Breaking the Bottleneck –

Solving the Storage Challenges of Next Generation Data Centers

---

An Isilon® Systems Whitepaper

January 2008

Table of Contents

1. Introduction

If it's true that information is the lifeblood of business, then it's also true that the corporate data center has become the heart of the enterprise. As computing power, networking and traditional NAS/SAN storage technologies matured in the late-1990's, centralized architecture, in the form of corporate data centers, received widespread adoption. However, as these data centers – even those built as recently as 5 years ago – have grown in complexity, scale and business importance, their efficacy has been severely challenged by an unprecedented avalanche of data in the form of digital content, unstructured data and other critical business information. As this well-documented crush of data continues to accelerate unabated (and with no end in sight), the need to remove bottlenecks associated with traditional data center architectures has reached a critical state.

Breaking the data center down to its three most elementary components gives us: Compute Power, the Network, and Storage. Recent advancements in the first two components have allowed for significant and important architectural shifts; from uncontrolled server sprawl to virtualized computing; from single gigabit network speeds to multiple gigabit transfer rates – all propelling the evolution of an improved centralized architecture. The final data center component, traditional NAS and SAN storage, has become the last bottleneck, stalling the full maturation of the centralized data center to its next and necessary evolutionary state.

This paper takes a holistic view of the data center, discusses each of these components individually, and explains how the crucial element of Storage has lagged behind advancements in Compute Power and the Network. It will also examine how an ongoing shift towards clustered storage architectures may help to drive storage forward and, perhaps, even enable it to leap-frog the other data center components in speed and performance. With the storage bottleneck removed, enterprises will be able to gain new insights, create meaningful IT operating leverage and transform information into business breakthroughs – truly realizing the potential of the next generation data center.

2. The Avalanche of Data

In the next few years IDC estimates a 6-fold expansion of what they call the "digital universe", from 161 exabytes in 2006 to an astronomical 988 exabytes in 2010, far exceeding expected storage capacity using current technology. The acute pain of this much data being created this fast is felt not only by those responsible for designing and managing corporate data centers, but has now reached into the business unit and executive levels of enterprises – catalyzing the impetus for change. However, according to some analysts, the onslaught has really only just begun.*[1]

What is perpetuating this massive flood of data? According to IDC it is primarily user-created files categorized as "unstructured data". From sound bites to presentations to digital images and videos, the ease at which data can be created, distributed and replicated is far beyond the imaginations of even a few years ago. However, this does not mean this data is unimportant to the corporate environment. In fact, it is quite the contrary. Most of these files are created by business users for business purposes and transmitted over business networks. IDC estimates that 80% of newly created (or replicated) unstructured data is traversing the corporate network and touching the corporate data center, if not actually being stored there. Just as crucial (if not more so), business systems themselves are churning out exponentially increasing amounts of

unstructured data in the form of CAD/CAM files, 3D modeling and simulation, VOIP recordings, massive digital images of hi-def surveillance, satellite, geospatial and scientific exploration, etc. Because this data is mostly unstructured, it does not fit easily into the structured systems upon which most IT organizations are built (with specific categories and rules of how, when and where data is transmitted and stored). The data is complex, large, and diverse yet still requires instantaneous access. This different type of data requires a different type of system than traditional technologies have afforded to this point.

The IT industry recognizes the need for a quantum leap in the evolution of technology to address these issues. In recent years, industry leaders have invested huge sums into R&D concerning these types of challenges and have experienced success in several areas. Though this is not meant to be comprehensive exposition, it is important to now take a closer look at each of these intertwined components and some of the major advancements which are propelling the evolution of the data center.

3. Compute Power

With Moore's law in full swing through the '90s, the need to address the growing mountain of data combined with the move from monolithic proprietary systems to inexpensive, powerful, commoditized hardware inadvertently led to the bane of IT environments everywhere, commonly referred to as "server sprawl". The costs and management headaches this problem represents are all too familiar. Hence, the dramatic shift toward server virtualization.

