

## Clustered Computing Architectures Fuel the Need for Clustered Storage

October, 2007



We are witnessing an architectural sea change away from monolithic systems to clustered and distributed based computing models. Over the past five years, technical computing applications have crossed over from scientific and national laboratories and have begun to find their way into commercial usage with increasing frequency. Compounding that effect, the popularity of server virtualization and web-based applications practically ensure that distributed and clustered based application architectures will be a mainstay of the enterprise data center going forward.

Furthermore, IT has clear economic incentive to deploy these clustered computing models. In an environment of tighter fiscal controls, IT is reticent to purchase high cost, monolithic servers for their most demanding data center computing needs. Server clustered architectures offer superior raw computing and scalability at a price unmatched by monolithic systems. As a result, for their most computational and highest scale applications, IT is increasingly choosing to replace monolithic systems with loosely coupled clusters of one to four CPU nodes running the Windows, Linux, or UNIX operating systems.

However, these distributed and clustered applications impose a new set of requirements on the underlying storage infrastructure. To that end, we believe that these types of clustered applications are reshaping how storage systems are designed and managed. Clustered storage has emerged as the logical response to meet the high I/O throughput, high capacity demands of a clustered, distributed server infrastructure. In the following profile, we will explore the ramifications of this computing trend on storage system design, and discuss the fundamental issues with monolithic file storage. With that as context, we will then posit four evaluation criteria that users should use to assess storage systems and compare how Isilon Systems, a leading provider of clustered file storage, stacks up against these criteria.

### Clustered Computing Architectures Crossover to Mainstream

One of the more interesting developments in the last five years is the rapid rise of clustered and distributed applications. Traditionally, to scale a system, IT would elect to scale up – purchasing a larger, more powerful

multiprocessor server to deliver higher levels of computing horsepower to the application. However, monolithic, multi processor servers have given way to scale out approaches for many applications in the data center.

Three examples of these clustered computing architectures are apparent in the data center today – web applications, consolidated,

**S O L U T I O N P R O F I L E**

virtualized server environments, and commercial technical computing applications. First, with the rise of the Internet, web based applications have replaced client server as the dominant application architecture today. Moreover, the Internet enables a level of scale and reliability previously unimaginable. Storage systems must be able to handle 1000s of concurrent users, expand seamlessly to support 100s of online TBs and meet 24x7 uptime requirements.

Second, through the use of server virtualization tools like VMWare, IT has consolidated underutilized servers and their applications onto a centrally managed pool of servers running multiple virtual operating environments. As the scale of these virtual environment increases, the performance, availability, and cost-efficiency of the storage infrastructure become crucial to enabling the next level of scalability of a virtualized server environment. Virtual server environments requires a low latency, high IOPS storage system to deliver overall rapid application performance.

Third, technical computing applications have crossed over from scientific research organizations to become core mission-critical applications responsible for key business functions. These clustered technical computing applications loosely couple potentially 1000s of cost-effective x86 servers together to amass tremendous amounts of computational performance to solve some of the most challenging problems. Examples of these applications abound across a wide variety of industries – econometric and Monte Carlo simulation

clusters in financial services, blast clusters in biosciences, seismic processing applications in Oil and Gas exploration, video rendering in movies and TV production, and CAD/CAM applications in manufacturing. Most technical computing applications are voracious consumers of storage capacity and I/O throughput – making monolithic systems a poor fit.

The common theme among these clustered computing applications is that they must share data and that the performance of the storage is a potential bottleneck to the overall responsiveness and throughput of the application or business process. In short, monolithic storage cannot meet the high I/O throughput demands, fault tolerance, and seamless capacity scaling capabilities that these new class of applications require. As these distributed and clustered applications have gained a foothold in corporate data centers, they have caused IT to rethink their storage backbone and pulled clustered storage into the heart of the data center.

