

SMARTER STORAGE MANAGEMENT

Fine-tuning
storage infrastructure
can help you
do more without
spending more.





Every passing year, companies produce more data than the year before. In a recent *InformationWeek* study, only a handful of companies experienced less than 10 percent growth in their storage needs last year; most companies are seeing 15 percent, 25 percent or even 50 percent increases.

At the same time, a struggling economy has led many companies to tighten their belts. While IT budgets don't appear to be dropping across the board, neither are they expanding. So the big question is: How do you do more without spending more?

The solution: A comprehensive effort to optimize infrastructure. This is one where storage can be increased or decreased as business needs dictate, redundant data that consumes excess storage can be eliminated and all available storage can be treated as one pool that can be tapped for maximum utilization.

Finding and eliminating that data, or better yet preventing its creation in the first place, offers another route to maximizing available storage without increasing budget. Leveraging a mix of cheaper slow storage and well-placed fast storage can provide more bang for the buck. And cloud-based storage is another way to ease storage woes and considerably improve management efficiency.

Growing Data

Although one cause of the tremendous growth of corporate data is simply the increased computerization of almost every aspect of business life over the past couple decades, the largest growth can be traced to two separate trends. The first is the increase in "unstructured" data — data generated by end users such as documents, presentations, images, spreadsheets and the like.

Chaotic and unpredictable, unstructured data grows at an effectively random rate and can come from virtually anyone in and even outside an organization. "Just a few years ago that wasn't the case," says Joe Disher, solutions marketing manager at Overland Storage.

"You used to have to be an expert in a specific application to create data," he says. "Today, all these things can be created on a mass scale by just about everyone."

The other trend shaping our growing storage needs is virtualization. "Let's say you have a server with four applications that you want to virtualize," says Dick Csaplár, an Aberdeen Group senior research analyst. "When you virtualize that server, now you have four copies of the operating system and the original copies of the application, plus the software from the virtualization layer that you didn't have before, all on one machine."

Adding new servers has become simple. However, adding the data storage capacity to sustain all those servers has become more complex.

The increased trend toward desktop virtualization also affects storage needs. Desktop virtualization offers many benefits, not the least of which is that it allows end-user files and data to be treated the same as other important corporate data.

Because it does not "live" on desktop or notebook hard drives, no separate security, backup and update policies are needed. Still, bringing user data into the data center means a place must be created to store it.

"The company looking to save on desktop operations through a virtual desktop initiative will likely have to build out more consolidated storage," says John Sloan, lead research analyst with the Info-Tech Research Group. "This is to support the virtual desktops that now reside on servers."

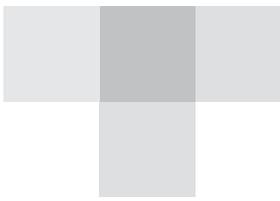
One thing is clear, simply throwing new drives at your data is not a viable long-term solution. It's simply too expensive and too inefficient. Although some of the solution lies in deploying technologies such as automated tiering, disk deduplication, storage virtualization and thin provisioning, the real fix rests in how you think about your data and the business processes around it.

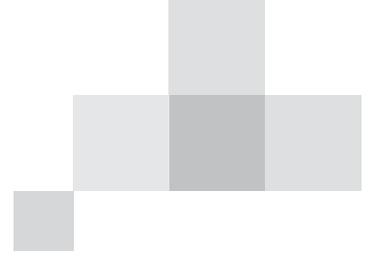
Virtual Solutions

One of the biggest recent trends in storage is the increased use of virtualized storage, tapping storage area networks (SANs) and thin provisioning from a central disk array. The concept behind storage virtualization is simple: Virtualization divorces the use of storage space from the physical location where data is actually stored.

So, for instance, a server might "see" a 500-gigabyte hard drive attached to it. But in physical space that storage is spread over two, three or even more storage devices, with a virtualization layer in between translating the reads and writes to the actual drives.

>>>





Storage virtualization offers a number of benefits. However, for IT departments looking to maximize their storage budgets, perhaps the most significant is the ability to recapture stranded storage capacity that would otherwise go unused.

“You don’t want to have these islands of storage from different vendors and in different locations,” explains Csaplár. “What you do is put a layer of storage virtualization software on top of it, and that allows you to link all these storage arrays into one large virtual storage platform.”

Using a SAN such as HP’s StorageWorks P4000 G2 SAN Solutions or the IBM System Storage SAN Volume Controller, you can install memory when and where it’s practical and assign it to servers and applications anywhere on the network, all completely transparent to the end user.

Once the virtualization layer is in place, new possibilities open up such as thin provisioning and automated memory assignment. These technologies can help maximize the usage of existing storage capacity even further.

full terabyte instead of the one-third terabyte of actual space they’ve been given) and will continue to try to write to the virtual disk.

Automatic provisioning offers a way around this by increasing the available storage space as needed. Overland Storage’s S2000 SAN appliance uses a technology called Auto Capacity Extension which, Disher explains, “gives the ability to provision storage as you need it, so it’s ‘just-in-time’ provisioning. Firms set up a policy so that when storage on such and such device reaches a certain percent, that storage automatically grows by X percent.”

Another benefit of storage virtualization is the ability to move data between physical drives without interrupting the performance of the device using that storage. It’s also possible, as server and client virtualization become more prevalent, to move virtual machines from one machine to another without worrying about copying the live data.