Server virtualization software brought major relief to server sprawl and, combined with powerful clustering technologies, brought us one step closer to realizing the next generation data center. Server clustering and virtualization allows an IT organization to pack more computing power into a smaller footprint through server consolidation (often at a 10:1 ratio), increase utilization from the average 10-20% per machine up to 80% (as virtual machines (VMs) share resources, each can run at a higher utilization rate), reduce power consumption (a function of the consolidation ratio), and often cut hardware and operating costs in half, all the while easing server management and increasing uptime. *[2]

However, a purely software-driven virtualization solution, with all the inherent benefits, still has limitations that affect most data center efforts. Traditional software VM solutions allow virtualization of x86 machines only, yet x86 technology and the associated operating systems were never built with virtualization in mind. This requires the virtual machine monitor (VMM) to either modify the OS or perform complex compute-intensive hand-off's between the VMM and the OS. Likewise, some operating systems are incapable of loading onto a VM at all, and traditional virtualization software does not support 64-bit operation of many operating systems, restricting them to 32-bit mode. *[3]

To address this issue, a major breakthrough has been made with hardware assisted virtualization technology, built specifically for the virtualized environment. Working with the major VMM software providers, Intel has optimized its line of multi-core processors for virtualization. With quad-core processor technology and 16 individual cores in a 4-socket machine, consolidation ratios of 24:1 or higher are common. And the advancement to 8-core is underway. But beyond simply being able to pack more VM instances into a single box and significantly increasing scalability, hardware assisted virtualization eliminates most of the aforementioned limitations of software-only solutions. Operating systems now have more direct access to the hardware, reducing the compute intensive translation previously required and producing near native performance benchmarks. Additionally, mixed 32-bit & 64-bit configurations are now possible, as well as cross platform solutions such as Linux running on Microsoft VMMs and Windows on top of Xen-based solutions.

In summary, due to software and hardware virtualization, coupled with powerful server clustering technology, the computing power available on today's industry standard components is limited only by monetary means.

4. The Network

The second core component of the data center, the Network, is logically divided into two realms: the internal backbone & LAN connections and the external wide-area connections. Ethernet and IP have seen a rapid and steady progression in switching speeds, from 100Mbps to now standard 1Gbps with 10Gbps also now being deployed. Additional new technologies, focused on improving performance and reducing latency, are emerging with the introduction of InifiBand (IB) which has seen a steady progression in switching speeds from 10Gbps to currently 20Gbps and soon 40Gbps. InfiniBand switching, powered by Mellanox silicon, is seeing large-scale adoption as a storage cluster interconnect and backbone technology, with some penetration into the front-end LAN as well. The advancements in both Ethernet and InfiniBand networking technologies bring the internal side of the network into the next phase of centralized computing.

The external wide-area connections, however, have been a veritable quandary. WAN speeds have been progressing, but in most cases not at a pace to keep up with business needs. The major problem, however, is not necessarily the actual WAN link speeds, it is the inherent characteristics of the applications traversing the network over TCP. In virtually all cases, merely increasing the size of the WAN link will not appreciably improve the performance of these applications. TCP was developed in the 1980s, and though it has proven impressively versatile and scalable, the networking protocols built on top of it and still in use today, such as FTP, CIFS (Common Internet File System), NFS (Network File System) and even proprietary protocols, were simply not created for and perform poorly over the WAN. *[5]

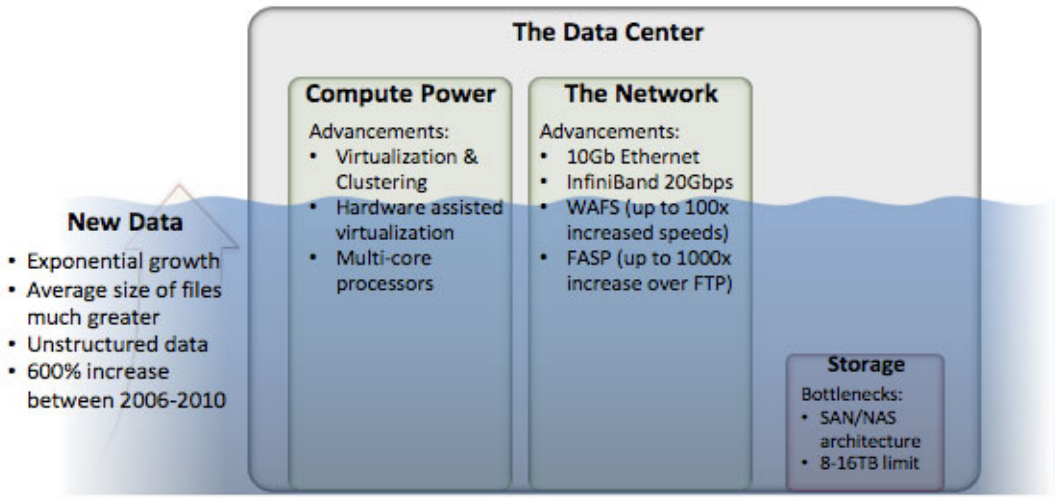
Given these inherent limitations, without a breakthrough technology, the current growth of data will easily choke the current pipelines. Enter Wide Area File Services (WAFS). WAFS focus is to solve this very issue. Two notable companies that have been making headway in this area with impressive metrics are Cisco with their "WAAS" product and Riverbed with "WDS". Both companies provide excellent solutions for WAN optimization and a slew of additional products to effectively tune and tone WANs to meet enterprises' exact needs. These WAFS products boast network utilization reductions of up to 95% and transfer rates of up to 100x current (with 5x to 20x standard). *[6]

