2007

AccessFlow, Inc.

Gary Lamb & Steve Kaplan

[VIRTUAL INFRASTRUCTURE STORAGE & PILLAR DATA SYSTEMS]

The exploding popularity of VMware's server virtualization has created a corresponding demand for shared storage. Traditional storage arrays, burdened by a legacy of addressing physical infrastructure requirements, fail to optimize the unique storage requirements of a virtual infrastructure. The Pillar Data Systems Axiom Storage System, presented in this paper, meets or exceeds the requirements of a next generation Virtual Infrastructure Storage System.

CONTENTS

Pillar SANs Optimize VMware Environments

By Gary Lamb, CTO AccessFlow and Steve Kaplan, CEO AccessFlow

INTRODUCTION:

IT shops are consolidating and virtualizing resources to reduce costs and to improve agility. VMware's Virtual Infrastructure enables IT to more fully utilize the capacity and computing power of the server and storage products they purchase, thereby fundamentally changing the computing landscape in forward-looking organizations. IT shops of all sizes are considering, or have implemented, Virtual Infrastructure for server consolidation, power and space savings, server high availability and to enable affordable disaster recovery.

Disruptive shifts in technology such as virtualization often occur not because of a single product or technology, but rather the combination of two or more technologies that create synergistic benefits. The triple storm of VMware Infrastructure 3, multi-core processors and storage area networks is an example of this phenomenon. VMware VI3, released in mid 2006, provided the robust performance and capabilities demanded by enterprise data centers. The multi-core processor was the enabling technology that both mitigated the latency in high context switching applications and that made it exceptionally economical to consolidate many physical servers onto a small number of multi-core ESX hosts. Storage area networks enabled the new VI3 features such as high availability, Consolidated Backup and Distributed Resource Scheduling while their acquisition cost significantly deluded within a virtual infrastructure. The result has been a rapid shift from a data center standard of direct attached to one of shared storage environments.

> *The Pillar Data Systems Axiom storage system demonstrates how a centralized storage system can satisfy all of the storage demands of a traditional computing infrastructure as well as provide optimized features and design that dovetails perfectly with the specific demands of today's virtual infrastructure.*

EMC originally purchased VMware with the intention of driving storage demand, which it certainly did. VMware also, however, is helping accelerate the commoditization of storage by incorporating functionality previously touted by SAN manufacturers as a differentiating advantage.

Server snapshots, migrations, and multipath software are examples of features that have become largely irrelevant in VMware's virtual infrastructure which provides these features through disk abstraction.

Most leading midrange storage array manufacturers have jumped on the VMware bandwagon by touting their products as "storage virtualization", but in fact continue to utilize the same legacy architecture long extant in the physical data center environment. A new generation of networked storage, specifically the Pillar Data Systems Axiom Storage System, is capitalizing on the disruptive shift in computing infrastructures with arrays designed to optimize the storage commoditization virtualization promotes. The Pillar Axiom enables a level of performance, scalability, reliability and value unobtainable with back-end subsystems designed for traditional infrastructures. In this paper, we discuss the more demanding storage requirements of a virtual infrastructure environment, the characteristics necessary for an *optimized Virtual Infrastructure Storage System* and how the Pillar Axiom meets or exceeds those requirements.

VIRTUALIZATION AND STORAGE:

X86 hardware virtualization has been around for many years, but originally tended to be relegated to non critical tasks such as test and development servers where availability and performance were less essential. The lack of high performance sub-systems and associated redundancy was a major impediment to widespread production system adoption of server virtualization.

Storage Area Networks (SANs) connecting servers with centralized storage have long been used to provide high performance and highly available consolidated storage for demanding or critical applications. The high cost of per server connections (HBA's, switch ports, and host agent software), often $3,000 - $4,000 per server, along with expensive fibre channel disks kept SAN installations the exception in most small to mid-sized IT shops. Even larger organizations deploying SANs typically had only a fraction of their servers benefiting from the performance, reliability and scalability SANs enable. The lack of widespread adoption helped keep the cost of entry high, making storage networks an expensive technology both to purchase and administer.

