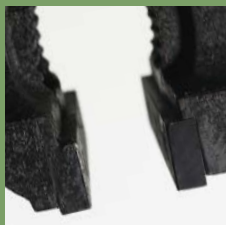


STORAGE

ESSENTIAL GUIDE TO

Efficient Storage Management



Use your storage resources more efficiently and improve data management through data reduction, tiering and capacity management.

INSIDE

- Efficient storage no longer optional
- Plan for primary dedupe
- Manage less data, and do it better
- Dynamic storage tiers
- Plan for capacity growth



Efficient storage is no longer optional

Getting the most out of installed storage systems is job No. 1 these days, and there are plenty of ways to make your storage more efficient.

TRADITIONALLY, storage array disk utilization rates at most companies hovered in the 30% to 50% range. That might have been fine when there was still plenty of floor space in the data center and data storage managers were flush with ample budgets. But the space is going—or gone—and the days of big budgets seem like distant memories while the amount of data we need to store is growing at an unprecedented pace. The old ways just won't work anymore.

If necessity is the mother of invention, she's been pretty busy lately coaxing along a number of key technologies that promise more efficient use of storage resources.

One of the first attempts at getting the most out of installed disk capacity was thin provisioning, with 3PAR credited as its pioneer and chief promoter for open storage systems. Although some storage vendors might have been reluctant to offer a feature that could put off purchases, most now support thin provisioning in one form or another. Thin provisioning may not help you avoid buying additional capacity, but it can push those purchases down the road to when prices have fallen further and newer technologies are available.

Data reduction in primary storage is another important storage efficiency innovation. Data reduction is actually an umbrella term that can include data deduplication, single instancing and compression. Riding the wave of its success in significantly paring down disk-based backup data, data reduction is beginning to show up in primary and near-line storage. Although the reduction ratios for primary storage aren't likely to approach the lofty numbers racked up by backup dedupe, there's still a strong appeal—and potentially huge savings—in trimming current data down to size.

Back when money grew on trees, storage managers may have adopted some less-than-parsimonious storage management habits, like letting not-

Data reduction is actually an umbrella term that can include data deduplication, single instancing and compression.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

so-critical data stay on expensive primary disk. You knew it didn't belong there, but who had the time to determine what belonged where? With a deluge of new data pouring in daily, the question of where information should reside has taken on a new urgency. Automated tiering systems attempt to address the issue with as little manual intervention as possible, helping to neutralize the "who has time" argument. In a handful of systems today, but spreading rapidly, automated tiering systems move data around dynamically based on the application it's attached to and its frequency of use.

In this Essential Guide to Efficient Storage Management, you'll find detailed information on these and other technologies that are designed to make better use of your data storage and help you become a better storage manager. 🕒