Another breakthrough technology in the WAN space, addressing the inherent problems in FTP, is Aspera's fasp solution for high-speed large file transfers. This revolutionary technology is far beyond typical TCP tuning and acceleration, providing increases of up to 1000x that of FTP – regardless of the transfer medium – with predictable and controllable flow rates.*[7]

Though it will take some time for organizations to adopt these breakthrough technologies, with such performance enhancements available, the WAN path is quickly being cleared for the continued avalanche of digital bits.

5. The Storage Crisis

IDC estimates that the digital universe will increase at a CAGR of 57% while, given traditional storage technologies, storage is only able to increase at CAGR of 35%.*[1] This means that starting last year in 2007, there is more data being created than there is storage space to hold it, and it looks like it will only get worse without a major breakthrough and widespread acceptance of new storage technology. But volume alone, though tremendous, is not the underlying problem. Without the proper technology in place to both store and utilize this flood of data, what was once the lifeblood of the business could become the very thing that paralyzes it. Therefore, the answer to the storage crisis is not how to stop the flood, or even how to simply make room for the data, but instead how to enable the utilization of this rich, complex, unstructured information in a way that propels business forward.



The last major shift in storage architectures was the introduction of network storage, or SAN and NAS, which abstracted storage from the server and brought relief to the myriad of problems traditional direct attached storage (DAS) presented. NAS became mainstream in the late 1990's and has remained virtually unchanged, aside from increases in speed and capacity. Why is that not enough? For the same reason that simply increasing processor power and adding more servers to the network in the '90s was not enough. In fact, the very same fundamental architectural challenges that led to server sprawl have led to "storage sprawl". Yet the remedy for server sprawl -- server virtualization -- has done little to alleviate the storage crisis. Some may even argue it has only exacerbated the problem. Server virtualization allows for multitudes more servers and applications per square foot in the data center, yet each server still requires its own slice of storage, leading to ever more volumes, LUNs, RAID arrays, and file systems, and therefore, acute pain in storage management. Beyond the storage sprawl problem, NAS technology was engineered primarily for small file data, not large, diverse, unstructured data. The architecture of NAS is simply not robust enough for the avalanche of unstructured data flooding the data center and does not meet the performance requirements of this data.

The well known 8-16TB limit of NAS and SAN file systems is at the crux of this storage crisis. Though seemingly ample when introduced, it has since led to the creation of "islands of storage" in nearly every large organization, sometimes requiring hundreds, if not thousands of separate file systems and directories, each representing a single point of management and of potential failure.

Organizing, navigating, and backing up these "system imposed" directories requires the allocation of precious IT resources and creates significant management headaches. In addition to the scalability limitations, when pushing NAS scalability to its limits, reliability and availability issues emerge. Yet even if the 16TB file system limits are somehow addressed through inventive ideas such as switch-based virtualization or veneer approaches like global namespace software, the underlying architecture and performance bottlenecks of NAS still persist – limiting the usefulness and value of new data.

What is to be done then? Interestingly, the answer can be found by looking at the advancements in the compute power component. If server virtualization and clustering quelled server sprawl and paved the way for major advancements in data center computing utilization, performance and scalability, then employing storage virtualization should correspondingly affect the storage crisis. As previously discussed, for server virtualization to really leap forward, a foundational change was required in both hardware and software. Likewise, truly breakthrough storage virtualization technology will need to incorporate both a hardware and intelligent software approach.

So, it would seem logical that a redesign of the data center's fundamental storage architecture is in order. Yet it would be errant to negate the advancements made in commoditized hardware by employing some rigid, monolithic, proprietary system, regardless of the promised performance gains. Therefore, the optimum solution would be one that utilizes industry standard hardware powered by sophisticated virtualization software in a way that overcomes, versus just covers up, the limitations of NAS for a truly virtualized, scalable, easy to manage, single unified storage platform with unmatched performance, scalability and flexibility.

