

Managing the information that drives the enterprise

# STORAGE

Vol. 8 No. 8 November/December 2009

# Hot Technologies

Capacity is up, budgets are down.  
We look at five technologies that  
will most likely make a difference  
on your storage shop floor in 2010.

p. 10

## ALSO INSIDE

- 5 Can consolidation backfire?
- 23 Inside the latest VTLs
- 34 Storage salaries on the rise
- 44 Are file systems old-fashioned?
- 51 Better backups with desktop virtualization
- 54 Power conservation zapped again

# STORAGE

sponsors | november/december 2009



## Can We Survive Consolidation?

- 5 EDITORIAL** Dedupe, server virtualization and data archivers are great tools to control storage capacity growth, but they treat the symptoms and don't provide true consolidation. Don't throw them out; use them better. *by* RICH CASTAGNA

## Hot Storage Technologies for 2010

- 10** In our annual assessment, we pick five technologies we think will impact your storage operations in 2010. Read how VMware backup, solid-state storage, thin provisioning, 8 Gbps Fibre Channel and data dedupe for primary storage can change how you manage storage. *by* STORAGE/SEARCHSTORAGE.COM STAFF

## Virtual Tape Libraries in Depth

- 23** Virtual tape libraries (VTLs) have been a relatively easy way to replace traditional tape libraries, but as other disk backup targets emerged, many thought VTLs would disappear. Now, with added features such as dedupe, they can be an attractive alternative to other disk target systems. *by* W. CURTIS PRESTON

## Salaries Rise as Storage Grows

- 34** In our exclusive annual survey charting the salaries and benefits of storage pros, many of them managed to see pay increases even as closings and layoffs sent some looking for new jobs. As data capacities grow, so does the need for dedicated storage pros. *by* ELLEN O'BRIEN AND RACHEL KOSSMAN

## Object Storage Gains Steam as Unstructured Data Grows

- 44** Object storage isn't a new concept in the NAS world, but some new products are bypassing traditional file system interfaces as an industry debate emerges about the best way to cope with unstructured data. *by* BETH PARISEAU

## Desktop Virtualization: Better Data Protection?

- 51 HOT SPOTS** Backing up desktop/laptop PCs has been a thorn in the side of storage managers. Virtual desktop infrastructure technology can ease the burden of data protection for PCs, but it may not be a fit for all users. *by* LAUREN WHITEHOUSE

## Energy Conservation Efforts Still Anemic

- 54 SNAPSHOT** Our Snapshot Survey reveals that "green storage" is still not top of mind for most storage managers. Some might be willing to spend more on systems that promise energy savings, but most are still dubious. *by* RICH CASTAGNA

## Vendor Resources

- 55** Useful links from our advertisers.





**Highest in-use and in-plan dedupe vendor  
in Fortune 1000 companies.**

*TheInfoPro Wave 10 Storage Study, November 2007*

**"Clearly the revenue and customer count leader."**

*The 451 Group, Data Deduplication Report, July 2007*

**"Delivering scalable deduplication to large enterprises."**

*The Enterprise Strategy Group, January 2008*

datadomain

Find out why Fortune 1000 companies are choosing  
Data Domain for disk backup, archiving and networked DR.

[www.datadomain.com](http://www.datadomain.com)

**data domain**  
#1 IN DEDUPLICATION STORAGE

1-866-WE-DDUPE | [www.datadomain.com](http://www.datadomain.com)

Copyright © 2008 Data Domain, Inc. All rights reserved. Specifications subject to change without notice. Data Domain, the Data Domain logo and Global Compression are trademarks or registered trademarks of Data Domain, Inc. All other trademarks used or mentioned herein belong to their respective owners.



# Can we survive consolidation?

*Consolidation tools have helped some storage shops keep a lid on storage growth, but a consolidation solution could very well become the problem.*

**BACK IN** the '90s and throughout most of this decade, a lot of IT shops must have had revolving doors leading into their data centers. Expansion, upgrades and technical refreshes kept the doors spinning as a constant flow of bigger-better-faster gear made its way in and the older, suddenly obsolete equipment was shoved aside. It's a legacy of sprawl that most IT managers are grappling with now, and storage systems are squarely in the sights of consolidation efforts. Unfortunately, most consolidation has the opposite effect.

Figuring out just how we got to this point may be instructive for future planning, if not particularly useful information for dealing with the issues now at hand. It's true that a good part of the problem may seem unstoppable—we live in a data-driven world where protecting intellectual property is approached with a kind of religious zeal, causing every jot and tittle that makes it into digital form to be safeguarded as if the future of the company hinged on its very existence. Sure, some stuff is important, but a lot of it was barely considered when conceived and will never, ever be read again. Compliance has also added a dose of paranoia to the mix, causing companies to hoard data as if it were some kind of get-out-of-jail-free card.

But in a twist that may seem to defy logic, I think the very technologies that are ostensibly designed to help untangle this mess and allow consolidation are either exacerbating it or poised to do so down the road. The thing is, these technologies—and how they're applied in data centers—tend to treat the symptoms without actually providing a cure.

Take a look at server virtualization, the current poster child for data center consolidation. On paper, the idea of bunching up a lot of virtual machines (VMs) on a few massive servers so that you can eliminate a slew of other physical servers makes a lot of sense. But in the real world (according to many of the IT managers I've heard from), that's the way these consolidation efforts started out, only to be undermined when

In reducing the amount of hardware on the floor, new issues and problems arose, with bottlenecks where they never existed before.

systems admins and users realized just how easy it was to spawn a new VM, a few VMs or a few dozen VMs. In most cases, shops ended up with more virtual servers than the number of physical servers they had before consolidation began. In reducing the amount of hardware on the floor, new issues and problems arose, with bottlenecks where they never existed before.

Data deduplication—that paragon of storage consolidation—looms as another potential consolidation paradox. At some point, it seems likely that its advantages will be outweighed by the infrastructure required to keep it doing its job with some measure of efficiency. As your backup volume grows and performance or capacity is threatened, you'll need to add more dedupe devices, which will increase your infrastructure and administrative burdens. This situation compounds itself as the dedupe setup expands in other directions, such as adding remote sites to the mix or replicating data among dedupe boxes. Maybe just adding more cheap disk to your backup environment wasn't such a bad idea after all.

Data archivers may also turn out to be detrimental despite their obvious benefits. When old or unused data is archived off online systems, it has to go somewhere, and that somewhere typically needs to be nearline storage (not tape), which can grow pretty fast.

My point isn't that there's anything inherently wrong with these technologies, it's just that they don't address the root of the problem. No matter how well these tools work and how impressive their results are, it's a good bet they're moving around and storing a lot of garbage.

Effective consolidation has to start with data reduction that accurately assesses stored data and separates the one-off and not-very-useful stuff from true intellectual property. That means you need to know the data you're storing, and the only way that will happen is if you're able to accurately classify it based on the data itself, not on scanty peripheral information about the data.

A few years ago, a handful of startups emerged with products that could classify data with varying degrees of effectiveness. They've since been bought out, gone out of business or morphed into e-discovery tools. It's not that they weren't useful or necessary, it's just that they came on the scene at a time when the need wasn't as apparent as it is today.

Hopefully, some storage vendors out there have recognized this opportunity and are busily cobbling together data classification apps that will actually help manage the information stored on disks and tapes instead of just shuffling around all those bits and bytes. ☺

---

Rich Castagna ([rcastagna@storagemagazine.com](mailto:rcastagna@storagemagazine.com)) is editorial director of the Storage Media Group.

*\* Click here for a sneak peek at what's coming up in the January 2010 issue.*

# Are you prepared?



Learn why over 22,000 businesses trust  
i365 EVault Disaster Recovery Solutions

Backup and Recovery

Bare Metal Restore

Remote Disaster Recovery

Data Replication

Disaster Recovery Assessments

Remote and Branch Offices

DOWNLOAD OUR DR WHITEPAPER



[www.i365.com](http://www.i365.com) | 1.877.901.DATA | [concierge@i365.com](mailto:concierge@i365.com)



# STORAGE

## COMING IN JANUARY

### VMware vSphere and Storage

The latest incarnation of VMware's leading virtual server software accommodates storage better than past versions, but does it eliminate all the storage allocation headaches associated with virtual machines? We take an in-depth look at vSphere and how it interfaces with storage, as well as how other server virtualization products handle storage.

### Fast SAS Challenges FC Drives

The newest SAS drives—SAS-2—offer speeds that challenge Fibre Channel (FC) for disk system performance supremacy, and storage arrays that support them are showing up more often. In this in-depth look at SAS-2, we describe its place in the data center, if it's a viable alternative to FC and if pairing SAS-2 with SATA for internal tiering is feasible.

### Quality Awards IV: NAS Systems

In this fourth edition of the *Storage* magazine/SearchStorage.com Quality Awards for NAS arrays, users rank the leading systems for initial product quality, sales-force competence, product features, product reliability and technical support. See if NetApp can repeat its double win from the last NAS Quality Awards survey.

**And don't miss our monthly columns and commentary, or the results of our Snapshot reader survey.**

#### TechTarget Storage Media Group

### STORAGE

**Vice President of Editorial**  
Mark Schlack

**Editorial Director**  
Rich Castagna

**Senior Managing Editor**  
Kim Hefner

**Senior Editor**  
Ellen O'Brien

**Creative Director**  
Maureen Joyce

**Editorial Assistant**  
Rachel Kossman

**Contributing Editors**  
James Damoulakis  
Steve Duplessie  
Jacob Gsoedl



**Site Editor**  
Ellen O'Brien

**Senior News Director**  
Dave Raffo

**Senior News Writer**  
Beth Pariseau

**Features Writer**  
Carol Sliwa

**Senior Managing Editor**  
Kim Hefner

**Associate Site Editor**  
Chris Griffin



**Site Editor**  
Susan Troy



**Site Editor**  
Andrew Burton

**Associate Site Editor**  
Heather Darcy

**Assistant Site Editor**  
Megan Kellet

**Features Writer**  
Todd Erickson

**Executive Editor and Independent Backup Expert**  
W. Curtis Preston

### Storage Decisions

**TechTarget Conferences**

**Sr. Editorial Events Manager**  
Lindsay Mullen

**Editorial Events Associate**  
Nicole Tierney

**Storage magazine**  
117 Kendrick Street,  
Suite 800  
Needham, MA 02494  
[editor@storagemagazine.com](mailto:editor@storagemagazine.com)

Subscriptions:  
[www.SearchStorage.com](http://www.SearchStorage.com)



# Cash for Disk Clunkers

Receive \$1,000 per TB  
for old disks

no limit\* to the amount you trade in

**The "Cash for Disk Clunkers" program has now been**

**extended until December 31, 2009.** This revolutionary

program—the result of our commitment to providing the best

storage value—means you get \$1,000 per terabyte (NO

LIMITS\*) toward the purchase of an equal amount of

capacity on a new Emprise 5000, Emprise 7000, or

Emprise 7000 Edge system, or VM Storage Solution

when you trade in your old, inefficient drives. No strings

attached and no waiting for checks in the mail. Don't

miss out, program ends December 31, 2009.



Simplified  
Data Storage  
Solutions

[www.Xiotech.com](http://www.Xiotech.com)



**DATA**  
disk allowance trade agreement

# Hot Storage Technologies *for 2010*

***VMware backup, solid-state storage,  
thin provisioning, 8 Gbps Fibre Channel  
and data deduplication for primary storage:  
Are these on your 2010 storage to-do list?  
If not, they should be.***

*By Rich Castagna, Todd Erickson, Chris Griffin,  
Ellen O'Brien, Beth Pariseau, Carol Sliwa, Sue Troy*

**HOT**—in reference to enterprise data storage technologies—can be interpreted in many ways. Hot technologies could be the stuff of dreams that engineers are cooking up in research labs—but that often takes years, if ever, for real products to emerge. You could also define hot as those emerging technologies that may still be on the cusp of maturity but can have a significant impact on current storage environments.



We favor the latter definition because we think you're more likely to be fighting the storage wars than *Star Wars*, and would like to be armed with the latest technology available. The five technologies we think will be hot in 2010 may be familiar, but they're still cutting edge while being advanced enough to be practical.

