Managing the information that drives the enterprise

Vol. 8 No. 2 April 2009

the state of **Replication**

STURAGE

Replication is more affordable, scalable and easier to implement than ever. But where's the best place to deploy it? page 10





ALSO INSIDE

- 4 Rethinking the cloud
- 8 Data virtualization turns less into more
- 19 How virtualization may affect your job
- 24 Power-smart disk systems
- 28 Backup on a budget
- 31 Online storage: Not your father's SSP

33 Do you dedupe?







 EDITORIAL Cloud storage vendors might find a silver lining to the current economic cloud as firms try to cut costs.
But enterprise storage shops are still wary. by RICH CASTAGNA

Data Virtualization Turns Less Into More

8 STORAGE BIN 2.0 Less physical data living and moving over our plumbing means we can have more virtual instances of that data, which drives even more value. by STEVE DUPLESSIE

Replication Alternatives

10 Data replication is great for protecting critical data and ensuring quick recoveries. Find out where you should deploy replication: in your array, network or servers. by JACOB GSOEDL

Virtual Jobs

19 Virtualizing servers, storage and even networks will change the face of IT and significantly impact the roles of storage professionals. by RICH FRIEDMAN

Power-Smart Disk Systems

24 Disk drive systems use more power than just about any other data center gear, but storage vendors are addressing this problem with a variety of technologies. by MATT PERKINS

Value-Based Data Protection

28 There's no skimping when it comes to data protection, but tight budgets mean that storage managers will be looking for solid value in their backup products. by RACHEL KANNER

Is Cloud Storage the Return of the Service Provider?

31 HOT SPOTS Cloud storage is a service, just like what storage service providers attempted to do years ago. But the technology has significantly evolved in the last decade and this time, cloud storage is real. by TERRI McCLURE

The Buzz About Dedupe

33 SNAPSHOT The storage industry is abuzz with dedupe. Approximately one-third of respondents to *Storage* magazine's latest survey use a dedupe product, but how easy will it be to convert the stragglers? by RICH CASTAGNA

Vendor Resources

34 Useful links from our advertisers.





2

SIMPLIFY OURSTORAGE TO IMPROVE OVERALL IT EFFICIENCY



INTRODUCING THE EQUALLOGIC[™] PS6000 SERIES! THE NEWEST ENTERPRISE-CLASS STORAGE SOLUTION FROM DELL.

Your storage solution should be simple, from initial purchase to scaling out your SAN. With a single, intuitive GUI for multiple group management, plus all inclusive features like Smart Copy for Microsoft[®] Hyper-V^{™*} and support for VMware[®] vStorage, the EqualLogic PS6000 Series iSCSI SAN streamlines and simplifies management while protecting your data. What's more, its patented "peer" architecture allows for on-demand scaling, enabling a true pay-as-you-grow iSCSI SAN solution to help improve overall IT efficiency.



SIMPLIFY YOUR STORAGE AT DELL.COM/PSseries/PS6000

* Availability expected by Q2, 2009



Rethinking cloud storage

Cloud storage vendors might find a silver lining to the current economic cloud as companies try to cut costs. But enterprise storage shops are still wary.

T WAS APPROXIMATELY a year ago that I wrote about cloud storage in fairly glowing terms; if I wasn't entirely on the bandwagon back then, I was pretty close to hopping on board. Six months later I was less enthusiastic about how realistic cloud storage was for enterprises. Now I'm convinced that if cloud storage ever does play a significant role in larger companies, it won't happen anytime soon.

Make no mistake, there are a lot of reasons to be optimistic about cloud storage even after it stumbled in its first incarnation some years back (see "Is cloud storage the return of the service provider?" p. 31). The services are clearly better than ever, more secure, easier to use and less expensive. So what's not to like?

For businesses with limited resources for dealing with the massive onslaught of data and the challenge of providing enough capacity to house it, cloud storage can seem like a good fit. These same businesses are also likely those struggling with backup (if they're doing it at all), so an online backup service might be a compelling alternative.

For those firms, often in the small- to mediumsized range, the perceived risks of using a cloud storage service—data security, availability and accessibility—are better than doing nothing or dealing with poorly maintained, inadequate systems.

But large enterprises have alternatives. If you have the infrastructure in place and recognize

Shipping data offsite makes compliance, which is already a pretty tough task, even more difficult.

the value that lurks within those tons of data, you'll find the means to accommodate it. Any big organization has to think beyond just housing its data. Compliance is an obvious issue. Most companies are bound to regulatory mandates, and some have to ensure compliance with dozens of regulations.

Shipping data offsite makes compliance, which is already a pretty tough task, even more difficult. As David Sengupta of Ferris Research recently pointed out in an email newsletter, "compliance regimes necessitate that certain customers need to have a clear handle on their data, be that for retention, legal discovery, search, or other purposes." Enterprises can't afford to lose those capabilities or have them limited

Δ



Cloud storage

5

in any way, especially when litigation or the threat of non-compliance looms.

Some cloud storage vendors say tapping into their services is like using a utility where you pay for what you use. It's true that these services typically charge on a usage basis: use a little, pay a little; use more, pay more. But there are some pretty big differences when you compare these services to familiar utilities, like buying electricity or water for your home or business.

Those services have been around forever and can reasonably guarantee a predictable level of service. And with utilities like water or electricity, the customer doesn't bear any risk. But if you can't get to or properly manage the data you've stored in the cloud, you could be a lot worse off than just missing the *American Idol* finale when your power goes out.

In our world, SaaS means "storage as a service," but the rest of IT defines the first "S" in that acronym as "software." That's a significant differentiator. The software cloud guys offer apps that you can tap into and run your own data through, often without actually moving that data out of your shop. Other software-as-a-service offerings are management oriented, with applications residing on distant systems that you can run to maintain and manage your own systems.

Those types of cloud services are likely to do very well, especially when you consider that the risks are much lower for the services' clients. And with so many applications that can take months to implement, cloud software services Some cloud storage vendors say tapping into their services is like using a utility where you pay for what you use.

can be attractive and economical alternatives. But storage is a different animal. On our twice-a-year purchasing surveys, respondents have shown tepid interest in backup-oriented cloud storage services; only 12% to 15% use them at all, and it's typically for special circumstances like backing up remote offices or laptop/desktop data.

Cloud software and cloud-based servers seem much more likely to make it into the enterprise mainstream, as storage cloud providers try to build on the inroads they've made in small- and medium-sized businesses while shoring up their offerings. Cloud storage vendors might see a silver lining to the current economic cloud with companies trying to avoid capital expenditures, but CAPEX consciousness alone won't get these services into enterprise shops. They'll have to build a track record first and, more importantly, a compelling enough case to convince enterprises that their data is safer in the cloud than in their shop. **O**

Rich Castagna (rcastagna@storagemagazine.com) is Editorial Director of the Storage Media Group.

