



Data
Mobility
Group

datamobilitygroup.com

White Paper

May 2005

Dianne McAdam

The Varieties of Disk-to-disk Backup Solutions



The Varieties of Disk-to-disk Backup Solutions

The backup-and-recovery tapes are still spinning in data centers all over the world. But as backup windows continue to shrink, due to the prevalence of online business functions, many businesses are looking for ways to add disk to their backup infrastructures to speed up the backup, and more importantly, the restore processes.

Responding to this demand, vendors have brought to market a variety of disk-to-disk (D2D) backup-and-recovery options. These solutions fall into two categories, depending on whether the disk is implemented as a disk target (the disk system appears to the host as disk, not tape) or as virtual tape (the disk system emulates tape). Then there are varieties of each of these two categories. Each variety has benefits and disadvantages and these differences will determine which will be the best choice for customers.

Tape-based Backups

Years ago, if you wanted to back up your data every night, the only choice was to back the volumes up to manual tape drives. Tape operators scurried from one tape drive to another, mounting and dismounting tape media, to keep the backup jobs running. Sometimes a job had to wait until the operator could find the right tape. In time, automated tape libraries replaced sneaker-clad tape operators. Fast robotics could locate and mount tapes within the libraries in seconds, faster than any human operator. Tape libraries also simplified management, since the capacity within a library can be shared across numerous backup servers. And bandwidth between backup servers can be shared when the libraries are attached to a SAN.

But once automated libraries had speeded up the backup process and made it possible to consolidate tape resources, customers started looking for solutions that would speed up the restore process as well.

Disk-based Backup Solutions

There are a variety of disk-based backup solutions. In some implementations, the disk systems appear to the operating system as disk. In other implementations, the disk system emulates tape drives and libraries. Customers evaluating disk-based options need to understand and weigh the benefits and disadvantages of each approach to determine which best meet their present and anticipated needs.

Disk-As-Disk

One segment of the D2D market—“disk-as-disk”—includes disk systems that do not emulate tape. Backed up data is written in disk, not tape, format. Disk-as-disk products are available in several varieties and may not support writing directly to tape drives and libraries.

Clones and Snapshots

One approach uses copying or snapshot functions. These copying functions can reside in the disk array, in an appliance within the SAN or as host-based software. Array-based functions must create the secondary copies on the same disk system and the customer must purchase a software license for each disk system. Appliance-based and host-based implementations can, if desired, create the secondary copies on another disk system. While array-based copies do not use channel bandwidth, appliance-based and host-based copying functions do add additional traffic to the network or channels and use server cycles. Some of these solutions create physical volumes, some create logical volumes, and some can create both types. A physical copy, or clone, is a separate copy of the volume and customers need to allocate storage identical in size to the primary volume for each copy required. Logical copies, sometimes called snapshots and sometimes described as copy on first write technology, only store the changed blocks and do not require a full volume copy.

Copying technologies have evolved and most can keep track of the changes that have occurred since the last copy function was initiated. Each subsequent copy request only copies the changed blocks, or deltas, reducing the time required to make an additional copy. These same techniques also improve restore time; only the changed blocks are copied back to the original volume.

It is important to note that this technology does not replace existing tape backups, since clones and snapshots are not designed to retain permanent backup copies. However, this technology enhances tape backups. The copied volume, not the primary volume, is now the source volume for the backup eliminating contention between online access of the primary volume and the backup process.

Use Existing Disk - “Create Your Own [D2D]”

In this approach, secondary disk (usually higher-capacity SATA disk) serves as the target device for backups. This is sometimes the lowest-cost solution, either because no special appliances or disk systems need to be purchased, or because customers can make use of their excess disk capacity. This approach also allows customers to make effective use of older disk systems that have been displaced by newer technology.



Native disk arrays have basic interfaces that allow IT administrators to configure disks; for example, administrators can set up LUNs and define RAID groups. However, native disks do not have facilities to monitor backup activity, so customers need to regularly monitor these disks to determine if they are running out of capacity or if the disks have become fragmented. Another disadvantage of this approach is that disk systems must be partitioned to support different operating systems and capacity cannot be shared across partitions. If the backup requirement for a particular operating system increases, then the disk system may have to be re-partitioned. A further disadvantage is that the backup applications must be changed to reflect the new disk target. A redesign of the process may also be required to optimize performance.

Specialized Appliances

Specialized software and disk appliances not only serve as backup targets but also provide options such as filtering (to eliminate redundant copies) and compression. Some vendors have developed appliances to address specific problems that have plagued IT administrators for years. For example, if the file system that is currently being backed up is corrupted, the backup will simply mirror the file system's corrupted state. One appliance runs consistency checks against the file systems to flush out integrity problems.

