



Caution: Data Overload Ahead

New tools can help companies free up storage space, reduce costs and rein in out-of-control data.

IT departments today face opposing forces when it comes to data storage. While budget belts tighten, the amount and types of data requiring storage are on the rise. Many IT departments just can't seem to keep up with exponential data growth.

According to a study by "InformationWeek," nearly half of 331 business technology professionals participating say their top storage concern is insufficient resources for critical applications. Compared to a year ago, more IT pros say they have insufficient budgets to meet business demands, insufficient tools for storage management and insufficient storage resources for departmental and individual use.

Meanwhile, the data growth rate isn't letting up either. In 2009, three-quarters of survey-takers reported administering more than 1 terabyte of data, while a quarter managed more than 100 terabytes. By 2010, the number of respondents exceeding 1 terabyte climbed to 87 percent, while 29 percent administered more than 100 terabytes.

The dilemma has storage managers looking for ways to optimize existing storage systems, as well as find affordable new options for storage management. Learn how storage experts explain today's storage optimization choices and how to decide what is right for your company.

STORAGE VIRTUALIZATION

This represents the pooling of physical storage from multiple network storage devices into what seems like a single repository that is managed from a central console. Storage virtualization is commonly used in a storage area network (SAN).

By disguising the actual components of the SAN, storage virtualization helps the storage administrator perform the tasks of backup, archiving and recovery more easily — and in less time. This also makes data recovery in time of disaster or disruption much easier.

Many mid-size companies are following the enterprise lead by implementing virtualization as a way to rein in a myriad of storage devices under the umbrella of a single set of storage applications for replication, mirroring or snapshots. Therefore, there's no need to learn multiple applications to achieve the same function.

"Many of them have waited for large companies to validate that it's a robust technology and that it's viable and some are getting to final testing phases," says Joe Disher, technical marketing manager at Overland Storage. Likewise, storage vendors like Overland have unveiled products that make storage optimization easier in a virtualized environment.

In February, Overland Storage announced its SnapServer SAN S2000, an iSCSI SAN platform for businesses with growing data. The solution enables companies with Windows, VMware, Hyper-V, UNIX, Linux or Mac OS X to take advantage of SAN capabilities that have traditionally been considered too sophisticated without specialized expertise.

The SnapServer SAN S2000 enables policy-based storage volume growth to occur in a just-in-time manner without IT intervention. It scales to 120TB and is a cost-effective solution for businesses with virtual server environments or storage consolidation requirements.

Disher adds that the SnapServer SAN S2000 combines the maturity and reliability of a SAN platform with the leading-edge features of virtualization. "Obviously SAN technology itself is considered an optimization step — a common storage platform to share among many servers or hosts," he adds.

It's important to note that virtualization adds an extra twist to the storage optimization challenges, especially for IT departments using VMware, says Kevin Gray, senior product marketing manager at EMC Corp. With VMware, "you cut up [chunks of storage] and hand it over to the VMware administrator — who is now applying that capacity at an on-demand basis. Unless the tools are in place to identify the relationship between the virtual guest and that capacity, managing the configuration becomes a bigger issue."

To that end, vendors such as EMC offer its Ionix ControlCenter family of storage software that can simplify and automate tasks such as discovery, monitoring, reporting, planning and provisioning for large, complex physical and virtual infrastructures.

HP points out that virtual environments require high levels of data availability and storage utilization plus the ability to scale nondisruptively as the business evolves. Its HP StorageWorks P4000 G2 SAN Solution (formerly Lefthand) touts that it can reduce costs for multisite, high-availability storage by as much as 50 percent, increase capacity utilization by one-third and deploy shared storage within a virtual server environment without disrupting business operations.

THIN PROVISIONING

In a shared storage environment, this is a method for making the best use of available storage. It relies on on-demand allocation of blocks of data versus the traditional method of allocating all the blocks up front.

This methodology eliminates almost all unused storage spaces and maximizes the use of each space, often leaving less than 10 percent of unused storage space, compared to traditional storage allocation methods where large pools of storage capacity are allocated to individual servers but remain unused. This traditional model is often called “fat” or “thick” provisioning.

“With thin provisioning, storage capacity utilization efficiency can be automatically driven up toward 100 percent with very little administrative overhead,” according to George Crump, lead analyst of Storage Switzerland, an IT analyst firm focused on the storage and virtualization segments.

“Organizations can purchase less storage capacity up front and defer storage capacity upgrades in line with actual business usage,” he points out in a recent column. “They can save the operating costs — electricity and floor space — associated with keeping unused disk capacity spinning.”

Several vendors offer thin provisioning solutions. For virtual storage environments, EMC’s Ionix ControlCenter Automated Resource Manager lets storage managers meet or exceed time-to-provision service level commitments by quickly fulfilling application requests for storage. The software simplifies storage allocation by enabling automated policy-based storage provisioning.

The software helps storage managers establish repeatable, auditable provision policies that identify free storage available from hosts, host bus adapters (HBAs), ports, arrays or devices. Add-ons include a performance manager that collects, correlates and graphically presents performance information for the storage infrastructure, and a SAN Advisor that simulates changes to the SAN design and foresees potential effects before they go to production.

AUTOMATED TIERING

Many storage vendors are beginning to offer automated tiering as part of their optimization tools. The technology allows the storage area to look at performance characteristics of the volume and by itself decide whether to put that data on a higher tier or a lower tier based on its importance.