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

30-second recovery for your Exchange, SQL Server and Windows file system data





Get a free BakBone stopwatch just for attending a live demo!

See for yourself

Attend one of our weekly 30-minute LIVE online demonstrations and witness

- Near-instantaneous Exchange recovery
- Exchange brick-level recovery without brick-level backups
- Fast and flexible any-point-in-time recovery for Windows file system data
- Easy set up for SQL Server protection

Register today! www.bakbone.com/30-second-restore

NetVault: Real-Time Data Protector" gives you near-instantaneous recovery

In as little as 30 seconds, your users are able to regain access to their Exchange, SQL Server and Windows file system data, significantly reducing downtime and its impact on your business.

Near-zero data loss

Don't miss a thing! NetVault: Real-Time Data Protector is *always on* – capturing continuous history for granular any-point-in-time recovery of your most important data.

Maximize revenues and cut costs to survive (and thrive) in tough times

Downtime and data loss generate high recovery costs and strip away revenue. By significantly reducing downtime and ensuring near-zero data loss, NetVault: Real-Time Data Protector is a must have solution for those who need to protect their revenue streams while reducing capital expenditures.



COMING IN MAY

Mr. Backup on Testing Dedupe

W. Curtis Preston, executive editor and independent backup expert, takes you step-by-step through the process of evaluating a data deduplication product. While vendors boast of impressive dedupe ratios, Curtis will show you how to test their systems to see what kind of results you can realistically expect in your backup environment.

Who's Buying What?

Twice a year, *Storage* magazine surveys its readers to see where their budget dollars are going. See how your peers plan to stretch tight budgets in 2009 while dealing with ever-growing data capacity demands. The survey, which is in its seventh year, also reveals developing trends and shows what storage-related technologies are gaining the most interest among your peers.

Case Study: Storage Tiering

The benefits of storage tiering can be profound, with significant hardware and software savings, and more efficient management of stored data. In this firsthand case study, you'll get the inside info on how a tiered storage environment was implemented at a large research facility.

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

STORAGE

Vice President of Editorial Mark Schlack

Editorial Director Rich Castagna

Senior Managing Editor Kim Hefner

Senior Editor Ellen O'Brien

Associate Editor Christine Cignoli

Creative Director Maureen Joyce Contributing Editors

James Damoulakis Steve Duplessie Stephen Foskett Jacob Gsoedl

Storage April 2009

TechTarget Storage Media Group

SearchStorage.com

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 Christine Cignoli



Site Editor Susan Troy



SearchDisasterRecovery.com

Site Editor Andrew Burton

Associate Site Editor Heather Darcy

Assistant Site Editor Chris Griffin

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

Subscriptions: www.SearchStorage.com

State of replication



Data virtualization turns less into more

Less physical data living and moving over our plumbing means we can have more virtual instances of that data, which drives even more value.

ATA VIRTUALIZATION, which is a virtual instance of data instead of a physical one, is perhaps the most interesting means of gaining efficiency throughout the entire spectrum of IT operations. But people often overlook that data is the reason why the infrastructure exists to begin with. Therefore, solving inefficiencies directly at the data layer will often have the most significant return on effort and investment. Deduplication, thin provisioning and snapshots are examples of techniques that dramatically reduce or eliminate the true physical issues associated with data, and continue to provide unfettered "virtual" access.

Consider data deduplication. Data is duplicated time and time again (pushing the capacity limits of all infrastructures and taxing operating processes), often for perfectly valid reasons. However, it has an exponential negative impact on the backup, recovery and disaster recovery (DR) processes we conduct. By default, most backups regularly create full images of data that might have been duplicated dozens of times. The positive gains of eliminating duplicate data inefficiencies can be felt throughout an entire organization.

For example, by backing up to a disk target with deduplication, IT gets an immediate primary benefit of speed, as disk enables a tremendous performance gain. When the data is deduplicated, it then takes up significantly less space at rest and during transmission. Therefore, data deduplicated at point A will require significantly less infrastructure (bandwidth, capacity, etc.) to store and transmit that "virtual" instance of data to point B and so on. Because typical backup dedupe rates can be Data deduplicated at point A will require significantly less infrastructure to store and transmit that "virtual" instance of data to point B and so on.

20:1 or higher, keeping virtual data on disk can dramatically improve the efficiency of recovery operations, bringing even higher service-level abilities and value back to the business. This new efficiency gain is multiplied by improving overall IT operations and allowing an already taxed staff to gain perhaps the most valuable of all operating assets—time.



STORAGE

Mandated or regulatory retention requirements of an email record are fairly easy to understand. But it's a tougher task to determine the most effective and economical means of complying with this mandate. The easiest way to comply is to apply the rule to the data, put it somewhere and never move it. That way, you can always find it when you need it. But that practice is often at odds with gaining efficiency and optimization. Email is a representation of data, and during its early life it may require availability that it simply will not need after a certain period of time; for example, when it becomes a fixed (persistent) data object that will never change. It therefore makes more sense to house that

email on the most cost-effective infrastructure platform attainable, which most likely isn't the originating platform. Just because retention and immutability are mandated doesn't mean that object is relegated to inefficient treatment forever.

The same benefits are realized and magnified as we apply this logic to "non-mandated" data. Every data object, regardless of form, will eventually become a persistent, nonchanging asset that will be infrequently accessed. Data in this stage, which represents the overwhelming majority of corporate data being managed, has radically different attribute requirements from when it was active and dynamic. Generally, it's safe to say that whether considered as data within an archive or simply data in a lower tier of infrastructure, once it stops changing and being heavily accessed, we can apply the same efficiency logic to it. Every data object, regardless of form, will eventually become a persistent, nonchanging asset that will be infrequently accessed.

By deduplicating this data, we can more easily and efficiently protect it, access it, secure it and store it without requiring many of the superhuman efforts our IT staffs are currently forced to provide.

Less physical data living and moving over our plumbing means we can have more virtual instances of that data in even more areas of our business which, in turn, drives even more value. This is a positive cycle that feeds and leverages off a basic truth: Less is better than more. Θ

Steve Duplessie is founder and senior analyst at Enterprise Strategy Group. You can see his blog at http://esgblogs.typepad.com/steves_it_rants/.

9

Replication alternatives

Data replication is great for protecting critical data and ensuring quick recoveries. Find out where you should deploy replication: in your array, network or servers.