Separating storage from computing in the virtual infrastructure opened up a whole range of new capabilities. It allowed the creation of Virtual Infrastructure processing clusters which act as a common unified execution environment for virtual servers, enabling dynamic mobility. Virtual machines centralized on high availability storage arrays could now be executed on any of the severs in the cluster enabling features such as live server migrations, high availability, guaranteed service levels, and simplified replication of virtual infrastructure to disaster recovery sites.

The new features in Virtual Infrastructure management systems aided by the dramatic increase in per server computing power rapidly launched enterprise virtualization technology into mainstream datacenter technology. The popularity of server virtualization and its requirement for networked storage has been largely responsible for the surge in mid-range storage array sales. Mid and small sized organizations along with departments of larger organizations purchasing their first storage network as part of a Virtual Infrastructure project have been a primary driver of this trend.

BARRIERS TO IMPLEMENTATION:

Server virtualization lowers the cost of fibre channel SAN implementation by slashing the requirement for fibre channel host bus adapters and switches. Yet SAN costs frequently still comprise the largest portion of a small Virtual Infrastructure project and prohibitively raise the cost of entry for many organizations.

While storage administration is simplified in a virtual infrastructure by no longer requiring a separate LUN to be carved out for each server, it still requires careful management. Large organizations typically have dedicated storage administrators or a storage management team. A mid-sized or small company purchasing its first storage network and shared array as part of a virtual infrastructure implementation often has little or no staff or expertise dedicated to storage networks. This skills gap further increases the cost of the virtual infrastructure and likelihood of service interruption caused by VMware / storage issues.

Storage related problems are the most common and often most complex support and troubleshooting issues with which VMware Technical Support must work. In addition to the technical peculiarities native to each storage vendor, compatibilities between all components of the storage network are crucial to the reliable operation of virtual infrastructure storage. These availability and reliability issues create substantial support overhead, and they don't even begin to address the complexity and experience needed to deal with SAN performance tuning or optimization issues.

VMware's VirtualCenter provides intuitive management interfaces and processes; however, the storage management layer has traditionally presented a steep learning curve or required expensive experienced storage administrators. The knowledge gap faced by companies implementing their first storage network drives up the cost, erodes the potential ROI of a VMware deployment and can add instability to the infrastructure due to SAN mismanagement.

Expansion costs are also a common pain point whether companies are growing existing SAN capacity for an initial VMware deployment or are expanding the virtual infrastructure. Traditional storage arrays often incur substantial licensing costs for both features and capacity. Additional licenses and even new SANs are required to resolve performance degradation created by the non linear performance characteristics of scaling capacity in traditional storage arrays.

VIRTUAL INFRASTRUCTURE STORAGE SOLUTIONS:

Given the functional inseparability of virtual servers and shared storage, the full potential and savings of virtualization can only be realized when the storage system is optimized and managed as a single functional unit within the virtual infrastructure.

The storage related barriers to virtual infrastructure implementation include acquisition and expansion cost, manageability, and complexity. To be considered a virtual infrastructure storage solution an array must exhibit:

- Intuitive Management
- Virtualized LUNs
- Linear performance through capacity scaling
- QOS & performance tuning
- Availability & reliability
- Cost – both acquisition and expansion
- Efficiency

INTUITIVE MANAGENT

Technology systems often work together and communicate only as well as the people who implement and manage them. Virtual Infrastructure management is dynamic. Traditional storage management is usually much more structured, controlled and deliberate. The ideal situation is to have the storage system managed as a part of the virtual infrastructure, extending the boundary of the virtual infrastructure to include the physical and logical storage.

Storage as an extension of a virtual infrastructure must be agile and manageable if it is to add to, and not detract from, the overall value and ROI of the virtualization deployment. Management features such as quality of service, data migrations, snapshots, and LUN management put storage administration on par with overall virtual infrastructure management.

VIRTUALIZED LUNS

Storage virtualization can provide the same management and flexibility benefits as server virtualization by creating virtual resources within a common physical resource pool. Virtual infrastructure storage systems should allow the creation and manipulation of virtual LUNs independent from management of the physical storage resources. Virtualized LUNs combined with intuitively designed management systems allow virtual infrastructure administrators the ability to manipulate virtual storage as easily as virtual servers and networks.