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

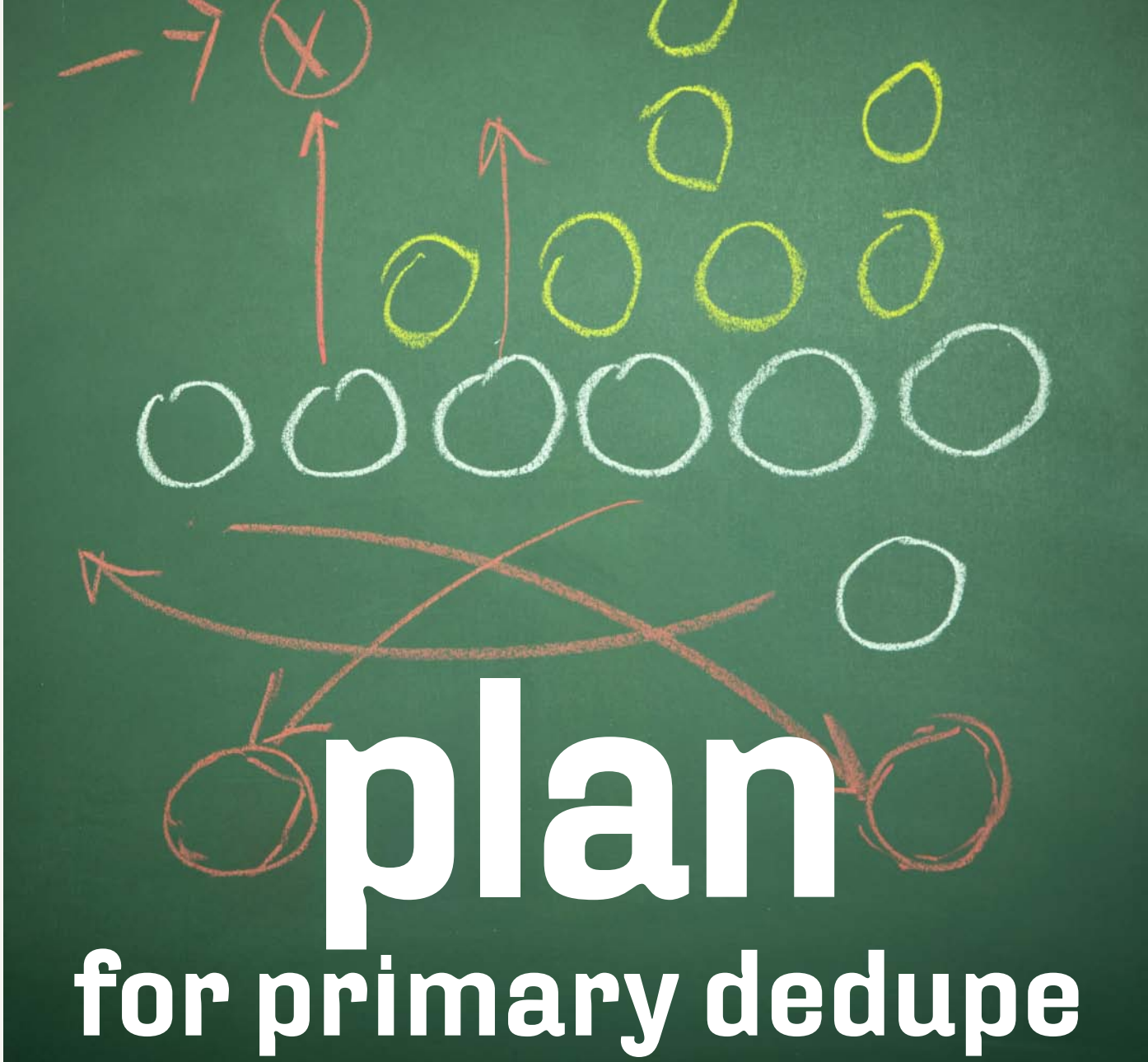
Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth



plan for primary dedupe

With demonstrated results for backup, data reduction techniques are now being aimed at primary storage. It's a hot market, but there are still plenty of wrinkles to iron out. By JEFF BYRNE

MERGERS AND ACQUISITIONS are alive and well in the data storage industry, particularly when it comes to storage capacity optimization (SCO). SCO technologies, such as data deduplication and compression, increase the utilization of primary, secondary and/or archived storage by shrinking the amount of stored data.

Despite being a relatively new market, there was a surge of SCO consolidation in July 2010 when Dell Inc. scooped up compression and dedupe vendor Ocarina Networks Inc. and IBM picked up real-time compression

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

supplier Storwize Inc. just 10 days apart. And don't forget EMC Corp.'s acquisition of Data Domain Inc. in 2009.

DEDUPE SLIMS BACKUPS

SCO technologies have traditionally had their greatest impact on backup storage, with data deduplication playing the leading role. Dedupe vendors can often reduce needed backup capacity by 90% to 95% or, put another way, increase effective backup capacity by 10 times to 20 times. But while traditional deduplication can pay big dividends in shrinking backup capacity requirements, it hasn't been as effective in primary storage environments. So for primary storage, the focus has shifted somewhat to compression-based technologies.

OPTIMIZING PRIMARY STORAGE WITH COMPRESSION

Primary storage has particular characteristics that make it difficult to shrink. Unlike backup storage, it doesn't consist of a lot of nearly identical data. And many primary storage environments are performance-sensitive and can't be slowed down by optimization processes.

Storwize's technology compresses file-based data in-line with little or no impact on application performance. The Storwize appliance sits between a NAS array (NFS or CIFS) and users of the data, and typically reduces the stored data footprint by 50% to 90%.

Storwize only works with file storage, and its compression algorithms aren't optimized for specific file formats. But what it does, it does well. For IBM, Storwize will work with its N series and SONAS NAS systems; it will also work with non-IBM NAS systems from EMC, Hewlett-Packard (HP) Co., NetApp and others.