### **The Problems with Monolithic NAS Designs**

Traditionally, all major NAS vendors have adopted a scale-up approach to delivering NAS performance and capacity. For their main filer lines, both Network Appliance and EMC use a single head design connected to a set of physical disk. This monolithic approach means that all I/O to a given file system must be processed by a single head that resides in the data path. To scale, NAS vendors sold bigger, more expensive monolithic NAS systems that could handle

**S O L U T I O N P R O F I L E**

greater I/O rates and capacity. However, this scale-up model creates three core problems:

**Problem #1: Difficult to Scale the Datapath** – A well tuned, single headed, traditional NAS system can deliver at best 250 MB per second of I/O throughput to a single file system. However, with consolidated workloads and new distributed applications that require more than 250 MB/s of throughput, traditional NAS systems hit a performance wall. In order to scale performance further, file data must be manually distributed across multiple filers. However, by doing this, administrators lose a key benefit – the ability to manage all file data under a single namespace. Therefore, to get the required performance, administrators must sacrifice manageability of their overall unstructured data environment – a jagged little pill for most to swallow.

**Problem #2: Large Capacity Systems are Costly and Unwieldy to Manage**– Each filer has limits to how much capacity can be accessible by a single system. Therefore, if data grows faster than expected, an administrator may need to do a “forklift” upgrade of the system and purchase a new larger, more expensive NAS system to meet his capacity needs. This “forklift” upgrade approach not only causes downtime as the system is upgraded and data migrated, but it also requires large upfront capital commitments, instead of small incremental purchases of capacity.

**Problem #3: No Economies of Management across Filers** – NAS filers are often commended for their ease of use and simplicity of operations. However, no

matter how easy the overall system is to manage, as the number of filers proliferates in a data center, the amount of administration necessary increases proportionately. Each filer introduces a separate namespace and set of file data that must be provisioned, protected, replicated, and maintained individually. Hence, a common user refrain is: “I loved my first filer, but hated my tenth.”

## **Clustered Storage Defined**

Clustered storage was born out of the pain and suffering from these fundamental issues with monolithic NAS storage. Through the use of advanced file systems, clustered storage systems aggregate the performance and capacity of a set of disparate controllers and physical disks to create a single scalable, fault tolerant networked storage resource. An outgrowth of this approach is that the system presents a single shared namespace accessible via common file access protocols, like NFS, CIFS and HTTP. The clustered storage system’s namespace provides a single file and directory structure for all unstructured data to be centrally managed.

From our vantage point, clustered storage offer four crucial advantages over traditional monolithic storage:

- **I/O Scalability** – The ability not only to support the most I/O intensive applications, but also consolidate multiple workloads onto a single scalable resource for all storage requirements.
- **Single Point of Management** – Superior management through

**S O L U T I O N P R O F I L E**

virtualization of physical nodes and storage resources under a single shared namespace.

- **Fault Tolerance** – The ability to survive multiple node, disk, and network failures, while continuing to provide access to the stored content.
- **Investment Protection** – The ability to grow the storage infrastructure and capacity over time, without the need to make large upfront purchases in systems that may be obsolete before they are deployed.

### **Four Storage Requirements Created by Clustered Server Computing**

After speaking with many end users who have deployed clustered server computing environments, there is a pattern of key criteria that any storage system must meet to be deployed alongside a clustered server computing environment. It is clear to us that this new class of application architecture requires a new level of scalability, performance, availability, and simplicity from the underlying storage infrastructure. We will now review the four evaluation criteria of storage systems best suited for deployment alongside clustered computing environments:

1. High I/O Performance
2. Fault Tolerance
3. Extreme Capacity Scaling
4. Turnkey Ease of Use

#### **Criteria #1: High I/O Performance**

Clustered server architectures place a heavy premium on high performance access to shared file data. Given the fact that 1000s of server nodes may potentially access shared data concurrently in a clustered application, storage systems can often be the bottleneck or gate to application performance. On the flip side, a well suited, high performance system can actually be the single biggest determinant of application performance in a clustered server deployment.

To that end, storage systems that are well suited for these clustered computing environments must be able to sustain high IOPS with millisecond response times. In many cases, clustered applications require I/O throughput rates in excess of 250 MB/s and oftentimes a large compute farm will need 1 GB/s of I/O throughput. A traditional, single headed NAS devices can only deliver approximately 250 MB/s of throughput, making them ill suited for the performance rigors of clustered-based computing architectures.

