

## STORAGE VIRTUALIZATION

Server virtualization from VMware, Citrix and Microsoft has taken the business world by storm, allowing organizations to consolidate tens or hundreds of hardware servers into one or a few hardware servers containing multiple virtual servers. The results are much lower hardware costs, more efficient use of server resources, power and cooling savings, flexibility and agility. Slightly less known to the world of small and medium businesses is storage virtualization, which does many of the same things for storage that server virtualization does for servers. Together, storage and server virtualization make a powerful combination that reduces costs and makes your organization much more agile and competitive.

Organizations of all sizes today are seeing their storage needs almost double or even triple every year as they digitize business processes, expand their Web presence, embrace disk hungry voice and video applications, and retain files and data longer to comply with regulations and legal pressures. Unfortunately, in the physical world, storage is usually tied to a single server or application. As a result, some storage is bursting at the seams and in constant need of an upgrade while other storage sits largely unused with no way to be shared. Configuring new storage for company applications requires a lot of manual work and technical expertise to master the quirks and configuration issues of individual storage systems from different vendors. Since migrating data to a new storage system is such a hassle, many organizations overprovision storage for many applications, resulting in a very inefficient infrastructure.

While there are many different types of storage virtualization solutions, the concept of storage virtualization is very simple and similar to that of server virtualization. Storage virtualization abstracts the logical aspect of storage from the physical, allowing you to pool and share large quantities of storage among several applications and servers, regardless of the physical hardware that lies underneath. It masks the underlying complexities of individual storage device configuration and management and puts all provisioning, management and allocation under a single storage virtualization management interface. The result is a single logical storage pool that you can slice, dice and allocate to applications at will.

The benefits of storage virtualization are similar to those of server virtualization. You use your existing storage much more efficiently. Storage becomes much more mobile and easier to manage. Disaster recovery tasks such as backup and replication are much easier with a single, highly mobile storage pool than individual physical devices, and many storage virtualization solutions come with advanced software solutions for migrating and protecting the virtual storage pool as well.

By itself, storage virtualization is a very effective way to cut costs and make provisioning resources for new and existing business purposes quick and easy. When combined with server virtualization, however, the benefits are magnified.

### Three Architectures

While the overall concept of storage virtualization applies to just about every solution, the placement of the virtualization engine can vary. Storage virtualization can reside on the storage host, on the network, or on the controller. It can be in band or out of band. It can be software, appliance or switch based. And it can be based on a storage area network, network attached storage or both. Each method has its pros and cons.

### Host-based Virtualization

Host based virtualization products are implemented through software residing on a single server or as a preconfigured network virtual storage appliance. Once known as volume managers, host based virtualization solutions present several physical drives to a single server operating system or application as a single storage pool that can be portioned to applications at will.

Host based storage virtualization can be an effective, simple solution for a relatively small environment with a few servers and usually has the advantage of working with many different storage devices and vendors. Many solutions have been tailored to support a virtual server environment with features like virtual machine storage failover. The disadvantage of this architecture is that it is limited to a single server. If you want to provide virtual storage to additional servers you must configure each of them separately. This is a cumbersome solution for larger environments.

### **In Band Network Appliances**

In band storage virtualization appliances sit on the network between storage systems and hosts virtualizing all the storage on the network. They are considered in-band because all storage traffic has to pass through them to get from hosts to storage systems and back. Host servers think they're interacting with a single storage unit, but they're actually querying the appliance, which then redirects the query to the appropriate storage device on the back end.

The fact that they can work with and manage all the storage on the network makes in-band appliances a better, more manageable solution than host-based virtualization for larger installations. You can implement high performance storage where you need it and lower cost storage where you don't and virtualize it all with a single interface. Rather than having to configure each host, you can use the virtualization management software to propagate a single change automatically across multiple storage units and hosts. Another advantage of in-band over host based solutions is that data can be migrated among storage systems without having to reconfigure the hosts involved. The appliance handles that for you.

