



# The Clustered Storage Revolution

---

Defining the Paradigm Shift to Clustered Storage

An Isilon Systems Technical Whitepaper

March 2006

## Table of Contents

<b>1</b>	<b>Introduction.....</b>	<b>3</b>
<b>2</b>	<b>Three Macro Trends Driving the Clustered Storage Revolution .....</b>	<b>4</b>
<b>3</b>	<b>A New Category of Storage: Clustered Storage .....</b>	<b>7</b>
<b>4</b>	<b>Clustered Storage Defined.....</b>	<b>9</b>
<b>5</b>	<b>Isilon Systems: The Leader in Clustered Storage .....</b>	<b>11</b>
<b>6</b>	<b>Conclusion .....</b>	<b>16</b>

## Figures

<i>Price history of hard disk products versus year of product history</i>	6
<i>Categories of Storage</i>	7
<i>2-way simple failover clustering</i>	7
<i>Namespace Aggregation</i>	7
<i>Differences between the types of Clustered Storage in how data is controlled</i>	8
<i>Scalable Distributed File System</i>	9
<i>Differences in types of Clustered Storage</i>	10
<i>Isilon IQ: OneFS</i>	11
<i>Isilon IQ: Network Architecture</i>	12
<i>Average Disk Rebuild Time</i>	13
<i>Isilon IQ: Linear Scalability Throughput</i>	14
<i>Isilon IQ: Infiniband Backend</i>	15
<i>Isilon IQ: Product Family</i>	15

## 1 Introduction

**“Clustered Storage** is becoming pervasive and is a major paradigm shift from previous generations of storage products, much like when CDs made records obsolete.”

~ Tony Asaro, Enterprise Strategy Group, Oct 2005

In 1962, Thomas Kuhn published a groundbreaking book entitled *The Structure of Scientific Revolutions*. He argued that progress of science is not gradual but a kind of punctuated equilibrium with moments of epochal change, much like our understanding of biological evolution. The computer industry experienced such a revolution when IBM introduced the standardized architecture of the IBM personal computer in 1981. In a huge departure from previous industry practice, IBM chose to build its computer from off-the-shelf components. As a result, the IBM personal computer architecture became the standard. This architecture came to displace more than just other personal computer designs, though — over the next few decades, the design for minicomputers and mainframes also changed to fit the IBM standard and began to be built from off-the-shelf components.<sup>(1)</sup>

The purpose of this whitepaper is to introduce you to a new paradigm shift that is currently taking place in the data storage industry: the movement toward **Clustered Storage** architectures. Distributed storage clustering for the data storage industry is in much the same position today that IBM was in 1981, poised to change the rules of the computer industry. *Clustered Storage architectures are changing the rules of how data is stored and accessed.* In this paper we will discuss the trends that clearly define clustered storage architectures as the future of data storage, detail the requirements of this new category of storage, and introduce the Isilon® IQ clustered storage solution which is the first to deliver on the promises of this paradigm shift.

## 2 Three Macro Trends Driving the Clustered Storage Revolution

The movement toward Clustered Storage architectures is being driven by three macro trends:

- Explosive growth of unstructured data and digital content
- Paradigm shift to cluster computing
- Proliferation of cheaper and faster industry-standard enterprise-class hardware



### **Macro Trend 1: Explosive Growth of Unstructured Data and Digital Content**

Today's competitive companies are facing a tremendous increase in the amounts of data used to conduct their everyday business, driven largely by the explosion of unstructured data. IT managers know that applications using and storing video, audio, images, research sets, and other large digital files and unstructured data are pushing the bounds of traditional storage system capacity and performance.

Pratt & Whitney is familiar with explosive growth of unstructured data. As the global leader in the design, manufacture and support of the world's leading commercial and military aircraft and spacecraft engines, they conduct exhaustive testing that generates many terabytes of engine test data, with each high-bandwidth test recording more than 100,000 samples per second. Cedars-Sinai

Cancer Research Center in Los Angeles, CA; a cancer research center that combines data from various sources, including clinical mass spectroscopy and genomic data, also knows the challenges of storing large sets of research data. For Cedars-Sinai, a single drop of blood generates more than 60 gigabytes of unstructured data for proteomic studies. Multiply this by the hundreds, if not thousands, of blood samples taken from the Center's patients, and the massive growth of unstructured data is quite evident. Finally, in 2004, Sports Illustrated, for the first time, instituted a 100 percent digital workflow at the 2004 Summer Olympics in Athens, generating over 250,000 digital images (with average image size between 18 to 24 megabytes) through 17 days of events. Extend this digital trend across the landscape of industries that use unstructured data and digital content, including media and entertainment, digital imaging, life sciences, oil and gas, manufacturing, and government, and the explosive growth of unstructured data is clear.

According to forecasts from the Enterprise Strategy Group (ESG), reference information will represent 58 percent of new corporate and government information by the end of 2006. ESG defines reference information as "digital assets retained for active reference and value." This includes, but is not limited to, electronic documents, CAD/CAM designs, historical documents, medical images, bioinformatics, geophysical data and voice data. ESG expects reference information to grow at a 92 percent compound annual growth rate (CAGR). Meanwhile, migration reference assets (i.e. data that is migrated from tape to disk-based storage resources) are expected to account for an additional 420 Petabytes of capacity during that time.