Data backup is still one of the toughest chores in most storage shops, and it got even tougher when server virtualization upset the balance of traditional backup practices. We predict virtual machine backup technologies, already in high gear, will shift even higher with enhanced and new products emerging. Borrowing from the backup world, data deduplication for primary storage systems will become more pervasive to help storage admins cope with spiraling disk capacities. And additional disk system efficiencies will be realized as more vendors offer—and more shops implement—capacity management tools like thin provisioning.

With solid-state storage, touted by many as the logical evolution from magnetic media, we might be sticking our necks out a bit. But we think the proliferation of new products, dropping prices and intense interest will result in many more deployments in 2010. Our final hot technology is far more evolutionary than revolutionary: 8 Gbps Fibre Channel (FC). Although storage array vendors have some catching up to do with 8 Gig, we think this is the year they'll do it.

**We predict that virtual machine backup technologies, already in high gear, will shift even higher with enhanced and new products emerging.**

### BACKUP FOR VIRTUAL SERVERS

VMware Inc. may rule the data center, but for storage administrators virtual server backup was just an afterthought as many companies embarked on server virtualization implementations. Virtual machine (VM) backup is still in its adolescence but maturing fast, with significant developments that should offer some relief for beleaguered backup admins in 2010.

Traditional backup software vendors were slow to respond to the specialized needs of VM backup. Still, many IT organizations stuck with their traditional backup apps for their VMs, which may have distracted those vendors who saw the prospect of selling multiple agent licenses.

But other technologies have emerged to better address the unique needs of virtual server backup. Source-side deduplication and continuous data protection (CDP) products are well-suited to virtual machine backup because they reduce the volume of backup data and therefore lessen the likelihood of I/O contention.

John Merryman, services director at Framingham, Mass.-based

GlassHouse Technologies Inc., sees source-side deduplication in products like CommVault Systems Inc.'s Simpana, EMC Corp.'s Avamar and Symantec Corp.'s NetBackup PureDisk as delivering "some pretty tight integration with the ESX environment from a backup perspective."

W. Curtis Preston, TechTarget's Storage Media Group executive editor and independent backup expert, agrees that both source-side dedupe and CDP are good approaches to VMware backup. They both follow an incremental-forever backup model that produces far less data than traditional backup tools.

VM-specific backup products, such as PHD Virtual Technologies' esXpress, Veeam Software's Backup and Replication, and Vizioncore Inc.'s vRanger Pro were designed from the ground up to handle VMware backup. Their advantages include per-socket rather than per-server licensing fees (though experts and users caution that it doesn't always equate to lower costs); and they enable recovery of the virtual machine disk (VMDK) image for greatly simplified disaster recovery (DR) preparedness, as well as recovery of individual files within the VMDKs. Traditional backup tools operate from within the VM, so they're adept at file-level restore but require multiple steps to restore entire VMDKs. And the VM-specific tools are adding deduplication capabilities.

These products are gaining traction. "[With these VM-specific backup tools] it's faster to recover, it's easier to recover and it's easier to move things around because everything's encapsulated," said Edward Haletky, a virtualization consultant and author of two books about VMware.

Nathan Johnson, manager of IT services at NAI Utah, a commercial real estate company in Salt Lake City, avoided traditional backup tools early on. His company implemented Veeam's Backup and Replication software at the same time it rolled out server virtualization. Johnson said he didn't consider a traditional tool "because of how convoluted VCB [VMware Consolidated Backup] was. It's gotten better, but I want something simple. If I get run over by a bus, I want someone from my company to follow the procedures that I've written so that it can come back up easily." (In vSphere 4, VCB has been superseded by new storage integration capabilities and VMware Data Recovery, which addresses some of VCB's limitations.)

Welch's, the Concord, Mass., grape juice company, took a different

"[With these VM-specific backup tools] it's faster to recover, it's easier to recover and it's easier to move things around because everything's encapsulated."

—EDWARD HALETKY,  
virtualization consultant and author



# WE CALL IT THE BEAST



But its real beauty  
is the way it brings storage  
environments to tiers

## SiliconFS™

Only from BlueArc

*Seamless, automated data migration across storage tiers*



SEE MORE

▶ BEAUTIES OF THE BEAST

**BLUEARC®**

route. George Scangas, manager of IT architecture, said the company initially used CommVault's Simpana to back up its VMs. "With traditional backup, if we had to restore files and folders within the virtual machine, that worked great. If we had to restore the entire virtual machine, that was a 50/50 shot," he said. The company now uses vRanger Pro to back up its virtual machines in combination with Simpana on nonvirtual servers. vRanger Pro backs up the VMDKs to disk, and Simpana includes that disk when it backs up physical servers to tape, a practice followed by many IT organizations.

The traditional backup vendors aren't sitting still. Hewlett-Packard (HP) Co. and Symantec, for example, are working on updates that promise to deliver end-to-end backup for VMware environments and nonvirtual servers. "With Symantec [Veritas NetBackup] and HP Data Protector getting into the market as strongly as they are, [PHD Virtual, Veeam and Vizioncore] have to start looking over their shoulder for their backup product," consultant Haletky said.

In 2010, VM backup won't disappear as a chore at many IT organizations, but better tools are emerging. A year from now, simpler and more effective VM backup processes should be within reach for most storage administrators.

## NOT QUITE HOT ... YET

**CLOUD STORAGE.** Truth is, cloud storage was already struggling for clarity in the marketplace before the October mess with Microsoft and Sidekick carrier T-Mobile. Maybe it wasn't as bad as it first seemed, but some users still lost data—not exactly a boost for the cloud storage cause. Still, cloud storage is getting plenty of good publicity, which has resulted in a lot of buzz, some new fans and a long line of experts predicting eventual success. But all of that hasn't translated to a prime-time slot for cloud storage.

**DISASTER RECOVERY (DR) TESTING SOFTWARE.** All the DR gurus keep warning us that solid DR strategies—and regular DR testing—should be top priorities, but these products just can't seem to get the respect they deserve. These apps have gotten some traction, but they're still viewed as luxury items at a time when storage pros are spending only on necessities. The bottom line is that DR testing software isn't likely to take off until budgets loosen up.

**FCoE.** Talk about Fibre Channel over Ethernet (FCoE) and it's easy to get smart people to agree on two things. Yes, it has tangible, proven benefits. No, they don't want to overhaul their data center to accommodate it. Despite all the chatter about FCoE, most storage arrays don't support it yet and most vendors aren't rushing to add it. Experts say FCoE won't heat up until 2011—and then the fun will really start, as storage and networking teams duke it out over control of the converged infrastructure.

**TAPE ENCRYPTION.** If you have tape media going offsite, encryption makes sense, right? Especially with encryption built into LTO-4/5 drives. But key management and the challenges of encrypting at the client remain obstacles. As the security pros like to say, if you lose your keys, you lose your data. Despite hardware and software technology improvements, tape encryption still can't squeeze its way into the spotlight.

## SOLID-STATE STORAGE

Flash memory has been around for decades, but it's only been in the last 18 months or so that the persistent solid-state storage medium has made its way into enterprise data storage products.

EMC introduced solid-state drives (SSDs) into its Symmetrix array in January 2008; following that, most major IT vendors, including HP, IBM Corp., Hitachi Data Systems and NetApp Inc., made some form of solid-state storage available in server and storage products. Smaller players like Compellent Technologies Inc. and emerging companies like Atrato Inc. have also incorporated solid state with software that automatically migrates data between flash- and disk drive-based tiers of storage.

Even with that level of activity, there's still substantial work to be done to integrate solid-state storage into the rest of the IT environment, particularly with SSDs, which typically consist of flash memory fronted by a disk interface. Other implementations, such as Fusion-io's PCIe cards, offer an alternative to the disk interface and reside in servers rather than disk arrays.

MySpace is familiar with the pros and cons of solid-state storage. The social networking site recently replaced all of the Serial Attached SCSI (SAS) hard drives in one of the massive server farms that serves its Web portals with solid-state devices from Fusion-io.

Although solid-state storage is generally thought of in terms of high performance, Richard Buckingham, vice president of technical operations at MySpace, said the big benefits were savings in power, cooling and server hardware. "Instead of eight \$6,000 servers, we can go with one \$2,000 box and the cost of the Fusion-io devices doesn't even make up the difference," he said. "The ROI is immediate."

Buckingham remains open to SSD as well as the PCIe cards, but said the technology hasn't yet proven to be mature enough in his internal tests for production deployments. "It seems like it would be a simple step to pull out one hard drive and put in another that's faster, but under our real-life workload we found that SSDs just didn't perform as well behind a drive interface," he said.

Buckingham also said MySpace won't be replacing its Fibre Channel storage-area network (SAN) infrastructure with solid state anytime soon. "SSDs have a bright future, and flash will almost certainly take over in the future," he said. "But the SAN infrastructure is something we've invested a lot of time and money in and won't be tearing out and replacing

**"Instead of eight \$6,000 servers, we can go with one \$2,000 box and the cost of the Fusion-io devices doesn't even make up the difference. The ROI is immediate."**

—RICHARD BUCKINGHAM,  
vice president of technical  
operations, MySpace

for a very long time," he said.

Jeff Boles, senior analyst and director, validation services at Hopkinton, Mass.-based Taneja Group, agrees that SSDs have some way to go before they become a fully integrated part of enterprise storage systems, and that much of that integration work will take place in 2010.

Boles said the market has taken a "massive step forward" in the last six months with systems that intelligently integrate solid state, providing a more efficient way to share solid-state capacity among hosts and automatically move data among multiple pools of storage media.

Boles cited IBM's addition of SSDs to its SVC storage virtualization device, new offerings from startups Avere Inc. and StorSpeed that offer granular automated tiered storage, and hints of products to come related to developments such as Texas Memory Systems' acquisition of storage virtualization player Incipient.

"This trend will carry forward in 2010," Boles said, but so far most automated tiered storage and storage virtualization devices handle moving data in and out of solid-state devices at the LUN or volume level, when the most efficient method would be at the block level. "It may be 2011 before we see solid-state storage applied in more unique ways at increased densities," he said.

**"It may be 2011 before we see solid-state storage applied in more unique ways at increased densities."**

—JEFF BOLES,  
senior analyst, Taneja Group

## 8 GBPS FIBRE CHANNEL

IT organizations haven't made a mad dash to get to 8 Gbps Fibre Channel, but they'll certainly move steadily in that direction as they refresh or add new host bus adapters (HBAs), switches and storage arrays. The pace will accelerate when the cost of the faster technology nears parity with the price of current 4 Gbps gear.

For instance, when Atomic Energy of Canada Ltd. (AECL) needed to increase the port count of its core switch infrastructure, it found the cost of new 8 Gbps 64-port switches from Brocade Communications Systems Inc. to be close to what it had paid for 4 Gbps switches the prior year.

Simon Galton, manager of IT infrastructure services at the Mississauga, Ont.-based company, said the decision to go to the higher speed switches was opportunistic rather than highly strategic, as AECL has no plans at this time to go to 8 Gbps in its HBAs and disk arrays. Because 8 Gbps FC is compatible with earlier generations of the technology, a fork-lift upgrade isn't required. You just won't get the full benefit of the higher speed until you have 8 Gbps capability across the board.

Moving to 8 Gbps can improve I/O response time and prove especially



# Basic backup won't keep up with your rapid data growth



“From the perspective of total business value, CA ARCserve is leading the pack of backup software solutions.”

~ Jon Toigo, founder Toigo Partners International

Cut your data challenge down to size with ARCserve® Backup 12.5.

With built-in data deduplication you can slash your data storage up to 80%, while significantly reducing storage costs. ARCserve has more features, scales better and retrieves more consistent data\*. Save time, save money.

To get your custom savings report, go to [arcserve.com/dedupe](http://arcserve.com/dedupe)

useful with bandwidth-intensive applications, such as backup and data warehousing, and for virtualized server environments.

