit Your Desktop?

Very few information workers need 40 Gigabytes of desktop storage for business files. If users are filling up these hard drives, check for MP3 files, downloaded images, video clips, family photos and computer games.

Another thorny issue in the remote world is conflict management, Jones says. If two salespeople in an office try to reserve the last widget on the shelf for one of their customers, for example, the salesperson who enters the order a split-second earlier reserves the widget. The other salesperson, meanwhile, instantly sees that widgets are now out of stock. The process is far harder to coordinate when salespeople enter orders on their handhelds but do not synchronize the devices with the central database server for hours. "You have to have smarts on the server end to automatically detect and resolve any conflicts that may have arisen," Jones says.

New Updates To Consider

As applications are developed specifically for wireless devices, IT departments will need to update all employees' remote devices with new versions of those applications as bugs are fixed or new features are added. The ability to know what version of an application is on a device goes far beyond today's idea of simple data backup and restore. What customers will need for wireless devices is the same kind of systems administration and management capabilities they have had for years on mainframes and other enterprise systems.

It's no surprise, therefore, that traditional suppliers of IT products and services that already sell high-end storage products to corporate IT departments are beginning to create products for the wireless market. In January, for example, Hewlett-Packard purchased Bluestone Software for its wireless application infrastructure, which is designed to make it easier to extend existing Web applications to wireless and handheld devices. Hewlett-Packard plans to combine Bluestone's software with its own connectivity offerings to provide, the company says, the same capabilities for wireless applications as it does for core business applications.

Going to the trouble of finding and installing remote backup tools can pay off in reduced support costs. After beginning regular backups of its laptops, O'Malley says, calls to Concert's PC support group from users complaining about lost or corrupted files "have almost been eliminated." Users now can restore the files themselves. "If they're in Sydney or they're in Amsterdam, they don't have to depend on calling somebody out of bed in the middle of the night. They can actually do much of their own repairs, because it's a very simple interface to restore and retrieve files," he says.

Concert, though, had the luxury of only needing to back up data from laptop computers. Most large corporations "want to be able to support the top three devices — Palms, PocketPCs and RIM pagers — and whatever else happens to be the 'device du jour.'" says Andrew Hogg, director of product marketing at synchronization software vendor Pumatech.

The OS Dilemma

Finding tools that can handle all those different platforms will be difficult for some time, as manufacturers continue to squabble over how to handle issues such

Meet The Future Storage Manager

Whether or not he intended to be one, John Guilz is a remote storage pioneer. The IT manager for North American operations at billion-dollar office-supply giant Esselte has seen his number of remote employees quadruple in three years, to the point where 40 percent of the people he supports are off-site.

Guilz today oversees a hard-disk army of more than 7.6 Terabytes. That, he says, is more than double what he had in 1998, and he projects it will grow at least one Terabyte a year for the foreseeable future. Only about 2.1 Terabytes of the current total capacity are on the 32 servers Guilz manages at headquarters, with the rest spread over hundreds of field locations.

Knowing that other corporate technology managers will likely be in the same position in just a few years, the best single piece of advice Guilz can offer is to improve asset management inventory procedures.

"With Y2K, we had to physically go around to get inventory on every machine," says Guilz, adding that the remote machines needed to be packed up and shipped to headquarters, analyzed, repacked and shipped back.

as mobile-friendly operating systems.

The ideal remote data management tool would work with all the popular platforms — laptops, Palms, RIM Blackberries and handhelds based on Microsoft's PocketPC operating system. Such a tool would automatically adjust for any speed connection, from a wired desktop cradle to slow wireless connections, as well as allow users to continue working with the information on their devices even when not linked to the corporate network.

This ideal tool also would make it easy to automate policies that govern which information would be sent to or from the proper mobile users at the proper time. "In remote backup, policy becomes one of the most important issues," says Tarkan Maner, vice president of corporate marketing for software vendor Computer Associates. "In a LAN environment, for example, your IT department may ask you to change your passwords every three months, whereas you may have to do this every two weeks if you are mobile and using your passwords in public places and over the public network called the Internet."

Security fears can cut both ways, adds Maner, pointing to some healthcare and government agencies that forbid remote employees from doing standard backups because of a lack of security on the backup device itself.

Stop Reinventing The Wheel

Far from that nirvana, some of today's tools are limited to only a few handheld platforms. Others are capable of sending either only to or from the remote device, while still others lack the capability to selectively update only important data. "We're seeing a lot of reinvention of the wheel," says Charles Lukaszewski, chairman and CEO of handheld management outsourcer Ubiquio, as impatient customers write replication software only to see those capabilities quickly matched by software vendors.

TIPS

Insist On 2-Way Or No Way

Consider tools with a two-way synchronization capability. You may need to send updated data, such as price lists, customer leads and revised forecasts, to remote users as well as pluck information from their systems.

For now, analysts recommend looking for remote data management tools that: work on a variety of mobile devices; can adjust how data is replicated to compensate for varying connection speeds; work "in the background" without help from the user; and can send as well as receive updates from the mobile device. Tools that centralize backup and replication rules at a central server can help the corporate IT group impose some order on the backup process. ■

Triangle Publishing Services Co. Inc.
The Best Strategic Content for Web, Print, Multimedia and Beyond

This storage special advertising section was prepared for *BusinessWeek* by Triangle Publishing Services Co. Inc.
The Best Strategic Content for Web, Print, Multimedia and Beyond
www.triangle-publishing.com ■ 617-244-0698

Triangle is a leading content provider of business and technology editorial for publishers and vendors. Triangle delivers print, Patent-Pending online and multimedia products, as well as a wide variety of publishing services.

Storage Section writers: Alan Radding, Robert L. Scheier and Worth Civils