---

**For Cedars-Sinai a single drop of blood generates more than 60 gigabytes of unstructured data**

---

So what does this mean for the IT manager? As unstructured content stores increase in size and complexity they are straining traditional storage systems, which were designed primarily for *structured data* with small file sizes and high transactions such as relational databases and email servers. *Unstructured data*, on the other hand, has unique characteristics for which traditional storage systems were not designed, including large file sizes and data volumes, high throughput requirements, read intensive access patterns and high concurrent file access.

For lack of better alternatives, many companies have attempted to meet their needs for unstructured data by extending traditional storage systems designed for structured transactional or text-based data. Even the newest NAS and SAN systems employ architectures with inherent limitations that lead to enormous management complexity due to “islands of storage”, limited scalability, performance bottlenecks, availability and high costs that are created when such systems are used for unstructured data. These limitations have triggered the need for a new storage architecture — one that is designed and optimized from the ground up specifically for unstructured data and digital content.

### ***Macro Trend 2: Paradigm Shift to Cluster Computing***

The second macro trend is the widespread adoption of clustered computing. Enterprise data centers have evolved from the era of “big iron” proprietary mainframes and symmetrical multiprocessing (SMP) servers to that of standards-based (using industry-standard hardware), clustered machines running Linux or Windows. The most dramatic evidence of this trend is the change in worldwide server revenues. Since the mid 1990s, approximately a quarter of this \$50 billion market, and a much larger fraction of units, has shifted from mid-range servers costing many tens or hundreds of thousands of dollars to smaller servers priced as low as \$2,000 to \$3,000 per server.<sup>(2)</sup>

The prime motivation for IT managers to adopt server cluster architectures is the higher levels of performance, reliability, scalability and overall workload management that can be realized by aggregating industry standard servers, all at a fraction of the cost of traditional bigger box solutions. No longer do organizations deploy a large database on a 200 processor box. Today, IT managers buy a cluster of off-the-shelf servers to put together one large, seamless expandable system.

---

***The Clustered Storage Revolution is extending this cluster trend from the server application world to the data storage world***

---

One example of the benefit of aggregation is a web-server farm. The availability, reliability and performance needed to meet the requirements of being live on the internet 24x7 while also using a cost-effective solution is best realized through server clustering. A server cluster farm streamlines internal processes by distributing the workload between the individual components of the farm and expedites computing processes by harnessing the power of multiple servers.

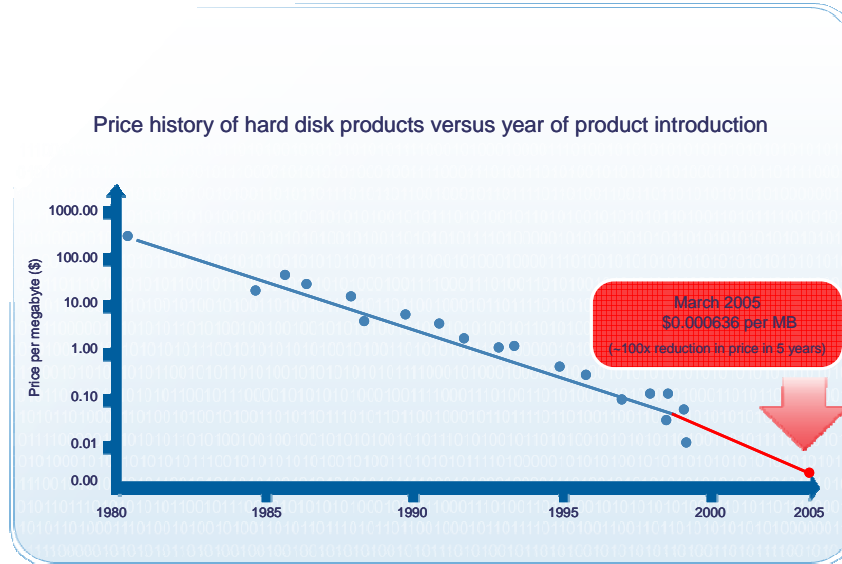
When one server in the farm fails, another can step in and assume the workload. Combining servers and processing power into a single entity was a practice once found only in research and academic institutions, but it is now becoming pervasive in the enterprise market as well. Today, more and more companies are using server clustering as a method to handle enormous amounts of computerization of mission-critical tasks and services.

The *Clustered Storage Revolution* is extending this cluster trend from the server application world to the data storage world. In the same way and for the same reasons that the server application world is turning to clustered architectures, the storage world has begun this major architectural shift as well.

### ***Macro Trend 3: Proliferation of Cheaper, Faster Industry-Standard Enterprise-Class Hardware***

The third macro trend driving the movement to clustered storage is a dramatic decrease in the price-performance curves of industry-standard hardware components. This trend is part of the continual movement toward the promise of Moore’s Law: over time, companies are getting higher computing power for a lower cost and realizing the economics of commodity hardware. The low cost of commodity hardware components has made the merits of clustered architectures affordable. Google is a prime example of how clustering has leveraged the price-performance curves of industry-standard hardware to realize industry-leading performance and reliability at a fraction of the cost of traditional, custom-built systems. On average, a single query on Google reads hundreds of megabytes of data and consumes tens of billions of CPU cycles. To handle this “high performance computing” workload, Google’s architecture features clusters of thousands of commodity class PCs, off-the-shelf components, with fault-tolerant software. This clustered architecture achieves superior performance at a fraction of the cost of a system built from fewer, but more expensive, high-end servers.