Ryan Perkowski, the SAN manager at a large financial institution, has justified 8 Gbps switch ports only for backups. His company purchased a pair of Brocade 5100 switches with three 8 Gbps ports each to link its disk/tape backups and Brocade DCX core.

But the connections between the host servers and the DCX are still 4 Gbps, as are the links between the DCX and storage arrays. Perkowski said he won't expand the 8 Gbps footprint until the firm's major storage vendor offers native ports. "There's no business need for it," he said. "We're having trouble saturating a 4 Gig link. I'm not going to buy stuff just to have it."

The pace of the shift from 4 Gbps to 8 Gbps has been slower than it was from 2 Gbps to 4 Gbps among the Fortune 1000, according to Robert Stevenson, managing director of storage technology at TheInfoPro Inc., a New York City-based research firm. He attributed the sluggish uptake, in part, to the economy's effect on IT spending.

Other contributing factors include the increasing interest in 10 Gigabit Ethernet (10 GbE) for file-based network-attached storage (NAS) or iSCSI SANs, as well as curiosity about Fibre Channel over Ethernet (FCoE). Any major FCoE adoption, however, will likely happen beyond 2010.

Meanwhile, 8 Gbps technology will likely see a marked uptick as the price gap with 4 Gbps continues to narrow. Seamus Crehan, vice president of network adapters and SAN market research at Dell'Oro Group, noted that 8 Gbps switch-side port shipments grew 50% quarter over quarter and became a majority of total Fibre Channel port shipments for the first time since the technology started shipping.

Also, 8 Gbps HBA port shipments doubled between the first and second quarters to nearly 60,000. Crehan cited the March launch of Intel Corp.'s Xeon 5500 (previously codenamed Nehalem-EP) server platform, which offers substantially higher server I/O throughput, as a major driver.

Robert Passmore, research vice president at Stamford, Conn.-based Gartner Inc., predicted that 2010 will be a big year for 8 Gbps FC when the majority of HBA, switch and storage array purchasers will go for the faster technology. "We're in the beginning of a very rapid transition," he said.

## THIN PROVISIONING

Thin provisioning has moved beyond its management and application issues of the past to become a must-have feature on many storage systems, and interest should only intensify in 2010.

Brian Garrett, technical director, ESG Lab at Milford, Mass.-based En-

**We're having trouble saturating a 4 Gig link. I'm not going to buy stuff just to have it."**

—RYAN PERKOWSKI,  
SAN manager at a large  
financial institution

terprise Strategy Group (ESG), said vendors have mostly worked out implementation and management issues related to defining separate logical pools and having to reserve capacity for thin-provisioned volumes. Garrett said thin provisioning works smoothly in most cases and is becoming a “feature check-off item” in the storage systems he evaluates.

The benefits of thin provisioning are evident, especially as tightening budgets bump up against ever-growing capacity demands. Releasing provisioned but unused disk capacity to a virtual storage pool and mak-

## REPORT CARD: GRADING OUR 2009 PREDICTIONS

**A**

### REPLICATION FOR DISASTER RECOVERY

We'll pat ourselves on the back for this one. We said server virtualization, more product alternatives and lower prices would put remote replication at the center of most disaster recovery plans. Replication's been around for some time, but it got a big boost in 2009 when it became a key feature in products like deduplication appliances and other backup targets.

**B+**

### 10 GIGABIT ETHERNET

Too soon to say 1 GigE is dead, but 10 Gbps Ethernet made steady inroads. For any company expanding or rebuilding its networks in 2009, 10 Gig was a no-brainer. The slowish pace of 8 Gbps Fibre Channel (FC) developments helped sway some users, as did the growing reliability of iSCSI storage systems. Prices are still on the high side but are decreasing—just not as quickly as we had anticipated.

**B+**

### STORAGE-AS-A-SERVICE (SAAS)

If we had predicted that the acronym “SaaS” would endure, we'd have to take an “F” on this one as “cloud” quickly became everyone's favorite buzzword. Cloud storage vendors are proliferating, the services can be far cheaper than in-house counterparts, and now there's even talk about “internal clouds”—so why shouldn't we get an “A”? Despite the progress, most enterprises still need more convincing before they launch their data into the cloud.

**B+**

### GLOBAL DATA DEDUPLICATION

We called global dedupe “a big deal” for users coping with siloed dedupe repositories. With dedupe deployments growing steadily, many users are learning just how big a deal it is. We'd give ourselves an “A” for calling that one but we have to shave a little off the grade because the dedupe vendors, with a few exceptions, have been slow to respond.

**B**

### SAS-2

Not quite the threat to FC that we had anticipated, 6 Gbps SAS still made some major inroads in 2009. It became the disk du jour for direct-attached storage in servers, the preferred interface of solid-state drives in arrays and a credible alternative to FC—just not at the rate we had expected.

**C**

### SELF-HEALING SYSTEMS

We said this prediction was “going out on a limb.” It turned out to be a fairly short limb, indeed. Although there were some new developments, self-healing systems haven't exactly caught on like wildfire. We were probably a year or two premature with this one.

ing it available to other applications can significantly increase utilization rates. John Michaels, chief technology officer at Maxim Group, a New York City brokerage firm, used his thin provisioned FalconStor Software Inc. IPStor and Network Storage System (NSS) units to increase his capacity utilization by 59.87%. Michaels said he “could see a difference right away.”

3PAR Inc. was a thin provisioning pioneer, rolling out the technology in 2003. Since then, most major storage vendors have jumped on the bandwagon: EMC’s Virtual Provisioning for Clariion, Symmetrix and Celerra systems; HP’s StorageWorks XP Thin Provisioning Software; IBM’s space-efficient virtual disks for its SAN Volume Controller (SVC); and NetApp’s FlexVol. And there are many others, including Compellent Technologies Inc.’s Dynamic Capacity software and DataCore Software Corp.’s SANmelody software, which converts standard servers, blades or VMs into virtualized storage servers.

User interest in thin provisioning is growing, too. In the 2009 *Storage* magazine/SearchStorage.com Storage Priorities survey, 14% of respondents said they had already implemented thin provisioning, 21% planned to deploy it by year’s end and 35% planned evaluations.

Mark Peters, an ESG senior analyst, noted that thin provisioning will continue to evolve as vendors add the capability to easily convert “fat” storage volumes to thin-provisioned volumes. Last October, 3PAR announced the release of Thin Conversion, a technology the company said will thin previously fat volumes. 3PAR also announced Thin Persistence to reclaim deleted thin capacity, and Thin Copy Reclamation to recapture unused virtual-copy snapshots and remote copy volumes.

Compellent and DataCore already offer fat-to-thin and reclamation technologies with their storage systems. As thin provisioning finds more and more users, other vendors will likely follow suit and upgrade their offerings to compete.

Initially, some storage vendors may have been reluctant to offer a technology like thin provisioning that could conceivably cut into their disk sales. But the successes of early entrants and eager acceptance of users persuaded them to follow suit.

## DATA DEDUPLICATION FOR PRIMARY STORAGE

The rate of growth of digitally stored information is putting many storage managers on the defensive as they struggle to address the operational risks and costs associated with unchecked data growth. In 2010, a variety of data-reduction technologies for primary storage, including deduplication, will provide some relief in hard-pressed storage shops.

“Business are finding that it’s taking a lot less time to reach that second terabyte or petabyte than it did to reach the first,” said Tory Skyers, a senior infrastructure engineer at a leading credit issuer. “Primary dedupe will allow any business to increase the density of data on their existing disks by at least twofold.”

A fixture in backup environments, dedupe can also be applied to



primary storage, thus helping to cut space, power and cooling costs. But primary dedupe won't yield the dramatic results common with backup dedupe.

Performance is another concern. "With backups, as long as the virtual tape loads and the backup works, everything is fine. With primary storage, performance isn't as cut and dry," said TechTarget's Preston. "If a restore [of a backup system] goes slowly, it's not the same as a system where you have thousands of people accessing files that they expect to open immediately."

The key to primary dedupe may come from finding the right balance between benefits and costs. "I'm looking to reduce my cost for storage and it's all about maximizing it with online compression and online dedupe," said Greg Schulz, founder and analyst at Stillwater, Minn.-based StorageIO Group. "Primary dedupe is not good for data that you're frequently working on, but it's good where you can trade time for money savings."

Both inline and post-processing dedupe can be applied to primary storage. For applications that can afford the performance hit, inline dedupe is perfect. If the data in those systems can be held in cache and then deduped before it hits a disk, fewer disks are required on the back end of the system, which ultimately cuts costs. "While inline is currently the slowest performer, I have a feeling with the advent of [solid-state storage] and larger inline caches, it's eventually going to catch up with post-process," Skyers said.

Some major storage vendors, including EMC and NetApp, are now offering primary data-reduction capabilities. NetApp's dedupe is built into its Ontap operating system. It works by storing the cyclic redundancy code (CRC) of every block written to storage, comparing the CRCs, and then eliminating and replacing any matching blocks with a pointer.

"NetApp is doing real dedupe and they're doing it essentially without a change in performance," Preston said. "When the actual dedupe process is running there's a change in the performance. But once the data has been deduped and you're just running your database or VMware, there's essentially no change in performance."

Ocarina Networks and Storwize Inc. also had early primary data-reduction entries. Ocarina's ECOsystem is an out-of-band appliance with software that's tuned to the data types associated with specific applications. Storwize's STN appliances work with NAS devices to compress and uncompress the data inline. Both of these startups have garnered a lot of attention that has led to partnerships with a variety of storage vendors. ☉

"I'm looking to reduce my cost for storage and it's all about maximizing it with online compression and online dedupe."

—GREG SCHULZ, founder and analyst, StorageIO Group



Grey Healthcare Group is one of the world's top five healthcare communications companies, with a global network that includes 43 companies in 22 countries. It has an extensive array of integrated multichannel digital and traditional marketing services in support of brand acceleration and sales.

## TOTALLY Open solutions form the roots of green IT

**"FalconStor VTL with de-dupe offers us the ability to optimize storage capacity and minimize not only our physical storage footprint but also our overall carbon footprint."**

In addition, the solution will improve data security to meet our future compliance requirements via its encryption capabilities."

- Chris Watkis, IT Director, Grey Healthcare Group

# Virtual tape libraries



## in depth

*By W. Curtis Preston*

For disk-based backup, VTLs have been a relatively easy way to replace traditional tape libraries. With added features such as deduplication, they can be an attractive alternative to other disk target systems.

**V**IRTUAL TAPE LIBRARIES (VTLs) are dead, right? Weren't they supposed to be temporary solutions that would be long forgotten once everyone started backing up to "real" disk? While that might be what the VTL naysayers had in mind, we're more than a few years into the VTL "fad" and many of the products are doing just fine.

What happened was that an industry segment morphed to encompass both VTLs and intelligent disk targets (IDTs), a segment that was ultimately validated when EMC Corp. acquired Data Domain for \$2.4 billion. We'll review some of the factors that led to the development of VTLs, the current state of VTL technologies and products (including the newer features they now offer), and then we'll end with a look into the future of VTLs and IDTs.



## WHY VTLs CAME ABOUT

The VTL/IDT market has become so overshadowed by the data deduplication craze that some people may have forgotten why the industry developed VTLs in the first place.

**TAPE WAS (AND IS) TOO FAST.** The core problem vendors were trying to solve with virtual tape libraries is the mismatch between the speed of tape and the speed of the disk drives, file systems and databases they're backing up. In approximately 15 years, the sustained throughput of open system disk drives has gone from approximately 4 MBps to 70 MBps—an increase of 1,700%. In roughly the same amount of time, the sustained throughput of open system tape drives has grown from 256 KBps (Exabyte EXB-8200 drive) to 180 MBps (LTO-4)—an increase of 70,000%.

**VTLs MADE THE UNFAMILIAR FAMILIAR.** For many backup administrators and their backup software, backing up to disk was a foreign concept. Knowing that progress in backup systems is an incremental process, VTL vendors felt they could take the unfamiliar (disk) and make it seem like an old friend (tape).

