Evaluating Storage Technologies for Virtual Server Environments

Russ Fellows
June, 2010

Technology Insight Series

Evaluator Group

enabling you to make the best storage decisions

Executive Summary

Within IT, some ideas have received significant marketing buzz and mindshare. Perhaps the most significant recent impact has come from server virtualization, which has touched nearly every IT organization over the past few years. While server virtualization can dramatically improve server utilization, it can also have just as big of an impact in other aspects of IT, namely data movement and storage technologies.

Combined with the challenges IT administrators are facing around high availability requirements, improved energy efficiency and ever-increasing complexity, there are no easy answers with current architectures. In order to solve these challenges, IT organizations must radically re-think how to build and deliver IT services cost effectively.

Evaluator Group believes there is a need for data center transformation[1], giving IT a new approach that is more flexible and allows IT departments to adapt to rapidly changing requirements. Server virtualization is the beginning of that transformation, and while it has improved server utilization, companies must move beyond these early gains and see the total benefits of savings across the data center.

To maximize the value of the new virtual infrastructures and minimize the impact on storage and networks, IT organizations must utilize I/O and storage systems that can address the need for new workloads imposed by server virtualization, along with providing high availability with architectures that can balance operational efficiency. This paper outlines important developments in storage for virtual server environments, along with considerations for executives and IT staffs tasked with implementing and managing this infrastructure.

Virtualized Data Centers

The new virtual IT data center must support multiple application and data workloads efficiently, while meeting performance requirements. IT administrators are looking to deliver the resources necessary – without over-provisioning equipment or experiencing performance bottlenecks.

The same trends described above that have been driving server consolidation are placing new demands on storage networks and storage subsystems. Current server technology supports multi-core processors, each capable of running 10 or more virtual machines simultaneously. This can result in over 1,000 virtual systems running in one rack, driving a massive increase in I/O density, performance and capacity requirements.

Along with the increase in density, virtual computing workloads are significantly more demanding than applications running in traditional environments. The combination of high density, driving high data rates, with more demanding workloads places a significant burden on the storage systems used to support these workloads. In order to reap the benefits of virtualization, it is imperative that I/O and storage systems are able to support these new, demanding workloads.

[1] Evaluator Group's "Data Center Transformation" - http://www.evaluatorgroup.com/data-center-transformation-presentation

The Impact of Virtual Servers

One of the biggest impacts of server virtualization is the effect it has upon storage. While virtualization has greatly improved server utilization, it has also increased the demands on storage. Traditional storage architectures have not kept up with the needs of virtual server environments.

Server virtualization improves server efficiency while raising the stakes to deliver performance and protect against data loss. Application and user performance metrics depend upon an infrastructure that delivers minimal delays. Moreover, server virtualization drives the need for low latency, high throughput storage access with minimal overhead, which is precisely where block storage systems excel.

Recently, a new class of systems has emerged that were designed to better support virtual server environments. Just as server virtualization allows resources to be used dynamically and reallocated as needed, so to do the new breed of storage systems designed for these environments.

By moving multiple workloads onto a single physical server, I/O rates have increased and become exceedingly random, while tolerances for delays have decreased. The combination of these three effects has led to a dramatic increase in the demands on the underlying I/O and storage system.

> ***Evaluator Group Comment: The effect of server virtualization on storage cannot be underestimated, or overstated. Storage used for virtual server environments must support simultaneous, random I/O workloads while providing high availability and performance, cost effectively.***

Storage systems have attempted to use cache to hide this problem, but, as the performance gap between CPU's and disk access rates increases, cache becomes ineffective. Traditional monolithic storage systems rely on cache as the primary method to deliver performance. However, cache is ineffective for random workloads, while adding cost. In contrast, midrange systems do not rely upon cache, but are unable to support a large number of disks or multiple controllers for balanced workloads.

What is needed is a modern storage architecture that provides the scalability of monolithic designs, with the cost effectiveness of midrange systems.

Storage Access Types

Storage access technologies include both block and file access methods. Block access requires the application or system utilizing the storage to organize the data, leaving the storage system to manage the data on the storage system. In contrast, file access adds a network file system, such as NFS or CIFS as a means of organizing and controlling access to the storage system.

Virtual machines used for server virtualization utilize their own, built-in filesystem to store and organize data within virtual machines. As a result, using a network filesystem adds an additional layer of to the storage I/O stack.