Leveraging enterprise-class industry-standard hardware leads directly into the trend toward clustered storage solutions. Consider that the price history and density of storage disk products (i.e. SATA) over the last 5 years has experienced more than 100x reduction in price per MB. (See Figure below) Coupled with dramatic declines in the cost of processors, memory and bandwidth, IT managers are now in a position to attain the full value of clustering via commoditized components in storage.



### What these Macro Trends Mean for Storage:

These macro trends point to three fundamental implications:

- The storage industry is undergoing a revolution
- Clustered storage is becoming the dominant new storage architecture
- Customers are reaping substantial business value and benefits from clustered storage

From big monolithic boxes to clustered architectures, storage is following the paradigm shift that has already occurred in the server application world. Driven by intelligent software and built on industry-standard hardware, clustered storage is a rapidly emerging new storage architecture. Customers understand that to address the explosion of unstructured data in the enterprise, it is clustered architectures that best deliver the breakthrough price/performance, reliability and scalability to meet their needs, all at substantially lower operating costs. *The Clustered Storage Revolution has begun!*

### rev-o-lu-tion

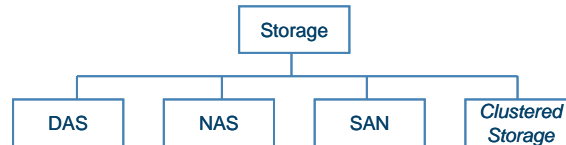
*Pronunciation:* "re-v&- 'lū-sh&n; *Function:* noun

*2 a :* a sudden, radical, or complete change... *d :* a fundamental change in the way of thinking about or visualizing something : a change of paradigm <the Copernican revolution> *e :* a changeover in use or preference especially in technology <the computer revolution>

*Merriam-Webster Dictionary Online 2005*

### 3 A New Category of Storage: Clustered Storage

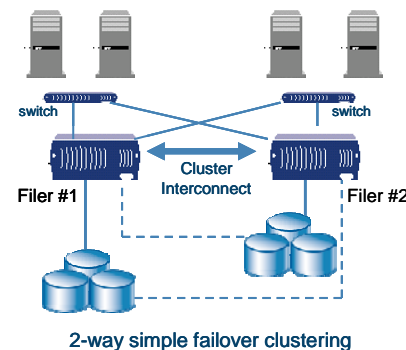
Direct Attached Storage (DAS), Storage Area Networks (SANs), and Network Attached Storage (NAS) are the typical storage approaches most IT managers envision when they talk about storage architectures. Today, a fourth approach to storage has emerged — *Clustered Storage*.



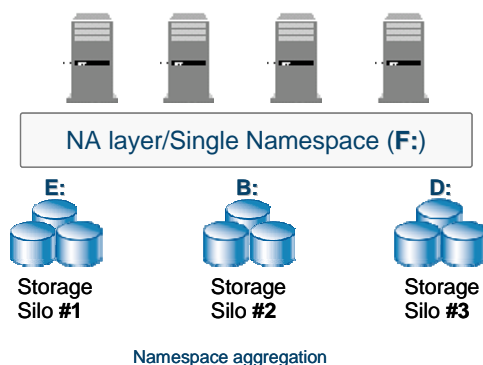
Clustered storage architectures have the ability to pull together two or more storage devices to behave as a single entity. Clustered storage can be broken down into three types:

- 2-way simple failover clustering
- Namespace aggregation
- Clustered storage with a distributed file systems (DFS)

**2-way simple clustering:** Historically in the storage industry, “clustering” meant active failover between a pair of redundant nodes (“node” is defined as server/controller head & disk). Although this approach is more accurately described as a redundant technique rather than a clustering technique, NAS vendors commonly refer to this as “2-way clustering.” 2-way clustering evolved out of the need to continue to improve fault tolerance and redundancy with legacy and traditional single-head storage architectures. Typically these solutions enable one controller head to assume the identity of the failing controller head, and allow the failed controller’s data volumes to continue to be accessed or written to by the new controller head. This solution approach has seen very limited adoption in practice. The inherent limited performance and scalability, small file system sizes, management complexity and relative high cost to achieve the high availability are the main limiting factors. Couple this with the explosive growth of unstructured data it becomes evident that these solutions will not meet the future requirements of the growing enterprise.



**Namespace aggregation:** These types of clustered storage solutions essentially present a single pane of glass, or veneer, which pulls storage management together. Such solutions can be purely software-based (i.e. software virtualization), or can be a combination of software and hardware (i.e. appliance and switch), and create a single namespace and cluster of storage resources that appear as one large pool of data management. Typically, these solutions enable “synthetic trees” that encompass



a cluster of NAS servers or storage devices, present the silos to a network user as one (a unified namespace) and park data on any given silo. In other words, they create gateways through which data from several different files and heterogeneous systems is redirected to be accessed from a common point. Solutions in this class can control laying out a file (striping data) across disk volumes to a specific silo — but not across the silos that make up the cluster — while still allowing data movement between tiers of storage with limited client interruption. While this architecture approach can sometimes be attractive on the surface from an initial cost standpoint the IT

administrator is still managing, growing and configuring “islands of storage” (heterogeneous silos of storage) but now with an additional virtualization layer. Ultimately, this solution approach creates higher complexity, higher management burden, and higher long term operational costs.