“SANs are an essential component of server consolidation,” says Info-Tech’s Sloan. “Virtual server infrastructures are better enabled when all the physical host servers share a common storage pool,” he adds. “This enables features like live VM [virtual machine] migration where a virtual server can be moved from one physical host to another with zero downtime.”

Finding the Right Medium

Behind all those virtual drives are still physical devices where data is actually stored. Therefore, careful consideration of what data should be stored on what kind of drive is another way to maximize the effectiveness of a flat storage budget.

While high-speed flash drives have begun to come into their own, less-expensive serial ATA drives and tape still offer the best value for most storage needs. And the savvy IT director can easily mix and match drive types to balance performance and cost using automated tiering.

“Where virtualization is all about maximizing utilization, data tiering is about lowest dollar per gig — or rather the right-size dollar per gig,” says Chris McCall, product manager for the HP storage product group.

Data that is rarely used but which needs to be kept for business historical or regulatory needs does not need to be on fast, easily accessible drives. It can be stored far more cheaply than, say, transactional data from your company’s sales database or customer relations tools.

Products such as HP’s P4000 or EMC’s CLARiiON SAN disk arrays use automated tiering to make sure that data is saved to the most cost-effective medium based on business uses and policies. For instance, EMC’s Fully Automated Storage Tiering (FAST) can identify data that hasn’t been accessed in a preset period and shift it to lower-speed (and lower-cost) storage. To the end user, nothing appears to have changed.

**GET EXPERT ADVICE
ON STORAGE
SOLUTIONS AT
CDW.com/storage**

Thin provisioning allows assigning the same physical memory space to several devices concurrently, essentially oversubscribing the actual disk space available. For instance, a single 1-terabyte drive might be provisioned to three separate servers.

This would normally require three separate 1TB drives but that might only be using 100GB at the moment. As the drive fills up, more memory can be assigned — but that’s on next year’s budget, or the year after, when most likely drive space will be cheaper.

Thin provisioning is effective at increasing disk utilization, but it does require a watchful eye. If the physical space allocated fills up, the servers themselves don’t know (since they see a

Storage in a Flash

With prices for solid-state storage dropping and storage capacities increasing, flash-based storage has finally become a viable option for many businesses. Although still more expensive than traditional hard drives, the advantages of solid-state storage make it ideal for specific applications where its speed and resiliency more than pay for the initial investment.

Flash storage has many benefits over spinning hard drives. For one thing, no moving parts means far lower failure rates. Flash drives also generate less heat than rotating drives do, and they consume less energy. But the biggest benefit by far is speed.

Whereas hard drives must spin-up and position the read head over the proper location on the disk, flash drives have instant seek times. They can read and write at several times the rate of even the fastest physical drives.

So when does it make sense to shell out for flash? For most applications, the speed gain may not be worth the price. You'll need to do the math to identify where the "bang for the buck" is great enough to pay off.

For instance, consider this example from Overland Storage's Joe Disher. "Let's say you have an online retail business and thousands of people are accessing your site at any given time. In this situation, having instant access to all the database elements is critical. That's your money."

Solid-state storage offers a remedy to this type of bottleneck. Wherever you have a highly transactional environment, supplying the application with flash storage instead of a spinning disk can be a lifesaver. Because flash is more durable than magnetic disk, it can also help prevent lost sales or other transactions due to disk failure.

According to Dick Csaplar, senior research analyst at Aberdeen Group, 50 percent of server outages are due to hard drive failure. Therefore, installing flash in mission-critical servers can greatly reduce downtime overall.

"Storage tiering allows IT managers to leverage high-performance flash drives for applications that need higher levels of performance," says Scott Delandy, senior product manager at EMC. "Doing that allows firms to get more performance and do it more cost-effectively than with physical disks."

But businesses can also take advantage of high-capacity drives such as 2TB SATA disks. Here you can identify where the inactive data is or the parts of the application that aren't heavily performance-oriented, and move that data to these high-capacity devices to significantly reduce the price of storage.

Cloud-based storage is also becoming increasingly useful, especially for backup and archival data. While the tools for adding cloud storage into the active storage infrastructure are rudimentary so far, in applications where speed takes a back seat to availability, cloud storage is becoming an important option.

For instance, storing backups in the cloud limits the speed of disaster recovery to the speed of an Internet connection. But having the data safe and secure offsite, when everything goes sideways onsite, is a much more important concern than being able to recover data at flash-like read speeds.

Whether storage is virtualized or not, and however storage is tiered, it goes without saying that storing less data is cheaper than storing more. This is where deduplication comes in.

Deduplication can be applied across the network or at the drive to remove repetitive blocks of data. These can include signature blocks in e-mail stores, images used in thousands of presentations across the corporation, even words that are used repeatedly in text documents — replaced with a pointer to a single copy.

"We find a lot of copies of data that get used over and over and are saved over and over again," says Delandy. "This is particularly in file environments and backup environments." Applying deduplication into the environment significantly reduces the amount of storage needed because it's finding all of that redundant data and eliminating it. ♦

STORAGE is now the single most power-hungry application in the data center, accounting for

32.3%

of corporate data center power usage on average.

Source: "Breaking Point: 2010 State of Storage,"
InformationWeek Analytics