

## WHITE PAPER

---

# Quantifying the Business Benefits of Scale-Out Network Attached Storage Solutions

Sponsored by: EMC

---

Randy Perry

Benjamin Woo

November 2011

---

### Executive Summary

The growth of Web, cloud-based, and other data-intensive applications has ushered in the era of "Big Data," forcing companies to find new ways to scale and manage their storage environment. IDC defines Big Data technologies as a new generation of technologies and architectures designed to economically extract value from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis. Increasingly, scale-out storage is becoming the solution of choice for organizations seeking to manage data growth seamlessly and cost-effectively in Big Data environments.

Scale-out solutions support data growth by simply adding new storage "nodes" when extra capacity and performance are needed. In contrast, the traditional scale-up architecture aggregates many individual disk drives behind one or two large controller servers to provide additional capacity. Because the multiple nodes in scale-out solutions behave like a single, larger file system, it is easier to manage data growth without having to hire additional IT specialists.

Also, as server prices drop and capabilities and capacities increase, scale-out storage solutions allow enterprises to replace large, costly equipment with a readily sized pool of inexpensive servers. Further, by combining and configuring many small, inexpensive storage components, scale-out solutions create an aggregate storage pool that can exceed the capacity of a single, traditional storage array.

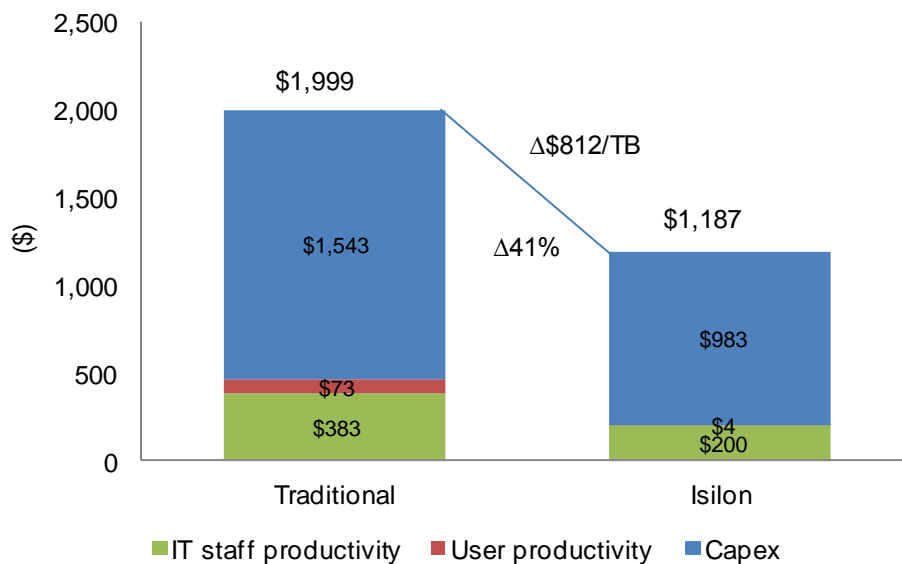
Scale-out solutions can provide more timely processing and greater resource utilization than scale-up systems because they allow companies to scale storage resources in line with datacenter demands and changing business needs. Also, as nodes are added, the processing power and storage I/O bandwidth increase, along with the storage capacity, so scale-out systems become increasingly faster and more powerful as capacity is added to the storage infrastructure.

Additionally, the simple, inexpensive nodes allow for granular expansion of storage resources, as needed, without disrupting business operations or risking the overbuying that can occur with scale-up systems. With scale-up solutions, each additional system increases the management burden and the likelihood of having to add IT staff. By avoiding these problems, scale-out solutions can hold down costs and make IT staff more productive. Also, by supporting data growth in a dynamic, nondisruptive manner, scale-out solutions ensure that system capabilities always remain as productive as possible.

To assess the business value of scale-out solutions in Big Data environments, IDC conducted in-depth interviews with organizations using EMC Isilon's scale-out storage hardware and software platforms. From the results, IDC was able to quantify the benefits the companies realized from their deployments. IDC found that, on average, the companies in our study were able to reduce their average annual cost per TB used by 41%, achieving a total benefit of \$1.6 million (or \$812 per TB in use as shown in Figure 1 — based on an average of 1.9PB per organization in the study).