**Clustered storage with a DFS:** The third type is *Distributed Clustered Storage*, the natural evolution beyond N-way simple clustering and namespace aggregation. Distributed clustered storage is a networked storage system that allows users to combine and add storage nodes, all of which access the same pool of data. These solutions reside directly on the storage layer with fully distributed file systems across any number of nodes/storage controllers. Since the software resides at the storage layer itself, it can fully control layout of data (data striping) across all the storage nodes that make up the cluster, down to the ECC error correction level for every chunk of data. This is in contrast to namespace aggregation/virtualization products that only direct the specific storage silo to which data is written. Intelligent software makes the nodes symmetric and distributed, so the cluster (made up of 3 or more nodes) works together as an intelligent unified team, with each node capable of running on its own and communicating with other nodes to deliver files in response to user needs. Each node in the cluster is a coherent peer, meaning each node knows everything about the other. Because of these characteristics, distributed clustered storage provides the highest levels of availability, reliability, scalability, aggregate throughput and ease of management when compared to any of the other solutions identified above.

Differences between the three types of Clustered Storage solutions in how data is controlled

	<u><b>File Layout</b></u>	<u><b>Data Striping</b></u>
<b>2-way Simple Failover Cluster</b>	File stored on one of the silos. In failover scenario, remaining controller assumes identity of failed head.	Stripes data blocks across a RAID group of disks within a specific storage silo.
<b>Namespace Aggregation</b>	Controls which storage silo receives the file. Presents and aggregates all namespaces of locations of files.	Stripes data blocks across disk within specific silos that make up a cluster.
<b>Clustered Storage with DFS</b>	<b><i>Distributes file layout across all storage nodes and disks. Presents one unified view of all files from every node. Single global namespace.</i></b>	<b><i>Breaks files into data blocks that are striped across all storage nodes and disk that make up a cluster.</i></b>

## 4 Clustered Storage Defined

When defining *clustered storage* solutions we find six common characteristics:

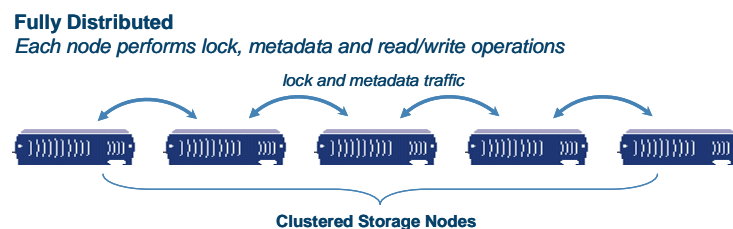
- Symmetric Clustered Architecture
- Scalable Distributed File System
- Inherent High Availability
- Single Level of Management
- Linear Performance Characteristics
- Enterprise Ready

**Symmetric Clustered Architecture:** The key design principle behind distributed clustered storage solutions is symmetry among the nodes which can be thought of as self-contained storage controller heads, 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, in contrast to traditional storage architectures deploying master server-based approaches where the storage nodes are not symmetric and are 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 grow resources seamlessly and enable the modular growth, or “pay-as-you-grow”, benefits of the storage system. When more memory, bandwidth, capacity, or drive actuators are needed, the cluster can be grown by simply adding additional 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 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 can seamlessly scale to hundreds of terabytes.

In the figure below, a fully distributed file system handles metadata operations, file locking and cache management tasks by distributing operations across all the nodes in the cluster. The distributed file system ensures proper locking behavior without reliance on a master server — eliminating performance bottlenecks by removing dedicated metadata servers — and maintains full cache coherency among all nodes, ensuring every read and write is guaranteed to retrieve the most up-to-date data.



**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. In fact, fully distributed cluster architectures can sustain multiple simultaneous drive and node failures and still be able to recover and continue operation. Moreover, high availability is “inherent” for distributed cluster 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.

Clustered storage architectures present unique *reliability* challenges: since these solutions use a wide variety of industry-standard hardware components and hundreds or even thousands of disks spinning under a single file system, meeting enterprise reliability standards requires new innovative technologies. To meet this challenge, complete clustered storage solutions have to be able to deliver very fast drive rebuild times to minimize windows of risk, provide proactive “self healing” capabilities to ensure all data is always available, and ensure the distributed file system is fully journalled (i.e. protected against failures during a write operation across the entire cluster) in all states of operation.

**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 minimum cluster size of three 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 drive letters and mapping applications to many separate “islands of storage”. For system administrators, this eliminates client-side management issues because all of the files belong under one drive letter or mount point.

**Linear Scalability of Performance:** 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, it maintains its coherency as one logical system and is able to aggregate across all resources; 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 stay 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.

**Enterprise Ready:** Distributed clustered storage solutions must be enterprise ready. Historically, clustered architectures were first deployed primarily in non-commercial research labs, not in mainstream commercial enterprises. In order to be part of a paradigm shift, though, the clustered solution must be ready for implementation into a commercial enterprise data center. Specifically, the solution must support standard network protocols and provide the tools that IT managers have come to expect.