Storage for Virtual Environments

The primary factors driving server virtualization are the desire to improve utilization and to centralize management of systems onto fewer physical resources. The choice between using block or file storage for virtual server environments is not necessarily an either - or - discussion. In some environments, file access can meet the objectives, and in other cases, block access is more beneficial.

Specifically, storage in large virtual server environments must have the following attributes:

- Highly scalable and reliable – able to support large, high density workloads
- Consolidated pooling of resources – delivering better utilization without contention
- Support high performance for random workloads, during degraded operation
- Automated Optimization – including automated use of thin provisioning and tiering

What is clear is that the underlying storage system must be designed to support highly random workloads, while supporting a scalable multi-controller design that delivers balanced I/O across the system, even during degraded operation. The I/O connectivity choice should be a secondary consideration, and not overshadow the importance of selecting a storage system designed for virtual workloads.

Optimal Storage System Designs

Storage systems have typically used caching to improve performance; however, caching alone does not solve this problem. In particular, random I/O workloads derive almost no benefit from caching. As a result, monolithic frame base architectures designed to accommodate mainframe workloads provide little benefit to this type of environment, due to their cache centric design. Monolithic systems are designed for mainframe workloads, and are not optimized for highly random open systems workloads.

A common approach to meeting the performance needs of random I/O streams is to utilize as many disk devices as possible. As the growth of storage capacity has outpaced advances in disk performance, the approach of simply adding more drives can often result in a massive degree of over-provisioning.

> ***Evaluator Group Comment: Large, cache centric storage systems were not optimized for virtual server workloads. An entirely new approach is needed for storage in virtual server environments – one that provides massive scale-out performance and efficient capacity utilization.***

While monolithic storage systems support a large number of drives, they were not designed to spread data across many drives efficiently, a technique known as 'wide-striping'. Other limitations of these systems designs has limited their ability to utilize thin provisioning, or actively move data between tiers of storage. Recently, these legacy systems have attempted retrofit these features into their designs but in the designs imposes overhead and inefficiency while still requiring massive amounts of cache.

Traditional midrange system designs have also suffered from many of the design limitations monolithic systems have experienced. Although they do not rely extensively on cache, midrange systems were not designed to support the high degree of I/O connectivity, or support high throughput during degraded operation, due to their dual-controller design limitation.

Solid-state storage has been available for decades, but until recently, these devices were very expensive, and storage systems did not integrate these technologies. With the advent of Flash technology, a new type of Solid State Data (SSD) device is now available to system designers. However, cost effective use of SSD's requires that systems are able to intelligently monitor data access patterns, and dynamically move unused data to slower tiers, while moving high access data to a higher performing tier.

Evaluator Group Comment: An ideal storage system designed for virtual workloads will support thin provisioning, wide striping, multiple SSD's all coupled with intelligent software to automatically protect and move data for optimal performance.

Without dynamic, sub-volume capability, solid-state storage is not cost effective, as the manual burden placed on storage administrators far outweighs any gains. Some systems that have attempted to retrofit this technology require service interruption and scheduling to move data between tiers, negating the dynamic nature of virtualized workloads.

Moreover, a storage system optimized for virtual workloads is free from mainframe RAID group and cache centric designs while supporting multiple controllers with a highly scalable system able to deliver high performance in all circumstances.

Storage Connectivity for Virtual Servers

There are multiple access methods and specific storage types available to support server and application workloads. Choices include using block storage protocols over fibre channel or iSCSI networks, or using file based networked attached storage.

Block Storage

There are many reasons an increasing number of environments are choosing block access storage for virtual server environments. These include broader industry support, better performance, and more flexibility of design choices. Additionally, most enterprise environments have a long history of using block storage over a SAN as the primary mechanism for managing and providing storage.

Currently, the two market leading server virtualization vendors provide far greater support for block-attached storage than other storage attachment choices. In particular, Microsoft's Hyper-V has significant restrictions on the use of file access storage. The most recent version, Server 2008 R2 Hyper-V does not support NAS attached storage at all.

Evaluator Group Comment: Block storage protocols have broader support by hypervisors, OS's and the applications running within them. Additionally, block storage provides the most configuration options, and unquestionable performance. For environments looking to deliver broad virtual-server application support, block storage is the best choice.

The use of block storage protocols has dominated the enterprise storage market due to several factors. First, applications requiring the highest performance are optimized for block storage, and block storage provides the broadest support for OS and applications. Block storage protocols are the most widely supported connectivity method by the two market leaders VMware and Microsoft.

