Pund-IT

PUND-IT RESEARCH

Weekly Review

-June 7, 2006-

In this issue:

- **The Real World Benefits of Storage Virtualization**

Pund-IT, Inc.
Hayward, CA
U.S.A. 94541

Phone: 510-909-0750
Fax: 510-886-4937
charles@pund-it.com
www.pund-it.com

The Real World Benefits of Storage Virtualization

By David G. Hill, Mesabi Group, LLC

Virtualization is the fashionable word in IT jargon these days, but does it have staying power? What really is virtualization and why/where/when does it matter? Virtualization is not reserved for storage; in the IT world any resource can be virtualized including servers, networks, operating systems, and applications. But the focus for today is storage virtualization.

The Many Faces of Storage Virtualization

Block-based storage virtualization is the first solution many people think of, but it is only one of many types of storage virtualization. The Storage Networking Industry Association (SNIA) has put together a storage virtualization taxonomy that breaks down storage virtualization as follows:

What is created — block virtualization; disk virtualization; tape, tape drive, tape library virtualization; file system/file/record virtualization; other device virtualization

Where it is done? — host-based, server-based, network-based; storage device, storage subsystems

How is it implemented? — in-band/ out-of-band

SNIA also has a shared storage model that ranges from level 1 storage devices to level 2 block aggregation, to level three file/record layer, to level four application. For simplicity's sake, the focus of this report will be only on two types of storage virtualization.

Block-based virtualization

File-based virtualization

(Virtual tape libraries are another important form of storage virtualization, but will not be covered in this report. To learn more about how a virtual tape library (VTL) supports virtualization, see "Virtual Tape Libraries Are Very Real" in the March 1, 2006 issue of Pund-IT Review.)

Welcome to the Block Party

Block-based storage virtualization is nothing new (Table 1). Xiotech has had block-based virtualization available for its midrange disk arrays from the time that it opened its doors in 1998. Three software-based companies — DataCore, FalconStor, and StoreAge — were among the pioneers of block-based storage virtualization. Today all four companies still utilize block-based storage virtualization as the core of their solutions, but, for the most part, do not emphasize the technology. Rather they focus on providing the data management services (such as copy, mirroring, and migration services) that are byproducts of the underlying technology, as well as some of the benefits, such as ease of use/management.

From the perspective of the major players in storage, IBM has the most mature and widely-accepted block-based virtualization product in its Storage Volume Controller (SVC). SVC (now on V4.1) has been in the market for three years and IBM reports that it has over 2000 customers. A SVC is a storage network virtualization appliance in contrast to the storage-controller-based virtualization approach of HDS and the switch-based network virtualization approach of EMC.

Hitachi Data Systems (HDS) is next in line with its TagmaStore Universal Storage Platform (USP). HDS says that it has 3000 units in the field that are *capable* of virtualization. The operative word is *capable*. The reason is that 60 to 65% of HDS customers use the USP as a standalone array due to the considerable capacity behind it. That means that roughly 35 to 40% of the customers use tier 2 and tier 3 storage external to the USP (in addition to internal storage). HDS feels that this external storage is probably modular (i.e. midrange) storage or perhaps older storage. The company's newer Network Storage Controller (model NSC55) is, in effect, a modular, rack mountable subset of the USP.

EMC is the first major vendor to deliver a pure switch-based network storage solution, Invista, which became generally available earlier this year. While Invista will require some time to gain mo-

mentum, EMC has a strong relationship with Cisco to use Invista on the Cisco MDS9000 in a Storage Services Module and the company works closely with Cisco's virtual SAN technology. EMC recently acquired Kashya, which is likely to be another example of EMC's successful program of R&D through acquisition, employing a Darwinian model for selecting small vendors that have survived as a result of having excellent technology. As a result of the Kashya acquisition, Invista will eventually be able to offer data management services, such as network-based snap functionality and remote replication.

Although not directly involved with delivering block-based storage virtualization directly, key storage networking companies including Brocade, Cisco, and McDATA support the process. Cisco works closely with both EMC and IBM. EMC's Invista and IBM's SVC can be coupled to the Cisco MDS9000 through service modules. Brocade and McDATA have recently been less vocal than Cisco in supporting block-based storage virtualization.