However, since they work in band, these appliances can create a performance bottleneck. Many in band product compensate for performance issues with data caching, but caching tends to increase the price of the technology. Another performance and fault tolerance solution is to place multiple in band appliances on the network, but that becomes expensive as well.

### **Out-of-band Appliances**

These are similar to in-band storage virtualization appliances, except they sit outside the network, acting as a sort of traffic cop. The host queries the appliance separately, which then tells it what storage system to access. This takes the bottleneck off of the network, but it requires configuring each host to query the appliance.

### **Virtualization Controllers**

Controller based virtualization puts the virtualization engine inside the controller of a storage array. It's an inexpensive method for virtualization, but it tends to be vendor specific, meaning that it only works with a single vendors' storage. The nice thing is that you have a single point of support. However, the scope of your virtualization efforts is limited compared with switch or appliance based storage.

### **Switch Based Virtualization**

As you might guess, switch based solutions sit inside a switch, usually a fiber channel switch located between hosts and storage on a storage area network, with much of the storage traffic passing through at wire or near wire speeds.

Switch based virtualization is the best performing and most scalable virtualization architecture available. However, there's a certain amount of vendor lock in here as well as once you have a switch with virtualization built in, it's difficult to move to a new vendor. Virtualization switches are also expensive compared with other architectural solutions and are most appropriate for high performance storage, since they don't do any of their own caching or other storage performance enhancing features. They also tend to have fewer extra features than appliances and other solutions, since there's little they can hold in memory.

As you can probably tell by now, no architectural solution is perfect and fulfills everyone's needs. The solution you choose depends on your virtualization needs, the size of your storage infrastructure, the scalability and performance you require, the types of applications you're looking for, and your budget. You also want to make sure whatever solution you choose works with your existing storage infrastructure, if that's what you need it to do. Finally, look carefully at availability and reliability features, as a problem with your storage virtualization solution may affect much or even all of your storage.

### **NAS vs. SAN**

Virtualization can be undertaken with either network attached storage or a storage area network or both, but it's usually not a good idea to combine the two in a virtualization strategy as the performance differences between NAS and SAN are pretty significant. Some solutions work with server attached storage as well. Most organizations virtualize one or the other and SANs usually win because of their performance advantages. However a small business that relies mostly on NAS can use NAS based virtualization quite successfully.

iSCSI is a fast block-based storage area networking architecture that runs over classic Ethernet, the same technology used in your traditional network, using traditional switches, etc., to provide small and medium businesses with many of the benefits of a SAN without the training and management overhead required by Fiber Channel. A small business looking for a high performance storage virtualization solution should definitely consider investing in an iSCSI storage area networks.

### **Benefits of Virtualization**

As with server virtualization, storage virtualization enhances your organization's efficiency and agility and makes storage protection infinitely easier than with traditional physical storage architectures.

- *Agility:* By breaking down the barriers between physical storage devices and creating a single storage pool and management interface, storage virtualization makes provisioning new storage for new company initiatives infinitely simpler. The management interface masks the underlying complexity of the physical storage devices, so you no longer have to deal with the individual quirks of each storage device. You don't even have to know which devices you're using. Instead, you can add or migrate storage simply by clicking on icons in a software application. Provisioning new storage is quick and painless, so you can respond to new business initiatives fast, much as you can with server virtualization.
- *Efficiency:* Intimately tied with agility is storage efficiency. Since provisioning and migration are so much easier, companies are no longer inclined to overprovision storage in order to prevent time consuming upgrades. In many organizations today only 25 percent of storage is actually in use, which is costly and inefficient. With storage virtualization that percentage often rises rapidly to around 80 percent or more.
- *Thin Provisioning:* Many storage virtualization solutions enable a feature called thin provisioning that makes provisioning even more painless and increases storage efficiency further. Thin provisioning allows file systems to pull new storage from a shared storage pool instantly at the moment they need to write to it, or based on thresholds you configure in advance. In many cases, thin provisioning actually lets you allocate more storage to your applications than is actually available in the pool, then purchase additional physical storage when necessary.
- *Performance:* Some storage virtualization solutions let you stripe data across multiple drives, drive arrays, and network storage devices, regardless of different physical storage brands and products. This can enhance performance tremendously for high performance applications such as video manipulation.
- *Easy Management:* Since storage virtualization pools storage from different devices and brands and presents it all under a single management interface, it makes managing storage infinitely easier. You only have to master one interface instead of multiple. Troubleshooting storage issues is infinitely quicker and easier as well.
- *Low TCO:* Storage virtualization reduces both the capital and ongoing costs of storage. Since you're no longer over provisioning storage, your initial storage capital costs are greatly reduced. And since provisioning and managing storage take less time and training your ongoing costs are reduced. Less physical storage also means lower power and cooling costs.