Specialized D2D appliances have many features not available with native disks. For example, appliances can help prevent backup problems by sending IT administrators a warning when disk capacity is reaching a critical threshold. Some appliances automatically invoke processes to move disk-resident backups to tape in order to free up disk space before an "out of space" condition can occur. Unlike native disk, many appliances have graphical interfaces that allow administrators to quickly determine the status of numerous backup tasks. Some backup appliances offer specialized functions, such as compression, that can save disk space and money. When additional capacity is required, many appliances make it easy to add disk arrays into the existing storage pool. These and other benefits usually come with a higher price tag, so customers need to compare the cost and capacity of native uncompressed disks with the cost and capacity of an equivalent appliance.

In general, "create your own" solutions and specialized appliances can reduce restore times. However, the backup processes need to be redesigned to write to disk targets rather than tape. And D2D solutions, by themselves, are not a disaster recovery solution. Customers should consider procedures for transferring their disk-based backup copies to physical tape for off-site storage.

Disk as Tape, or Virtual Tape

With virtual tape implementations, the disk array responds to software commands just as if it were a tape drive, hence the name virtual tape. That gives virtual tape products an important benefit: They require few, if any, changes to the existing backup infrastructure.

Usually, the disk array and software can be easily plugged into existing backup products. The normal process of directing the backup stream to tape is simply redirected to disk.

Vendors offer different implementations of the virtual tape solution. Some only supply software, some ship appliances with software already installed within their disk arrays, and others have tape drives or libraries coupled with their disk arrays and software.

Software-only Solutions

For certain customers, there are distinct advantages to a “software-only” solution in which a vendor supplies the software while the customer supplies the disk hardware. Virtual tape software, which runs on a server situated between the hosts and the disk system, makes the attached disk look like a tape drive or tape library to the operating system. IT organizations can add virtual tape software to an existing disk controller within their environment to build a virtual tape appliance. Customers with excess storage capacity can put that disk storage to good use.

Disk/Software Solutions

Other disk-as-tape solutions integrate virtual tape software and disk (usually a SATA array) into an appliance which ships with the software already installed. Disk/software solutions resemble software-only in that, in both cases, the virtual tape software responds to operating system commands as if the attached disk devices were tape drives, not SATA disk storage. But, unlike software-only products, disk/software products don’t support heterogeneous disk. They only work with disk systems provided by the software vendor. The software is developed to run on a specific storage platform, which allows vendors to optimize performance or implement additional features, such as hardware compression.

Some solutions emulate tape drives; others emulate tape drives and libraries. Customers who have automated functions within libraries will lose that functionality if the virtual tape software does not support tape library emulation.

Some of these solutions do not support direct connections to physical tape drives. In such cases, customers implementing software-only and disk/software solutions can continue to use existing tape drives and libraries by adding an additional step to the backup process. The backup application that initially creates the disk backup image can be used later to copy, or clone, the image to physical tape for longer-term storage or off-site disaster recovery. The cartridges remain under the control of the backup application. One caveat: This extra step consumes additional I/O bandwidth and backup server cycles.

Virtual Tape Support for Physical Tape

While all virtual tape solutions emulate tape drives, not all solutions can write directly to physical tape drives. Some virtual tape solutions have tighter integration with physical tape drives (in libraries or autoloaders), and have software to manage the disk and tape media. Backup data is written to disk and then migrated to tape over dedicated paths at user-defined times.

With integrated solutions, the capacities of the disk and tape layers are not presented separately but merged into a single pool. The backup application writes to the total capacity of the disk and tape storage without actually knowing where the backup copy resides. The virtual tape software keeps track of the physical location of the backed up data. Capacity can be easily increased by adding more tape cartridges to the library.

Some integrated disk/tape solutions migrate disk images to tape automatically, without notifying the backup application. In these cases, the physical cartridges remain under the control of the virtual tape application, not the backup application. Customers who plan to remove tapes from the library for disaster-recovery purposes must determine if the tape can be used outside the control of the virtual tape software or must be returned to the library before it can be read.

Recommendations

Which combination of benefits and disadvantages makes the most sense? The answer depends on your environment and your budget.

Small customers with one or two backup servers and very limited data growth have a multitude of choices since any of these solutions will work for them.

Budget-conscious customers will be rightly drawn to the “create your own” approach, but careful analysis is required. Native disk arrays need to be carefully monitored and managed. Disk space can become fragmented and backups can quickly consume the remaining available space. Factor in the additional time the storage administrator has to put in to manage these issues, and the solution may not end up being as much of a bargain as expected. “Create your own” solutions work out best for customers with few backup servers and limited data-growth.