**SCALABLE.** Demand for VTLs has been driven by the needs of large enterprise customers. With a tape library, they had hundreds or thousands of tapes and dozens or hundreds of tape drives, and they could just throw all their backups at this big tape library and it would sort it out. To use disk, however, they would need to manage and load balance their backups across dozens to hundreds of discreet disk systems.

The VTL solved this problem by presenting disk as large tape libraries, something they were already familiar with. In various ways, VTL vendors made dozens of individual disk arrays look like one or more tape libraries that could scale to almost limitless levels.

**SHAREABLE.** Because backup software already knew how to share tape libraries, they could easily share VTLs. Instead of

## WHY DEDUPE AND FC DISK DON'T MIX

Fibre Channel (FC) is essential to the enterprise and data deduplication is important as well. But the only way to get both in a backup appliance is to buy a virtual tape library (VTL). Why don't they just make a deduplicated logical unit number (LUN) that's accessible via FC? The short answer is that it's a lot harder than it sounds. Giving you a LUN allows you to pick your own file system, which the appliance would then need to support. Windows, Linux, Solaris, HP-UX, AIX, MacOS, etc., all have their own completely incompatible file systems. The IDT vendor would have to test deciphering all of the various backup formats on all of the various file systems as well. Think of *that* test matrix.

But GreenBytes Inc. has gotten close. It's about to offer an iSCSI deduplicated LUN with its GB-X Series of storage appliances.



using extra-cost sharing software (such as Symantec Corp.'s Veritas NetBackup Shared Storage Option or EMC Corp.'s NetWorker Dynamic Drive Sharing Option), you could create as many "tape drives" as you needed to give each backup server its own tape drives, while dynamically sharing the VTL. And if you have multiple backup applications that refuse to share, a VTL can be carved into separate virtual libraries.

## PRODUCT SAMPLER: VTLs AND IDTs

Use this list of virtual tape library (VTL) and intelligent disk target (IDT) vendors to get started on your product evaluations. Whatever else is included in your evaluation process, be sure to test your leading candidates. Many of these products are relatively new, and without doing a proof of concept you won't know if they'll actually be able to do what they claim in your environment.

PRODUCT	SOFTWARE SUPPLIER	VTL/NAS	DEDUPE	GLOBAL DEDUPE	INTEGRATED TAPE	TAPE VIRTU-ALIZATION	OST SUPPORT
Copan Revolution	FalconStor Software Inc.	VTL	✓	✓	✓		✓
EMC/Data Domain DDX Series	Data Domain Inc.	BOTH	✓				✓
EMC Disk Library 1500/3000	Quantum Corp.	VTL/NAS	✓		✓		✓
EMC Disk Library 4000	FalconStor	VTL	✓		✓		✓
ExaGrid EX Series	ExaGrid Systems Inc.	NAS	✓	✓			
FalconStor File-interface Deduplication System (FDS)	FalconStor	NAS	✓	✓			
FalconStor Virtual Tape Library (VTL)	FalconStor	VTL	✓	✓	✓	✓	✓
Fujitsu Eternus CS	Fujitsu	VTL			✓	✓	
GreenBytes GB-X Series	GreenBytes Inc.	NAS/iSCSI	✓				
Gresham Clareti Storage Director	Gresham Storage Solutions Inc.	VTL			✓	✓	
Hewlett-Packard (HP) D2D Backup System Series	Hewlett-Packard (HP) Co.	NAS	✓				
HP Virtual Library System	Sepaton Inc.	VTL	✓	✓			
IBM ProtecTIER	IBM	VTL	✓	✓			
NEC Hydrastor	NEC Corp. of America	NAS	✓	✓			
NetApp NearStore VTL	NetApp Inc.	VTL	✓		✓		
Overland REO Series	Overland Storage Inc.	VTL					
Quantum DXi-Series	Quantum	BOTH	✓		✓		✓
Sepaton S2100 Series	Sepaton	VTL	✓	✓			
Sun StorageTek Virtual Tape Library	FalconStor	VTL	✓	✓	✓		✓



# Storage Virtualization Gets Real.

## WOULD YOU LIKE TO GET MORE OUT OF YOUR EXISTING STORAGE ASSETS – AND FEEL SAFE DOING IT?

Hitachi Storage Virtualization transforms your multi-vendor storage assets into a common pool of shared resources. It lets you reclaim, utilize and optimize “lost” space.

Reinvigorate your infrastructure and get more out of what you already have with Hitachi Storage Virtualization.

### Download the new Storage Optimization Kit now and you'll learn how to:

- Maximize existing storage utilization and move data seamlessly across storage tiers
- Increase storage performance and application availability
- Manage all virtualized storage—regardless of vendor—from a single interface

[\*\*Download Now\*\*](#)



### *The Storage Optimization Kit includes:*

- The “6 Essential Strategies For Economizing Your Storage” eGuide that helps you make the most of your shrinking IT budget
- An ESG white paper that examines proven tools for implementing Intelligent Tiered Storage
- A case study on a Fortune 500 company's success with virtualization and tiered storage

Hitachi Storage Virtualization  
**Reclaim, Utilize, Optimize.**

**FRAGMENTATION ISSUES WITH FILE SYSTEM DEVICES.** VTLs also avoid the fragmentation issues associated with backing up to file systems. They solved this problem using proprietary file systems that wrote data contiguously.

## STATE OF THE VTL INDUSTRY

VTLs came about to address specific backup issues. Let's look at how they've progressed in those areas they were supposed to fix.

**SCALABILITY.** Scalability isn't just an issue for big enterprises; it's also necessary to meet the needs of small- and medium-sized businesses (SMBs). When the VTL market was in its

early days, there were very few products that could scale well for either of these segments. But times have changed, and there are now several products that scale both up and down.

With some notable exceptions—Copan Systems Inc., IBM Corp., NEC Corp. of America and Sepaton Inc.—all VTL/IDT vendors offer products for SMBs. Companies with less than 20 TB of data to back up each night can choose from a number of products—some less than

\$5,000—that offer a lot of the same functionality available in high-end products. Offering products to the SMB market before they're deemed bulletproof typically spells failure, so the arrival of these SMB virtual tape libraries and intelligent disk targets is a sign that vendors have done a good job of working out any kinks in their products.

Midsized enterprises with 20 TB to 40 TB to back up each night can choose from almost every vendor. To back up that kind of data you need a system capable of handling 500 MBps to 1,000 MBps. Almost every vendor listed in the “[Product sampler: VTLs and IDTs](#)” (p. 25) sidebar has a product with that capability.

The high end of the enterprise (companies with 40 TB or more to back up every night) has fewer products to choose from. Users with that much data to back up connect large servers to a Fibre Channel storage-area network (FC SAN) and back them up using local-area network (LAN)-free backups. The last thing those users want to do is send those backups over IP; therefore, a product targeting this market segment must have FC as a transport.

Another reason why there are only a few products appropriate for this market is the lack of global data deduplication in some products. A user with 100 TB to back up each night needs 2,300 MBps throughput. They won't want to (nor should they have to) create and maintain three separate 33 TB backup collections that'll back up to three devices that can only

There are sub-\$5,000 VTLs and IDTs that offer a lot of the same functionality available in high-end products.

handle 40 TB per night each. They need a single system that can handle this load over FC without splitting it into multiple backup collections. There are only a few companies with products capable of doing that: FalconStor Software Inc. and Sepaton (and their respective OEM partners Copan Systems and Sun Microsystems Inc., and Hewlett-Packard [HP] Co.). The aggregate throughput of NEC's Hydrastor is actually much higher than 2,300 MBps, but it doesn't yet offer Fibre Channel as a transport. If you need this kind of throughput over FC, but don't need deduplication, EMC, Fujitsu and Gresham Storage Solutions Inc. have products that can help.

Noticeably absent from the list is EMC/Data Domain. Their fastest FC-based VTL runs at 900 MBps. Data Domain's DDX "array" boasts a number much higher than that, but it's actually 16 separate DDR units in the same rack that aren't integrated as far as deduplication goes. Data Domain doesn't support global deduplication, although the company has said it's on its roadmap. However, there's been no indication as to when this feature may become available.

**EASE OF USE.** VTL and IDT products range from the "ridiculously easy" to use to "so hard you can't believe it passed any kind of functionality testing." But most are relatively easy to use. Still, ease of use varies considerably, so you should definitely test with any products you're considering.

**INTEGRATION WITH BACKUP APPLIANCES.** All VTLs and IDTs can be backup targets for just about any backup software product on the planet, and most can also replicate their data to another VTL/IDT. But few products today integrate with the backup software so that it knows about replicated copies and can use them for restores and copies to tape.

Symantec's NetBackup OpenStorage (OST) API offers one solution to this problem. With this API, the disk target isn't addressed as a virtual tape or a file system; the backup job is named and passed to the target, and the target stores it however it wants to. Once the backup is stored on the target, NetBackup can tell the IDT to replicate the data; when the replication is done, the IDT tells NetBackup. So, NetBackup is aware of the replicated data and the replication process, and can use it to create a tape copy. The process yields an onsite copy, an offsite disk copy and an offsite tape copy without anyone ever touching a tape. Today, only Data Domain, FalconStor and Quantum Corp. support this API—and only FalconStor supports it via Fibre Channel; Data Domain and Quantum use IP as their transport.

CommVault Systems Inc. has a similar feature that works with net-

The aggregate throughput of NEC's Hydrastor is actually much higher than 2,300 MBps, but it doesn't yet offer FC as a transport.



work-attached storage (NAS)-based IDTs (but not VTLs). A media agent watches a directory that you're replicating to and looks for changes. It communicates with the CommServe (the main backup server) and tells it about the other copy, resulting in both copies being available for restores. If this other media agent were located offsite, you could then use that replicated copy to create an offsite tape copy of your replicated backup.

HP also offers this capability for its Data Protector software and the HP Virtual Library System (VLS). The product is similar to CommVault's, except it uses a completely separate Data Protector backup server (with its own catalog) to watch for newly replicated virtual tapes. Once those tapes are detected, it asks the other Data Protector server for its catalog information. Both servers can then use those virtual tapes, which would allow creation of a tape copy of the replicated backup.

#### **SOFTWARE-BASED VTLs VS. VTL APPLIANCES.**

Because all appliances are just servers running software, the difference between a software VTL and an appliance is more a matter of packaging than a technical issue. It comes down to preferences: prepackaged or build your own. Most VTLs and IDTs are prepackaged, but there are some exceptions, such as the software-only versions of FalconStor's and Gresham's products.

You may also opt to buy a virtual tape library/intelligent disk target with its disk already attached or choose to add your own. In the latter case, options include software-only products or gateway products such as those offered by Data Domain and IBM.

**INTEROPERABILITY WITH TAPE LIBRARIES.** A VTL may provide a direct connection to and integration with a physical tape library. The appeal of this feature has diminished with the increased interest in data deduplication. VTL-tape library integration made it easier to stage data from disk to tape to save space on expensive disk. But with deduplication, there's less need to do this. Products that integrate with physical tape are available from FalconStor, Fujitsu, Gresham, HP and Quantum.

**IDTs VS. VTLs.** Whether you should back up to a file system device or a virtual tape library truly boils down to personal preference. If you want FC as a transport, your choice is easy; if you want a scalable, deduplicated system, only VTLs offer that today.

File system-based devices have two advantages over virtual tape

Because all appliances are just servers running software, the difference between a software VTL and an appliance is more a matter of packaging than a technical issue.



# DeDupe<sup>SG</sup><sup>TM</sup>



## Deduplication has just been redefined. *The only dedupe with "SG" — Speed with Green*

First generation deduplication systems focused on **saving storage**. Nexsan and FalconStor are redefining what you expect from a deduplication solution by delivering an ultra-efficient system that **saves** more than just capacity. DeDupe SG is faster to **save time** on backup and restore, greener to **save energy** between backups, simpler to **save management** resources and more dense to **save space**. Introducing DeDupe SG<sup>TM</sup>, the highly efficient deduplication storage system.