In contrast, Ocarina Networks compresses data using an out-of-band, post-process approach, reading and compressing stored data and then writing the smaller files back to storage. Ocarina's technology is content-aware so its optimization is tailored to the particular type of content. Overall, the capacity savings afforded by Ocarina are on a

Storwize's technology compresses file-based data in-line with little or no impact on application performance.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

scale similar to those delivered by Storwize.

Dell has offered (or shown interest in) several competing storage capacity optimization technologies and is now integrating Ocarina's technology throughout its storage portfolio. Before the acquisition, Ocarina introduced a software-based deduplication product that can be embedded in other vendors' storage arrays, but Dell is likely to discontinue that OEM business. That would leave Permabit Technology Corp. as the only vendor that will provide an embeddable SCO solution to other storage system vendors. Looking ahead, we believe Quantum Corp. might also OEM an embeddable SCO solution, combining its StorNext file system with its dedupe capabilities.

THE HOLY GRAIL: END-TO-END OPTIMIZATION

So what's driving this wave of storage capacity optimization market consolidation? All the leading data storage companies are in need of effective SCO solutions, and they're scurrying to acquire technology from the rapidly dwindling ranks of independent SCO vendors. Typically, vendors accumulate technologies to shore up or fill gaps in their SCO portfolios, and then try to piece the technologies together into a coherent whole. More often than not, the results are collections of poorly integrated, often incompatible point solutions.

Users are the biggest losers in this game. Consider this scenario. A storage manager selects a deduplication product for backup and successfully reduces backup storage by 90%. So far, so good; but what comes next? Maybe they now want to move inactive chunks of deduplicated backup

data to an archive where it can be searched, used for e-discovery and so forth. But that can't be done without rehydrating the deduped data back to its original file format, thus losing the benefits of the original deduplication effort.

Now suppose our storage manager decides to use a compression

All the leading data storage companies are in need of effective SCO solutions, and they're scurrying to acquire technology from the rapidly dwindling ranks of independent SCO vendors.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

app to optimize primary storage. Once again, there are some limitations. In some cases, compressed files will need to be rehydrated before they can be moved among different storage tiers. And the same could happen if you want to dedupe the compressed data during backup—you may have to rehydrate the data.

That might not seem like such a big deal, but rehydrating consumes networking and CPU resources, as well as the disk capacity to store the data once it's rehydrated. Rehydration may also result in a loss of information where data has been deduplicated or compressed using a “lossy” algorithm.

There's currently no easy way for users to “knit together” diverse SCO solutions to enable stored data, once optimized, to retain the benefits of that optimization throughout its lifecycle. Instead, you'll have to be content to choose the best optimization solution for each storage tier (backup, archive and primary) and put up with the inefficiencies caused by data crossing those boundaries. It's not likely to get better in the short term as vendors focus on building and differentiating their own proprietary stacks while interoperability standards fall by the wayside. All of that makes planning a capacity optimization strategy for the next three years to five years pretty tough.

We recommend you try to stick with a single vendor that's in the process of integrating its various storage capacity optimization technologies; it's probably the best opportunity to achieve something close to end-to-end optimization at some point. Several vendors are pursuing end-to-end strategies, but none is yet delivering on the promise.

Over the long term, we think your interests would be better served by standardized, cross-vendor solutions that enable interoperability among different SCO solutions. With a standard way to communicate among optimized systems, it might be possible to migrate or re-tier optimized data without the overhead of rehydration. In addition, vendors would be able to use their optimization technology outside of single storage boxes or cross-license their technology as part of a richer SCO ecosystem. But right now, standardization seems unlikely.

Still, we can be optimistic that the best days of the capacity optimization market lie ahead. Several major vendors we've spoken with say they understand the ultimate goal even as they're amassing pieces of the puzzle. ☺

Jeff Byrne is a senior analyst and consultant at Taneja Group. He can be reached at jeff.byrne@tanejagroup.com.

Make storage
more efficient

Plan for
primary dedupe

Manage less
data, better

Dynamic
storage tiers

Plan for
capacity
growth

Manage less data, and manage it better

Improve storage efficiency through data reduction techniques and capacity optimization using data deduplication, compression, snapshots and thin provisioning.