*Clustered Storage* is the natural evolution of storage to meet the changing needs of the modern enterprise and its unstructured data growth. The table below summarizes the differences between the types of Clustered Storage solutions generally available on the market today:

	<i>Symmetric Architecture</i>	<i>Scalable File System</i>	<i>Inherent HA</i>	<i>Single Level Mgmt</i>	<i>Linear Performance Scalability</i>	<i>Enterprise Ready</i>
<i>2-way Simple Failover Cluster</i>	No	No	No	No	No	Yes
<i>Namespace Aggregation</i>	No	Limited, Yes	No	Yes	No	Limited
<i>Clustered Storage w/DFS</i>	Yes	Yes	Yes	Yes	Yes	Yes

## 5 Isilon Systems: The Leader in Clustered Storage

Isilon Systems® is now delivering its fourth generation of fully distributed clustered storage solutions and is the clear leader in this emerging category. Isilon's award-winning family of Isilon IQ products consists of high-performance clustered storage systems that combine an intelligent distributed file system with modular industry-standard hardware to deliver unmatched simplicity and scalability. Isilon IQ was designed for unstructured data and for use in data-intensive markets such as media and entertainment, digital imaging, life sciences, oil and gas, manufacturing and government.

### ***Isilon IQ: Scalable Distributed File System***

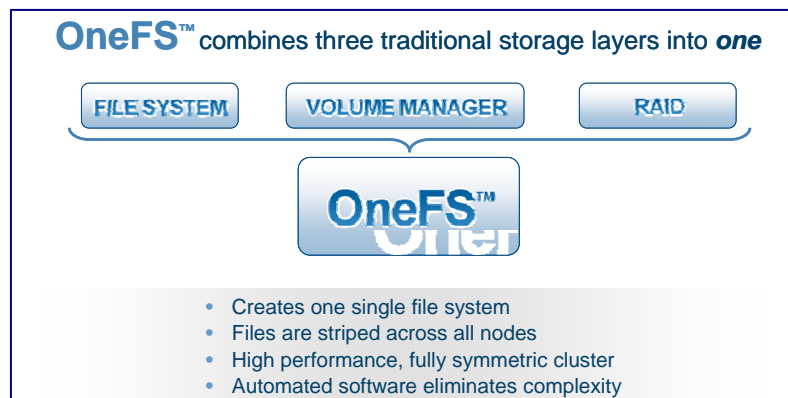
At the heart of Isilon's clustered storage solution is Isilon's OneFS® patented distributed file system. It combines the three layers of traditional storage architectures — file system, volume manager and RAID — into one unified software layer, creating a single intelligent fully symmetrical file system that spans all nodes within a cluster. OneFS provides a single point of management for large content stores, faster access to large content files, inherent high availability, the ability to easily scale a single cluster's capacity, up to 7 Gigabytes per second of total throughput and hundreds of terabytes of capacity, all from a single network file system.

OneFS uniquely stripes files and meta data across multiple storage nodes within a cluster, an improvement over the traditional method of striping content across individual disks within a single storage device or volume.

This fully distributed approach enables Isilon to deliver break-through performance, scalability, availability and manageability.

OneFS provides each node with knowledge of the entire file system layout and where each file and parts of files reside. 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. Instead, OneFS provides the user with the ease and simplicity of managing a single NAS head with scalability, performance, and flexibility that exceeds SAN systems.



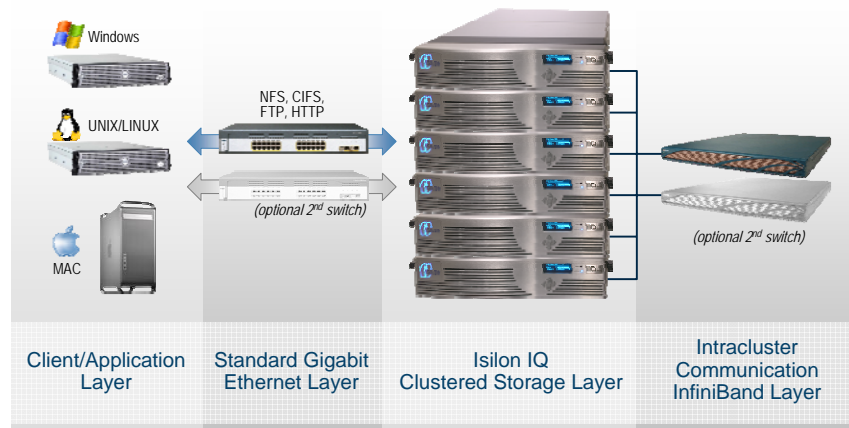
### **Isilon IQ: Symmetric Architecture**

Each Isilon IQ cluster consists of anywhere from three to 42 Isilon IQ nodes. Each modular, self-contained Isilon IQ node contains disk capacity along with a powerful storage server, CPU, memory and network, all in a self-contained, compact, 2U rack-mountable system. As additional Isilon IQ nodes are added to a cluster, all aspects of the cluster scale symmetrically, including capacity, throughput, memory, CPU and network connectivity. Isilon IQ nodes automatically work together, harnessing their collective power into a single unified storage system that is tolerant of the failure of ANY piece of hardware, including disks, switches or even entire nodes.