LINEAR PERFORMANCE SCALING

One vital, but often underappreciated requirement for virtual infrastructure storage systems is the need for true linear performance / capacity scaling. Organizations usually start with a small deployment with plans to increase utilization over time. As their IT staffs grow comfortable with, and appreciate, the benefits of virtualization, they accelerate its deployment faster than expected including the virtualization of larger and more demanding workloads making performance predictability and consistency crucial features.

Virtual infrastructures have a very different storage profile than most traditional workloads. Capacity expansion is driven by adding virtual servers. Maintaining multiple virtual disk files in each LUN combined with the resource time sharing inherent in virtualization results in increased random disk I/O with each virtual server added to the infrastructure. Traditional mid-range storage arrays with shared loops and storage controllers for all disk shelves often see resource contention as load and capacity increase. Disk level performance tuning cannot overcome the limitations of congestion on shared controllers and data paths. A *Virtual Infrastructure Storage System* must provide the ability to scale performance linearly with virtual server growth.

Traditional SAN manufacturers tend to tout their storage array controllers as being active/active. This is true from the standpoint that the controllers are both running and that the LUN can technically access either controller, but in practical circumstances, will only access the source I/Os from the primary controller. In the event of failure, the standby interface source I/Os are then accessed. **Figure 1** shows a drawing of a typical traditional storage array serving a virtualized server environment. In this case, two front side interfaces and one active back end interface cut the number of I/Os that the front side can handle by half because of the active/passive relationship.

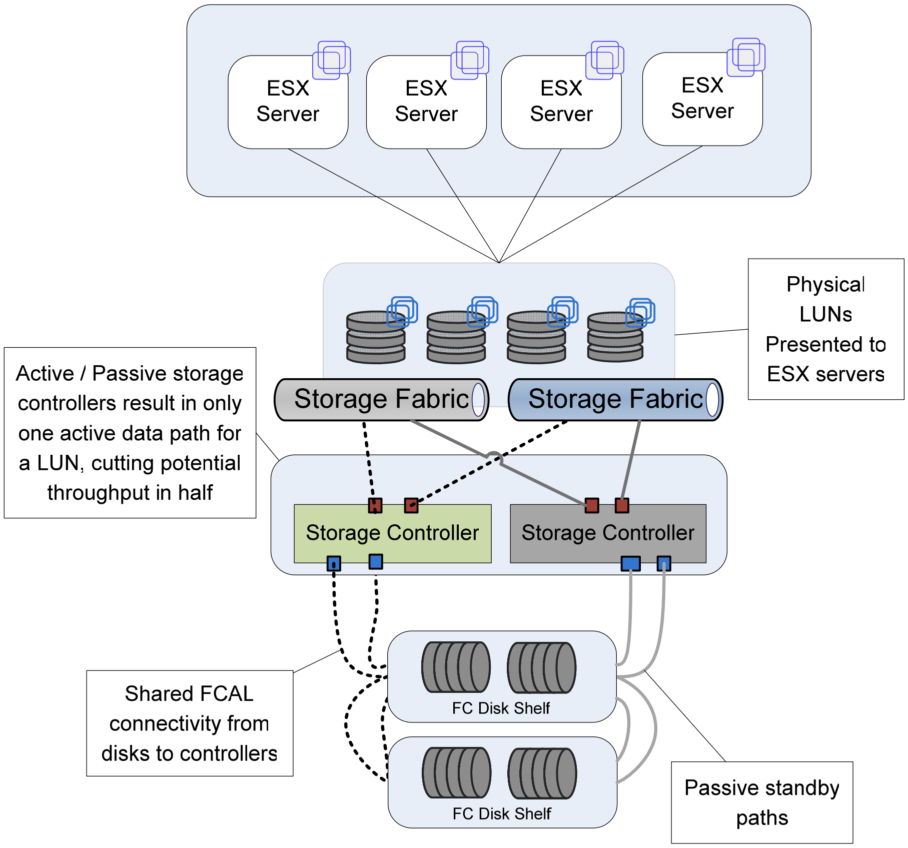

Figure 1: Traditional Storage System actively utilizing only one shared data path between disks and controllers. LUNs are passed through to ESX servers without QOS or virtualization.

In a physical world, this configuration is adequate because on average less than 20% of the physical servers of the IT enterprise are attached to the SAN enterprise. In a virtual infrastructure, however, where 80% - 90% or more servers are attached to a SAN, this architecture is flawed in that it doesn't enable linear performance as the organization scales.