**SAVE STORAGE**



**SAVE TIME**



**SAVE ENERGY**



**SAVE RESOURCES**



**SAVE SPACE**



See how DeDupe SG is faster and greener  
at [www.nexsan.com/dedupesg-now.php](http://www.nexsan.com/dedupesg-now.php)

libraries: what happens when your backup software expires a backup, and simultaneous read and write support.

When an IDT deletes a file, it automatically reclaims the space. But a VTL has no idea that the tape it's holding has expired. A workaround is to manually re-label tapes when they expire. When the VTL sees a new label being written to the tape, it knows it can throw away the rest of the data on that tape.

File system devices support simultaneous read and write, but VTLs don't. If a backup is writing to one virtual tape, another process can't read that tape to do a restore or copy. But this only happens if you're backing up and restoring/copying at the same time—probably a rare occurrence that can be made even less likely by using smaller virtual tapes.

### NEW FEATURES OF VTLs AND IDTs

Virtual tape libraries and intelligent disk targets continue to evolve; here are some of the areas where these products are developing.

**DATA DEDUPLICATION.** The biggest game-changing feature has been deduplication. It changes a VTL from a disk staging device with only a few days of backups (due to the cost of disk) to a device that can affordably hold all onsite backups. And dedupe built the IDT market; without dedupe, an intelligent disk target is truly just a NAS filer.

Deduplication can reduce backup size by 10:1 or 20:1 without significantly affecting the performance of restores and copies from disk to tape. But not all data dedupes well. Applications such as imaging, audio, video or seismic processing systems generate new data

every time they run, so there's little detectable duplication. Dedupe systems also use compression, but not all data compresses well either.

There are other significant differences among target dedupe systems (VTLs/IDTs). The IBM ProtecTIER product, for example, has a single-stream restore speed limitation of approximately 90 MBps. Although Quantum has made significant progress with restore speed, the restore speeds from their "block pool" (i.e., deduped data) are still nowhere near those possible when restoring from the last few backups stored in native format. Sepaton's dedupe system is backup product-specific, and the firm has yet to release support for CA ARCserve Backup, CommVault Simpana, EMC NetWorker and Symantec Backup Exec, among others. And the lack of global deduplication from some of the

Deduplication can reduce backup size by 10:1 or 20:1 without significantly affecting the performance of restores and copies from disk to tape.

major vendors (e.g., Data Domain, NetApp and Quantum) means that users must continue to slice their backups into chunks that are manageable by a single appliance.

**DEDUPLICATED REPLICATION.** Deduplication also makes replication much more affordable and feasible. Without dedupe, you might need 10 times to 100 times more bandwidth to replicate a full backup. With dedupe, a typical full backup only stores and replicates 1% to 10% of its native size.

**TAPE CONSOLIDATION AND VIRTUALIZATION.** Some vendors, notably Fujitsu and Gresham, tend to use the term tape virtualization rather than VTL. They see tape virtualization as a way to enhance your continued use of tape while removing many of tape's limitations, especially if you want to use tape as a long-term storage device. If you store data on tape for multiple years, you're supposed to occasionally "retension" your media and move backups around to keep all the bits fresh. Updating your tape technology is another issue: What do you do with the old tapes and drives?

A tape virtualization system solves these issues by employing what's often referred to as a hierarchical storage management (HSM) system for tape. Newer backups are stored on disk; older backups are stored on tape. When you buy new tape drives and bigger tapes, you simply tell the tape virtualization system that you want to retire the older tapes and they're migrated to the newer, bigger tapes by stacking the smaller tapes onto the larger tapes and keeping track of which "tapes" are stored on which tapes. If the backup application requests a bar code that's been stacked onto a bigger tape, the system loads the appropriate tape, positions to the point in the physical tape where the requested "tape" resides, and the application doesn't know the difference.

## THE FUTURE OF VTL TECHNOLOGY

Virtual tape library technology continues to develop and expand, but just being a VTL may not be enough anymore. With so many users replicating backups offsite, the industry must find a solution to the challenges posed by using replicated backups. Unfortunately, in the near term we're likely to see more product-specific approaches such as Symantec's NetBackup OpenStorage and HP's Data Protector/Virtual Library System.

There have also been predictions that as data deduplication becomes more pervasive in backup software, the need for intelligent disk targets will be reduced. But that's only likely to happen if source deduplication software products can address their restore speed limitations, which were designed to back up remote sites. As such, their restore speeds are slow (10 MBps to 20 MBps). Unless that changes, there will continue to be a market for high-speed disk targets. ☉

---

W. Curtis Preston is the executive editor for SearchStorage.com and an independent backup expert.





**NORTHERN STORAGE SUITE**

# CONTROL YOUR USERS, CONTROL YOUR COST

Your users create the majority of unnecessary data in your environment. For them, storage seems infinite. For you, it's a constant battle against time, space and escalating costs.

**Northern Storage Suite** gives you complete insight into your storage environment through comprehensive reporting. It allows you to take control of your users; encourage deletion, prevent storing of unwanted file types, establish storage rules and makes users responsible for their storage through on-demand, interactive user reports.

With **Northern Storage Suite** you will reclaim otherwise wasted capacity, flatten usage growth curves and postpone new hardware investments.

Learn more about how Northern Storage Suite can help you control your users and your cost.

Watch a webcast: [www.northern.net/StorageMag](http://www.northern.net/StorageMag)



NORTHERN – STORAGE MANAGEMENT SINCE 1995. TO US IT'S SECOND NATURE.

[WWW.NORTHERN.NET](http://WWW.NORTHERN.NET) / [SALES@NORTHERN.NET](mailto:SALES@NORTHERN.NET) / 1.800.881.4950



**NORTHERN**



# SALARIES RISE AS STORAGE GROWS

*Storage professionals in many industries managed to see pay increases last year even as company closings and layoffs sent some looking for new jobs. As data grows, so does the need for dedicated storage professionals, according to our annual Salary Survey.*

*By Ellen O'Brien and Rachel Kossman*

**THE AVERAGE ANNUAL SALARY** for storage professionals jumped approximately 3.5% in 2009 and respondents expect it to grow another 3.8% in 2010, according to our annual *Storage* magazine Salary Survey. In a year of layoffs and losses, recessions and rebounds, our annual survey found that storage professionals across many vertical industries and geographic regions managed to secure pay increases.

The [average annual salary](#) reported this year by our 363 respondents was \$85,869 vs. the \$82,915 they said they earned in 2008. For 2010, these same respondents predict they'll earn an average annual salary of \$89,065.

Many of those surveyed reported accepting new jobs in 2009 that didn't include a raise, and placing greater value on job security and benefits. Indeed, the average 2009 salary reported by this year's respondents was slightly lower

than the \$86,573 average salary reported by those who participated in our 2008 Salary Survey.

Storage professionals who responded to our survey reported longer days, stricter budgets and fewer staff as a result of the economic chill that has tied up businesses from Florida to Canada.

"The headcount freeze is on," said Jean Veronneau, director of storage at Manulife Financial Corp., a financial services company based in Toronto. "You don't demand things the way you used to."

At age 60, Veronneau is a storage enthusiast with 15 years of experience dedicated solely to storage technology. "I'm not going into anything else," he said. "One [reason] is that I'm very good at it, and two is that I enjoy it."

Veronneau is among the many former private consultants and contractors who jumped at a full-time job when it became available during shaky economic times a couple of years ago.

"I took a hit," he said. "I took a cut compared to some of my contracting [income] but there was a challenge. I was going to be an architect for a new EMC solution when I started here. We started out with a contract, which was great, and then they offered me a full-time job."

Judging by our survey results, Veronneau made the right move. The financial services industry he works in reported average annual salaries of \$93,798, which is a figure topped only by the \$94,513 reported by respondents in the IT services industry. Ranked third was media and publishing, where annual salaries averaged \$90,577. These three vertical industries were also atop the salary chart in the 2008 survey.

Veronneau noted that the cost of living in Toronto compares to some of the priciest U.S. cities. Still, our salary survey shows Canada ranking third to last in annual salaries vs. U.S. regions, with an average of \$78,042. Only the Midwest and Southeast regions of the U.S., with average salaries of \$77,336 and \$76,102, respectively, fared worse. The Mountain region had the highest annual salaries with an average of \$120,300. Last year, the Pacific region recorded the highest average salaries at \$96,141; this year, the region ranks fourth with \$94,375.

**Q: Which region pays the most to storage professionals?**

**A:** That would be the Mountain region, according to our survey. At the other end of the regional salary spectrum is the Southeast region, which ranked last.

**Q: In which region can storage pros expect the biggest year-end bonuses?**

**A:** Optimists in the Pacific region predicted they would receive a healthy average bonus of \$12,719. The New England region, despite being known for Yankee frugality, was a close second at \$12,343. Northwest region respondents had the most modest expectations, anticipating an average \$2,224 bonus.

# THE SUN UNIFIED STORAGE SYSTEMS

It's time to change your storage economics

	Sun Storage 7410	NetApp FAS 3170	EMC NS80
Data Services*	Included	Fee Based	Fee Based
Flash Optimized	Yes	Not Available	Not Available
DTrace Analytics	Yes	Not Available	Not Available
Power Consumption	2524w	6426w	6975w
RAID Rebuild	Minutes	Hours	Hours
Space	12RU	45RU	43RU
Grand Total	\$221,000	\$894,000	\$554,000

\*Sun's Data Services Include Snap/Clone, Restore, Mirroring, Replication, Compression, Thin Provisioning, CIFS, NFS, HTTP/FTP and WebDAV.

Radically change the way you  
manage storage with breakthrough  
cost, speed, scale, and ease.

Learn more about the Sun Unified Storage Systems at  
[sun.com/unifiedstorage](http://sun.com/unifiedstorage)



*The Network is the Computer™*



## LEARNING CURVES AND CERTIFICATION CLOUD

From Veronneau's viewpoint, storage is a dynamic, cutting-edge technology, and it can be a challenge to keep current. "What I've seen is that the storage skill sets really vary," he said. He's a fan of vendor-specific certifications.

"It's expensive to stay on the bleeding edge, but you have to stay up to date," Veronneau said. However, he admitted that he let his SAN certification with EMC Corp. lapse recently. "It had to be renewed, but I've been too busy to go out and get it," he said.

In an EMC shop, Veronneau said, every six or nine months customers get an announcement letter with a "90,000-foot view of what's happening. But you have to drill into that." He believes his EMC certifications over the years have helped get his resume noticed, and he looks for some certifications when making staffing recommendations.

Much has been made of the value—or lack of value—of certifications in the storage industry. A survey by Vero Beach, Fla.-based Foote Partners LLC, released in the second quarter of this year, showed the average salary increase for holders of a single certification was at 7.5% of base pay this year vs. a peak of 8.6% in 2001.

According to our 2009 Salary Survey, there was no significant benefit to having one or two certifications, but having three or four certifications did translate to higher salaries for our respondents. Those without any certifications averaged an annual salary of \$85,277, slightly higher than respondents with one or two certifications. But storage pros with three certifications reported average salaries of \$96,053. At four certifications, the average salaries peaked at \$96,400, but then dipped to \$82,858 for those with five or more.

In the Foote Partners survey, several storage-specific technology certifications did buck the trend and manage to rise in value in 2009, including the EMC Technology Architect (EMCTA) and the EMC Proven Professional Certification program.

Mike Horvath, 38, is one of three data storage administrators at Fifth Third Bank in Cincinnati, and a relative

- ➔ **Q: Which industry boasts the highest salaries for storage professionals?**  
**A:** IT services was most generous, with salaries averaging \$94,513. Financial services placed a close second with \$93,798.
- ➔ **Q: If you manage more capacity, do you make more money?**  
**A:** Once you move past 1 TB, yes. Our survey showed annual salaries of \$73,646 for storage pros at shops managing between 1 TB and 9 TB. But the average jumps to \$82,666 for those managing between 10 TB and 99 TB. Above 500 TB, the average spiked to \$104,441.
- ➔ **Q: Does managing a bigger storage budget translate into a bigger salary?**  
**A:** Yes. Those managing budgets of \$500,000 or less showed an average salary of \$75,327, while respondents overseeing budgets of more than \$10 million earned an average annual salary of \$126,000.

storage newbie, with 15 years of IT experience but just two in storage. In his opinion, the lack of documentation, the need for new application knowledge depending on the environment and the learn-as-you-go aspect of storage management mean certifications quickly lose value.