By CAROL SLIWA

DATA REDUCTION via simply deleting unwanted or unneeded data is the most effective way to reduce a company's data footprint and, in turn, its energy needs. The hard part is often getting end users to cooperate. When deletion isn't a viable option, IT shops can turn to technologies such as data deduplication, compression and thin provisioning to promote storage efficiency and possibly better power consumption.

"The real savings in energy efficiency comes in managing the data," said Christine Taylor, an analyst at Taneja Group in Hopkinton, Mass. "The less data you have to house, the less energy you're going to have to use."

Before turning to data reduction technologies, Taylor advises IT shops to delete as much data as possible and to set up automated data retention so they'll be able to safely eliminate or copy data to tape at the end of their pre-set retention periods. Classification mechanisms can also help to weed out unnecessary data.

If deletion hinges on user participation, IT shops might need to come up with some creative approaches or subtle forms of pressure. The Friedrich Miescher Institute (FMI) for Biomedical Research in Basel, Switzerland, for instance, decided to make the data archiving process a bit more difficult for its researchers.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

When FMI’s scientists had the ability to acquire data directly into the institute’s archive, they sometimes saved large quantities of unnecessary data, such as the results of failed experiments. Now they must first acquire the data locally and then move it to a “scratch area” to analyze it before they’re able to shift it to the archive.

“It’s work to get the data into the archive, so they only take what they absolutely need,” said Dean Flanders, head of informatics at FMI. “We really want to make sure they’re not just dumping things in there that they’ll never look at. We don’t want to store junk.”

"It's work to get the data into the archive, so they only take what they absolutely need. We really want to make sure they're not just dumping things in there that they'll never look at. We don't want to store junk."

—DEAN FLANDERS,
head of informatics, FMI

When the challenge of deletion proves too great, the following data reduction techniques can achieve data reduction and/or boost storage efficiency, which can have an impact on energy consumption:

DATA DEDUPLICATION. Deduplication technology shrinks the data footprint by removing redundant data at the file or sub-file level. It’s most common use is with backups and archives, but it’s increasingly gaining acceptance with primary storage.

The leading vendors in the backup/archive realm all offer data dedupe capabilities. NetApp Inc. is the top proponent in the primary storage space and claims more than 8,000 customers have licensed its free deduplication technology.

COMPRESSION. Data deduplication eliminates redundant data; compression reduces the size of every piece of data based on algorithms that have been around since the 1950s. Compression can be done standalone or in conjunction with data deduplication.

Leading storage vendors generally offer some form of compression. IBM promotes compression of primary storage since its acquisition of Storwize Inc. Dell Inc., with its Ocarina Networks Inc. acquisition, and EMC Corp. offer a combination of data deduplication and compression for primary storage.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

SNAPSHOTS. Snapshots are point-in-time copies of files, directories or data volumes that are especially helpful in the context of backups. Some systems save space by copying only the changes and using pointers to the original snapshot.

“A couple years ago, we used to make full-volume copies all the time to make backups more efficient. You quickly took a logical image of a backup, took a snapshot and used it to run your backups again, and that used to require the same amount of capacity as the volume that you were making a copy of,” said Brian Garrett, vice president of the Enterprise Strategy Group (ESG) Lab in Milford, Mass. “With snapshots, you can reduce that capacity required to make a logical copy of the volume, which can reduce the amount of capacity you need to deploy and the amount of spinning drives and, therefore, the energy you’re using.”

THIN PROVISIONING. IT shops can improve their storage utilization rates by using thin provisioning to allocate storage on demand, rather than allocating it up front, which often leads to overprovisioning and lots of underused or wasted disk space.

“In essence, what you’ve created” with thin provisioning, Garrett said, “is a pool of spinning drives that’s energy optimized to meet your needs.” ☉

Carol Sliwa is the Features Writer for *Storage* magazine and SearchStorage.com.

Make storage
more efficient

Plan for
primary dedupe

Manage less
data, better

Dynamic
storage tiers

Plan for
capacity
growth



Dynamic storage tiers

Find out why you should deploy dynamic storage tiering, learn about the various DST product categories and discover how to choose the DST product that's right for your environment.