QUALITY OF SERVICE & PERFORMANCE TUNING

Virtual infrastructure management systems such as VMware's VirtualCenter supply robust quality of service features to maintain desired service levels across different priority workloads running on shared hardware. VMware's *Distributed Resource Scheduler* provides granular effective controls over server CPU and memory resource shares and limits; however, there is no effective way to manage disk access performance levels from within the

> *To maintain true quality of service within a virtual infrastructure the storage systems must be able to differentiate and manage performance across different virtual server workloads.*

standard virtual infrastructure management systems.

To maintain true quality of service within a virtual infrastructure, the storage systems must be able to differentiate and manage performance across different virtual server workloads. For storage quality of service to be effective, performance must be managed both at disk access, and in the storage controllers. Virtual infrastructure administrators should be able to manage storage quality of service as easily and non-disruptively as CPU and memory QOS administration.

Managing performance on traditional arrays is typically done at the physical disk RAID level and is not easily or non-disruptively changed.

AVAILABILITY & RELIABILITY

Storage area networks have always been designed for high availability and reliability, especially the enterprise class systems. While high availability is also inherent in the design of midrange storage systems, the degree of high availability is often less than desired for a system hosting hundreds to thousands of virtual servers−in some cases hosting an organization's entire computing infrastructure. Storage systems with shared RAID/storage controllers for all disks and large fibre channel raid sets often take in excess of 12 hours to complete a disk rebuild. Disk rebuilds in traditional mid-range arrays must balance the speed of rebuild against the impact on RAID and I/O operations in the rest of the array. The traditional architecture leaves an uncomfortably long vulnerability to multiple disk failure as well as performance degradation during disk rebuilds.

A virtual infrastructure storage system must provide enhanced availability and reliability if it is to be used to host large numbers of production servers. Short disk rebuild times, the ability to withstand multiple disk failures, and non-disruptive component replacement and capacity expansion are a must for a viable virtual infrastructure storage system.

COST – ACQUISITION AND EXPANSION

To drive the full value of a virtual infrastructure, affordable, scalable, and easy to manage storage must be available. This allows an organization to implement a smaller initial virtual infrastructure that will easily scale as the organization's experience and comfort level with server virtualization increases. Storage arrays most commonly deployed to support virtual infrastructures are mid-sized arrays from large traditional storage vendors. These *first generation* mid-sized arrays have inherited much of the complexity and obscurity of their larger, older enterprise storage arrays; their limited capability and higher cost of management detracts from their value as virtual infrastructure storage.

EFFICIENCY

An increasingly important driver of virtualization is the power savings it enables. A recent EPA study found that data centers accounted for 1.5% of the total electricity consumed in the U.S. in 2006. It is not uncommon for a data center to utilize up to 100 times the power of the office building in which it is housed.

While VMware enables dramatic power and space savings through server consolidation, traditional storage arrays consume large amounts of both resources while generating copious amounts of heat. Organizations looking to create a more efficient datacenter need to consider the power and heat footprint of the entire virtual infrastructure including storage. A storage system for virtual infrastructures should be designed to maximize heat, power, and space efficiency in the datacenter.

PILLAR DATA SYSTEMS AXIOM:

Larry Ellison (of Oracle fame), recognized the costs and limitations in existing storage systems and foresaw the approaching commoditization of storage. He founded Pillar Data Systems in 2001 and is the primary investor. His support provided Pillar with the luxury of taking 2 ½ years to analyze current and future needs and create a new storage system architecture. Pillar's design incorporated desired features and functionality gleaned from hundreds of interviews with users of traditional storage arrays. By blending lessons from the past with common technologies used in uncommon ways, Pillar was able to create a compelling solution which increases the value, not the cost, of virtual infrastructure.

The revolutionary Axiom storage system is higher-performing, more scalable and lower priced than competing products. The Axiom storage system is well suited for many applications, particularly virtual infrastructure. The following sections address the unique value proposition created by the combination of VMware Virtual Infrastructure running on Pillar Axiom Storage Arrays.

Axiom Architecture – creating commodity storage:

The Axiom sets itself apart when it comes to supporting virtual infrastructure in both hardware and software design. It integrates seamlessly with virtual infrastructure by utilizing componentized and commodity based storage hardware and software.