Another consideration for storage connectivity is the application and OS environment the virtualized servers will be supporting. Microsoft Cluster Services (MSCS) is a common OS environment used to provide a highly scalable and reliable platform for mission critical applications. Currently, MSCS is only supported SAN storage. Two of the most common applications deployed on Microsoft OS environments include Exchange and SQL server, which are also supported only using block connected storage.

Microsoft's 2008 R2 Hyper-V supports block storage only, due in part to the use of Cluster Shared Volumes (CSV) storage, which requires block access storage.

Another design consideration is the use of RDM or Volume Redirection to support easily moving a volume from a physical environment, to a virtual environment, or from virtual to physical. This can facilitate transparent migration of physical to virtual server environments. This feature is also beneficial if problems occur in either the physical or virtual environment that requires moving to the other type of server environment. An additional benefit is the ability to run a workload in a physical environment at a primary site for performance, while supporting that same application in a virtual server environment at a remote disaster recovery site. This is a common scenario to provide a lower cost DR facility without dedicating a large number of servers for application standby support.

Both Microsoft's Hyper-V Volume Redirection and VMware's RDM require the use of block storage to support this feature.

All of these design choices exist only when using block access storage. Thus, these and other factors continue to lead many organizations to choose block storage systems as the best means of delivering storage for their server virtualization projects.

File Storage Access

Files can provide organizational information, and a convenient method for accessing data. In virtual server environments, the underlying access method is hidden from both the operating system and applications. Therefore, file access to storage appears no different than block access. From the applications perspective, there are no benefits, although there may be performance implications.

One benefit of file access for virtual environments is that files provide a convenient method of managing all of the data for a virtual server. Each file corresponds to one virtual server instance or a point in time of a virtual server. Although there is overhead involved, the visibility of the files can provide easier administration for environments that do not need the advanced features available with block storage.

Evaluator Group Comment: File accessed storage may be appropriate for some virtual server environments. These include non-production environments, those not requiring Microsoft Hyper-V or VMware RDM support for advanced failover capabilities. Production applications requiring strict performance and failover capabilities are not well suited for file or NAS access to storage.

Moreover, environments that will not run Microsoft Hyper-V, or mixed virtual and physical server environments may not require block storage access. Virtual environments without these requirements may be candidates for file-accessed storage.

Requirement	Block Access	File Access
Microsoft Hyper-V R2	Yes, supports Hyper-V R2 in all cases	No support
Native, multi-path HA access	Yes	Typically yes, with NFS and IP
VMware RDM	Yes	No
Microsoft Pass-through LUN	Yes	No
Low CPU Overhead	Very low	Varies, up to 5X higher than block
Low latency I/O access	Very low	Low, typically 10% - 40% greater
Access to latest features	Yes (New features developed for block protocols first)	No (features may add support for file protocols at a later date)

Table 1: Block vs. File Access Choices for Virtual Server Environments

VMFS vs NFS

Often, the question of whether to use block or file storage arises out of considerations for how system state information is stored for virtual server hypervisors. In VMware environments specifically, system state information is stored in a file, known as a "vmdk", or virtual machine disk. These state files are typically only accessed by one physical system at a time. The system that has exclusive access to the vmdk is the system hosting the virtual OS environment. Each system state is stored in unique and individual vmdk files.

In order to move a virtual system from one physical server to another, access to the system state must also be transferred. This is where the use of shared access to storage becomes important. In the case of block storage, shared access is provided by a cluster file-system, which for VMware is known as VMFS (Virtual Machine File System). An alternative way of providing shared access is to use NFS, which has historically provided this feature.

Evaluator Group Comment: Both VMware and Microsoft Hyper-V have chosen to utilize block storage as the primary method for virtual machines to utilize storage. Both of these VM's are optimized for block storage, and do not share storage with other systems. In contrast, network file access is optimized for data sharing between systems, an event that rarely occurs. Thus, block storage is the best choice for operational efficiency for virtual server environments.

Thus, both VMFS and NFS provide all the needed features to allow exclusive access of a vmdk image to one hypervisor, yet allow access to be changed when a virtual system is moved. The choice between which method is best has several questions. The first is which access method is optimized for virtual server environments. Another consideration may come in the form of licensing fees.

Cluster filesystems provide high-speed sharing of data. Specifically, VMware has designed VMFS as a high-speed cluster access method for use in high-speed SAN environments. Additionally, VMFS is included with a license of ESX server. In contrast, file-based NAS systems are not optimized for high-speed access, and may also require an additional license fee.