By JACOB GSOEDL

THE LIST of dynamic storage tiering (DST) options is growing quickly, but product capabilities and purpose vary greatly. So it's important to identify your objectives, and the problems you want to solve before you begin an evaluation. Your objectives may include:

- Lowering overall data storage costs
- Overcoming storage performance problems
- Improving compliance
- Incorporating cloud storage into your storage infrastructure

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

To help keep you from going astray during the evaluation process, remember the central purpose of dynamic storage tiering: to lower overall storage costs by keeping data on the appropriate storage tier, with active data on fast storage and more static data on slower tiers. To put data in its appropriate place, DST products—or data movers—use an analysis facility that looks at data characteristics such as I/O, activity, access patterns, last accessed time and predefined policies, as well as a data migrator that relocates data based on the outcome of the analysis.

Today, there are three primary reasons to deploy dynamic storage tiering, as well as DST product categories that address each reason:

- Maximizing the use of expensive fast storage, especially solid-state drives (SSDs)
- Coping with the rapid growth of file-based data
- Incorporating cloud storage services

When looking for the dynamic storage tiering category that best aligns with your objective, you can choose from array-based, storage virtualization-based and file system-based DST products, as well as from dynamic storage tiering products from the cloud.

ARRAY-BASED DST PRODUCTS

With a few exceptions, such as Compellent Technologies Inc. (now part of Dell Inc.), array vendors did little to provide dynamic storage tiering capabilities within their arrays until recently. Users who needed dynamic storage tiering had little choice but to complement their arrays with third-party products such as F5 Networks Inc.'s ARX appliances and Symantec Corp.'s Storage Foundation software.

While EMC Corp.'s Fully Automated Storage Tiering (EMC FAST) for its Symmetrix, Clariion and Celerra arrays has gained a lot of attention, other array vendors have either released or are working on storage tiering mechanisms for their arrays.

Because storage tiering is so closely linked to storage systems, you should first determine to what extent your array vendor supports dynamic storage tiering. Without question, DST options and features of your existing array vendor (if available) are preferable because they're typically less expensive, usually more tightly integrated with the array and you'll only have to deal with a single vendor. On the downside, array-based DST products usually only work within the same array family or within arrays

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

from the same vendor. So, if you have a mixed storage environment and want to dynamically tier across heterogeneous storage devices, you need to look at vendor-agnostic DST approaches.

STORAGE VIRTUALIZATION-BASED DYNAMIC STORAGE TIERING PRODUCTS

Storage virtualization with dynamic storage tiering capabilities is one approach to achieving vendor-agnostic DST. Storage virtualization products are perfectly suited to take on storage tiering responsibilities. Sitting between servers and physical disks, and mapping virtualized logical blocks to physical blocks on disk spindles, it seems like a natural extension to add tiering decision logic to the mapping process—a notion also held by several storage virtualization vendors.

Sitting between servers and physical disks, and mapping virtualized logical blocks to physical blocks on disk spindles, it seems like a natural extension to add tiering decision logic to the mapping process—a notion also held by several storage virtualization vendors.

In late 2010, IBM's SAN Volume Controller (SVC) added Easy Tier sub-LUN automated tiering that automatically moves data at a more granular level between disk tiers and SSDs. FalconStor Software Inc.'s Network Storage Server (NSS) SAN Accelerator identifies hot spots on magnetic disks and caches them in 4 KB, 8 KB or 1 MB pages in solid-state drives. Hitachi Data Systems added automated sub-LUN tiering to its Virtual Storage Platform (VSP) arrays with Hitachi Dynamic Tiering in late 2010. And Hewlett-Packard (HP) Co.'s acquisition of 3PAR came shortly after 3PAR launched its Adaptive Optimization volume-level automated tiering.

FILE SYSTEM-BASED DST PRODUCTS

The other group of DST products that usually supports storage tiering across heterogeneous storage systems is the file system-based products. While array-based and storage virtualization-based products typically operate at a LUN or block level (products with sub-LUN support),

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

file-system-based DST products operate at the file-system-level and move files instead of LUNs or blocks among tiers.

If your problems are managing growing file stores and ensuring that you have enough suitable types of disks to house them, a file system-based dynamic storage tiering product might be your best solution. A key characteristic of file system-based dynamic storage tiering is a single namespace that spreads across heterogeneous file storage; this enables the transparent movement of files between different tiers of storage, while maintaining the file path regardless of the physical location of files and folders.