**FIGURE 1**

Average Annual Cost per TB in Use



Source: IDC, 2011

## Introduction

While the scale-out approach can be applied to direct attached storage and storage area networks, it is more commonly associated with network attached storage (NAS) systems, which use servers, disks, and management software to serve files over a network. Scale-out NAS systems become a cluster of NAS nodes, which allow for capacity and performance to scale in combination or independently while maintaining a single system image.

A key component of scale-out solutions is the software that enables the nodes to be interconnected and referenced as a single object by storage administrators. With such solutions, the administrator's responsibility is simply to manage the data and not the storage hardware. The software also provides all the data services IT professionals have come to expect from their storage solutions, including snapshots, replication, and data protection.

Customers should be wary of traditional scale-up architectures that attempt to add on scale-out capabilities. Often, these scale-up architectures cluster together multiple separate nodes and add an abstraction layer to give a view of an aggregated namespace. However, the underlying limitations of separate volumes and multiple file systems still exist and must be actively managed.

With EMC Isilon scale-out storage solutions, companies can scale capacity from 18TB to more than 15PB in a single file system and scale performance to more than 85GB per second. The Isilon OneFS 6.5 operating system supports three series of scale-out NAS platform nodes — S, X, and NL — as well as data management, data protection, data replication, data retention, load balancing, and other software capabilities.

The Isilon S200 scale-out NAS platform node is intended to speed time to market and maximize return on investment with ultrafast primary storage for mission-critical, highly transactional and random access file-based applications. It scales with extremely demanding workflows and applications and delivers high performance in a small footprint. The platform uses enterprise solid state drives (SSDs) to accelerate namespace-intensive metadata operations and file-based storage workflows.

The companion Isilon X200 scale-out NAS platform node is a versatile, midtier platform designed specifically to accelerate high-concurrent and sequential-throughput applications while reducing cost and complexity. The company's highest-density platform, the IQ 108NL, also has the most aggressive cost per TB. Intended for nearline archive applications, it scales from 100TB to 15.5PB in a single pool of storage with global namespace, eliminating the need to manage multiple volumes.

The Isilon OneFS 6.5 is the company's sixth-generation operating system. It provides three layers of traditional storage architectures — file system, volume manager, and RAID — in one unified software layer, creating a single intelligent file system that spans all nodes in a cluster. The OneFS 6.5 provides integrated support for SSDs within SmartPools, a data management feature that allows users to combine pools of different node types in the same cluster, all in a single file system, and to set policies that automatically migrate data among nodes with different performance and capacity characteristics. As a result, critical SSD resources can be used to cost-effectively improve metadata performance across a wide range of workflows through global namespace acceleration (GNA).

The company has also enhanced disaster recovery and business continuity with the release of SyncIQ 3.0 data replication software, which provides automated high-speed replication of millions of files and terabytes of data over the WAN and LAN for disk-to-disk backup, disaster recovery, and content distribution. Integrated snapshots quickly determine which data has changed between syncs, reducing the time needed to protect the same amount of data.

### ***The Benefits of Scale-Out Efficiency***

IDC recently conducted a study to quantify the business benefits of scale-out NAS solutions with a survey of companies using EMC Isilon's scale-out NAS platforms. Specific figures for savings and costs came from interviewing IT managers at companies of different sizes and from various industries that have deployed EMC Isilon solutions.

For the survey, IDC interviewed IT managers at eight midsize to large organizations with an average of 1,052 employees (see Table 1) and asked a series of questions regarding their storage strategy and how migrating to Isilon's scale-out solution impacted their organization. These organizations have very complex and extensive storage needs, as evidenced by their average of 6.1TB per IT user, which is easily three to four times that of the average organization. Several organizations are content service providers. When you include a large public library and a research lab, you have a group of customers who are very familiar with the realities of Big Data. They also have already transitioned to NAS as their primary architecture. These customers chose Isilon primarily for its scalability and efficiency.