Headway is indeed being made in this arena with the advent of "clustered storage". True clustered storage applies a fully symmetric architecture and distributed file system (not simply N-way load balanced failovers) which eliminates nearly every limitation imposed by NAS and SAN and paves the way for the clearly defined avalanche of data that continues to challenge enterprise data centers. Just as server virtualization was a revolutionary new way to structure the compute component of the data center, clustered storage is an entirely new storage architecture, not simply a patch or fix to an out-dated model.

When defining clustered storage solutions five common characteristics emerge:

- Symmetric Clustered Architecture
- Scalable Distributed File System
- Inherent High Availability
- Single Level of Management
- Enterprise Ready

Symmetric Clustered Architecture: The key design principle behind clustered storage is symmetry among the nodes (which can be thought of as self-contained storage controller heads with disks, CPU, memory, and network connectivity). The tasks the cluster must perform are distributed uniformly across its members, enhancing scalability, access to data, performance and availability. This is in contrast to traditional storage architectures deploying master server-based approaches where the storage nodes are not symmetric and are inherently limited in scalability and performance.

Even as more nodes are added to the cluster, it still has one logical brain. Regardless of the number of nodes in the solution, there is still only one logical system. Fully symmetrical clustered architectures are able to grow resources seamlessly with a "pay-as-you-grow" model. When more memory, bandwidth, capacity, or drive actuators are needed, the cluster can be grown by simply

adding nodes to the cluster, which maintains its coherency as one logical, dynamically expandable system.

Scalable Distributed File System: The enabler of this architectural approach is a distributed file system (or operating system) that can scale to be a very large pool of storage or single network drive. Distributed file systems maintain control of file and data layout across the nodes and employ metadata and locking semantics that are fully distributed and cohesively maintained across the cluster, enabling the creation of a very large global pool of storage. A single network drive and single file system should be able to seamlessly scale to hundreds of terabytes.

Distributed clustered storage solutions have the unique capability to scale all performance elements in a near linear fashion. When more nodes/controllers of memory, processing, disk spindles and bandwidth are added, the cluster maintains its coherency as one logical system and is able to aggregate across all resources, thereby achieving linear scalability of performance with each additional node. In order to achieve this linear scalability of performance, it is critical for each node to remain in sync with all other nodes in the cluster. As a result, more robust solutions typically employ very high-speed intracluster interconnects to ensure low-latency between the nodes and real-time synchronization of the cluster.

Inherent High Availability: A distributed clustered architecture by definition is highly available since each node is a coherent peer to the other. If any node or component fails, the data is still accessible through any other node, and there is no single point of failure as the file system state is maintained across the entire cluster. Moreover, high availability is “inherent” for distributed clustered architectures, meaning that unlike traditional storage systems, where an IT manager would have to purchase additional software and expensive redundant hardware in order to achieve high availability, clustered storage solutions achieve high availability by the very nature of the fully symmetrical architecture.

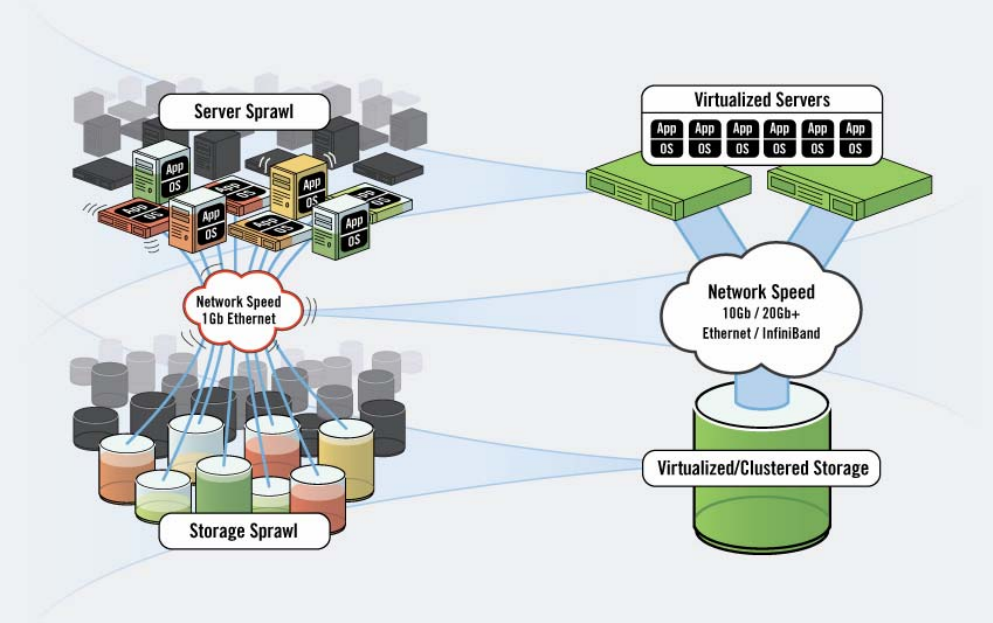
Single Level of Management: Distributed clustered storage solutions provide a single level of management regardless of the size of the file system and number of storage nodes added to the cluster, making it as easy to administer a cluster size of a few nodes as it is to manage a cluster of several hundred nodes. Complete clustered storage solutions automate traditionally manual tasks, including the load balancing of client connections across nodes in the cluster to ensure optimal performance and the automatic re-balancing of content when new nodes are added to the cluster to scale capacity and performance.