In a fully distributed architecture, it is critical for each node to stay in sync with all other nodes in the cluster. Isilon IQ storage nodes use either Gigabit Ethernet or high-speed, low-latency Infiniband

switching fabric for inter-cluster communication, synchronization and all intracluster operations. This enables each node to share information with every other node on the system, so that each storage node acts as a fully coherent peer with complete understanding of what the other nodes are doing.

Isilon IQ Network Architecture:



OneFS keeps the nodes synchronized by using a distributed lock manager, coherent caching and a remote block manager that maintains global coherency throughout the entire cluster. It is this global coherency through each node that eliminates any single point of failure for access to the file system. Any node in the cluster can take a write or read request and each node presents the same unified view of the entire file system. All nodes in the cluster are “peers”, so the system is fully symmetrical, eliminating hierarchy and inherent bottlenecks.

### Isilon IQ: Inherent High Availability

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 true way to ensure data integrity and eliminate single points of failure is to make all nodes in a cluster peers. Because each node in an Isilon IQ is a peer, any node can handle a request from any application server to provide the content requested. If any one node were to go down, any other node could fill in, thereby eliminating any single point of failure.

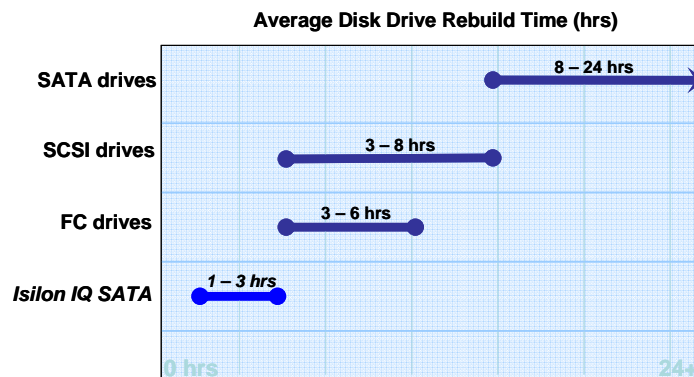
**Multi-failure Support.** With Isilon IQ, customers can withstand the loss of multiple disks or entire nodes without losing access to any content. OneFS's unique FlexProtect-AP feature utilizes Reed Solomon “n+1 and n+2” ECC (error correction code), parity striping and mirrored file striping (from 2x to 8x) that spans multiple nodes within a cluster. These policies can be set at any level, including cluster, directory, sub-directory, or even at the individual file level. Additionally, these policies can be changed at any time from a simple WebUI — even while the system is in production and fully available. *With Isilon, all files are striped across multiple nodes within a cluster, no single node stores 100 percent of any file, and if a node fails, all other nodes in the cluster can still deliver 100 percent of the files without interruption.*

Specifically, “n+2” double ECC error correction allows for multiple failures of disks or even nodes within a single cluster and file system. Each file is striped across multiple nodes within a cluster, with two parity stripes for each data block. Unlike the “n+1” single parity, if a second failure were to occur during this rebuild, all data would still be fully available because the data was originally striped with double ECC error protection. In contrast, the same scenario in a traditional storage system using RAID5 would result in a data loss with no chance of recovery. Isilon engineers estimate that the mean time between failures (MTBF) in n+2 RAID is over 100 times the MTBF in single-parity RAID approaches. Isilon IQ is the only clustered storage solution to offer this level of data protection across a single file system in a clustered architecture.

**Industry Leading Drive Rebuild:** In the event of a failure, OneFS automatically re-builds files across all of the existing distributed free space in the cluster in parallel, eliminating the need to have the dedicated “parity drives” typically required with most traditional storage architectures. OneFS takes advantage of the cluster by leveraging all available free space across all nodes in the cluster to rebuild data. By utilizing this free space while also drawing on the multiple processors and compute power of the cluster, data can be rebuilt five to ten times faster when compared to traditional architectures.

The time that it takes a storage system to rebuild data from a failed disk drive is critical to the data reliability of that storage system. With traditional storage systems, the rebuilding process already takes many hours; a trend that is steadily worsening as hard drive densities continue to increase – now at 500GB per drive. With the advent of terabyte sized disks expected within the next 24 months and the creation of larger and larger single volumes/file systems, traditional storage systems will require up to 24 hours or more to recover from a disk failure. During that time, such traditional storage systems are vulnerable to additional disk failures which will cause data loss and downtime.

Since Isilon IQ is built on a distributed architecture, it leverages all spindles and hardware within the cluster to their maximum capacity in order to reconstruct data from failed disks. Because Isilon IQ is not bound by the speed of any particular disk, Isilon’s systems are able to recover from disk failures extremely quickly. Depending on drive density, disk failures within an Isilon IQ cluster can be rebuilt in as little as one to two hours. When compared to fiber channel and SCSI disks, which can take upwards of 8 hours, or other ATA disk drives, which can take anywhere from 8 to 24 hours to rebuild a single failed disk drive, the advantage of Isilon’s architecture is apparent. By delivering industry-leading drive rebuild times, Isilon IQ offers a more reliable storage system that is also more resilient and much less susceptible to multi-failure scenarios.



Isilon IQ used Maxtor 250GB Serial ATA drives for the benchmark; disk was 87% full, with file system sizes from 1 – 10MB. Note: Industry average rebuild times are highly variable and depend on many factors. The averages shown above are ranges representing optimal to moderate use cases using drives between 160 – 250GB.