The Axiom hardware architecture is based on a modular approach that provides redundancy, and truly linear capacity and performance scalability. This is essential if a storage system is to maintain capacity / performance scalability along with ESX cluster scaling.

Figure 2 shows the Axiom storage system as the foundation for a virtual infrastructure. Axiom has three components, the *Pilot*, the *Slammer*, and the storage *Bricks*. This architecture separates system management, the input / output engine, and the disks/storage management functions.

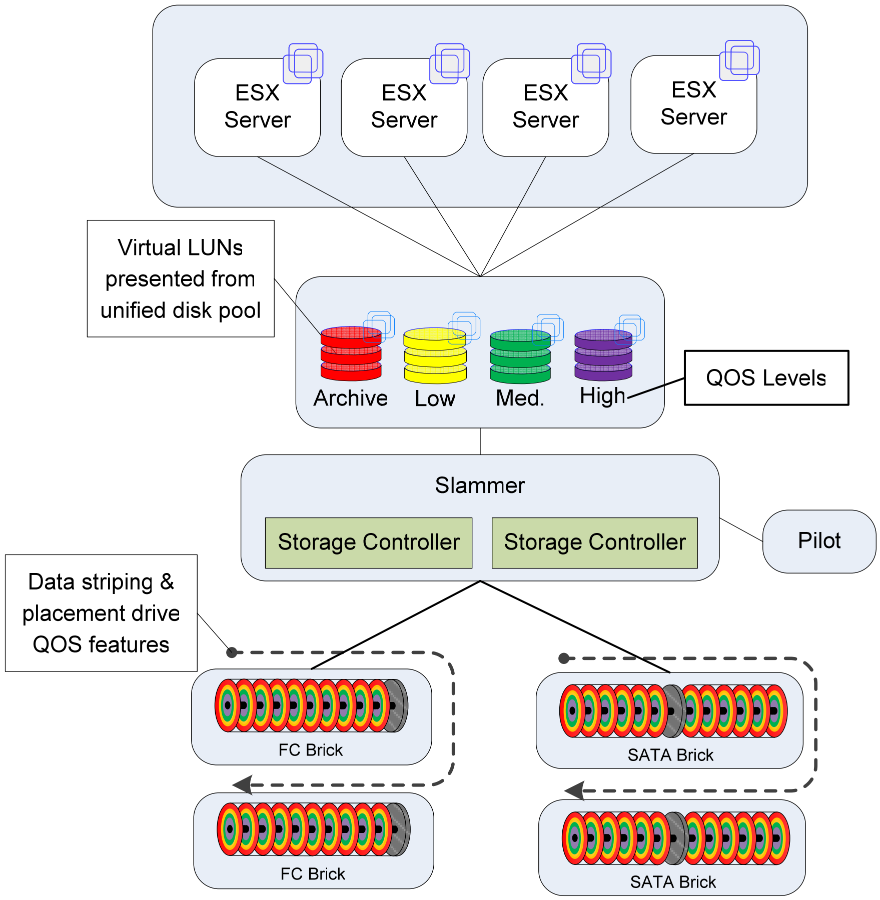

Figure 2: Pillar Axiom Architecture depicting data striping across 24 disks of each type using data placement QOS. Layered Physical LUNs are virtualized and presented to the Virtual Infrastructure

INTUITIVE MANAGEMENT

Intuitive Management has been identified as a primary requirement of a virtual infrastructure storage system. The Axiom Storage System has been designed with the same type of intuitive task based management interfaces as VMware's VirtualCenter. The pilot controller provides a central interface for managing the entire consolidated storage pool.

> *A Virtual Infrastructure Storage Solution is optimized and managed as a single functional unit within the virtual infrastructure.*

As an example, creating and allocating a LUN on the Axiom is a simple process of answering a series of guided questions in the Pilot interface. These interfaces require no additional storage training and can be easily managed by existing virtual infrastructure administrators. Removing obscurity from storage management reduces ongoing operational costs and the likely hood of service interruption due to configuration error.

VIRTUALIZED LUNS

Axiom storage integrates into a virtual infrastructure by presenting LUNs virtualized from a common unified pool of storage. Virtualizing storage presentation from the physical disk pool replicates the simplicity and flexibility native to virtual infrastructures. Including the storage layer creates a unified virtual infrastructure where virtual servers, storage, and networks are all created and managed in real-time.