A single file system spanning the entire cluster simplifies the management of the environment and eliminates the task of navigating through many volumes and network drives as well as the painful mapping of applications to many separate “islands of storage”. For system administrators, this eliminates client-side management issues because all of the files belong under one single file system, volume or mount point, again, not because of the veneer of a namespace aggregation approach, but as a function of the actual architecture.

Enterprise Ready: Distributed clustered storage solutions must be enterprise ready in order to function well in large-scale data center environments. In order to be part of the next generation data center, clustered storage solutions must support standard network protocols and provide advanced enterprise software tools for customizing and addressing specific business needs. The ability to plug into existing data center ecosystem players, such as networks, backup/restore infrastructures and heterogeneous application/OS environments is paramount. Furthermore, effective clustered storage solutions must provide a suite of software applications and functionality to ensure the proper data protection, data management and data access expected of an enterprise class storage system.

6. Breaking the Bottleneck

VMware fueled the revolution in virtualized server architecture; Intel is blazing the way in computing power; Mellanox, Cisco and Riverbed, along with a host of other inventive companies, are clearing and broadening the network paths – all are aimed at preparing the enterprise data center for the current and future flood of file-based data and other critical business information. In the same manner, Isilon Systems has been pioneering revolutionary new clustered storage solutions in order to break the last bottleneck, reverse storage sprawl and bring data center storage architecture into alignment with the advancements already evident in the compute and network elements of the data center.



Isilon's clustered storage solution (Isilon IQ) – first to market and now in its fifth generation – is clearly the market and technology leader. In fact, Isilon IQ is the only hardware/software combination of truly clustered storage using industry standard components that will completely address storage sprawl and enabling businesses to utilize and leverage the well-documented avalanche of data heading toward the data center.

Isilon has defined a revolutionary new clustered storage architecture with the purpose of creating not just a bigger, faster, more resilient system but with the underlying goal of sustainable linear scalability. Our current solution achieves 20GBps of aggregate throughput and is scalable to more than 2.3 petabytes within single file system. As technology continues to advance, our systems are built with inherent flexibility and forward compatibility, meaning no rip-and-replace just because business needs change or hard-drive densities increase.

The Key Ingredient: OneFS® Operating System/Distributed File System

To address the scalable distributed file system requirement of clustered storage, Isilon built and delivered the revolutionary OneFS operating system, which combines three layers of traditional storage architectures – the file system, volume manager and RAID – into one unified software layer. This creates a single intelligent, fully symmetrical file system, achieving up to 20GBps of total throughput, which spans all nodes within a cluster.

File striping in the cluster takes place across multiple storage nodes versus the traditional method of striping across individual disks within a volume/RAID array. OneFS provides each node with knowledge of the entire file system layout and stripes files across nodes within a single cluster. Accessing any independent node gives a user access to all content in one unified namespace, meaning that there are no volumes or shares, no inflexible volume size limits, no downtime for reconfiguration or expansion of storage and no multiple network drives to manage. This fully distributed approach enables the breakthrough technology required to meet the performance, scalability, availability and manageability demands of the next generation data center.

Inherent High Availability & Reliability

Traditional file systems use a master/slave relationship to manage multiple storage resources. Such relationships have intrinsic dependencies and create points of failure within a storage system. The only way to truly ensure data integrity and eliminate single points of failure is to make all nodes in a cluster peers. Because each node in an Isilon cluster is a peer, any node can handle a request from any application server to provide content. If any particular node were to fail, any other node could fill in, thereby eliminating any single point of failure.