**TABLE 1**

Demographics

Category	Average
Employees	1,052
IT end users	842
IT staff	50
TB per employee	6.1
TB per storage management staff	500
Total storage (TB)	1,849
Storage — NAS (TB)	1,750
Average percentage of NAS	97%
Annual growth rate in storage	28%

Source: IDC, October 2011

Because these organizations are storage heavy, they are keenly focused on increasing storage efficiency. We looked at four critical metrics associated with storage efficiency, comparing their position before and after deploying the Isilon solution. At 74% utilization, they were already well above the average of 50% typical of most companies using traditional storage architecture. However, Isilon was able to boost utilization to 87%. Other key efficiency metrics were more dramatically improved (see Table 2).

**TABLE 2****Performance Improvements**

	Traditional	Isilon	Savings	Change (%)
Increased utilization	74%	87%	13%	18
TB per storage manager	261	500	240	92
Recovery speed (hours)	34.0	4.1	29.9	88
Time to provision new capacity (hours)	182.5	20.0	162.5	89

Source: IDC, October 2011

To determine the actual business benefits of scale-out NAS solutions, IDC translated the change in efficiency metrics into financial savings. IDC grouped the benefits into the three major categories of cost savings (capex), increased IT staff productivity, and improved end-user productivity and normalized the results by expressing them in dollar savings per TB of storage in use. This allows companies with various storage capacities to extrapolate the results to their own organization.

#### **Cost Reduction (Capex)**

Capex refers to the direct cost for storage and the costs to house storage assets (facilities and power). Capex reduction is driven by Isilon's much more efficient use of storage compared with traditional storage. Table 3 shows that for every \$1.00 spent on raw storage, Isilon delivers \$0.72 worth of storage in use compared with \$0.58 for other traditional storage. One key reason is that the utilization rate of Isilon storage is typically much higher than that of traditional storage. With traditional storage, IT administrators often have to set aside significant discretionary reserves for snapshots and performance headroom so that their storage can continue to operate. In addition to storage costs, they were able to drop their related facilities costs by 40%, saving \$42 per TB. Facilities savings were driven by the more efficient head architecture, which requires less rack space. Depending on the architecture being replaced, Isilon could use anywhere from 25% to 75% fewer racks.

Companies also reduced their power requirements by 13%. The companies in our study also paid less per TB for their Isilon storage, which, when combined with the efficiency advantages, means that they lowered their storage costs by 37%, saving \$1,527 per TB. This advantage means that companies can budget more accurately for their storage requirements and more efficiently scale to meet new or fluctuating demand.

**TABLE 3****Scale-Out Efficiency**

	Traditional	Isilon	Savings	Change (%)
Total TB raw	1.0	1.0	–	–
Total usable (TB) available	0.80	0.84	0.04	5
Total actually used (TB)	0.58	0.72	0.14	25
Average cost per TB raw*	\$2,503	\$1,917	\$587	23
Average cost per TB usable	\$3,005	\$2,232	\$773	26
Average cost per TB actually used	\$4,129	\$2,602	\$1,527	37

Note: The data in the table is the average over three years for comparable tier 2 storage.