LINEAR PERFORMANCE THROUGH CAPACITY SCALING

The Axiom maintains the linear performance scalability required of a virtual infrastructure storage system. Consistent performance through capacity scaling has been architected into all components. The one-to-one relationship between the hardware RAID controllers and RAID groups in the Bricks along with the massive data cache available in the Slammer units create an exceptionally scalable solution. A virtual infrastructure can scale from dozens to thousands of virtual servers with predictable performance that guarantees virtual machine service levels.

Unlike traditional storage arrays, the Axiom 500 provides true active/active controller interfaces enabling the LUNs to access the source I/Os on any controller. **Figure 3** shows an example of a 4 X 8 Axiom 500 slammer - there are 4 front side interfaces and 8 back end interfaces that attach bricks in a switch fabric. This enables any LUN (and therefore ESX Server) to access any front side interface. The whole solution scales linearly as more ESX servers, bricks and slammers are added.

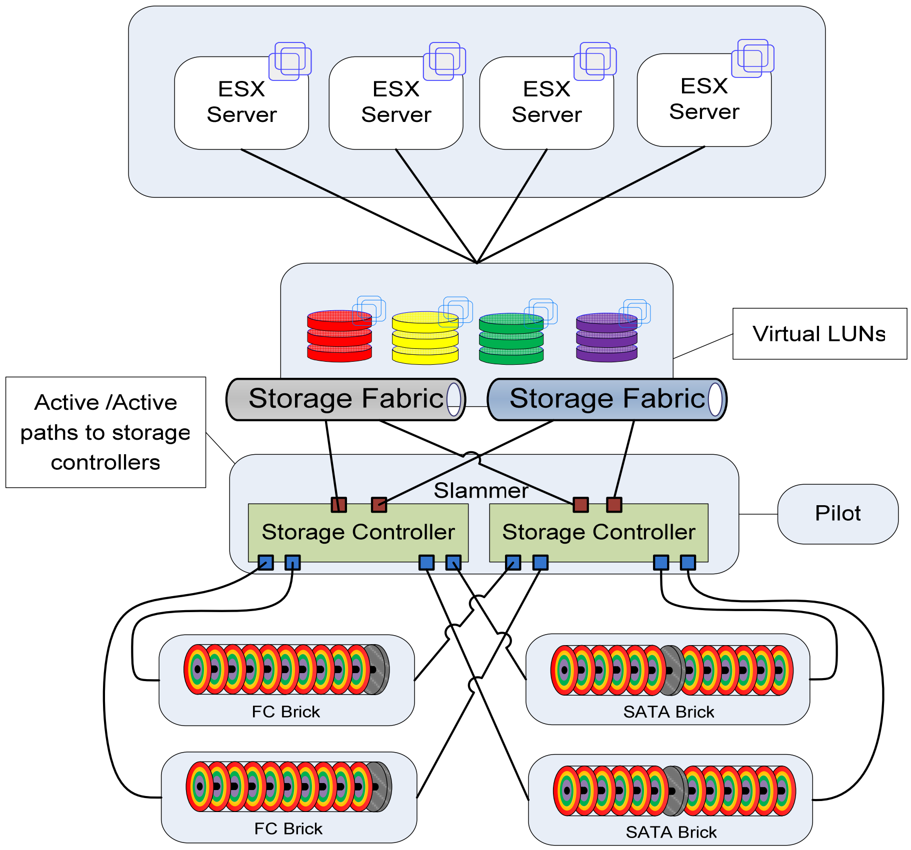

Figure 3: Axiom system utilizing multiple dedicated switched data paths from disks to storage controller doubles available throughput and increases scalability of the Axiom used as a Virtual Infrastructure Storage System

QOS & PERFORMANCE TUNING

At the core of the Axiom system is a QOS engine that guarantees and manages LUN performance tiers. The Axiom's robust QOS features complete the CPU and Memory QOS management provided by VMware's Distributed Resource Scheduler. While DRS will move virtual machines based upon CPU and memory capacity contention, it cannot adjust for disk based performance issues.