**Self-Healing Capabilities:** OneFS constantly monitors the health of all files and disks and maintains records of the smart statistics (e.g. recoverable read errors) available on each drive to anticipate when that drive will fail. When OneFS identifies at risk components, it preemptively migrates the data off of the “at risk” disk to available free space on the cluster in a manner that is both automatic and transparent to the customer. Once the data is rebuilt, the user is notified to service the suspect drive in advance of actual failure. This feature provides customers with confidence that data written today will be stored 100 percent reliably, bit-for-bit correct, and available whenever it is needed. No other cluster solution today provides this level of data protection reliability.

### Isilon IQ: Single Level of Management

Isilon IQ creates a single, shared pool of all content within the cluster, providing one point of access for users and one point of management for administrators. *Today, Isilon has tested and supports growing a single network drive to over 500TB.* Once an Isilon IQ cluster is established, users can connect to any storage node and securely access all of the content within the 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.

As a distributed file system, OneFS eliminates captive server-attached storage and creates substantial improvements in the efficient viewing, sharing, and allocation of resources. Users can enjoy instant access to previously inaccessible content and administrators can dynamically add and reallocate

content when capacity needs increase. The result is faster deployment of new business applications and the ability to access and share content anywhere on the network.

One of the key benefits of OneFS is the ease with which it allows users to add both performance and capacity to an Isilon cluster without downtime or application changes. System administrators simply plug in a new Isilon IQ storage node, connect the network cables and turn it on. The cluster automatically detects the newly added storage node and begins to configure it to become a member of the cluster. In less than 60 seconds, a user can grow available capacity and grow the single file system by 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 Isilon IQ also enables customers to incorporate new technologies in the same cluster, such as adding a node with higher-density disk drives or more Gigabit Ethernet ports for higher performance.

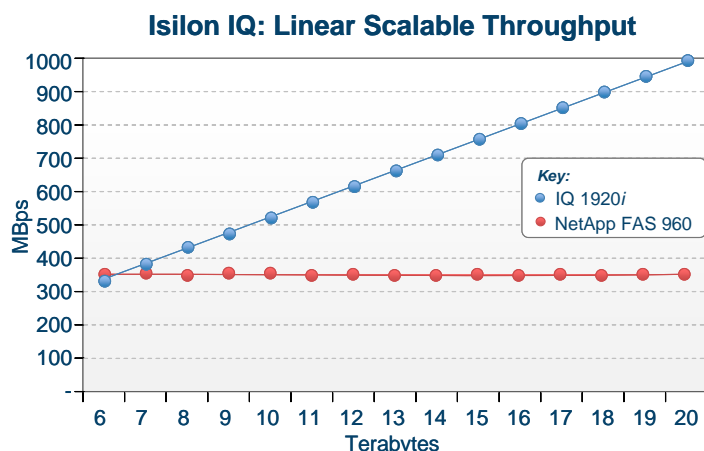
Finally, OneFS automates several advanced features that for traditional storage solutions are manually intensive operations. Two of these include Isilon's AutoBalance and SmartConnect features.

**AutoBalance:** When a system administrator adds a new storage resource, the common next step is to manually migrate content from an existing storage device to the new one in order to balance capacity across resources. Isilon IQ delivers automated content migration when scaling and totally eliminates the need for business application outages. Using its AutoBalance feature, a new storage node can be added to an Isilon IQ cluster in less than 60 seconds. As soon as the node is turned on and network cables are connected, AutoBalance immediately begins to migrate content from the existing storage nodes to the newly added node across the cluster interconnect back-end switch, re-balancing all of the content across all nodes in the cluster and maximizing utilization.

**SmartConnect:** Another OneFS automation feature is SmartConnect. The SmartConnect feature enables the balancing of client connections across a cluster of Isilon IQ storage nodes to provide optimal utilization of the cluster resources. By providing a single virtual host name, SmartConnect makes it easy for IT administrators to manage client connections. SmartConnect applies intelligent policies (i.e. CPU utilization, connection count, throughput) to simplify the connection management task by automatically distributing the client connections across the cluster based on the defined policies to maximize performance.

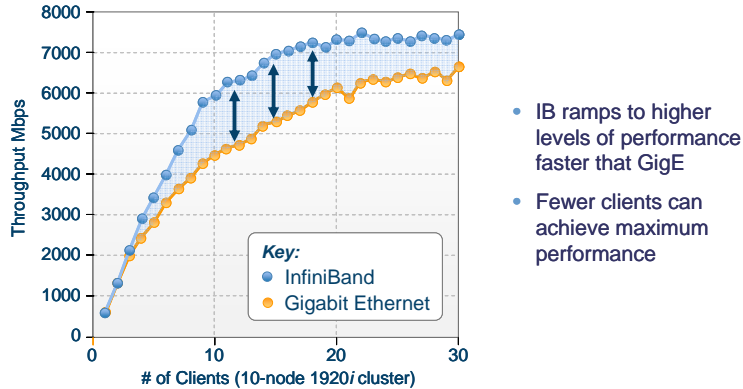
### Isilon IQ: Linear Scalability in Performance

One of the key benefits of OneFS is the ease with which it allows users to add both performance and capacity to an Isilon cluster in a near linear fashion. See Graph below. Unlike other storage systems that communicate below RAID at the physical disk level, OneFS controls the optimal placement of files directly on the disk and dramatically improves performance of the disk subsystem when delivering data. Each addition of an Isilon IQ storage node or Accelerator increases memory, CPU power, journal space and disk spindles. A new Isilon IQ node equips the aggregate of the cluster with approximately 700 megabits per second of available throughput that scales linearly, allowing customers to easily meet increasing bandwidth needs.