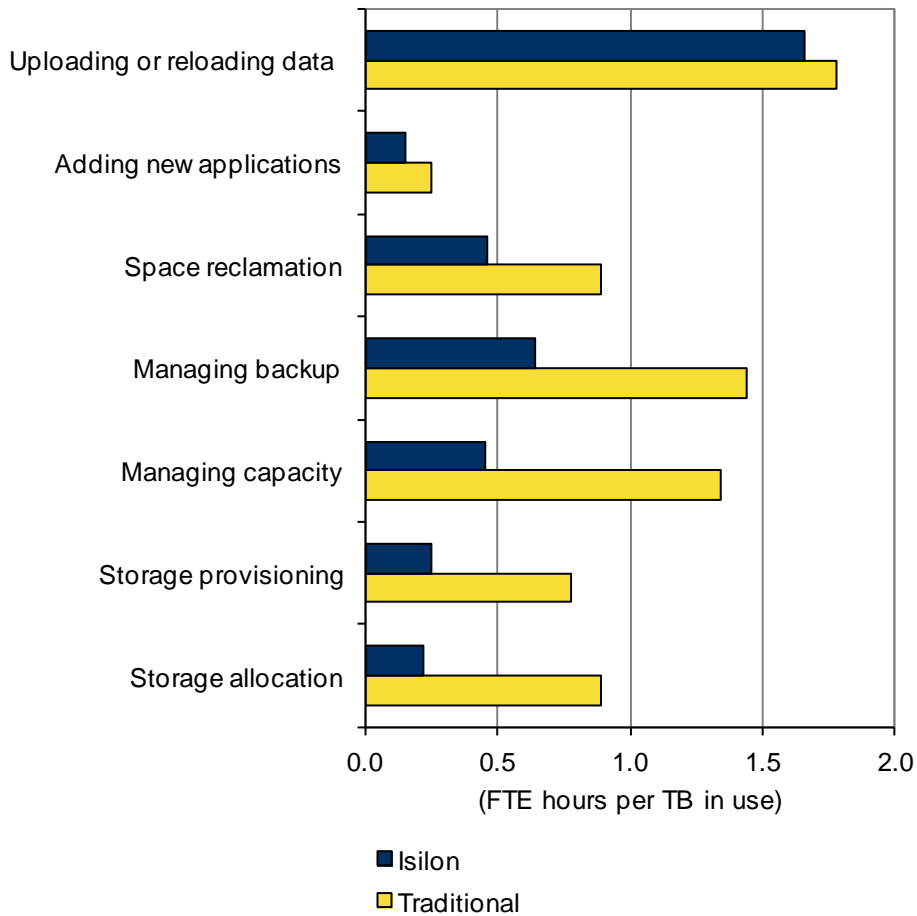
Source: IDC, October 2011

**IT Productivity**

For companies to remain competitive, IT managers and their staff must use their time more effectively. Besides reducing costs, gains in IT staff productivity can free up staff to implement new initiatives more rapidly, helping to create a competitive edge. In the era of Big Data, companies have to deal with ever-increasing storage needs driven by the continual automation of business process (new applications), increased regulation, and the proliferation of new content (social networking). IT productivity gains mean that companies do not have to grow their staffs to meet new challenges. By efficiency gains that may come from managing capacity more efficiently, faster provisioning, and reducing the time spent on backup and recovery operations, companies avoid increasing their operations costs. Totaling the IT productivity gains shown in Figure 2 reveals that IT productivity gains reduced the operating cost per TB by 3.5 FTE hours or 48%.

**FIGURE 2**

IT Productivity



Source: IDC, 2011

**End-User Productivity**

When users are unable to access stored data and applications, their productivity may be severely impaired, so increasing storage availability can generate significant savings. Minimizing business disruptions during upgrades to storage capabilities also helps to increase user productivity. To quantify end-user productivity gains, IDC multiplies the increased availability of the data and applications that end users need to do their jobs by a loaded annual salary scaled by a productivity factor (in this case, 68%, which is the aggregate average for companies interviewed). The primary driver for increasing user productivity was reducing downtime. The companies in our study reduced downtime by 95%, and four companies eliminated it altogether. One company had zero downtime — both before and after implementing Isilon. The net result was that companies in our study realized savings of \$138 per user, which translated to \$69 per TB (see Table 4).

**TABLE 4****End-User Productivity**

	Traditional	Isilon	Savings	Change (%)
Downtime hours per year	16.09	0.86	15.24	95
Hours per year per user	5.13	0.27	4.86	95
Cost of downtime per user	\$146	\$8	\$138	95
Cost allocated per TB in use	\$73	\$4	\$69	95

Source: IDC, October 2011

**Deployment Drivers**

In the interviews, the companies gave a variety of reasons for deploying the EMC Isilon solutions.

Ease of management and scaling was important to a leading provider of renewable-energy risk analysis because the company had no one dedicated to storage. "The software makes managing storage very easy," a company spokesperson said. "Someone who is not very tech savvy can run a sophisticated and powerful storage system."

Since the deployment, the company has added 300TB of storage without having to hire anyone. "We would have needed two more people for storage with the old system we had," the spokesperson said. With its single namespace, the EMC Isilon solution has also simplified storage allocation. "We used to have several meetings a week involving four or five people to determine who gets what and where. We just don't have those meetings anymore."