The storage area is now keeping track of the performance characteristics of that data sample. And it can take corrective action — either move it upstream or downstream — based on what policies the administrator defines.

What type of storage environment would benefit from automated tiering? Crump says it is a must-have if a business has a performance profile that can take advantage of the performance it can deliver.

“If you need to broadly increase storage input and output across multiple applications, then automated tiering can provide significant value,” he adds. “If your applications aren’t pushing your current storage platform, putting faster storage in won’t make that big of a difference.”

But on the other hand, “An increasing number of storage managers are at that wall,” he adds. “Thanks to virtualization, online applications and growing databases, automated tiering may be a must-have.”

Most vendors, in some shape or form, offer automated tiering, says Ashish Nadkarni, principal consultant at storage consultancy GlassHouse Technologies Inc. “The question is what environment you have and what vendors have you locked into. It’s more evolutionary than revolutionary. So you shouldn’t just dump your vendor solution for this one incremental piece.”

The promise of automated tiering is that it will remove one of the challenges and time consuming tasks from storage managers, which is data placement, Crump adds. "For many organizations, it may be the only practical way to fully leverage all the new tiers of storage."

DATA DEDUPLICATION

This is a method of reducing storage needs by eliminating redundant data. Only one unique instance of the data is actually retained on storage media, such as disk or tape. Redundant data is replaced with a pointer to the unique data copy.

With data deduplication, only one instance of the attachment is actually stored. Each subsequent instance is just referenced back to the one saved copy.

Data deduplication lowers storage space requirements, which saves money. Efficient use of disk space also allows for longer disk retention periods, which provides better recovery time objectives for a longer time and reduces the need for tape backups. It also reduces the data that must be sent across WANS for remote backups, replication and disaster recovery.

"With online data deduplication, we just have to keep in mind it has the potential to impact storage performance," Nadkarni says. "So it's like thin provisioning. You have to be careful about what kind of compression ratios you have when you do a data dedup online, and how does rehydration of that duplicated data affect performance?"

Overall, thin provisioning gives more tangible benefits in the immediate term than data deduplication unless it's in fileshare or VMware environments, Nadkarni adds. "That's where you know there is a lot of duplicated data."

STORAGE IN THE CLOUD

Cloud storage, another virtual technology where companies save data to an offsite storage system maintained by a third party via the Internet, is another optimization option in its early stages. Data in a "public" cloud, which is shared with other users, can be accessed from any location that has Internet access and can serve as a useful business continuity and disaster recovery tool.

Many companies are looking at public cloud storage as a destination for tertiary copies of data rather than primary copies, Nadkarni says. Private clouds, which give a company its own dedicated storage, on the other hand — "can bring virtualization within their data center," he adds. "But they are a little tougher for proprietary data right now.

"Keep in mind, you certainly can find yourself locked into a vendor environment if you go with a private cloud strategy," he says. "So be careful building out a cloud unless you know exactly what kind of investment and lock-in it's going to be."

At the Burton Group in Salt Lake City, Utah, senior storage analyst Gene Ruth says clients are considering cloud storage, but they are proceeding with caution. "I've heard plenty of people try to pick and choose what might be an interesting application for cloud storage," he says. "But it's an emerging market and not a done-deal."

GETTING STARTED

Taking the first steps toward storage optimization involves first determining your goals. "What are you trying to solve?" Nadkarni says. "Is it a cost problem or a performance issue? There should be some blueprint."

Second, determine how the end result of the optimization will be measured. "Can you say that you saved X amount of dollars over a period of time or be able to now accommodate more assets onto the same storage frame without buying more processing power," he adds. "You need to be able to justify that, or you're wasting a whole lot of man-hours doing nothing."

Finally, find out if other parts of the organization have already made or are planning to make investments in virtualization, virtual desktops or asset refreshes, for instance. Those kinds of things can be partnered into this effort and can save money.

Matching Goals to Optimization Solutions

Whether the goal is to save money, get a grip on out-of-control storage or invest in new technology, there is no shortage of optimization options. Principal Consultant Ashish Nadkarni at storage consultancy GlassHouse Technologies, Inc., offers tips for determining the right storage solution for your goals.

Goal: Saving Money — Companies can save money on storage by reducing run rate, such as through data deduplication. “If you buy 10TB of storage every year and are able now to stretch that over a year and a half — you’re making more efficient use of storage. Even if your budget is stagnant, you’ll be able to afford more or use the same money elsewhere.”

Goal: Getting Data Under Control — If your data is all over the place, you need data classification, which can be aided with automated tiering, Nadkarni says. Start by looking at where all data is stored and come up with a model for classifying data as higher tier, mid-tier or lower tier.

Then define the performance characteristics that should be used to measure the tier and what kind of data goes on it. “That data qualification is the first step,” he says. “Once you’re done with that — then you do your data mobility initiatives — how you’re going to move from one location to another.”

Goal: Freeing Up Resources — If you have the money to invest, you would want to invest in technology that allows you to do automated tiering, Nadkarni says. “It frees up administrators’ time and keeps data optimized. This is a new area that vendors are jumping on.”

LET CDW'S DATA STORAGE SPECIALISTS HELP YOUR BUSINESS CONSOLIDATE STORAGE INFRASTRUCTURE AND FREE UP I.T. RESOURCES FOR MORE STRATEGIC INITIATIVES. CALL 800.800.4CDW TO TALK TO A SPECIALIST TODAY.

071410