By Jacob Gsoedl

ATA REPLICATION as a means of data protection has seen continuous and increasing adoption since it first emerged in storage systems after the first World Trade Center bombing in 1993. Over time, it has evolved into an indispensable component of disaster recovery (DR), as well as for operational backup for applications that require shorter recovery point objectives (RPOs) and recovery time objectives (RTOs) than what traditional tape backups can offer. Firms are also adopting data replication for remote- and branch-office data protection; in a hub-and-spoke architecture, branch-office data can be replicated back to central data centers, thus eliminating unwieldy tape-based backup procedures at the branch sites.





The growing adoption of replication services has been driven by a wide array of data replication products, more lower cost replication offerings, faster and less-expensive networks, and an overall maturing of the technology itself. "Replication-based data protection is among the top three priorities of 60% of our clients, which is very different from only a few years ago," said Tim Bowers, global product manager, storage services at EDS, a Hewlett-Packard (HP) company.

NOT ALL REPLICATION IS EQUAL

At a macro level, data replication copies data from one storage location to one or more other local or remote storage systems. But venture beyond that basic task and you'll find that data replication products vary in several key aspects:

LOCATION: One of the main differentiators among products is where replication occurs. The replication service or software can reside on the storage array, in the network or on the host (server). Array-based replication has been dominating the replication market up to now.

"We did a recent study that shows that in 2007, 83.7% of worldwide revenue for storage-based replication was done using array-to-array replication, followed by hostbased replication with 11.5% and networkbased replication with 4.8%," said James Baker, research manager, storage software at Framingham, Mass.-based IDC. But according to the same study, both host- and networkAt a macro level, data replication copies data from one storage location to one or more other local or remote storage systems.

based replication are catching up. Host-based replication is expected to grow at a compound annual growth rate (CAGR) of 18.2% until 2012, while a CAGR of 15.4% is anticipated for network-based replication. Both are expected to expand significantly faster than the 10% forecasted annual growth for array-based replication.

MODE: Replication can occur synchronously, where data is written to the primary and secondary storage systems simultaneously; or it can be performed asynchronously, where data is replicated to replication targets with a delay. In synchronous replication, the primary storage system only commits I/O writes after the replication target acknowledges that data has been written successfully. Synchronous replication depends on sufficient bandwidth and low latency, and supported replication distances range from 50 km to 300 km. It's typically used in applications where zero RPOs and RTOs are required, such as high-availability clusters and mission-critical applications that demand 100% synchronicity between the primary and target systems. Conversely, asynchronous replication writes data to the primary array first and, depending on the implementation approach, commits data to be replicated to memory or a disk-based journal. It then copies the data in real-time or at scheduled intervals to

Cloud storage



STORAGE

replication targets. Unlike synchronous replication, it's designed to work over long distances and greatly reduces bandwidth requirements. While the majority of array- and network-based replication products support both synchronous and asynchronous replication, host-based replication offerings usually only come with asynchronous replication.

TYPE: Replication products can replicate blocks of data on volumes or logical unit numbers (LUNs), or replication can be performed at the file level. With the exception of network-attached storage (NAS), which can support both block- and file-based replication, array-based replication products usually operate at the block level. The same is true for networkbased replication products. In contrast, most host-based replication offerings operate at the file-system level. Block-based replication is platform-agnostic and will work seamlessly across various OSes. File-based replication products are very much operating system-specific and the majority of available host-based replication products are written for Windows. Unlike file-based replication, block-based replication products have no knowledge of the attached platform, file system or apps, and depend on auxiliary services like snapshots for any type of application integration. As a result, most storage arrays with replication support also provide snapshot capabilities that are more or less integrated with the file system and key apps like Exchange and SQL Server databases.

DATA REPLICATION TRENDS

These data replication-related trends are gradually changing data protection and disaster recovery.

- The use of replication-based data protection is increasing.
- Replication-based data protection is merging with traditional data protection, and traditional backup products are increasingly able to manage replicas and snapshots along with backups.
- The integration of replication and applications to provide application-consistent recovery is on the rise.
- Replication-based data protection is becoming an important option for protecting virtualized server environments.
- As storage-as-a-service (SaaS) and cloud-based computing become more prevalent, hosted replication offerings will become more common.
- Because replication relies on available bandwidth, wide-area network (WAN) optimization offerings from the likes of Blue Coat Systems Inc., Cisco Systems Inc., Citrix Systems Inc., F5 Networks Inc., Juniper Networks Inc., Packeteer (now a Blue Coat company), Riverbed Technology Inc. and Silver Peak Systems Inc. are used to complement replication products to preserve valuable WAN bandwidth.

State of replication



ARRAY-BASED REPLICATION

In array-based replication, the replication software runs on one or more storage controllers. It's most prevalent in medium- and large-sized companies, mostly because larger firms have deployed higher end storage arrays that come with data replication features.

With more than 15 years of history, array-based replication is the most mature and proven replication approach, and its scalability is only constrained by the processing power of the array's storage controllers. "Customers scale replication performance in both our Clariion and Symmetrix arrays by distributing data replication across a larger number of storage processors," explained Rick Walsworth, director product marketing replication solutions at EMC Corp.

With the replication software located on the array, it's well suited for environments with a large number of servers for several reasons: it's operating system-agnostic; capable of supporting Windows and Unixbased open systems, as well as mainframes (high-end arrays); licensing fees are typically based on the amount of storage rather than the number of servers attached; and it doesn't require any administrative work on attached servers. Because replication is offloaded to storage controllers, processing overhead on servers is eliminated, making arraybased replication very favorable for mission-critical and high-end transactional applications.

The biggest disadvantage of array-based replication is its lack of support

		TYPE OF REPLICATION	
	Array based	Host based	Network based
Support of heterogeneous environments	Low; only works between similar arrays	High; storage-agnostic and works with network- and direct- attached storage	High; storage array- and platform-agnostic
Performance and scalability	Depends on the storage array; very good in high-end arrays	Good; workload is spread across servers; limited scalability because of manage- ment challenges	Very good
Cost	Requires similar arrays; high entry cost; expensive for a large number of locations	No hardware required; low entry cost; cost rises proportionally to the number of servers	Requires intelligent switches or inline appli- ances; high entry cost; expensive for a large number of locations
Complexity	Medium to high	Low	Medium to high
Replication modes	Synchronous and asynchronous	Asynchronous	Synchronous and asynchronous
Predominant replication type	Logical unit number (LUN) or volume block-level based	File-system based	LUN or volume block-level based

COMPARING DATA REPLICATION METHODS

SearchStorage.com

of heterogeneous storage systems. And unless the array provides a storage virtualization option—as Hitachi Data Systems does for its Universal Storage Platform (USP)—array-based replication usually only works between similar array models. Besides a high degree of vendor lock-in, entry cost for array-based replication is relatively high, and it could be particularly expensive for companies that have to support a large number of locations. In general, array-based replication works best for companies that have standardized on a single storage array vendor.

