Mesabi Group

Commentary

August 17, 2005

Tape Virtualization — You Need to Know More Than You Think You Do

Ah, tape virtualization that is simply another name for disk-based backup. Well, not really. Of course, disk-to-disk backup is a typical form of tape virtualization, but it is not the only variant. Understanding the different choices that you have available to you is key to determining which choice best fits your needs, and in what situations.

Virtualization Attracts Attention

Virtualization is one of the in-fashion words in the IT industry. All kinds of virtualization — including server, disk, and tape — are being bandied about. Virtualization sounds simple — the logical abstraction of a physical resource. Virtualization hides back-end complexity; for example, it can allow you to add disk or tape resources without having to bring an application down.

Today IT professionals talk about and debate virtualization, but when the transition in future years to virtualization is complete, then the whole discussion will be moot. That is because virtualization will be a seamless background technology that will simply be there and not generate the attention that it does today.

The Many Faces of Tape Virtualization

However, trying to best understand when and where to use virtualization now sometimes continues to cause confusion — and that confusion is perhaps worst with "tape virtualization."

The confusion starts because we use the word "tape" loosely to refer (a) to the physical tape cartridge as a piece of media upon which the information is actually stored, (b) to the tape drive into which a piece of media is inserted, and (c) to robotics in the form of a tape library, which contains both a number of tape drives and slots in which pieces of tape media can be stored when not in active use in the library.

The confusion continues because disk drives are often used to augment or complement tape virtualization.

The result of this loose nomenclature is that it is not clear what is being virtualized (tape, drive, or library), and what is used to implement virtualization (tape or disk or both).

Tape Virtualization in a Nutshell

To avoid this confusion, let us look at the practical choices available to the storage strategist seeking to virtualize. The three

26 Country Lane
Westwood, Massachusetts 02090

Telephone: 781 326 0038
Fax: 781 326 0038
www.valleyviewventures.com

basic types of tape virtualization today are:

- *Virtual tape* — disk is used as a cache to concatenate datasets in a manner that most efficiently uses the capacity of a tape cartridge.
- *Virtual tape library (VTL)* — disk is used to emulate a physical tape library.
- *Tape library virtualization* — tape drives and tape slots can be allocated dynamically rather than having fixed assignments.

Technically, writing to disk as a single virtual tape drive is a fourth alternative. That alternative may be appropriate for some smaller configurations, but is not easily scalable to larger configurations.

Each of these tape virtualization choices solves a different problem (Table 1). Virtual tape has been used primarily for writing mainframe datasets so that more of the capacity of tape cartridges is used. However, with the increase of size in open systems tapes, there may be a need to consider virtual tape technology for open systems tapes as well.

The focus of a VTL is to attempt to improve the backup/restore process. It does this by using disk as a replacement for a tape library.

Tape library virtualization is still an emerging technology and is about the implementation of software virtualization intelligence in the tape library itself to improve the sharing of both tape drives and tape slots.

Table 1: Tape Virtualization Choices

Type	Description: What/ Where?	Description: Chief Benefit
Virtual Tape	Temporary disk workspace organizes data for writing to tape	More efficient use of tape cartridges
Virtual Tape Library	Disk storage is used to emulate a tape library	Increased reliability of restoration, shortened backup times
Tape Library Virtualization	Flexibly allocates the tape drives and tape slots of a physical tape library	More efficient use of tape library resources

Source: Mesabi Group August 2005

IT administrators should determine first what problem they are trying to solve and then apply the appropriate tool. A closer look at each type of tape virtualization should help you in that decision-making process.

Virtual Tape Virtualizes the Media

Recall that virtual tape refers to the virtualization of a piece of media rather than the virtualization of the tape drives that go into a tape library.

Virtual tape has a longer history than virtual tape libraries and the use of virtual tape is a commonplace practice on mainframe systems. On the mainframe, the process of writing datasets to tape often left the tapes with a lot of empty space. With virtual tape, multiple datasets are concatenated on disk and then written to tape. Open systems typically have not had the same issue with empty space, but some efficiencies can still be achieved with open systems

26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.valleyviewventures.com

tape, so virtual tape is now available for open systems as well.

Virtual tape is primarily an asset utilization and ease of management benefit play. Virtual tape achieves indirect benefits for data protection by minimizing the number of tapes that have to be restored, which leads to fewer chances for restoration problems.

A VTL Uses Disk to Emulate a Tape Library

Today's backup/restore software is designed to minimize the impact upon the existing policies, practices, and procedures of an IT organization. Standard backup/restore software packages can target disk as well as tape. A virtual tape library is software that runs on a disk array to emulate a tape library.

A VTL adds in the cost of the virtual tape library software in addition to the cost of the standard backup/restore software. However, simply retargeting standard backup/restore software from disk to tape requires that each backup job be manually retargeted to disk. That is not true of a virtual tape library.

If the number of backup jobs that have to be changed is manageable, straight disk-based backup may be a feasible alternative. A second concern is that there might be a two-terabyte (TB) file system limitation, which would apply to straight disk-based backup, but not to a VTL. Thus, two potential benefits of VTL in some situations are integration (saving administrator effort) and scaling.

A more complex backup environment and/or large capacity backup requirements (say, six TB or greater, as a rough measure) would tend to favor a VTL. Otherwise, straight disk-based backup might be a reasonable choice, because it requires no additional software to maintain.

Virtualizing the Real Tape Library Itself

Tape libraries tend to suffer from the same "sharing" problem that storage area networks (SANs) with block-based storage virtualization suffer from. A SAN's drives are partitioned so that each server can see a configured set of logical units. The drives are not "shared" in the sense that multiple servers can use the same logical units.

The same is true of tape drives and tape slots in a tape library. The allocation of servers to tape drives/slots tends to be fixed, with rigid partitioning.

What you would like to do is create logical tape libraries that are larger than the number of physical libraries that actually exist.

Dynamic partitioning enables storage slot pooling and flexible drive assignments. That enables re-configuration of tape drives and storage slots as need without having to take the tape library offline.

Obviously, only one application can be using a tape drive at a particular time. However, use of the tape drive/slot can be "time-sliced", with each server/application using the tape when another does not need it, allowing faster backup and restore for a server/application parallelized across tape drives. Sharing by fixed partition, by contrast, puts an added burden on the administrator to "guess" a priori the most efficient partitioning, leading to reconfiguration problems down the road.

26 Country Lane
Westwood, Massachusetts 02090
Telephone: 781 326 0038
Fax: 781 326 0038
www.valleyviewventures.com

Putting storage virtualization intelligence in the tape library itself is not yet commonplace among tape library vendors. However, such intelligence can lead to not only more efficient use of the tape library resources, but may also alleviate the need for a virtual tape library, as it could lead to more timely completion of backup jobs.

Conclusion

Getting the most out of your tape assets is important. Tape virtualization is the secret sauce that helps you do that, but virtualization comes in different flavors. You have to choose the flavor (or flavors) that best meet your needs. Virtual tape helps you get the most out of your tape media. A VTL can reduce backup times and improve the reliability of data restoration. Tape library virtualization can help you get the most out of a tape library. The choice is yours.

David Hill

Analyst Name: David Hill
Topic Area: Storage Management

Mesabi Group LLC
26 Country Lane
Westwood, MA 02090
www.mesabigroup.com

Mesabi Group LLC is an affiliate of Valley View Ventures that aims to provide thought leadership and sound advice to both vendors and users of information technology. .

Phone: (781) 326-0038
email the author: davidhill@mesabigroup.com