“What I found is that there are two kinds of people—one that was book smart and one that knew the job,” Horvath said. A single father with a three-year-old daughter, Horvath has carved out a long-term plan for his career that relies on excelling at storage.

“The next 15 years will be spent digging deep into storage,” he said. “Finding out everything I can about allocation, storage management. I intend to go into consulting because I’ve done it before.”

A resume that’s chock full of dedicated storage experience does mean fatter paychecks for storage professionals, according to our survey. College graduates who spent more than five years specializing in storage did reap the rewards. The average salary for those graduates with three to five years dedicated storage experience was \$88,717. That jumped to \$95,440 for college graduates with six to 10 years of dedicated storage experience. At 10 years, salaries rose to \$105,921.



**Q: Do salaries grow as a company’s storage expands?**

**A:** Yes. The biggest jump was seen in shops where storage grew by 41% or more last year. Shops that didn’t experience any storage growth in 2009 reported an average salary of \$82,250; at companies where storage grew by 41% to 50%, the average salary was \$91,588.



**Q: Do professional certifications pay?**

**A:** Not one, not two, but three certifications in storage can apparently contribute to higher salaries. Respondents without any certifications had an average salary of \$85,277, slightly higher than respondents with one or two certifications. But storage pros with three certifications reported an average of \$96,053. The sweet spot for certifications was four, with an average salary of \$96,400, dipping to \$82,858 for five or more certifications.

## THE NEW VALUE OF BENEFITS

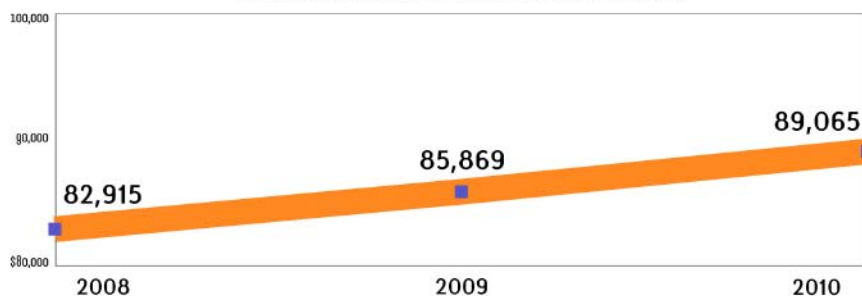
Storage professional Troy Downing began job hunting in a tough economy, just as his employer of 13 years was being acquired. Within a month of searching, Downing landed a position at Rain and Hail L.L.C. in Des Moines, Iowa, where he has been a member of a two-person storage team for the last two years. Today, Downing earns a smaller base salary than he did at his longtime employer, but his top-notch benefits package makes up for the difference.

“Maybe my base pay doesn’t look as good, but our [healthcare payments] are zero. Zero. There’s no premium,” Downing explained. “Knowing what I was paying in the past ... that’s worth \$12,000.”

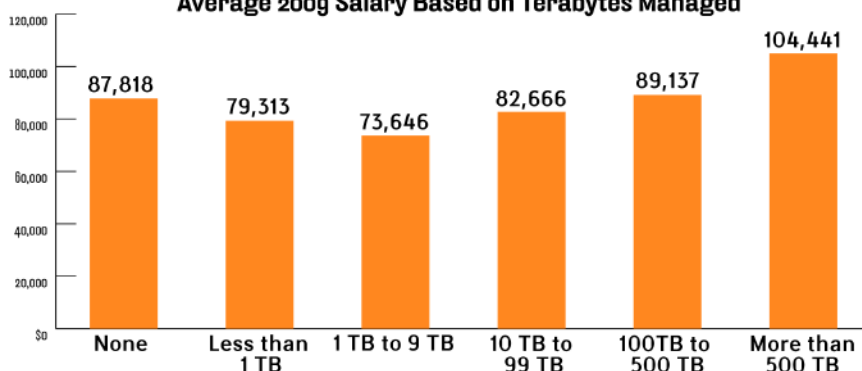
Downing’s new benefits plan means that he pays only a small fee for

# 2009 STORAGE SALARY SURVEY BY THE NUMBERS

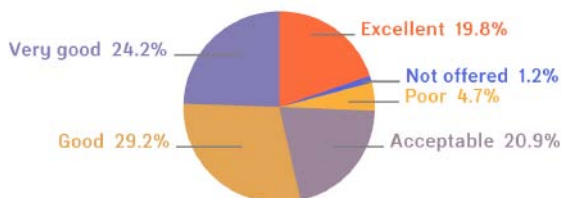
**Average Salary Over 3 Years (2008-2010)**



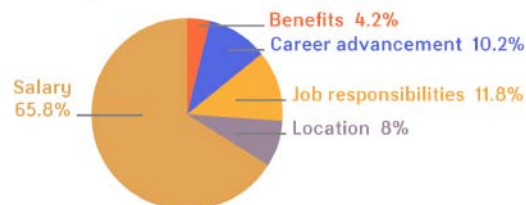
**Average 2009 Salary Based on Terabytes Managed**



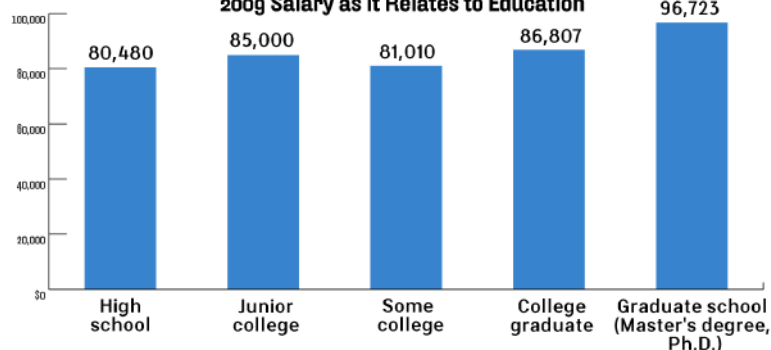
**Rating the Quality of Health Benefits**



**Most Important Factor in Job Choice**



**2009 Salary as it Relates to Education**



office co-pays, and there's no healthcare withdrawal from his paycheck. He also has a flex spending account of \$1,200 to offset costs that aren't otherwise covered.

With the rising cost of healthcare at the center of national debates, many storage professionals said they're [placing greater emphasis on their benefits](#) than before.

However, satisfaction with overall benefits packages dipped in 2009 vs. 2008. In 2009, 64% of respondents rated their company benefits as good, very good or excellent vs. 73% last year. And in 2009, more of our respondents reported working at companies that reduced

benefits. In 2008, 18% said their benefits packages were reduced vs. 30% in 2009, reflecting a widespread trend at companies looking to keep their doors open and their margins appealing to Wall Street.

"A lot of my friends in town took pay cuts—a lot of mandatory pay cuts," Downing said. "Watching that, I feel fortunate I ended up where I ended up."

## EXPANDING STORAGE, BIGGER PAYCHECK

For Downing, another draw in changing jobs was the opportunity to work on server and storage virtualization projects. "I'd like my two primary focuses to be storage and virtualization," he said. At Rain and Hail, Downing is also part of a dedicated, albeit small, storage team, which is a selling point for many of our respondents who said they believe dedicated storage teams are essential to a company understanding the overall value of storage.

In general, the likelihood of having a dedicated storage team increases with the size of the company, according to our survey. At companies with revenue of less than \$50 million, only 20% of our respondents reported having a dedicated storage group in place. At companies with revenue between \$501 million and \$1 billion, that number rose to 34.7%. And a whopping 68% of companies with revenue ranging between \$5.1 billion and \$10 billion had dedicated storage teams, which suggests that larger companies may be putting more emphasis on and committing more resources to their storage systems.

Downing has seen storage capacity double to just under a petabyte in

➔ **Q: Are all those years of education truly worth it?**

**A:** Yes, diplomas can really translate into dollars. High school graduates in our survey earned an average of \$80,480, while college graduates earned \$86,807. Those who earned a Master's degree or a Ph.D. reported an average annual salary of \$96,723.

➔ **Q: Does dedicated storage experience mean more money?**

**A:** Specialize in storage for more than five years and it starts to pay off. College graduates with three to five years dedicated storage experience earned \$88,717; six to 10 years yielded an average of \$95,440. More than a decade on the front lines plus a college degree kicks the salary average up to \$105,921.



the last 18 months, growing at a rate of approximately 10% to 15% a year. Repeating a trend that appeared in our 2008 survey, the current survey showed storage salaries growing as installed storage capacity does.

Those who [managed between 1 TB and 9 TB](#) had an average salary of \$73,646, while the average salary for those who managed between 10 TB and 99 TB spiked to \$82,666. Above 500 TB, the average salary jumped to \$104,441. Survey respondents who reported no storage capacity growth in 2009 received average salaries of \$82,250. For those whose shops experienced capacity growth of 10% to 40%, there were incremental salary increases. Above 40% growth, however, there was a more significant rise: respondents at companies whose storage grew between 41% and

## WHY I LOVE/HATE MY JOB

**Storage professionals who participated in our annual *Storage* magazine salary survey weighed in on the best and worst aspects of their current jobs, from technology to tight budgets and free donuts.**

### Eight reasons to **love** my job

- 1) Learning curve: "The variety of the day-to-day operations and the continual learning."
- 2) Hands-off: "I create my own projects; there's no micromanagement."
- 3) The challenge: "Achieving the near impossible."
- 4) The perks: "Free donuts on Friday."
- 5) Coworkers: "Working with good people who have passion and dedication to their jobs."
- 6) Costume changes: "I get to wear lots of different hats."
- 7) Home office: "Flexible work hours; being able to work from home."
- 8) Pride: "Being the 'go-to' guy."

### Eight reasons to **hate** my job

- 1) Too demanding: "Too much pressure on one person—me."
- 2) Tight budgets: "Growing so rapidly—budget cannot buy what we need."
- 3) Typecast: "Not much advancement opportunity."
- 4) The bottom line: "Poor benefits and mediocre pay."
- 5) Spread too thin: "Too many areas under my responsibility to be an expert at any."
- 6) Time crunch: "Lack of time to plan and complete projects—never given time to properly plan and implement technologies."
- 7) The big three: "Office politics, paperwork and bureaucracy."
- 8) Finally, free bagels, but "have to pay for cream cheese."

50% reported average salaries of \$91,588 vs. \$84,226 for those at firms with storage growth between 31% to 40%.

When it comes to salaries and company size, our 2009 survey revealed that only storage professionals at the very largest of companies—those with more than \$10 billion in revenue—saw any dramatic increases. Those who worked at companies with revenue of \$10 billion and beyond reported average salaries of \$99,588—nearly \$3,000 higher than last year's average for this group.

### COMMUTES, KIDS AND JOB SECURITY

At a time when many retirement funds are dwindling, Eric Hess, a 35-year-old operations engineer at the College of American Pathologists in Northfield, Ill., said his 401(k) is a key factor in his job satisfaction.

The non-profit organization contributes 15% of each employee's total pay into a 401(k) every six months, regardless of whether the employee is also contributing. That 401(k) benefit, Hess said, "is probably the most valued out of the college." While Hess' healthcare premiums didn't rise last year, he said, they have doubled since he arrived at the college approximately seven years ago.

Hess is a member of the newly formed work-life balance committee at the college. "They want to know if any team members are dissatisfied," Hess said of the committee, which consists of eight people and meets periodically to field employee requests and address employee concerns.

Echoing a sentiment expressed by many of our respondents, Hess said storage technology keeps him engaged and committed. "From the time I have been here—the last 7 years—we've gone to 1 Gig switches to 8 Gig switches. That shows you the growth and the potential that has happened," he said.

Jeffrey Lawrence, a contractor and senior systems engineer at the Asymmetric Warfare Group at Fort George G. Meade in Maryland, recently accepted a position at the Pentagon in Washington, D.C. Married with five children ranging in age from one to 19, Lawrence is looking forward to the stability of a government job.