Almost all vendors of midsized to high-end arrays provide a replication feature. The replication products of these leading array vendors have made significant inroads and gained market share:

• EMC Symmetrix Remote Data Facility (SRDF) for both synchronous and asynchronous replication, and EMC MirrorView for synchronous and asynchronous replication of Clariion systems.

• Hitachi Data Systems TrueCopy for synchronous replication and Hitachi Data Systems Universal Replicator software for asynchronous replication.

• HP StorageWorks XP Continuous Access and Continuous Access EVA for both synchro-

The biggest disadvantage of array-based replication is its lack of support of heterogeneous storage systems.

nous and asynchronous replication for HP XP and EVA arrays.

• IBM Corp. Metro Mirror for synchronous replication and IBM Global Mirror for asynchronous replication.

• NetApp SnapMirror for synchronous and asynchronous block-based replication, and NetApp SnapVault for file-based replication.

Even though these replication products are similar in many aspects, a close technical analysis reveals subtle differences. For instance, the efficiency of the handshake between primary and target storage systems used during synchronous replication greatly impacts the distance a replication product can support. "Metro Mirror is able to write data to the target system with a single handshake, enabling it to support distances of up to 300 km," said Vic Pelz, consulting IT architect at IBM. That distance goes well beyond the 50 km to 200 km cited by other storage vendors.

Differences can also be found among asynchronous replication implementations. While EMC buffers data to be replicated in memory, IBM Metro Mirror tracks changes with so-called bitmaps, continuously transmitting changes and periodically re-synchronizing the source and target to ensure they stay in sync. On the other hand, Hitachi Data Systems uses change journals stored on disk in its Universal Replicator software.

"The combination of disk-based change journals that are pulled by the replication targets instead of pushed by the source, makes it extremely resilient, capable of automatically recovering from elongated disruptions,"

Cloud storage



said Christophe Bertrand, senior director, solutions and product marketing business continuity at Hitachi Data Systems. "Because changes are pulled by replication target arrays, valuable processing cycles are offloaded from primary arrays to secondary target arrays."

HOST-BASED REPLICATION

In host-based replication products, the replication software runs on servers so, unlike array- and network-based replication, it doesn't depend on additional hardware components. That makes host-based replication the least-expensive and easiest replication method to deploy.

"Deploying host-based replication only requires installing the replication software on source and target servers and you are ready to go,"

CHOOSING A DATA REPLICATION SOLUTION

1. Selection of a data replication method should start with a business impact analysis to determine required recovery time objectives (RTOs) and recovery point objectives (RPOs).

2. For applications that can't accept data loss (RTO equals zero), synchronous replication is required. Heed latency in synchronous replication because it will drag down application I/O performance. If there is any risk of latency or unreliable bandwidth, or for replication beyond certain distances (50 km to 300 km), asynchronous replication is the way to go.

3. Besides the replication mode, application performance can be impacted by the replication platform. Host-based replication competes with applications for valuable processor, memory and I/O resources.

4. Have a clear understanding of the bandwidth requirements, impact on bandwidth cost, and how data replication will impact other applications and users. Clearly understand and take advantage of replication features related to bandwidth such as compression, bandwidth throttling and configurable bandwidth usage depending on the time of day. Consider wide-area network (WAN) optimization devices to preserve bandwidth.

5. Replication products that support heterogeneous environments can substantially reduce cost by supporting less-expensive or legacy arrays. They also limit vendor lock-in.

6. The disadvantage of vendor lock-in of array-based replication is offset by the advantage of close integration between replication and the storage platform and easier support, eliminating the risk of finger-pointing in multi-vendor configurations.

SearchStorage.com STORAGE

noted Bob Roudebush, director of solutions engineering at Double-Take Software Inc. It's well suited to work in heterogeneous environments, supporting the widest range of storage options that include both network- and direct-attached storage. While most products support Windows, Linux and Unix support is more tenuous and, therefore, platform support is clearly one of the critical evaluation criteria when selecting a host-based replication product.

On the downside, host-based replication adds processing overhead to servers and the installed replication software carries the risk of introduc-

ing unknown behavior. "For critical and high-end application servers, IT managers tend to favor array-based replication over host-based replication because it keeps server resources dedicated to the app and doesn't expose it to potential bugs or flaws in the replication software," said Lauren Whitehouse, an analyst at Milford, Mass.-based Enterprise Strategy Group. Furthermore, licensing costs and system administration duties increase proportionally with the number of servers, giving both array- and network-based replication an advantage in environments with a large number of servers. In addition, visibility in host-based replication is typically limited to source and target servers. This is very different from the centralized architectures of array- and network-based replication offerings that enable a more holistic view into the replication infrastructure.

"Deploying hostbased replication only requires installing the replication software on source and target servers and you are ready to go."

> –Bob Roudebush, director of solutions engineering, Double-Take Software Inc.

The target markets for host-based replication products are typically small- to medium-sized businesses (SMBs) that can't afford more expensive replication alternatives, enabling them to deploy data protection and disaster recovery architectures that, until a few years ago, were only seen in larger firms. CA, Double-Take, InMage Systems Inc., Neverfail Inc. and SteelEye Technology Inc. are some of the vendors that have enabled smaller companies to deploy replication-based DR and data protection at a fraction of the cost of array- and network-based replication. Although each of these products replicates data from one location to another, they differ in features such as efficiency, bandwidth throttling, management, high-availability failover capabilities, platform support and application integration. Only a thorough product evaluation will reveal which product offers the best fit for a given environment.

In addition to these standalone offerings, backup software vendors are integrating host-based replication into their backup suites with the hope of expanding their reach into the lucrative remote- and branchoffice data protection business.



"We see a convergence of DR and data protection, and consider replication to be a feature and not a standalone product," said Marty Ward, senior director, product marketing for the Data Protection Group at Symantec Corp. Most backup software vendors are already offering hostbased data replication options for their backup suites; some examples include BakBone Software Inc.'s NetVault: Real-Time Data Protector; CommVault Continuous Data Replicator (CDR); EMC RepliStor to complement EMC NetWorker; Symantec Backup Exec Continuous Protection Server (CPS); and Symantec NetBackup PureDisk with a deduplication option, as both a standalone product and a NetBackup option.