- *Automation:* Data center automation is a new software category that is coming into its own with the advent of storage and server virtualization. With data center automation software, new resources can be provisioned to applications automatically when they are needed, say at peak use points of the day, then removed when they are not. Users can actually use some automation solutions to provision their own server and storage resources for new projects or test configurations without even having to call on IT.

### **Business Continuity and Disaster Recovery**

Perhaps nowhere are the positive effects of storage virtualization felt more than with disaster recovery. The ability to provision new storage almost instantly without disrupting business systems and applications and causing downtime helps businesses continue functioning and performing on a 24/7 basis. Implementing disaster recovery solutions such as backup, replication and snapshots is infinitely easier with a single or a few storage pools, than configuring these functions for each individual physical storage device. The effortless mobility of virtualized storage also vastly simplifies these functions. You have fewer or no single points of failure and advanced disaster recovery functions such as mirroring and replication can be implemented across different physical devices and storage brands. Since you don't need the same brand of storage across a virtualized disaster recovery solution, you can save a lot of money by recycling older storage or using less expensive storage to hold replicated or backed up data either on site or at your alternate DR site.

In addition to the basic advantages of storage virtualization, many virtualization solutions come with their own disaster recovery software that makes it even easier to implement an advanced DR strategy. By combining advanced disaster recovery functions like snapshot and replication with backup to disk and tape, these solutions can provide you with near continuous data protection and no need for a backup window. The result is that organizations can implement advanced, bulletproof disaster recovery strategies with much less cost and effort than before.

### **Storage and Server Virtualization: Perfect Together**

It's a no brainer that server and storage virtualization make a combination that is more powerful than its parts. The ability to provision and migrate servers effortlessly across an IT environment is simply not as powerful if it is not combined with a similar storage capability. With server and storage virtualization you can provision new servers and storage in minutes, then migrate both across devices and networks as necessary.

Many companies configure server virtualization templates that reserve a certain amount of disk capacity by default, a potentially wasteful practice in terms of storage. With thin provisioning capabilities offered by many storage virtualization solutions, you can have new virtual machines pull storage from the pool as its needed instead, a much more efficient use of storage.

Virtual servers packed onto a single physical server share network connections and IO, which can impact performance, but by taking advantage of performance enhancements available with storage virtualization you can reduce or eliminate those impacts. Packing several virtual servers into a single physical server also creates a single point of failure, but with storage virtualization disaster recovery capabilities such as replication and snapshots you can reduce the risk of lost server dependent data and the virtual machine images themselves, which can be stored on disk and easily recovered in the event of a disaster.

Storage virtualization has countless benefits in terms of efficiency, speed and storage TCO, but it can always have its downside. Once you've implemented storage virtualization across your data infrastructure, it can be difficult to roll your storage back to a pre virtualized state and small management errors can impact large quantities of storage. That's why it's important to start small, with a single array, storage group or department to work out all the kinks. Disaster recovery is always a good place to start with storage virtualization or virtualized storage for a collection of virtualized servers. Once you understand the technology and how to deploy it most successfully, you can spread virtualization across your network to reap its full benefits.