Lawrence took rising gas prices and a shorter commute into account. The annual salary at his new job will be lower than his current contract rate. However, Lawrence, 39, will be commuting on the train and the government provides a \$250 monthly stipend for employees who use public transportation.

Heading into 2010, the key factor driving Lawrence is job security, he said. "As a civilian contractor, it's a very vulnerable time," he said. "But now that I'm going to the government side I feel a lot safer. That's a major factor for me." ☉

---

Ellen O'Brien is the senior editor at *Storage* magazine and site editor for SearchStorage.com and SearchStorage.co.UK. Rachel Kossman is a student intern/editorial assistant at TechTarget.

# Ensure you're dealt a winning hand.

CIO Decisions Ezine provides CIOs and business strategists with expert and peer insight "from the trenches" on pressing issues decision-makers face every day, including:

- Project and Portfolio Management
- Virtualization
- Business Process Management
- Outsourcing
- And much more!

**Subscribe Today!**



**CIO DECISIONS**  
ezine



# Object storage gains steam as unstructured data grows



*Object storage isn't a new concept in the NAS world, but some new products are bypassing traditional file system interfaces as industry debate emerges about the best way to cope with unstructured data.*

*By Beth Pariseau*

**IN THE AGE OF WEB 2.0**, the cloud and the digital content explosion, enterprise data storage managers are reevaluating how they store unstructured data as vendors roll out new object-based storage systems designed to offer simplified management and more scalable metadata schemes.

Unstructured data is expected to far outpace the growth of structured data over the next three years. According to the "IDC Enterprise Disk Storage Consumption Model" report released last fall, while transactional data is projected to grow at a compound annual growth rate (CAGR) of 21.8%, it's far outpaced by a 61.7% CAGR predicted for unstructured data.

"There are going to be extreme amounts of data as things like digital video and mobile networks grow; in five years, pretty much every phone will be 'smart,'" said Robin Harris, senior analyst at StorageMojo. "All of us storage geeks agree on that, and different people are beginning to visualize what that kind of growth needs in terms of storage infrastructure."



## THINK APIs, NOT FILES

Traditional hierarchical file systems organize data into “trees” consisting of directories, folders, subfolders and files. Files are a logical representation of blocks of data associated with an application and are the most familiar means of working with data. Network file system interfaces like NFS and CIFS are well-understood, standardized methods of conveying the logical groups of blocks from a storage repository to an application.

A problem arises, however, when a traditional file system, which has a theoretically limited number of files it can address in a single directory and tracks only simple metadata, runs into massive repositories of similar files.

“File systems make less sense over time as the amount of data grows,” StorageMojo’s Harris said. “Architecturally, it makes more sense for each file to have a unique 128-bit ID and use an Internet-like system for locating that file; a URL points to an address and there are files at that address, and object-based storage interfaces are essentially operating on the same principle.”

With an object ID replacing a file name, more extensive data can accompany an object than the simple “created,” “modified” or “saved on” fields available in traditional file systems. Thus, detailed policies can be applied to objects for more efficient and automated management.

Without NFS or CIFS to serve up files to applications, object-based storage systems need to replace that layer of abstraction between raw blocks of data on disk and files that applications can recognize. Today’s object-based systems use standard APIs such as Representational State Transfer (REST) and Simple Object Access Protocol (SOAP), or proprietary APIs to tell applications how to store and retrieve object IDs.

## NEW OBJECT-BASED STORAGE PRODUCTS TARGET THE CLOUD

For companies like Amazon, Flickr, Google or YouTube, whose intellectual property and differentiation comes from offering Web-based applications, programming their own interfaces isn’t such a big deal. But for companies with dozens or hundreds of applications, cobbling code to make each app work with object-based storage is likely to be an onerous and uneconomical task. There are, however, some storage vendors that offer pre-built but flexible architectures that do the job.

Caringo Inc. was first to position a content-addressed storage (CAS) system for nearline rather than archival storage, where CAS systems like

With an object ID replacing a file name, more extensive data can accompany an object than the simple “created,” “modified” or “saved on” fields available in traditional file systems.

EMC Corp.'s Centera (designed by the same engineers who later founded Caringo) historically played. In May 2008, the company claimed that its CASTor product can take the place of a file system or global namespace in traditional clustered storage products. CASTor runs CIFS or NFS using a file system gateway that can also be clustered (although no global namespace is available on the gateway), as well as HTTP access natively. According to the company, CASTor can be installed on nearly any x86 hardware with direct-attached storage (DAS).

EMC entered the market in November 2008 with its Atmos system, which it dubbed cloud-optimized storage (COS). Atmos uses object-based metadata to allow users to set policies that determine where to store data, which services to apply to it, and how many copies should be created and where they should be stored. REST and SOAP Web services are built in, as are capabilities such as replication, versioning, compression, data dedupe and disk spin-down. Users don't have to set up file systems or assign logical unit numbers (LUNs); during setup, they simply answer a few questions to set policies.

DataDirect Networks Inc. announced Web Object Scaler (WOS) in June 2009, and was expected to ship the system before the end of 2009. EMC said Atmos can scale to multiple petabytes and billions of files, but DataDirect Networks said WOS can handle more than 200 billion files and 6 petabytes (PB); the company

## PEACEFUL COEXISTENCE: OBJECT MEETS FILE

**"WE HAVE A LOT OF LEGACY STUFF"**—we want to use objects for scalability of medical image archives long term, but we're not a Web 2.0 company that can start fresh with a database and objects. Meanwhile, almost any computer system on the planet can connect through CIFS and NFS," said Michael Passe, storage architect at Boston-based Beth Israel Deaconess Medical Center.

Passe is working with EMC Corp. engineers to get file access into Atmos. "They're helping us push forward the file protocol side, but there's significant work to do with Samba to connect to Windows systems via CIFS," he said.

While managing objects will become a necessity down the road, Passe said Atmos' commodity hardware and scale-out architecture has appeal right now. "We went from Centera, at \$8 per raw gigabyte of data, to Atmos, at less than a dollar per raw gigabyte," he said. "Even if it makes four copies for data protection, it's still only \$2.80 per raw GB."

Connecting Windows systems via CIFS and Samba to an object-based system is fairly esoteric. However, Brent Welch, Panasas Inc.'s director of software architecture, said that Version 4.1 of the NFS standard will include support for connecting via the pNFS client to file, block or object-based storage systems, potentially easing integration of object-based storage into enterprise environments with legacy data like Passe's.

also claims a performance advantage over Atmos because its system holds object metadata in memory on its server nodes. Atmos metadata is partitioned and stored in a collection of databases spread across many disks in the system.

Cleversafe Inc. brought its dsNet Object Store out of the beta testing phase in September 2009. Cleversafe's SliceStor storage nodes can break a single file into as many as 11 pieces for redundancy, creating a hash that's appended to each slice for reconstruction. Cleversafe provides built-in encryption and previously offered the product with a block-level iSCSI or WebDAV interface. It's offering APIs for object-based access to the dsNet based on the Java software development kit (SDK) or using REST.

More recently, NetApp Inc. cloud czar Val Bercovici revealed in a blog post that the company best known for network-attached storage (NAS) will also be offering a native object storage interface "in the not too distant future."

### THE OBJECT DEBATE

Paul Carpentier, Caringo's CTO and co-founder, invented CAS as founder of FilePool, which became Centera after it was sold to EMC in 2001. Carpentier has become perhaps the most outspoken

proponent of object-based storage systems as a replacement for file systems altogether. "It's a heated debate," Carpentier said. "Personally, I'm very convinced we've stretched the hierarchical thing way too long."

Carpentier argues that file systems were originally built to allow concurrent access to smaller groups of objects shared among a few users. But now, he said, there's a "mismatch between prevailing use cases [for unstructured data] and how those systems work. Ninety percent to 95% of us don't need a storage system with concurrent locking for reference information."

Carpentier noted that the management of file systems is too meticulous to be practical at petabyte scale. "Some products create a virtualization layer that presents a global namespace, but there might be 20 underlying file systems you have to manage individually, and sooner or later the Web 2.0 business model bumps into an impossibility," he said. Furthermore, at scale, "backup just doesn't cut it anymore, you need live replication."

Object interfaces decouple data from the underlying disk hardware in a way file systems can't keep up with, said Cleversafe CEO Chris Gladwin. "With objects, there isn't a size limit or a concept of drive size; there's just a single namespace that can theoretically encompass all the hard drives on the planet."

"It's a heated debate. Personally, I'm very convinced we've stretched the hierarchical thing way too long."

—PAUL CARPENTIER,  
invented CAS as founder of FilePool

One EMC and NetApp user said he agrees with this point of view. “I feel really strongly that the file systems we have today are not all that great. In the mainframe days, you could include attributes with a file to help manage them,” said Tom Becchetti, a veteran storage professional who asked that his company not be named because of organizational policy. “With file systems, if you need to manage some files differently from others, you do it in separate server buckets today.”

That runs counter to the consolidation going on with server virtualization, and Becchetti said object-based storage “could be a key enabler to grow the virtual [server] world, where an object isn’t a file but a VMDK [virtual machine disk file]. It could mean I could share a VMDK between more physical servers than is possible with today’s file systems, and protect it on a grander scale with policy-based management, where I could say anything with ‘P’ in the VMDK name should be protected this way vs. anything with ‘D’ in the name.”

Still, even in some of the most demanding environments, users said file systems can get the job done.

Speaking on a recent Wikibon.org conference call, Eugene Hacopians, senior system engineer at The California Institute of Technology (the academic home of NASA’s Jet Propulsion Laboratory), said the 2 PB of storage in his environment, comprising billions of 5 KB to 25 KB files, still runs mostly on traditional storage systems from Nexsan Technologies Inc.

But that’s been a matter of timing, project lifecycles and budget rather than technical preference. “We have looked at [object-based storage] and are considering it for newer projects,” Hacopians said. “It’s difficult to convert to new technology and fork out additional money when you’re in the middle of trying to deliver on a project.”

"I feel really strongly that the file systems we have today are not all that great. In the mainframe days, you could include attributes with a file to help manage them."

—TOM BECCHETTI,  
veteran storage professional

## DIFFERENT PRODUCTS FOR DIFFERENT USE CASES

Another viewpoint maintains that file vs. object doesn’t have to be an either-or proposition. NetApp and EMC, for example, have both expressed this point of view.

“If there are limits to traditional file systems, we’re not running into them today,” said Peter Thayer, director of marketing, midrange products at EMC. “It’s more a matter of application-centric use cases in Web 2.0 requiring additional metadata than running out of gas in the traditional file system space today.”



John Hayden, EMC's CTO of NAS engineering, added that if users require shared read/write access to the same files, "you'll get more horsepower out of traditional file systems today in terms of performance."

NetApp's Bercovici echoed that outlook. NetApp continues to roll out file system-based products, most recently its Ontap 8 operating system, which will support scale-out. However, "if you need to support millions, hundreds of millions or billions of similar objects, like medical images, storage interfaces are just overhead," he said. "You don't want to create LUNs, folders and permissions; you just want a single scalable directory."

Some users find a combination of products works best for different needs within the same environment. At the Johns Hopkins University Bayview Research Campus Center for Inherited Disease Research, data processing for genetic research processing is done using clients attached to a 72 TB Isilon Systems Inc. clustered NAS system, but once data passes from being actively shared among researchers to being kept as reference information, it's moved to Caringo's CASTor object-based system.

"Isilon provides a large shared file system to support desktop data analysis for the computers that drive instruments in our lab," said Lee Watkins Jr., the center's director of bioinformatics. It's important to have file-locking capabilities and the ability to manage permissions across both Windows and Linux OSes in this environment, though Watkins said this can often carry management headaches.

"We have very large files people need access to from Linux, Mac OS X and Windows desktops, some reading, some writing, and we have to decide how to balance throughput to the different [Isilon] nodes—which file system is going to mount to each node," he said.