OneFS has further increased the availability of Isilon's clustered storage solution by providing multi-failure support of n+4. Simply stated, this means an Isilon cluster can withstand the simultaneous loss of an unprecedented 4 disks or 4 entire nodes without losing access to any content – and without requiring dedicated parity drives. No other storage vendor or technology delivers up to n+4 data protection, which is the key enabler to safely building 1PB+ single file systems. Additionally, self-healing capabilities greatly reduce the chances of a production node failing in the first place.

Single Level of Management

An Isilon cluster creates a single, shared pool of all content, providing one point of access for users and one point of management for administrators with a single file system or volume of up to 2.3 petabytes. Users can connect to any storage node and securely access all of the content within the entire cluster. This means there is only a single relationship for all applications to connect to and that every application has visibility and access to every file in the entire file system (per security and permission policies, of course).

Linear Scalability in Performance & Capacity

One of the key benefits of an Isilon IQ cluster is the ease with which it allows administrators to add both performance and capacity to an Isilon cluster without downtime or application changes. System administrators simply insert a new Isilon IQ storage node, connect the network cables and power up. The cluster automatically detects the newly added storage node and begins to configure it as a member of the cluster. In less than 60 seconds, an admin can grow a single file system by 2-12 terabytes. Isilon's unique modular approach offers a building block, or "pay-as-you-grow", solution so customers aren't forced to buy more storage capacity than is needed up front. Unlike existing systems, the modular design of an Isilon cluster also enables customers to

incorporate new technologies in the same cluster, such as adding a node with higher-density disk drives, more CPU horsepower or more total performance.

Another enabling technology that allows an Isilon cluster to reach breakthrough linear scalability of performance is the use of InfiniBand as the high–speed, low-latency intracluster interconnect. A backend InfiniBand switch provides an Isilon IQ cluster with the right performance and low-latency characteristics to keep the nodes in sync, allowing for optimal overall cluster performance.

Enterprise Ready

Now in its fourth generation, Isilon has delivered on the features that meet the requirements for integration into the commercial enterprise and the next generation data center. Isilon clusters are built to work in a wide array of existing environments without the use of any proprietary tools or protocols. Industry standard file-level network protocols (i.e. NFS, CIFS, FTP, HTTP, SNMP, NDMP) allows Isilon IQ to easily interoperate with existing systems. Additionally, Isilon supports the standard enterprise Windows and UNIX authentication and permission schemas that are standard requirements for deployment consideration.

7. Summary

There is no question that the data center and the IT professionals who manage them are experiencing, and will continue to experience, an unprecedented explosion of information in the form of file-based and other unstructured data. The data centers of yesterday are simply not equipped to handle this massive influx without incorporating recent breakthrough technologies in computing power, networking speeds and, now, storage.

Isilon Systems is the leader in the development and implementation of clustered storage solutions for next generation data centers. Underpinning Isilon's focus and innovations in clustered storage is a belief that storage should not be a problem, but an opportunity to propel critical enterprise information back into the bloodstream of business. With the three core elements – Compute Power, the Network and Storage – now aligned, the corporate data center is entering into its next and necessary evolutionary state. By embracing these advancements, enterprises are well positioned to gain new insights, transform information into business breakthroughs and create IT operating leverage, all while dramatically reducing cost and complexity – truly realizing the potential of the next generation data center.

About Isilon Systems

Isilon Systems is the worldwide leader in clustered storage systems and software for digital content and unstructured data, enabling enterprises to transform data into information and information into breakthroughs. Isilon's award-winning family of IQ clustered storage systems combines Isilon's OneFS® operating system software with the latest advances in industry-standard hardware to deliver modular, pay-as-you-grow, enterprise-class storage systems. Isilon's clustered storage solutions speed access to critical business information while dramatically reducing the cost and complexity of storing it. Information about Isilon can be found at **http://www.isilon.com**.

References:

*[1] IDC Whitepaper “Expansion of the Digital Universe” March 2007
*[2] www.vmware.com/solutions/consolidation/consolidate.html
*[3] www.intel.com/business/technologies/virtualization.htm?iid=technology_virtualization+rhc_server
*[4] www.cisco.com/en/US/netsol/ns708/networking_solutions_solution_segment_home.html
*[5] www.broadnets.org/2004/workshop-papers/Pathnets/03_TCPHighSpeedWAN-SylvianRavot.pdf
*[6] www.Riverbed.com/solutions/
*[6] www.cisco.com/en/US/products/ps6870/index.html
*[7] www.asperasoft.com/technology/index.html