Products in this category include Avere Systems Inc.'s FXT Series of tiered NAS appliances, F5 Networks' ARX Series, Hewlett-Packard's StorageWorks X9000 scale-out NAS and Symantec's Storage Foundation. All of these products have a data analysis facility to determine which files are active and a data mover to relocate files to an appropriate tier according to predefined policies; with the exception of the HP StorageWorks X9000, they work across NAS devices and file systems from different vendors. Products in this category provide file virtualization and, as with all storage virtualization products, you need to clearly analyze their impact on scalability, reliability and redundancy.

If your problems are managing growing file stores and ensuring that you have enough suitable types of disks to house them, a file system-based dynamic storage tiering product might be your best solution.

DYNAMIC STORAGE TIERING PRODUCTS FOR CLOUD STORAGE

Cloud storage service providers like AT&T, Amazon Simple Storage Service (Amazon S3), Nirvanix Inc., Rackspace, Zetta Inc. and a rapidly growing number of others have been promoting a vision of infinite storage without the need to manage an on-premises storage infrastructure.

Startups Nasuni Corp., StorSimple Inc. and TwinStrata Inc. offer cloud storage gateways with dynamic storage tiering capabilities. The Nasuni Filer cloud gateway is a software-based NAS head that runs on virtualized infrastructure (VMware ESX) to provide a virtualized file system

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

with some data in cache (on premises) and the rest in the cloud, very similar to file system-based DST products. Like traditional NAS products, it provides snapshot capabilities.

The StorSimple Hybrid Storage Appliance for primary cloud storage is optimized for Microsoft applications. It performs primary data deduplication and compression, and automatically tiers data across solid-state and SATA drives. StorSimple's appliance consists of a mix of tiered hardware, including a DRAM-based write cache, Flash SSDs from Intel Corp. and SATA hard disk drives. It connects to hosts using Gigabit Ethernet iSCSI.

TwinStrata's CloudArray handles block data. CloudArray's cache gives users local storage and automatically replicates volumes to cloud storage for high availability. When there is a disruption or outage, data can be accessed on-site, off-site or in the cloud.

GRANULARITY

An important evaluation criterion for DST products is the granularity at which data is moved between tiers. While some products require moving an entire LUN from one storage tier to another, others support sub-LUN-level data movement.

The bigger the price-performance difference between tiers, the more important sub-LUN-level data movement support becomes. Most file system-based DST products move files between tiers and, as a result, their app is limited to file-based content, making them ill-suited for structured data stores. To efficiently support continuously changing large files, such as an Exchange database and SQL databases, you need to look at products that support sub-file-level data movement support, such as Symantec's Storage Foundation.

While some products require moving an entire LUN from one storage tier to another, others support sub-LUN-level data movement.

DEGREE OF AUTOMATION

Prior to the adoption of solid-state storage in arrays, manual migration of LUNs between storage tiers was the prevailing data movement

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

method. For storage tiering with SSD and storage clouds, however, fully automated storage tiering has become a must; being able to keep active data in the fastest tier, as close as possible to real-time, is pertinent for tiering with SSD and cloud storage.

It's also important to assess the tools provided with dynamic storage tiering products to set up policies, schedule data movements and analyze data. They're clearly more important for products that only support data movement at a LUN level; understanding the impact of moving a LUN from one storage tier to another prior to actually moving it becomes more relevant because many applications may be attached to it and affected while the data is moved.

As a rule of thumb, the more granular data movement is between tiers, the more important automation becomes; because small chunks of data are relocated at all times, manual intervention is no longer an option. Conversely, if data is moved at a very coarse level, such as at LUN or volume levels, impact analysis and controlled data migration options become more relevant. ☉

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

Make storage
more efficient

Plan for
primary dedupe

Manage less
data, better

Dynamic
storage tiers

Plan for
capacity
growth

CHANNEL SPIN:

Plan for capacity growth

Find out the steps to storage capacity planning in a virtualized server environment, including factoring in storage efficiency techniques and planning for growth.

By GEORGE CRUMP

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

S **TORAGE CAPACITY** planning is a difficult task for IT administrators, even when they're dealing with the nonvirtualized physical server world where every server has a single app and a single storage volume it connects to. The chore doesn't get any easier in the virtualized server environment, with multiple physical hosts connecting to the same volume and each of those hosts possibly having a dozen or more applications.

Complicating this is the typical overprovisioning of the physical server environment, which can easily make its way to the virtual world during

the migration that occurs when establishing the virtual environment.

So how can you help your customers plan for storage requirements in a new virtualized server environment?