Once data passes into the archive stage, Watkins said it's more important to be able to access the data and metadata quickly when it's needed. "We also produce a tremendous amount of data. It can be between a terabyte and 3 TB per day," he said. For Johns Hopkins, writing an application to access the Caringo storage through an API "was pretty simple," according to Watkins. "We can move files around on the back end and not worry about addressing and where it is, and it doesn't matter what operating system is requesting the file."

"You don't want to create LUNs, folders and permissions; you just want a single scalable directory."

—VAL BERCOVICI,  
NetApp's cloud czar

## COMBINING FILE PROTOCOLS WITH OBJECT STORES

File and object aren't necessarily mutually exclusive ideas even within the same system. In fact, several existing scale-out NAS systems already

have object stores underlying a file interface, including BlueArc Corp.'s Titan, Panasas Inc.'s ActiveStore and ParaScale Inc.'s Hyper-scale Storage Cloud.

"Objects are kind of an overloaded term," said Brent Welch, Panasas' director of software architecture. "Different people define it differently, but it's essentially a container for data that serves as a building block for higher-level storage systems." The Panasas distributed file system knits together NFS with an underlying object store to meet the scalability demands of high-performance computing.

Systems like CASTor and Atmos essentially peel back the network protocol layer and let the application interface directly with the object store. Some products, like BlueArc's Titan, also allow administrators to search using more detailed object-based metadata schemes, though end users in the environment access the system through NFS.

James Rainey, BlueArc's executive director of strategic technology, said BlueArc has allowed some partners to integrate applications directly into the object store using a proprietary API, and they're considering opening up that API for more general use.

Some enterprise users are looking to ease object-based systems for archival data into their environments by putting together standard file-based access with one of the newer object storage systems built on commodity hardware. BlueArc stores file system and object metadata in proprietary field-programmable gate arrays—FPGAs—and Panasas uses a proprietary NFS client (see "[Peaceful coexistence: Object meets file](#)" p. 46).

Despite the efforts to meld object and file systems, StorageMojo's Harris predicts the debate over files and objects will continue. "There has been a low-level religious war going on for quite some time," he said. "File systems have been a key technology for decades, but we're rapidly reaching the point ... where it doesn't make sense to tie data to a specific disk drive attached to a specific path name anymore." ☉

---

Beth Pariseau is senior news writer for SearchStorage.com.

Systems like CASTor and Atmos essentially peel back the network protocol layer and let the application interface directly with the object store.



# Desktop virtualization: Better data protection?

*Virtual desktop infrastructure technology can ease the burden of data protection for laptops and desktops, but it may not be a good fit for all types of end users.*

**Q** **F ALL THE DATA** your company owns, data residing on desktops and laptops is often the least protected. Why? The distributed nature of endpoints makes it difficult to centralize and consolidate backup, and since desktop/laptop data exists outside the confines of the data center, backup administrators often don't see its protection as their problem.

Virtual desktop infrastructure (VDI) technology can address this problem by bringing data that would otherwise live on end-user devices into the data center.

VDI products enable the centralization of entire personalized end-user desktop operating environments so that they can be efficiently accessed, managed and protected from a central location. This allows organizations to reduce operational costs, improve service levels, and satisfy compliance and information security requirements, all while maintaining an identical—and in some cases, improved—end-user experience.

One caveat is that, as with server virtualization, desktop virtualization will have an impact on IT infrastructure. Server, storage and networking will all be impacted. Enterprise Strategy Group (ESG) research shows that nearly two-thirds (64%) of current VDI users have made some form of new storage purchase to support their implementation, since data that used to reside on users' PCs is stored on data center hardware in a VDI environment. And VDI isn't a one-size-fits-all solution. Some groups of users aren't well-suited to it, so different measures should be put in place to protect their data.

**VDI products enable the centralization of entire personalized end-user desktop operating environments so that they can be efficiently accessed, managed and protected from a central location.**

## **DISTRIBUTED COMPUTING'S BACKUP PROBLEM**

Most IT organizations today give short shrift to protection of PC data. According to recent ESG research, only 26% of nearly 500 midmarket and enterprise IT respondents said that all of their desktop PCs are backed up,

and only 18% of organizations back up all of their laptops. Furthermore, 24% of respondents said they have no data protection process at all for their desktops, and 29% don't back up their laptops. With desktops and laptops increasingly carrying business-critical data, the expenses incurred for a system loss or failure are much greater than simple hardware replacement—most notably, lost end-user productivity during downtime and more time spent reconstructing lost data.

Many IT organizations take proactive steps to centrally administer backup, ranging from manually copying files to a network share that's integrated in automated server backup processes to directly backing up desktops/laptops via server-based backup client agents. The former could create gaps in protection, while the latter introduces challenges due to the sheer volume of devices and the required software licensing to protect them.

Other companies take a different approach, allowing desktop/laptop users to do it themselves. Users might manually copy files or use a standalone PC-based backup product to automatically back up data to a local storage device, such as CD/DVD, USB drive or memory stick. Alternatively, some leverage backup software-as-a-service (SaaS) to enable automated backup of data to a third-party location. In these situations, copies of corporate data proliferate outside the custody and control of the IT department (and sometimes the user), potentially introducing additional risk to the organization.

While most IT organizations don't adequately protect end-user data, the hardware that contains that data is susceptible. With an average hard drive failure rate of 2% to 4%, a company with 500 laptops could have as many as 20 of these devices experience a disk crash. In addition, the portable nature of laptops makes them an easy target for theft/loss and prone to damage from being mishandled/dropped.

In spite of these dangers, some IT organizations don't see the risk of data loss outweighing the costs of desktop/laptop backup storage capacity and operational overhead. Moreover, many organizations cite a lack of business or legal requirements mandating data protection and simply procure, configure and re-image replacement hardware and let users worry about data reconstruction.

**With an average hard drive failure rate of 2% to 4%, a company with 500 laptops could have as many as 20 of these devices experience a disk crash.**

### **VDI's BACKUP REMEDY**

VDI enables a user's complete desktop environments—including OS, profile, applications, user data and customizations—to be deployed as a self-contained package, remotely accessible from anywhere. Administra-



tive and management tasks are, therefore, streamlined and centralized.

While you're not likely to implement VDI just to solve the PC backup problem (a host of difficult desktop computing challenges are driving its adoption), moving PC images to the data center puts them under the umbrella of the data center's data protection policies, processes, infrastructure and operational staff, which enables more efficient backup and recovery. And as an integrated component of server backup, desktop and laptop data can benefit from advanced features such as data deduplication.

**VDI could be too "locked down" for certain classes of users, such as knowledge workers.**

It's important to note that VDI isn't a solution for everyone. VDI could be too "locked down" for certain classes of users, such as knowledge workers. Organizations generally use or expect to use VDI for distributed workers such as remote employees and telecommuters, as well as for task workers in roles such as data entry and call center (who may require a more limited desktop environment). And while they're not ideally suited for VDI because their use of technology extends beyond a traditional set of limited tasks, VDI for knowledge workers offers benefits against the leakage of sensitive company information.

However, those user profiles that aren't a good fit for VDI will remain vulnerable without an alternative desktop/laptop backup strategy. Organizations adopting an "out of sight, out of mind" attitude regarding end-point protection could leave themselves open to risk. ☹

---

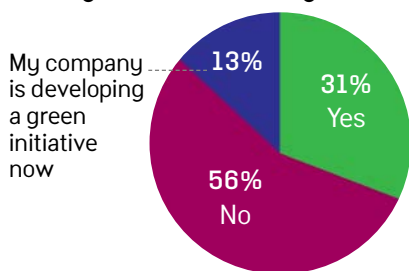
Lauren Whitehouse is an analyst focusing on backup and recovery software and replication solutions at Enterprise Strategy Group, Milford, Mass.

# Energy conservation efforts still anemic

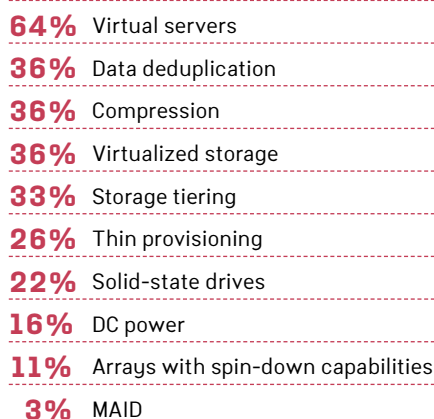
**IT'S NOT** a black-and-white decision when considering green storage. In our latest survey, only 31% of respondents said their companies have green initiatives or commitments in place, a drop of a few points from last year. While the dip is discouraging, those working toward power conservation report better results: 15% (vs. 12% last year) said their efforts exceed their expectations, while 36% report being right on target (a big jump from the 24% noted a year ago). There might also be a low-energy light at the end of the tunnel, with 16% of those surveyed (vs. 6% last year) expecting their companies to commit to green storage initiatives within a year. More than a third of respondents said they're using data deduplication, compression and virtualized storage to conserve storage energy; however, the biggest increases compared to last year were for thin provisioning (26% vs. 15%), solid-state drives (22% vs. 15%) and DC power (16% vs. 8%). Overall, 59% implemented these newer techs to help curb equipment power usage. Still, money is green, and while 53% were willing to shell out "a little more" for a green storage product, 33% of those surveyed said they wouldn't pay more.

—Rich Castagna

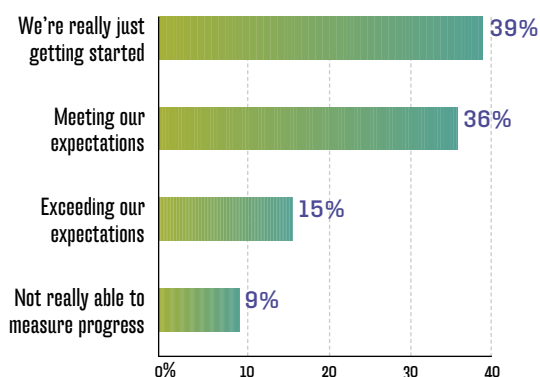
Does your company have initiatives or commitments in place from corporate management to address green storage issues?



Which of these technologies have you implemented?



How would you rate your power conservation efforts so far?



**31**  
Percent of respondents who said that green features are very important.

**"Green needs to bring a positive ROI—better than traditional solutions—to the table. Just being the 'right thing to do' will not fly in a hard-dollar ROI review."**

—Survey respondent

## Check out the following resources from our sponsors:

### BlueArc Corp., page 13

[SiliconFS™: The BlueArc® Filesystem](#)

[ESG White Paper: Enabling Enterprise Class Features for the Mid-Range: Focus on BlueArc Mercury](#)

### CA, page 17

[ARCserve Backup® r12.5 Trial Download](#)

[Business Continuity Planning IT Survival Guide](#)

### Data Domain, page 4

[Data Domain – Data Deduplication Center](#)

[Deduplication School](#)

### FalconStor Software, page 22

[Demartek Report: FalconStor Virtual Tape Library \(VTL\) and Deduplication Performance](#)

[Demo: FalconStor Virtual Tape Library Deduplication](#)

### Hitachi Data Systems, page 26

[Tiered Storage and Virtualization in the Real World](#)

[eBook: 6 Essential Strategies For Economizing Your Storage](#)

### i365, A Seagate Company, page 7

[Smart Move: Spend Less While Improving Data Protection](#)

[12 Best Practices for Data Backup and Recovery](#)

### Nexsan Technologies, page 30

[No More Tiers: Reduce Storage Costs with an Age-in-Place Strategy](#)

[If it's Just a Disk ... Why the Reliability Gap Between Storage Vendors?](#)

### Northern Parklife, page 33

[Northern Storage Suite - The Movie: See what NSS can help you with in six minutes](#)

[View our webinar for an in-depth look at what Northern Storage Suite can do for you](#)

### Sun Microsystems Inc., page 36

[Sun Storage 7000 Unified Storage Systems: The Core of an Accelerated MCAE Workflow](#)

[Sun Microsystems Open Storage and SSD Resource Center](#)

### Xiotech Corp., page 9

[The New Foundation of Storage - Intelligent Storage Element \(ISE\) Technology](#)

[Virtualization with Xiotech's Intelligent Storage Element \(ISE\)](#)