The main advantage of combining traditional backups and replication is the ability to manage replicas and back-

ups within a single tool. Aside from their host-based replication options, backup software vendors have been working on integrating their backup suites with leading storage arrays and network-based replication products to enable customers to manage all replicas and backups with the same tool.

"Just like with Continuous Data Replicator, array-based replicas of supported arrays are integrated into the backup application index and catalog, allowing users to restore an array-based snapshot by simply right-clicking it within our application," said Brian Brockway, vice president The main advantage of combining traditional backups and replication is the ability to manage replicas and backups within a single tool.

of product management at CommVault. Similarly, Symantec's Veritas NetBackup is integrated with more than 40 arrays and virtual tape libraries (VTLs), and EMC NetWorker offers tight integration for EMC's RecoverPoint network-based replication product.

NETWORK-BASED REPLICATION

In network-based replication, the replication occurs in the network between storage arrays and servers. I/Os are split in an inline appliance or in a Fibre Channel (FC) fabric; the I/O splitter looks at the destination address of an incoming write I/O and, if it's part of a replication volume, forwards a copy of the I/O to the replication target. Network-based replication combines the benefits of array-based and host-based replication. By offloading replication from servers and arrays, it can work across a large number of server platforms and storage arrays, making it ideal for highly heterogeneous environments. Most network-based replication products also offer storage virtualization as an option or as part of the core product.

Contemporary network-based replication offerings are either inline appliances or fabric based. With inline appliances, all I/Os need to pass through the replication device. Technically, the appliances terminate all

Cloud storage



SearchStorage.com

incoming I/Os and initiate new I/Os that are forwarded to the primary and, in case of write I/Os, to replicated storage targets. The inline approach has been plagued by performance and scalability issues. The poster child for inline appliances is IBM's SAN Volume Controller (SVC). A scalable architecture and plenty of cache have not only enabled it to overcome performance and scalability limitations but, aided by the simplicity of the inline appliance approach compared to the more complex fabric-based implementations, it has become one of the successes in the network-based replication and virtualization market.

In fabric-based replication products, the splitting and forwarding of I/Os is performed within an FC fabric. By taking advantage of FC switching and the separating data and control path, it's the best performing and most scalable approach. The majority of fabric-based replication products run on intelligent switches from Brocade Communications Systems Inc. and Cisco Systems Inc. Even though both Brocade and Cisco offer Data Mobility Manager (DMM) for local data center replication, third-party vendors like EMC and FalconStor Software Inc. offer more advanced fabricbased replication products that run on Brocade and Cisco intelligent switches. A case in point is EMC RecoverPoint, which provides fabricbased, asynchronous continuous data protection (CDP) with application integration that's on par with commensurate host-based CDP products. Despite obvious benefits, fabric-based replication has seen lackluster adoption. "Switch-based replication and virtualization have been overhyped, but there are people who are working on it and over time it will become more common," said Greg Schulz, founder and senior analyst at Stillwater, Minn.-based StorageIO Group.

LSI Corp.'s StoreAge Storage Virtualization Manager (SVM) straddles the line between inline appliances and fabric-based products that depend on expensive intelligent switches. The combination of SVM and LSI's Data Path Module, which plugs into existing Fibre Channel switches to perform switch-based forwarding and eliminates the need for intelligent switches, combines the simplicity of IBM SVC with the performance and scalability benefits of a split-path architecture. HP seems to concur, and is offering the LSI product as HP StorageWorks SAN Virtualization Services Platform (SVSP) to complement its host- and array-based replication offerings with a network-based replication and virtualization product.

Even though the market share for array-, host- and network-based replication will shift over time, there will be appropriate places for all three approaches. While each has its own set of advantages and short-comings, specific environments and situations will best determine where replication should occur. \odot

Jacob Gsoedl is a freelance writer and a corporate director for business systems. He can be reached at jgsoedl@yahoo.com.

Backup on a budget

Power-smart disk systems

Virtualizing servers, storage and even networks will change the face of IT and significantly impact the roles of storage professionals.

By Rich Friedman

IRTUALIZATION IS DRAMATICALLY changing how servers, storage and networks are configured and managed. But virtualization technologies aren't just changing storage environments; they're rapidly changing the nature and scope of storage jobs. Virtualization creates more interdependencies across different technology domains, which can instigate political turf wars over who should architect, divvy up and manage storage resources. Should it be the storage, network, server or VMware administrator?

"With server virtualization, storage administrators are, in some ways, ceding control of large blocks of storage to be allocated and managed by those responsible for VMware," said James Damoulakis, chief technology officer at GlassHouse Technologies Inc., Framingham, Mass.



FCOE COULD TIP BALANCE

Traditionally, Fibre Channel (FC)-based storage-area networks (SANs) have been, by and large, under the autonomous control of storage administrators. But Fibre Channel over Ethernet (FCoE) could disrupt that arrangement. As FCoE begins to take hold over the next two years to four years, even more questions about roles and responsibilities will arise. It's possible that the storage staff could be squeezed out, with network administrators finally taking control of the storage network.

"I'm betting that it will happen over time, but it will probably be a long time," GlassHouse's Damoulakis said. "With the continued use of Fibre

Channel, even over Ethernet, the specialized knowledge relating to zoning and LUN [logical unit number] masking will probably be left either to the storage admin or to a liaison role that straddles both storage and networking."

Today, the storage administrator dictates the type of network supporting the storage environment, said Bob Laliberte, an analyst at Enterprise Strategy Group in Milford, Mass. However, "it's easy to imagine that changing," he added. "If FCoE takes off, will the deployment of Ethernet switches be controlled by storage or networking companies?" And who controls the budget—the storage team or the networking group? The most likely scenario, he said, will be the elimination of siloed IT groups replaced "With the continued use of Fibre Channel, even over Ethernet, the specialized knowledge relating to zoning and LUN [logical unit number] masking will probably be left either to the storage admin or to a liaison role that straddles both storage and networking."

> –James Damoulakis, chief technology officer, GlassHouse Technologies Inc.

by hybrid IT groups, on either a project or permanent basis, staffed by members with expertise in networking, storage, servers and virtualization.

DIFFERENT PERSPECTIVES

There's no question that people with different skill sets within IT view the problems and their solutions differently. "Network engineers have a different point of view for I/O [than storage admins]," said Tom Becchetti, senior storage and Unix engineer at St. Jude Medical Inc. in St. Paul, Minn. "The only time the network folks are concerned about latency is when there is a distance involved. Most of the time for local traffic they are only concerned about bandwidth. Merging networking and storage doesn't work."