1 INVENTORY THE SYSTEMS TO BE MIGRATED

The first step in storage capacity planning for a customer's server virtualization environment is to create an inventory of the systems that will be migrated to the virtualized server environment. This can be done by manually auditing each server to determine how much capacity each one is using or employing a third-party capacity monitoring tool that will automate much of that work for you. Many of these tools can be sold to the customer or, in the case of an audit, licensed on a per-use basis.

2 DETERMINE WHAT STORAGE EFFICIENCY TECHNIQUES WILL BE USED

Once the data is collected, the next step is to factor in any storage efficiency strategies the customer would like to apply. There are typically three data reduction techniques that can be successfully applied in a server virtualization environment, and customers are more receptive to these technologies during a virtualized server project. With the first two techniques—data storage compression and deduplication—actual results will vary, but virtualized server environments usually have a high degree of redundant data in the form of operating system data; space savings of 50% or more aren't uncommon.

The third type of data reduction technique is writable snapshots, sometimes called cloning. With a writable snapshot there's essentially a golden master of a server image. That server master is then snapshotted and the new snapshot volume is assigned to another server. Unlike typical snapshots, which are read-only, this type is read/write, which means its configuration can be customized for a particular virtual instance. The only storage capacity required is the capacity to store the changed segments, eliminating the need to potentially store terabytes of redundant data.

Another feature, thin provisioning, will also improve utilization rates by making sure the virtual machine (VM) only consumes the space it uses. One caveat for resellers: Not all migration tools will "migrate thin." Instead, they'll migrate a server image block for block and write data marked for deletion in those blocks as if it were valid data, break-

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

ing the thin provisioning model. Make sure the storage system your customer is using can detect these zeroed pages of data and not write them during migration.

CALCULATE THE IMPACT OF STORAGE EFFICIENCY TECHNIQUES

With the decisions around these four factors that affect storage capacity planning determined, you can intelligently plan how much capacity your customer will require. In the past, it was acceptable to add up the total (used and unused) capacity the customer had, add 50% to that number to accommodate growth and then propose the solution. But today's budgets won't allow for that approach, and the capacity influencers above can curve the actual need so far down that your delivered capacity will look silly if you use the old model.

When recommending capacity, you should first set expectations with the customer that while you will try to make a more accurate calculation than just doubling what they have, there are a significant number of variables, more than in the past.

In addition, you'll need to analyze the information you've captured about current storage capacity use. Look for existing data that will be redundant in the new environment. For example, if there are 50 Windows servers and your customer will use either cloning or deduplication, the capacity requirements of the operating systems in those servers shouldn't be included in the calculation 50 times, though more than once makes sense. Even within like servers, look for redundant application suites. For instance, are there redundant database servers or email applications? If compression will be applied, look for and measure data sets that have very compressible data in them. Databases are typically very compression-friendly, sometimes saving as much as 90% or more.

Finally, look for a migration tool that is thin-aware and a storage system that can detect blocks of storage that have been zeroed out. If these are in your solution, budgeting the actual capacity in use, less the amount to be saved from the storage efficiencies discussed above, should provide a fairly accurate capacity estimate.

Don't be surprised if the actual capacity of a new storage system is less than the capacity of the system it's replacing. Advise your customer of the technique being used to calculate the new capacity and make sure they have extra budget in their back pocket in case there's a miss; but if the guidelines above are followed, there shouldn't be.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth

4

PLAN FOR GROWTH

Beyond the initial capacity plan, it's important to plan for growth. Even with all the storage efficiency capabilities available today, virtualized server environments are infamous for growing incredibly fast. Make sure both you and your customer understand what a storage upgrade will look like and the ramifications of installing that upgrade. If upgrades of one system are particularly limiting, that's an opportunity to propose an alternative system that may be more upgradable.

Capacity planning is more complicated than ever. Not only are there multiple virtual servers contained in a single server, there are also more tools to control storage growth. You have to be careful to factor this into any proposed storage solution. While it's extra work, it gives you a chance to separate yourself from the competition by showing the customer that you're willing to take the time to do it right, which eventually leads to savings. ☺

George Crump is president of Storage Switzerland, an IT analyst firm focused on the storage and virtualization segments.

Make storage more efficient

Plan for primary dedupe

Manage less data, better

Dynamic storage tiers

Plan for capacity growth