#### **Criteria #2: Fault Tolerance**

Fault tolerance and availability become paramount concerns in a clustered computing environment. For example, web applications must be available 24x7 in this global business environment. A storage failure where data is inaccessible for any sustained period of time can result in lost revenue, damage to reputation, or lost employee productivity. Moreover, a storage failure during a long running technical computing job might mean the loss of days or weeks worth of processing and force a corporation to restart the process again from the beginning. Given that many technical

**S O L U T I O N P R O F I L E**

computing applications today are business-critical, a storage failure can result in significant delays in terms of time to market and put an organization at a competitive disadvantage.

Users must evaluate storage systems on two primary dimensions – resiliency to disk failures and resiliency to controller failures. To that end, users should scrutinize storage systems to understand how they handle cascading failure scenarios and how long it takes recover and rebuild a disk once a corruption has been detected.

**Criteria #3: Extreme Capacity Scaling**

Since many of these clustered server applications are voracious consumers of storage capacity, users place a high degree of importance on the ability to incrementally scale the underlying storage non-disruptively and ensure that data is striped across all available controllers and spindles for maximum performance. In fact, it is commonplace to find 100s TB storage environments in technical computing application environments, even 1 PB deployments are not out of the question.

With high growth rate in storage capacity, users demand a storage system that gives them investment protection. They want to be able to “pay as they grow,” but also be assured that they can meet their storage capacity requirements for today and in the future. In this context, two critical capabilities of a storage system are simple, non-disruptive expansion of storage capacity and the ability to manage all storage capacity under a single namespace for economies of administration.

**Criteria #4: Turnkey Ease of Use**

With traditional storage systems and tools, storage management can be a laborious, error prone process. In fact, the entire storage administration paradigm does not scale well when administrators are asked to manage 100s of TBs if not PBs of storage. In short, storage administrator needs systems that will give them economies of management by automating repetitive tasks and enabling them to manage more storage than they could before.

Administrators require a system that is intuitive to learn and streamlines arduous storage management tasks in a matter of minutes, not hours or days. For example, given the frequency that capacity is provisioned in a clustered application environment, the speed and ease at which capacity can be added and be made available to a fast growing application is crucial. Therefore, time to usable capacity becomes a key operational metric in a clustered application environment.

In short, users in these clustered computing environments need storage systems that are easy to install, easy to provision, self tuning, and managed as a single unit, even if it is modularly expandable.

**Spotlight on Isilon**

Isilon has emerged as a leader in the clustered storage market by riding this market shift toward clustered server computing architectures and the requirements that they engender for highly scalable, high performance storage systems. Isilon has evolved its product by focusing on solving the problems of a high scale capacity,

**S O L U T I O N   P R O F I L E**

high performance environments. It originally focused on the most challenging storage problem in the digital entertainment market, but has subsequently broadened its product and focus to cover all cluster computing applications and markets. For example, Isilon focused on applications like video rendering and video production before branching out into a number of industries and applications which rely on file-based data as a core component of their business, including oil and gas exploration, life sciences and drug discovery research, manufacturing, telecommunications, the federal government, large scale Web 2.0 applications and businesses. As clustered server computing has crossed over to the core enterprise, so has Isilon.

An Isilon IQ clustered storage system features a set of clustered platform nodes interconnected on the backend by InfiniBand. The Isilon product line is comprised of the Isilon IQ 200, 1920, 3000, 6000, 9000, and 12000 platform nodes and the EX 6000, EX 9000, and EX12000 expansion nodes. The IQ200, 1920, 3000, 6000, 9000, and 12000 nodes scale both throughput and capacity in a linear fashion. The Isilon EX 6000, 9000, and 12000 allow users to cost-effectively add capacity without adding I/O throughput. Conversely, the Isilon Accelerator node allows users to cost-effectively add I/O throughput without adding capacity.

At the heart of an Isilon clustered system is Isilon's OneFS file system – a distributed file system that presents a single namespace of file data across the entire cluster, performs volume management and RAID protection

across the cluster. The namespace presented by OneFS is accessible and shared via industry standard protocols such as NFS, CIFS, HTTP, and FTP. Moreover, Isilon features robust data management capabilities such as quotas, snapshots, automated load balancing, and asynchronous replication that can be used across the entire cluster. As a result, Isilon systems appear as a single unified pool of file storage to clients, even though the system is composed of multiple modular units.