For example, "before I arrived here, they [the network and storage

groups] decided to use a BREAKING DOWN IT SILOS Contoural Inc., advises companies to do

the following:

• RECOMBINE stovepipe IT infrastructure organizations (server, storage and network) into a single management organization with specialists in these areas and others, such as virtualization and cloud computing.

SearchStorage.com

• SPEND the time and money to crosstrain everyone so they can reapply their experience and skills in this new world. Storage folks, for example, must know a good bit about server virtualization or their skills will have much less value.

• BRING the mainframe, security and records management folks to the party. They all have essential insights, and failing to give them a seat at the table would be a critical loss.

the data practice group at Mountain View, Calif.-based Contoural Inc., "network and storage pros have critical insight into their areas," and continuing to isolate them in specialized silos "would be a tragic loss akin to what happened when open-systems folks decided not to pay attention to the lessons of the mainframe generation."

Foskett added that new technology is changing the status quo. "Right after the data center is virtualized and [the] IT infrastructure is recombined, applications themselves will fundamentally transform, demanding a merger of the current IT infrastructure and IT applications groups," he said. "This could all come within five years, or it could be delayed or diverted by organizational infighting and intransigence."

DBA JOBS ARE CHANGING, TOO

dedicated NetApp filer with

[Microsoft] Exchange," Bec-

chetti said. "The network con-

nection, with a separate VLAN

entire network infrastructure. The disk response time was

so bad, it caused outages for

Exchange. In many shops, a

network I/O isn't monitored

story short, they converted the attachment to Fibre

Channel and the problem

was solved. The network

was fine."

as closely as a disk I/O. Long

group stated that the network

Becchetti's position is very storage-centric. According to Stephen Foskett, director of

[virtual LAN], was part of the

an iSCSI attachment for

The roles of database administrators (DBAs) and their working relationship with storage administrators are also changing. Mike Shapiro, distinguished engineer at Sun Microsystems Inc., agrees: "I'd like to see the conversation between DBAs and a typical storage group get beyond 'I need a LUN of size *x* gigabytes,' which is how it works today."

Traditionally, DBAs and storage administrators are often in conflict regarding storage provisioning and how to resolve storage performance/cost issues. GlassHouse Technologies' Damoulakis views this as a "natural tension and, as long as it is managed properly through open communications and an effective demand planning process, it's not a

HOW MANY TERABYTES SHOULD A STORAGE ADMINISTRATOR MANAGE?

Jusт ноw мucн storage should a single administrator be responsible for? It depends, especially in an interdependent IT environment.

James Damoulakis, chief technology officer at GlassHouse Technologies Inc., Framingham, Mass., said this isn't a particularly meaningful metric. Of course, as disk capacity grows, the number continues to change. More importantly, the biggest factors relating to TB/admin are the complexity and dynamism of the environment. Very complex environments—those with lots of different kinds of devices, complex architectures and so forth—by their nature require more people to manage them. Likewise, a fast-changing environment with many provisioning requests and high volumes of data movement also increases staffing needs.

"These are the areas we always look at first in terms of making staffing efficiency recommendations," Damoulakis said.

huge problem." However, he added, "the interesting challenge will be, given changes like virtualization, thin provisioning and solid-state storage, can the storage group build a strong enough argument to alter the tradition-based, risk-averse DBA mindset?"

Greg Schulz, founder and senior analyst at StorageIO Group in Stillwater, Minn., said that the more DBAs learn about storage, servers and networking, the better they can convey what they need and why; likewise, the more server and storage professionals can learn about apps and the needs of DBAs, "the better they can work together to drive efficiency and boost productivity instead stepping on or inhibiting production."

Simply put, uncertainty and change, if managed appropriately, will lead to new ways of doing things and an increase in productivity. As John Seely Brown, previously chief scientist at Xerox Corp. and now a visiting scholar and advisor to the Provost at the University of Southern California's Annenberg School for Communication, wrote in a recent published interview: "Many breakthroughs today come between disciplines, where multiple disciplines work together." **O**

Rich Friedman was formerly senior editor at *Storage* magazine. He's currently riding his bicycle from St. Augustine, Fla., to San Diego.

What's it like to feel 100% confident that your backup and restores will actually



Visit the SearchStorage.com "Advanced Guide to Backup" today: www.SearchStorage.com/backup_guide



The Web's best storage-specific information resource for enterprise IT professionals

TechTarget Storage Media



STORAGE



Power-smart disk systems

Disk drive systems use more power than just about any other data center gear, but storage vendors are addressing this problem with a variety of technologies.

By Matt Perkins



AST FALL, Adaptec Inc. kicked off its Green Power Initiative by introducing a new capability for many of its RAID controllers called Intelligent Power Management. The product was made available with the company's Series

5 and Series 2 Unified Serial RAID controllers, and was designed to help configure storage systems to reduce power consumption by up to 70% without forfeiting performance.

"What we decided to do was to address the whole storage package," said Suresh Panikar, director of worldwide marketing at Adaptec. "Our initiative said that from now on our products will have a flavor of power management." Storage administrators can configure banks of drives to operate at different power levels at different times, and set the drives to spin up and down according to usage patterns.



Adaptec is one of several storage providers eyeing green initiatives in the data center. Vendors such as Enhance Technology Inc., greenBytes Inc. and LSI Corp. are also trying to make their marks on the green storage map, as have larger vendors like Dell Inc., EMC Corp. and NetApp. Most vendor initiatives so far involve adding energy-efficient enhancements rather than reengineering products.

"We're making some refinements to our product lines in light of the need for reduced power, cooling, space and the whole greening of the data center," said Larry Freeman, NetApp's senior marketing manager of storage efficiency. "Thin provisioning is key in power consumption." In addition to thin provisioning, NetApp software features like data deduplication and snapshots are designed to save capacity in VMware environments. In October 2008, the company said its multiprotocol arrays were 50% more space-efficient than other systems supporting VMware. NetApp even pledged to make up the difference for free to users who didn't see an improvement.

NetApp has also embraced solid-state drives (SSDs) with its Performance Acceleration Module (PAM), a solid-state secondary cache that Freeman said will eventually become a flash implementation. "We see the benefit of using secondary memory to enhance performance without adding more disk drives," he said.

SLOW, SLOWER AND SLOWEST

THE IDEA OF slowing down or stopping disk drives when they're not in use was popularized by Copan Systems with its Revolution line of disk-based backup systems that use massive array of idle disks (MAID) technology.