The other enabling technology that

allows Isilon IQ to reach break-through linear scalability of performance is use of Infiniband as the high-speed, low-latency intracluster interconnect. A backend Infiniband switch allows the Isilon cluster to experience nearly zero latency in keeping the nodes in sync, allowing for optimal overall cluster performance. In fact, Isilon testing has shown that this enabling technology allows an Isilon solution to obtain much higher performance, much more quickly, than with a GigE backend interconnect. Isilon is the first and only clustered storage solution to utilize Infiniband as a clustered storage interconnect, and today over 90% of Isilon customers deploy this option.

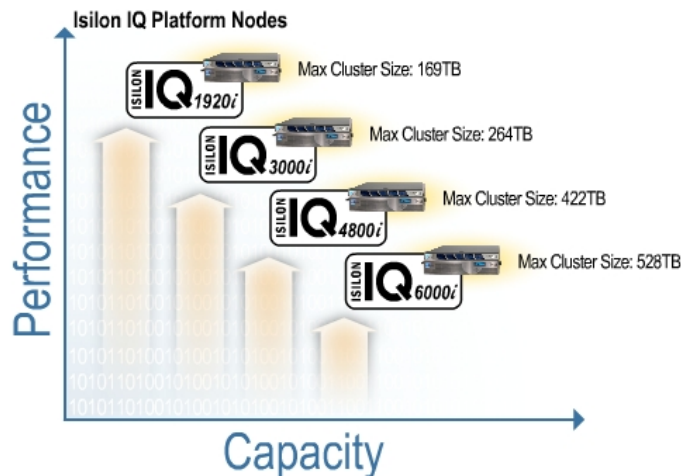


### Isilon IQ: Enterprise Ready

Now in its fourth generation, Isilon IQ has delivered on many of the features that meet the requirements for integration into the commercial enterprise. Isilon IQ is 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. In short, customers seamlessly deploy Isilon IQ in their existing data centers right next to their traditional storage systems from vendors such as EMC and Network Appliance.

### Isilon IQ Product Family

The Isilon IQ clustered storage product line addresses enterprises' full spectrum of storage needs – from the highest performance tier-1 applications to tier-2 enterprise archive, disk-to-disk backup and disaster recovery. The Isilon product line is comprised of the Isilon IQ 1920, 3000, 4800 and 6000 platform nodes and the Isilon EX 6000 and Isilon IQ Accelerator extension nodes. This flexible product line satisfies customer's varying capacity and performance needs for all unstructured data and enterprise file-based information. In addition, Isilon's unique TrueScale™ technology enables customers to scale both performance and capacity linearly or independently based on application or workflow needs.



## 6 Conclusion

*There is a revolution well underway in the storage industry – the movement to Clustered Storage architectures.* This technology shift is driving huge business benefits:

- Reduces storage costs: Costs 40-60% less than traditional storage solutions to own and operate
- Increases workflow productivity: Get up to 5x more work done with existing staff and resources
- Increases IT operating leverage: Manage 10x more storage with existing IT staff
- Unlocks new revenues: Create and distribute more products – faster

Adoption of Clustered Storage solutions is increasing at an exponential pace. World class companies such as Pratt & Whitney, Kodak Easy Share Gallery, NBC Sports, Lexis Nexis, Cedars Sinai, Sports Illustrated, American Title, and Kelman Technologies are reaping the benefits of Clustered Storage in production today. The time is now... *Join the Clustered Storage Revolution!*

About Isilon Systems:

Isilon Systems is at the forefront of the paradigm shift to Clustered Storage architectures and is now delivering its fourth generation of fully distributed clustered storage solutions. Its award-winning family of Isilon IQ products consists of high-performance clustered storage systems that combine an intelligent distributed file system with modular industry-standard hardware to deliver unmatched simplicity, scalability and availability. Isilon IQ was designed for unstructured data and for use in data-intensive markets such as media and entertainment, digital imaging, life sciences, oil and gas, manufacturing and government. The Seattle-based company is financed by Lehman Brothers Venture Partners, Sequoia Capital, Atlas Venture and Madrona Venture Group.

For more information visit us at [www.isilon.com](http://www.isilon.com).

© 2001-2006 Isilon Systems, Inc. All rights reserved. Isilon, Isilon Systems and OneFS are registered trademarks, and TrueScale and SyncIQ are trademarks, of Isilon Systems, Inc.

For more information, contact Isilon Systems at:

Isilon Systems, Inc.  
3101 Western Avenue  
Seattle, WA 98121  
Toll-Free: 877-2-ISILON  
Phone: 206-315-7602  
Fax: 206-315-7501  
Email: [sales@isilon.com](mailto:sales@isilon.com)

Footnotes:

1. Tim O'Reilly, Open Source Paradigm Shift article, [www.oreilly.com](http://www.oreilly.com), June 2004
2. Ben Eiref, RTC magazine, InfinBand Enables the "Wire Once, One Wire" Data Center article, October '05