Storage I/O Efficiency

Often, power and cooling requirements of storage systems are left as an afterthought. In order to provide the massive random I/O workload performance required, systems have traditionally relied on large, cache centric designs with a large number of disk drives. However, these designs impose a high burden on space, power and cooling, along with the expense inherent with this solution.

> ***Evaluator Group Comment: IT department can reduce their expenses by shifting to storage systems that are designed to deliver consistent high-performance, with limited, power and cooling. Intelligent use of SSD's couple with high-density design can provide real cost savings, particularly for virtual servers.***

Although SSD's appear expensive, they are more efficient in terms of cost per I/O rate. SSD's are also cost effective in terms of power per I/O operation, and even power total capacity. Thus, the use of SSD devices can dramatically improve the overall efficiency.

3PAR Advantages for Virtual Server Environments

As outlined, the demands placed on storage systems supporting production applications in a virtual server environment are extreme. It is important that these storage systems are designed to support many simultaneous random workloads, while maintaining performance. In addition, overall system reliability is critically important, due to the high amount of application data residing on the storage system. These features must be delivered cost effectively, without over provisioning. Finally, it is important that the management and administration of the entire system is well integrated, including between the server virtualization and storage layers.

Some design points for 3PAR Utility Storage that benefit a virtualized environment include:

- Massively-parallel, clustered system
- IOPS spread across all system components
- Fine-grained virtualization, wide-striping
- High volume IO performance with lower cost SSDs
- Existing caching architecture for tiering
- Variable size IOs for SSDs for better performance
- 3PAR's true active – active controller architecture
- Ability to utilize Round Robin volume access provides load balancing across multiple paths
- 3PAR's, zero detect, along with thin conversion and stay thin technologies.
- A vCenter plug-in that provides effective storage mapping
- Recovery Manager software specifically designed for VMware to provide hundreds of vm consistent
- snapshots

One way to achieve the necessary performance with highly random I/O workloads is with solid-state storage, or specifically flash-based SSD's. However, storage systems have not been optimized to dynamically leverage these devices. Many vendors have simply added support for SSD's without factoring in the cost, performance and management aspects this feature adds.

Evaluator Group Comment: 3PAR has addressed the limitation other storage vendors have with SSDs. By coupling Adaptive Optimization with a high number of lower cost SSDs, IT users are able to achieve their goals of meeting performance objectives and lowering their total costs.

The design of all 3PAR systems provide optimal performance to the random I/O workloads driving today's growing application mix. Monolithic frame based systems were instead optimized for legacy mainframe applications. New features, such as wide striping, storage pooling and thin provisioning were added later to these systems, imposing multiple layers and inefficiencies. In contrast, features such as thin provisioning and wide striping were the primary design points for 3PAR, not added as an afterthought to a legacy architecture.

Feature	Benefit
Massively-parallel, clustered system	Scale-out design optimized for random I/O workloads
I/O spread across entire system	Elimination of performance bottlenecks due to application hot spots
Fine-grained thin virtualization	Foundation for wide striping and automated data tiering
Multiple SSD support	Balanced system design optimizes utilization of high-value resource
VMware SRM support	Disaster recovery for VMware environments
VMware I/O optimization	Enhanced performance via queue depth and multi-path I/O drivers
Application integrated data protection	3PAR recovery manager provides PIT copy under VMware control
Adaptive Optimization[2]	Optimizes multiple storage tiers, including drive and RAID type
Automated tiering of data	Delivers optimal performance with minimal administration
Fat-to-thin and thin maintenance	Improves storage efficiency by eliminating allocation waste
VMware EZT volume, without writing zeros	Only 3PAR can provide an EZT volume without writing zero's. This provides the highest performance, fault tolerance while staying thin
Virtual server integrated management	Management integration with a vCenter plug-in
Autonomic Groups	Simplify and reduce the time it takes to provision storage in Virtual Server Environments
FAST RAID	Hardware accelerated RAID 5 and RAID 6 allows capacity savings with performance close to RAID 1 levels
Persistent Cache	Provides an option to gracefully recover from controller node failures without going intro write-thru mode. Ability to preserve service levels in Virtual Server Environments especially important due to the large number of consolidated workloads

Table 2: 3PAR InServ Feature – Benefit Overview

[2] "Automated Storage Tiering" 3PAR – Evaluator Group whitepaper by John Webster, March 2010

Simplifying administration is increasingly important with the addition of another layer to the IT stack. The addition of server virtualization can complicate overall system administration unless the management of all resources is integrated. Through a vCenter plug-in for VMware vSphere and MMC plug-ins for Microsoft Hyper-V, administration of storage is both simplified and integrated with server management.