In Copan's arrays, disks are shut down completely when they're not needed; and because only a handful of disks may be active at one time, you can cram many more disks into a single array. Running all of the densely packed disks would suck power and generate excessive heat, but MAID significantly reduces power and cooling requirements.

Other vendors use similar approaches, but some don't shut down disks completely. By reducing power to disks or parking their read/write heads, power consumption is reduced, but the disks can spin up to operational speed relatively quickly when needed. For example, EMC Corp. offers policybased drive spin-down in its Clariion CX4 arrays and Disk Library backup systems that puts inactive drives in sleep mode.

Nexsan Technologies Inc.'s AutoMAID, available with its SASBeast and SASBoy arrays, also uses policies to spin drives down in three stages—each stage offers greater power savings but also requires more time for the drive to get up to speed again.

Copan, EMC and Nexsan are just a few examples of disk system vendors that slow or stop disks to save power; many array vendors offer similar capabilities.



Other vendors are also focusing on ways to save power. GreenBytes' Cypress NAS appliance for archiving long-term data aims to cut energy consumption up to 80%, while Enhance Technology's products feature virtualization and auto spin-down RAID controllers. And last May, EMC Corp. unveiled virtual tape libraries (VTLs) featuring data dedupe. Moreover, EMC's Clariion CX4 storage arrays, which Dell also sells, released with built-in spin-down capabilities.

SEVEN WAYS TO CUT DISK SYSTEM POWER

As one of the few mechanical devices in the data center, disk arrays draw substantial power to keep disks spinning and make cooling systems work harder to dissipate the heat they throw off. Here are seven ways to cut a disk system's power requirements:

- 1. **THIN PROVISIONING.** Because thin provisioning ensures that applications only consume the capacity they truly need, you don't end up with a lot of spinning, nearly empty disks.
- 2. SPIN DOWN. In arrays used for secondary applications or those that use disk intermittently, drives can be slowed down or stopped when not in use.
- **3. SOLID-STATE DRIVES.** Solid-state drives (SSDs) use a fraction of the power required by magnetic-media disks and produce far less heat, but the power savings may be offset by the cost of SSDs.
- **4. BIG DISKS.** By using higher capacity disks (especially 1 TB, 1.5 TB or 2 TB drives) you can store more data using fewer disks.
- **5. SMALL FORM FACTOR DISKS.** For apps that require more performance than what SATA delivers, 2.5-inch SAS drives use up to 50% less power than 3.5-inch high-performance drives.
- **6. DATA DEDUPLICATION.** Store less and you'll need fewer disks and thus save on power. Dedupe is one of the most effective ways to cut capacity.
- **7. VIRTUALIZATION.** Like thin provisioning, storage virtualization helps you make better use of the storage you already have installed. So while you may not cut your power bill, you can keep it from rising.

SearchStorage.com

As for Dell, most of its energy-efficient focus has been on utilization, tiering, spin down and solid-state drives.

"If you had some set of data that requires very high performance but not a lot of data, solid-state drives may be very economical in power consumption," said Eric Schott, director of product management at Dell. "It's also about floor space consumed and heat output, because for every dollar you're spending to power the equipment, you're spending a dollar to cool the equipment."

But the environmental benefits aren't the only things on vendors' minds when pushing energy-efficient products. Economic benefits are deemed equally as important.

"The messaging is around 'Go green to save carbon footprint," said Greg Schulz, founder and senior analyst at Stillwater, Minn.-based StorageIO Group and author of *The Green and Virtual Data Center*. "Most IT organizations don't have a carbon footprint issue. What they have are issues to sustain business growth, economic enhancement and productivity."

On the management side, the notion of intelligent management—like that of Adaptec—is something Schulz said can be expected to continue.

"We've seen many vendors adding intelligent management as a second feature, just like a RAID level," Schulz said. "You're seeing vendors bring that to the table, but in a granular approach: Use it when needed."

Adaptec's software allows drives to be operated in three power states, including normal operation with full power and full RPMs; standby, with low power that spins at a lower RPM during idle times; and power-off.

Dell's Schott said intelligent management can bring forth several benefits that otherwise aren't there. "It's very important for a storage array to be energy efficient," he said. "But it's also helpful if the applications could also clue us in more. It's one thing if the storage arrays are doing it in the background, it's another if the application is telling you."

While some vendors aren't sure to what extent the demand for energy efficiency will impact their storage products, most are sure that products aimed at conserving power will continue to gain momentum.

"Some of this greenness is much more of a philosophy over the work that we do, rather than just a product feature or release feature," Schott said. "And you can always do better. Even when you reduce and improve a few things, you're not done." Θ

Matt Perkins is an assistant editor for TechTarget's Enterprise Applications Media Group.

Cloud storage

Value-based data protection

There's no skimping when it comes to data protection, but tight budgets mean that storage managers will be looking for solid value in their backup products.

By Rachel Kanner

S THE ECONOMY SLOWS, efficient data protection has become an even greater concern for companies. But as storage managers focus on doing more with less, finding the right data protection solution has never been more challenging.

Marty Ward, senior director of product marketing, Data Protection Group at Symantec Corp., said finding appropriate data protection has become more



urgent for two reasons: a general focus on saving money and companies no longer have the budget to hire more people.

In the past few years, storage vendors have expanded their data protection portfolios. For example, Symantec built Veritas NetBackup to bundle features, including data backup, replication, virtual tape libraries (VTLs) and space-saving snapshots, into one data protection product. "We created interfaces into all of these different data protection technologies, whether they come from us or others, so you can have one management platform

Backup on a budget



Users who will add new options or functions to their current backup apps* to manage it all," Symantec's Ward said. According to Peter Eicher, manager of product marketing at FalconStor Software Inc., users are interested in the company's new MicroScan data replication technology, which Eicher said uses a different approach than other data replication tools. MicroScan

SearchStorage.com

scans important data and identifies the new data to replicate, rather than scanning it all, which frees up more storage space.

After facing disasters such as water damage, and ready for an upgrade, John Michaels, chief technology officer at Maxim Group LLC in New York City, chose MicroScan with block-level, delta-based replication because of its cost effectiveness. "There's a whole other value in the system that I didn't realize when I bought it," he said. "I was only able to back up seven

servers because of monetary reasons, but what I realized was that I had some pretty powerful servers that were being underutilized. Now I'm protecting over 20 virtual servers."

Technologies like virtualization provide specific challenges. With the quick adoption of server virtualization over the past few years, it has become more common to put mission-critical applications on VMware; however, this requires a different set of data

protection solutions. For Michaels, the FalconStor product helped him to protect both virtual and physical servers. Users are also finding niche solutions and deciding how they can manage their virtual environment differently than their physical environment. That's where FalconStor also comes in. "Their single data protection solution will manage all of your virtual and physical environments," Michaels said.

