

# Clustered Storage

© Isilon® Systems

[www.isilon.com](http://www.isilon.com)

## Abstract

The paper introduces the reader to a new paradigm shift that is currently taking place in the data storage industry: the movement toward Clustered Storage architectures. Clustered Storage architectures are changing the rules of how data is stored and accessed. This paper discusses 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.

## 1 Three Macro Trends Driving the Shift to Clustered Storage

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.

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.

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 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.

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.

## 2 Clustered Storage

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.

**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.

**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.

**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.

**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 intra-cluster 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.

### 3 Isilon® Systems 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.

#### 3.1 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 10 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 breakthrough 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.

#### 3.2 Isilon IQ: Symmetric Architecture

Each Isilon IQ cluster consists of anywhere from three to 96 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 intra-cluster 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.

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.

### 3.3 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 ECC (error correction code), parity striping (from n+1 to n+4) 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. 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.

**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.

**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.

### 3.4 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 up to 1,000TB (1 PB). 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 client connection load balancing and dynamic NFS failover and failback of client connections across storage nodes to provide optimal utilization of the cluster resources. Without the need to install client side drivers, administrators can easily manage a large and growing number of clients and rest assured that in the event of a system failure, in flight reads and writes will successfully finish without failing. 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.

### 3.5 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 intra-cluster 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.

### 3.6 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.

## 4 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. And Isilon Systems is at the forefront of the paradigm shift to Clustered Storage architectures.