The 3PAR plug-in for vCenter is a management plug-in that allows identification of 3PAR volumes and a mapping of VM's to storage resources. In addition, 3PAR Recovery Manager provides the ability for 3PAR point in time disk snapshots, under VMware control, to protect and recover VMDK's and enables file level restores within Virtual Machines. The3PAR Replication Adapter supports VMware vCenter Site Recovery Manager (or SRM) for remote disaster recovery coupled with 3PAR Remote Copy.

Virtualized applications often impose different I/O workloads than more traditional compute stacks. Requirements include the need for low latency access, with minimal CPU processing overhead and high reliability. Due to the way that virtual server hypervisors schedule resources to maximize system utilization, I/O workloads are now interleaved. As a result, consolidation of virtual workloads transforms sequential access patterns into random workloads.

Evaluator Group Comment: 3PAR is able to utilize the standard VMware plug-in while supporting round-robin access to volumes across multiple controllers. This capability can significantly reduce or eliminate I/O bottlenecks and hot spots, while using a standard driver stack within virtual systems.

VMware has recently added a feature for some storage vendors to support an adaptive queue depth, which can also provide performance gains to individual virtual systems dynamically, without reconfiguration of OS, I/O interfaces, storage networks or the storage system.

VMware supports three different formats for its VM disk images: thin, thick or eager zeroed thick (EZT). VMware recommends EZT for highest performance and fault tolerance. 3PAR is able to support all three formats effectively. EZT offers the best performance and security by creating VMs and zeroing out the space when the volume is created. However, EZT can create a performance bottleneck by requiring a heavy write workload on the virtual zerver. 3PAR's can detect writing zero's and efficiently zero out the volume without requiring host writes.

Additional 3PAR Storage Considerations

Virtual server environments are notorious for having a high degree of redundant data. With hundreds or even thousands of virtual machines stored on one storage device, each virtual server is often assigned far more storage capacity than they require. By using thin provisioning to delay storage use until it is required, the storage savings can be significant.

While thin provisioning is supported on many of today's disk subsystems, not all of these methods work well with virtual servers' usage of storage. This is due to hypervisors writing zeros to unused space, which negates any benefit of using thin provisioning. In 3PAR's case, unique 'zero-detection' hardware is able to detect and remap zeros into a thinly provisioned volume.

Additionally, 3PAR's "stay-thin" technology allows systems to return allocated space to the storage system pool when possible, thus increasing the effective use of thin provisioning. 3PAR's thin provisioning, thick-to-thin volume conversion, and thin maintenance all work to eliminate wasted storage space. Allocated but unused space is one of the largest inefficiencies inherent with typical deployments of virtual server environments.

Still another issue is how SSD drives are used within the system. In most cases, only a few drives are used due to their high cost, which concentrates I/O workload onto one controller. This can easily overwhelm most midrange systems, which are not designed for unbalanced workloads. In contrast, 3PAR's InServ systems are designed to use a relatively high number of smaller capacity SSD drives, distributed throughout the 3PAR system. This results in well-balanced I/O across the entire storage system. Only 3PAR utilizes this approach to the use of SSD's among major system vendors.

Summary

The need for data center transformation is driving server virtualization and consolidation, which in turn is increasing server utilization. This same trend is driving storage users to choose scalable systems that are able to meet their performance requirements cost effectively. In particular, meeting the high random I/O performance needs of these new, virtual server environments requires a new generation of storage system.

Simply stated, IT executives and business professionals alike are looking to reduce their costs and lower their risk of missing the required service levels. IT administrators understand the limitations of existing storage systems, which are designed for either sequential access mainframe applications or for the modest needs of midrange applications. Old designs, constrained by a reliance on expensive cache or limited to dual controllers, cannot meet the requirements of virtual servers cost effectively.

The next generation storage systems are designed to deliver high random I/O performance, in a scalable modular architecture that can grow with the new virtualized workloads. In order to achieve the efficiency, performance, and cost objectives, it is clear that few storage systems are designed to accommodate these workloads.

Data center optimization is critical to improving IT efficiency. As data centers transition to support both IT as a service and cloud computing models, the ability to deliver the necessary efficiencies depends upon choosing the right storage architecture. Storage systems optimized to deliver high availability and performance in virtual server environments is an important aspect of this strategy.