If the IT organization has been plagued with file systems integrity problems, then specially built software and disk appliances that provide file system integrity checks can go a long way to solve that pain point.

Organizations that have very little update activity will find that appliances that provide compression and factoring (which eliminates redundant copies) can save a large amount of storage.

Virtual tape solutions are particularly suited to customers with limited or overextended IT staff because they require few, if any, changes to the current backup processes. Some customers have already installed backup applications that do not support writing disk-based backups effectively; for them, virtual tape is the best solution.

Virtual tape solutions are also best suited for complex, high data-growth environments. For example, administrators of environments with many different flavors of UNIX and Windows find that large virtual tape appliances which emulate multiple tape drives (and libraries) can assign a virtual library and several virtual drives, or several tape drives, to

each environment. The result: consolidation without additional complexity. Customers with high data-growth rates can easily and inexpensively scale by adding more cartridges to the library. Customers with three or more backup servers will benefit because virtual tape solutions allow resources, such as tape drives, to be easily shared.

One caveat: Some backup applications charge customers according to the logical presentation of devices the application sees rather than on a per TB basis. For these applications, if more virtual tape drives are defined than existing physical drives, the cost of the backup application may increase.

Conclusions

The backup-and-recovery process, although still a necessary evil, has come a long way. The wide variety of disk-based backup products on the market offer good solutions to customers who take the trouble to make an informed choice. “Create your own” disk-based backup solutions are well suited to smaller, less complex, budget-constrained environments; disk-as-disk appliances improve on native disk by providing additional features, potentially worth their extra cost; virtual tape solutions require the fewest changes to overall data management and backup infrastructures. ■

There is no right data protection solution for every company. In fact, different applications within a company require different levels of protection. HP has now added an additional tier of data protection to their existing disk-based backup products. This new virtual tape product, HP StorageWorks 6000 Virtual Library System (VLS), gives customers many new options when choosing the best data protections suite for their data.

Initially, HP will offer two virtual tape models, the VLS 6105 and the 6510. The 6105 will support 3 to 6 TBs of raw capacity and a maximum throughput of about 225 MB/sec. The 6510 has double the capacity (5 to 12 TBs raw capacity) and maximum throughput (450 MB/sec).

The Virtual Library System is built using industry-standard Proliant servers and MSA 20 enclosures with RAID 5-protected SATA disks. Both models of the Virtual Library System can emulate up to 64 HP tape drives (such as the LTO-Ultrium) and 16 tape libraries (such as the ESL E-series).

One of the most appealing features of virtual tape products is that they require few changes to the backup application. Customers who already directly back up streams to HP’s LTO-Ultrium tape drives, for example, can easily add the VLS 6000 system with few, if any changes to the backup application. However, HP has taken it a step further by designing a GUI interface that is so easy to use that the virtual tape library can be configured in under ten minutes. Since the VLS 6000 is a physical disk system, HP utilizes LUN mapping and masking to isolate the activity of the backup servers from disruptive events in the SAN.



Backups are directed to the VLS 6105 and 6510, where the data can be compressed to save space. Later, a request is sent to the backup server to migrate the disk-based copies to physical tape. The tapes can be retained within the library or moved out of the library to a remote location. We expect that in future releases, HP will provide an option of tighter integration with their tape drives and libraries. With such an option, the VLS would become, in effect, disk cache for the automated tape library. Backups would be created on the disk cache and later moved automatically from disk cache to physical tape under the control of the virtual tape software.

HP customers now have many ways to protect their data. EVA and XP customers can use Business Copy to create clones and snapshots. Customers with excess disk capacity (such as low-cost MSA arrays) can use these disks as targets for the backup streams. In fact, HP can also provide the backup software with HP OpenView Storage Data Protector 5.5 Advanced Backup to Disk. And, for those applications with less stringent Service Level Agreements, or for off-site storage, HP offers several different tape drives and automated libraries. ■

Copyright © 2002-2005 Data Mobility Group, LLC. All Rights Reserved. Reproduction of this publication without prior written permission is forbidden. Data Mobility Group believes the statements contained herein are based on accurate and reliable information. However, because information is provided to Data Mobility Group from various sources we cannot warrant that this publication is complete and error-free. Data Mobility Group disclaims all implied warranties, including warranties of merchantability or fitness for a particular purpose. Data Mobility Group shall have no liability for any direct, incidental, special or consequential damages or lost profits. The opinions expressed herein are subject to change without notice.

Research sponsored by Hewlett-Packard Company. All other brands, products, or service names are or may be trademarks or service marks of, and are used to identify, products or services of their respective owners.