Utilizing this QOS approach allows high performance levels to be configured for high utilization production systems such as database and exchange servers. Medium quality of service levels may be utilized for standard production or back-end services including print servers, file servers, and infrastructure monitoring systems. Lower quality of service LUNs can be utilized for test systems, archive disks, template servers and other less utilized and response sensitive systems.

Pillar's unique design allows LUN QOS levels to be non-disruptively changed for short term or permanent changes in data access priority. This approach to managing a range of service quality across commodity disks is far more effective and economical than storage systems that manage performance by using expensive high performance disks for all virtual servers, or use high capacity SATA drives with no way to manage performance within the multi terabyte RAID sets.

AVAILABILITY & RELIABILITY

Pillar uses high availability and redundancy at all levels to create a storage system able to satisfy the reliability and availability required by virtual infrastructures as they scale into the hundreds and thousands of virtual servers.

> *A virtual infrastructure storage system must provide enhanced availability and reliability if it is to be used to host large numbers of production servers.*

As with QOS, the Axiom allows data protection levels to be set at the virtual LUN level. This enables efficient resource utilization with standard (RAID 5), double, and triple data protection levels to match virtual machine service levels. These options stripe LUN data across all RAID groups multiple times creating protection which can withstand loss of multiple Bricks without data loss.

Long disk rebuild times on traditional arrays left virtual infrastructures at risk for double disk failure and extended performance impact. The Axiom overcomes these vulnerabilities by using a dedicated controller for each RAID group allowing very fast disk rebuilds with no system wide performance degradation.

COST – ACQUISITION, ADMINISTRATION, AND EXPANSION

Pillar has created a storage system that increases the value of virtual infrastructures dramatically. The Axiom system has a low level of entry but can scale as large as 768 TB in one system.

Pillar does not have hidden software license charges. All features in the Axiom storage systems are included with the hardware. Quality of service, snapshots, file replication and multipath functionality are included with the system. Scaling the system with additional bricks or Slammers does not incur additional licensing fees.

Using the Axiom with a virtual infrastructure eliminates the need for dedicated storage management staff, a savings that often exceeds the cost of a SAN purchase itself.

EFFICIENCY

The Pillar Axiom 500 is the most efficient storage system available today, saving up to 75 percent of the energy required to power and cool most storage infrastructures. This efficiency is created by the Axiom's ability to consolidate the capacity and tiers of what traditionally would be multiple arrays into a single system driving greater disk utilization. The ability to provide FC drive performance utilizing high capacity SATA disks dramatically reduces drive counts drives compared to traditional arrays. This architecture creates a highly efficient storage system that adds to not detracts from the energy and space savings inherent in virtual infrastructures.

NEXT GENERATION STORAGE MEETS VIRTUAL INFRASTRUCTURE

Networked storage systems are the most complex and expensive part of a virtual infrastructure. The physical servers have traditionally been the main functional unit and focus of datacenter operations. Virtual infrastructures have commoditized servers to the point of being generic execution containers for virtual servers which can run on any of a number of physical servers in the infrastructure. Virtual Infrastructure management systems such as VMware's VirtualCenter provide an abstracted management and control layer above the physical servers rendering them generic virtual machine containers. VirtualCenter provides for management of virtualized storage resources for the virtual machines; however there has always been considerable configuration, management, and tuning of storage resources at the physical layer for presentation to physical servers or VirtualCenter.

This has rendered the storage as both the most important and most complex subsystem in virtual infrastructures. The story changes dramatically when virtual infrastructures are paired with a next generation storage system designed and optimized to overcome the limitations and shortcomings of the past.

This paper has presented the advantages of a *Virtual Infrastructure Storage System* in contrast to traditional mid-range storage systems when paired with VMware's Virtual Infrastructure. By meeting or exceeding the requirements of *Virtual Infrastructure Storage Systems* with superior manageability, efficiency, performance, reliability and value, Pillar Data Systems has created the first fully functional *Virtual Infrastructure Storage System*.

AccessFlow and Pillar

AccessFlow is both a VMware Premier Partner and Gold VMware Authorized Consulting Partner. AccessFlow is also an authorized Pillar reseller and partners with Pillar in providing optimal storage solutions for demanding VMware environments. Gary Lamb can be reached at glamb@accessflow.com and Steve Kaplan at skaplan@accessflow.com