Table 1: A Sampler of Block-based Storage Virtualization Vendors

Vendor	Product	Product Focus	Technology Foundation
Broad based			
EMC	Invista	Initially, EMC is focusing on the use of virtualization for data migration.	Switch-based network storage virtualization
HDS	TagmaStore Universal Storage Platform (USP) TagmaStore Network Storage Controller (NSC55)	HDS views that block-based virtualization is controlled from the high-end enterprise level down	Controller-based virtualization across disk arrays.
HP	XP12000/ XP1000 StorageWorks 2000 Storage Virtualization System (SVS200)	HP does not emphasize the virtualization capabilities of the top-end products Focuses on data management services, such as data migration	Top-end products are equivalent to USP products from HDS The SVS200 is a virtualization appliance based upon the top-end products.
IBM	Storage Volume Controller (SVC)	IBM emphasizes virtualization. SVC supports a large number of heterogeneous platforms and continues to support more data management services, such as remote mirroring	Appliance-based network storage virtualization.
Sun	StorageTek 9990, etc. Sun StorageTek 6920	"Virtualize Everything" is one of the four steps Sun is advocating.	Sun's top end enterprise products are HDS USP products The midrange products are based upon its Pirus acquisition

Focused			
DataCore	SANsymphony SANmelody	Targeted for storage consolidation and automation Helps small to midsize companies expand storage capacity	Software-based.
FalconStor	IPStor	IPStor is a Swiss army knife of data management services; virtualization was the original core	IPStor is a software platform that has been extended beyond plain virtualization
StoreAge	SVM (Storage Virtualization Manager)	Data management services, such as copy, mirroring, and migration, are emphasized.	SVM is an out-of-band SAN appliance for providing virtual volume management and storage management across all SAN storage.
Xiotech	Magnitude	Midrange disk storage arrays that focus on ease of use, efficiency, and lowering overall costs.	Homogenous hardware virtualization of its own arrays
Supporters			
Brocade	SilkWorm Fabric Application Services	Downplays the virtualization word and focuses instead on volume management and data replication	Outgrowth of its Rhapsody acquisition
Cisco	MDS 9000	Provides a platform on which EMC and IBM, among others, can provide block-based virtualization; also active in SAN virtualization	Third-party storage vendors, notably EMC and IBM, can tie into the MDS 9000 through service modules.
McDATA	Application Services Module (ASM) with McDATA directors or switches	Focuses on virtualizing the whole fabric, not a single chassis.	An ASM is a 24-port, 1 U high module that attached to any port of a McDATA director or switch.

Source: Mesabi Group, June 2006

Open the Files

File virtualization makes files location-independent, helping users deal with the fact that NAS scalability can become an issue because a single filer can run out of capacity. File virtualization can be conducted in more than one way, but global namespace is one approach that is commonly used. Global namespace can pool storage across file systems to create a single logical entity. Files can then be physically migrated between servers without users ever noticing that anything is happening. Users still see files as if they are in the same location, but physically the files may have been moved. As a result, unutilized storage can be reclaimed, new capacity can be added non-disruptively, IT administrators have easier jobs, and users don't have to worry about disruption to their business activities.

Still, a global namespace is not always needed. For example, ONStor provides a NAS gateway approach that concentrates users into a single managed pool of storage that also allows for servers to be consolidated.

Table 2: A Sampler of File-based Storage Virtualization Vendors

Vendor	Product	Product Focus	Technology Foundation
Acopia	ARX Family	Enables file server consolidation, tiered storage, and non-disruptive file migration.	Uses the term Adaptive Resource Switch (ARX) to describe its file-switching system that provides file virtualization and global namespace
Attune Networks	Maestro File Management	Focuses on network file management using a NAS virtualization intelligent switching appliance	File virtualization and global namespace; software works with Microsoft OS
Brocade	Tapestry File Services — Tapestry StorageX	Focuses on the services that Tapestry StorageX provides including file sharing, file migration, and server consolidation	Tapestry StorageX uses a global namespace to provide a single, logical view of distributed files.
EMC	Rainfinity	Focuses on file-server-related capacity management, tiered storage management, performance management, and storage consolidation	Uses a patented architecture called the Global File Virtualization Platform.
IBM	SAN File System	Focuses on simplifying file and data management in a SAN by consolidating file systems across Unix, Windows, and Linux servers.	The SAN File System architecture is based upon what IBM called Storage Tank.
NeoPath	File Director	Focuses on tiering, consolidating, migrating, and seamlessly scaling file storage.	A network file management appliance that creates a virtualization layer between clients and servers
NetApp	Data ONTAP GX	Will focus on scalable performance for compute intensive applications managed as a single entity with a global namespace at an unannounced GA time.	The other shoe has finally dropped. NetApp has announced the combination of Data ONTAP with SpinOS, distributed systems technology acquired with its purchase of Spinnaker Networks some time ago.
ONStor	Bobcat Series NAS Gateway	Enables server consolidation	Use of a NAS gateway allows a number of servers to share the same pool of storage