When it comes to buying data protection software, the same people are still buying, but now they're trying to buy smarter. Eric Burgener, senior analyst and consultant at Hopkinton, Mass.-based Taneja Group, offered some advice for those trying to save money. "It's surprising how much data sits in storage without being accessed again, but continues to be backed up every time data is backed up," he said. "If you



can identify the 70% to 80% of information that needs to be backed up and move it onto a less-expensive but slower storage device, reductions will be huge. Secondary platforms don't perform as fast, but [they] are designed to store lots of data very cost effectively."

Rachel Kanner is a former assistant editor at SearchStorage.com.



Special Report: Online Backup

Have you looked at

Backup lately?

Visit the SearchDataBackup.com "Online Backup Special Report" to find out how online backup has evolved: www.SearchDataBackup.com/online_backup

Sponsored by:

IRON MOUNTAIN



The Web's best storage-specific information resource for enterprise IT professionals













Is cloud storage the return of the service provider?

Cloud storage is a service, just like what storage service providers attempted to do years ago. But the technology has significantly evolved in the last decade and this time, cloud storage is real.

HE TERM CLOUD STORAGE has been thrown around by so many vendors and industry pundits that it has just about lost any meaning. One vendor says cloud storage and means a service, another says it and means software infrastructure, while yet another vendor means hardware infrastructure. When asked who the cloud storage vendors are, we could name just about anyone in storage because just about every vendor has a cloud strategy and is providing at least one piece of a solution. At the end of the day, however, cloud storage is a service, just like what storage service providers (SSPs) in the earlier part of the decade attempted to do. While the SSP model didn't work in 2002, the technology has significantly evolved; this time, cloud storage is real.

There are many market drivers for cloud storage, and they are pretty much the same ones that existed in 2002. Just like death and taxes, data growth is a sure thing, even with the economy falling off a cliff. And the economics work. In a 2008 research survey of 516 IT executives at midsized companies, 30% of those surveyed cited a lack of physical space in the data center as a top challenge. In another survey of 504 large enterprise storage buyers, 28% of respondents cited running out of power and cooling capacity as a major challenge. Buying capacity as a service is a lot less expensive than building a new data center to accommodate space, power and cooling demands. And labor costs are reduced when storage administration tasks are outsourced.

If all of these factors were in play in 2002, why didn't the SSP market thrive? The dot.com bust was only a small part of the equation. Storage as a service (SaaS) was—and continues to be—a good idea; it was just a little before its time.

The following factors were inhibitors to storage service provider growth:

BANDWIDTH COST. Bandwidth cost and availability was a major market inhibitor. A T1 line is only 1.5 Mbps and, in many cases, users needed much more. Many were looking for Fibre Channel (FC) connectivity over optical networks, usually two FC connections and a Gigabit Ethernet



(GbE) one. That amounts to somewhere around 2.5 Gbps or 3 Gbps, which translates to an OC-48 connection. Monthly network fees for the bandwidths and distances required to support the model were exorbitant. Pure availability of network bandwidth was an issue. In some locations, you couldn't get connectivity or the last mile would cost a small fortune.

USING THE WRONG STORAGE PLATFORMS. The original SSP model was to take the massive arrays offered by storage vendors like EMC Corp. and Hitachi Data Systems and leverage them as consolidation platforms. The data was absolutely safe in big iron arrays and some economy of scale could be realized through multitenancy—hosting lots of different customers on a single array—but the big iron platforms from EMC and Hitachi weren't designed to support these environments. No matter how many tenants were housed in a single array, the break-even economies of scale didn't work.

TARGETING THE WRONG APPLICATIONS. Rather than going after less-used persistent data or remote storage as an archive tier, SSPs focused on offsite primary storage for any and all applications. They ignored latency issues associated with supporting I/O-intensive applications remotely. The solution was to put storage PoPs everywhere, which was an extremely expensive proposition.

Today, the Internet reaches every corner of the world, effectively creating a flat global network with few, if any, barriers to connectivity. The combination of wide-area network (WAN) acceleration and ubiquitous network connectivity allows business to be conducted anywhere. On the platform front, scale-out, commodity-based platforms that provide massive scalability, parallel data transfers and economies of scale not just for hardware, but for ease of use and management, are available. And, today, the application profiles that can withstand latency associated with storing data remotely are better understood. Now cloud storage can be part of a storage tiering model for persistent data.

The consumer and small office/home office (SoHo) markets, along with Web 2.0 businesses, will continue to be early adopters. Larger enterprises will proceed with caution, as there is a high degree of risk aversion in this segment. It's more likely that large enterprises will deploy purpose-built private clouds for bulk storage of persistent data and for archive—the move to disk-based archive clouds within the four walls of IT has been in process for some time. Eventually, even large enterprises will look to the cloud as a storage utility and an integrated part of a storage tiering model. Cloud storage is still in its early growth stage, and it will take a long time to evolve to become core to enterprise IT. But this time, the technology has come far enough to make the dream of a storage utility a reality. ⊙

Terri McClure is a storage analyst at the Enterprise Strategy Group, Milford, Mass.

The buzz about dedupe

If the storage industry had a word of the year award, it would go to "dedupe." Like with so many other IT technologies, the task is to determine if dedupe's hype or happening. For data deduplication, it seems like it might be a little of both. In our latest survey, we found that 31% of respondents currently use a deduplication product—not bad, considering that it's still a new technology. But how easy will it be to convert the remaining 69% of non-believers? Forty-four percent are dedupedeniers, who cite diverse reasons for their non-implementation: 30% don't use disk for backup and 29% just don't see a need. In addition, approximately one-third of the comments we received cite the cost of deduplication as a prohibitive factor. As for the great where-to-put-it debate, 50% of our dedupers chose products that do inline deduplication, 35% opted for post-processing systems and 10% didn't know what they had. What the heck, if it works who cares? —*Rich Castagna*



"Due to the economy and such, deduplication has had to take a back seat to more pressing matters. Once money is available, we will seriously investigate deduplication technologies." —Survey respondent

Check out the following resources from our sponsors:

BakBone Software Inc., page 6

BakBone Software

Protecting Microsoft Exchange Servers with NetVault: Backup

Dell Inc., page 3

Solution Brief: Auto-Snapshot Manager/Microsoft Edition For Hyper-V On Demand Webcast: vStorage enabled SANs; Emerging Requirement for the Virtual Data Center

State of replication