We will now turn our attention to assessing how Isilon compares against the four criteria that we previously established as important for clustered server computing deployments.

**Criteria: Isilon & High I/O Performance**

Isilon provides a modularly scalable clustered file storage system that can deliver up to 10 GB/s of I/O throughput – more than enough horsepower for most clustered applications today. Furthermore, Isilon also provides the ability to scale both capacity and I/O performance independently. This gives users flexibility if their application needs more I/O throughput than capacity or vice versa. Isilon allows users to add IQ Accelerator nodes to a cluster. IQ Accelerators do not add additional capacity to an existing cluster, but allow a user to scale I/O throughput access. This scaling flexibility allows users to personalize their Isilon deployment to meet their performance, capacity, and cost requirements – ensuring that end users maintain an ideal price to performance to capacity ratio at all times.

From our vantage point, Isilon has proven its ability to scale the data path given the

**S O L U T I O N P R O F I L E**

presence of numerous referenceable high scale, high I/O throughput deployments. End users are the ultimate barometer as to whether a product works as advertised in real world applications. Isilon can point to a litany of success in digital entertainment, oil and gas, life sciences, web 2.0 and other clustered computing markets.

**Criteria: Isilon & Fault Tolerance**

One of the hallmarks of a clustered storage system is its ability to recover from multiple failures. An Isilon system provides built in resiliency for up to four simultaneous node failures—N+4 protection. Moreover, Isilon ensures that all data remains 100% available and is rebuilt to any free space throughout the cluster—capacity is not lost to “recovery drives” and failed drives can be rebuilt in less than one hour. In short, recovery and fault tolerance has been engineered in from the ground up and the recovery process is automatic once a failure occurs. The Isilon cluster will detect, and proactively rebalance or rebuild the system in the background while data continues to remain online and accessible.

**Criteria: Isilon & Extreme Capacity Scaling**

Isilon allows users to adopt a “pay as they grow” model to scaling their capacity. Users can start with a small cluster of storage and incrementally grow their system up to 96 nodes and to 1.6 petabytes of capacity in a single cluster. Moreover, Isilon allows users to grow this capacity under a single namespace. Therefore, users can create up to 1.6PB file systems, enabling users to have a single global view and management point of all their storage. This point cannot be understated. A single namespace

dramatically simplifies overall storage management and allows users to get economies of administration in these extreme capacity scaling environments.

As mentioned, provisioning can be a laborious and repetitive process for storage administrators managing 100s of TBs of fast growing capacity. Isilon allows an administrator to add new nodes to their existing cluster in as little as 60 seconds, enabling the rapid and painless deployment of storage capacity.

**Criteria: Isilon & Turnkey Ease of Use**

Isilon has developed a reputation throughout the industry for maniacal focus on designing simplicity and ease of use into its products. For example, end users have reported having 100 TB Isilon system up and running in 10 minutes once racked and cabled. This speed of installation and configuration speaks to the simplicity of the user experience. Traditional storage system can take days to weeks to setup and properly configure before a single file system is presented. Isilon has invested heavily in making sure that users can get plug and play capacity insertion without complex reconfiguration. This simplicity of operation differentiates Isilon from traditional storage systems and many of its clustered file storage brethren.

**Taneja Group Opinion**

We are witnessing the emergence of clustered architectures at both the server and storage level. Although clustered computing environments are farther along in their progression and maturity, we believe that there is a symbiotic relationship between clustered server computing and clustered

**S O L U T I O N P R O F I L E**

storage. In fact, we think one needs the other and vice versa. At present, clustered server computing applications are crossing over into the mainstream and driving the requirements for storage infrastructure. This transformation is still in the early stages, but we believe it to have profound ramifications on storage system design and end user purchasing patterns in the future.

Clustered storage is making the leap with these clustered applications into the heart of the data center. Isilon recognized this transformation early on and has designed and developed a clustered file storage system to take advantage of this transformation. If you're in the market for a high I/O throughput, extreme capacity scaling storage system, we think that Isilon should be on your short list.

---

*NOTICE: The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.*