Source: Mesabi Group, June 2006

Generically, Virtualization Is a Good Thing

So why and where do businesses need storage virtualization? In essence, IT organizations face increasing challenges:

- *Service levels* — more applications run 24x7 with zero tolerance or either unplanned or planned downtime.
- *Management complexity* — not only are there too many servers, operating systems, storage systems, and switching systems, but too many management consoles are required to manage

them; this is coupled with budget pressures and staff management issues (such as cross-training, retention, and hiring)

- *Underutilized storage assets* — SNIA reports that non-virtualized disk has only a 30% to 50% utilization rate and tape only a 20% to 40% utilization rate; CIOs must dread the question from CFOs "Now tell me again why you want more storage when ...?" whenever they go to the CAPEX budget well for more storage dollars.

So guess what the promised benefits of storage virtualization are?

- *Improved service levels* — significant reduction in both planned and unplanned downtime
- *Reduction in management complexity* —
  - Simplifies the storage policies and procedures
  - Puts in place an architecture for the future that is more scalable, flexible, and secure
  - Enables on-demand dynamic provisioning without disruption
  - Improve the delivery and quality of services, such as replication and migration
- *Improve the utilization of storage assets* — scarce IT funds can be deployed to other pressing demands

But Specifically...

So with that activity in storage virtualization among vendors and all the potential benefits of storage virtualization, has storage virtualization achieved its goals? The answer is — overall — no. Despite the large number of block virtualization customers reported by IBM, the market does not seem to have lived up to its expectations. A recent discussion between two CIOs illustrates the dichotomy — actually schism — that is affecting the storage virtualization market today. One of the CIOs sang the praises of virtualization and cited the benefits listed above. Another CIO was unmoved, and saw no reason to move forward. For him the benefits were chimeras.

A recent InfoStor survey confirmed that schism. 51% of the respondents reported that block virtualization implementation is not in their plans. By definition, this half will either not implement virtualization or will at best be late adopters. However, the remaining 49% of respondents (those who are considering virtualization – and are therefore early but not real early adopters), have not yet deployed virtualization solutions.

Moreover, the rate of adoption varies by type of storage virtualization. File virtualization is still early in the game and has a lot of potential. Block-based storage virtualization is the technology under the microscope and under the gun, but a wise first step is not to have unrealistic expectations. The biggest assumption of many (since vendors began working with SNIA years ago) was that block-based storage virtualization could be used across heterogeneous arrays and thereby in effect commoditizes the underlying disk arrays.

But a continuing problem is that if Vendor A virtualized Vendor B's storage along with its own, the owner could no longer use data services, such as remote mirroring, that were tied to Vendor B's storage. The argument from Vendor A that its own data services could be used instead are quite unconvincing to most IT executives who have already spent time, money, and training in the existing products.

The second problem is that, by definition, virtualization means spreading data physically across storage volumes. If volumes for a mission-critical system happened to cross vendor boundaries, that would be a no-no. Why? Because if the data were spread across volumes of both Vendor A and Vendor B and if Vendor B's RAID 5 array took two disk failures before the first disk could be rebuilt, all the data would be lost. Yes, the data might eventually be recovered, say from archived tape, but the finger pointing process would not be fun. Owners of mission critical applications and processes tend to prefer one-neck-to-choke philosophies.

Mission Accomplished?
Block-based storage vendors have, wisely in our view, reset their sights on providing data services, such as migration, with virtualization in the background, not foreground. The challenge is how compelling are these services especially in light of the fact that there may be alternative approaches to accomplishing the same end?

Will storage virtualization (in all its forms) ever be so deeply embedded in the storage infrastructure that it is universally ubiquitous and no longer simply talked about? Obviously, the adoption of virtualization will take time, but enterprise storage customers understand that virtualization is a one-way trip. Once a particular form of virtualization has been adopted, there is no turning back.

This sounds like and is a big step for any organization, but how many drivers using automatic transmissions really want to go back to using a manual clutch? Perhaps some, but not many. From our perspective, the potential technical and business benefits of virtualization will eventually help it to prevail, but success will come via a process of erosion, not by a big bang.

The Mesabi Group (www.mesabigroup.com) aims to deliver high-quality insights on long-term strategies for mining mission- and business-critical data for cost-effective, low-risk competitive advantage. We helps organizations make their complex storage, storage management, and interrelated IT infrastructure decisions easier by making the choices simpler and clearer to understand.