Scalability and manageability, along with performance and cost, were key reasons for a provider of cloud-based services to choose the EMC Isilon solution. "Scalability was number one, followed by the cost for capex," a company spokesperson said. "Performance was number three and manageability number four."

Performance is related to scalability, the spokesperson noted. "As we get new customers, the data volume increases and we just add nodes. With other vendors, when you hit the wall, you need to rip and replace, which would be very disruptive in our 24 x 7 environment."

The easy scalability of the EMC Isilon solution also means the company can add customers of any size with the confidence of being able to handle the extra data volume. "If we had gone with the other vendors, we would have had to pass on some opportunities, especially with the biggest companies. Over the past year, we've added three or four customers we wouldn't have been able to go after with any confidence — and that's \$3 million a year in revenue."



The EMC Isilon solution also helped a medium-sized data protection firm go after large customers. "As we grow, we can add a node in 60 seconds, which means we can take on large customers and also be more responsive to existing customers," a company spokesperson said. "Without the ability to respond this quickly, I think we'd probably lose 15 or 16 customers a year."

The EMC Isilon solution also helps the company avoid capital spending until the extra storage capacity is needed. "We can't afford to spend on unused capacity," the spokesperson said. The company has also been able to save because of the easier storage and application management. "That's because of the simpler interface," the spokesperson explained. "You can get in, do what you need to do, and get out without wading through screens and wizards."

The easy scalability of the EMC Isilon solution and the ability to store all its images on a single file system were the main selling points for a medical institution devoted to cancer treatments. "With the other vendors, I'd have to segregate the images, which convinced me I had to have Isilon," the organization's spokesperson said. "Also, I would have needed more people with the other systems." The institution can now add storage faster and purchase high-performance instruments knowing that there will be enough storage capacity. "We can do bigger science now and generate bigger data," the organization's spokesperson said.

The EMC Isilon solution was an obvious choice for a software development firm dealing with very large data mapping sets. "A given data set would exceed the volume limits of other vendors, so we would have needed dedicated staff to carve up new volumes," a company spokesperson said. "With Isilon, there are none of these limits."

The EMC Isilon solution was also less expensive than its competition. "The pay-as-you-grow feature conserves capex, but the place where you really save is in management," the spokesperson said. "With Isilon, storage provisioning takes maybe an hour a year. Before it took three to four hours a week. Storage allocation used to take five to six hours a week, and now it takes six hours a year."

Performance was another key reason for choosing the EMC Isilon solution. "Every time we add capacity, we add performance — and we don't have to rip and replace; we just keep growing it," the spokesperson said. Because of the improved application performance, user productivity has increased by 30% to 40%, the spokesperson said. "We're much more competitive because of that, and we're winning more jobs. Our three biggest competitors have all bought Isilon [solutions] since we did."

An online service provider selected the EMC Isilon solution several years ago when it needed to make some fundamental changes to its storage infrastructure. "Storage is a core part of our business," a company spokesperson said. "We needed a solution that would allow us to manage petabytes of storage with a tiny staff and scale easily so we didn't have to worry about creating and managing volumes or managing a bunch of other things that create costs. The other vendors couldn't do it like Isilon."

"Isilon also integrates things into their software stack that you would otherwise have to spend time building and integrating," said the spokesperson. "That takes development and test time, and not just for the original deployment, but ongoing."

When a major public library compared storage vendors, EMC Isilon was the easy winner. "None of the other vendors came close on the amount of storage provided for the price," a spokesperson said. Easier management was also a factor since the library has no true storage administrator. The solution was also conducive to virtualization. "It lets you present hundreds of terabytes as one big volume to the user," said the spokesperson. "Also, as you scale out, it adds CPU cycles and network bandwidth."

A multimedia service provider chose the EMC Isilon solution to provide a single, shared storage platform to allow its customers to create, edit, and distribute content at higher capacities and speeds. "We did a comprehensive vendor review and found that Isilon's road map tied in nicely with what we were doing," a company spokesperson said.

The prior system couldn't scale to keep pace with the company's business growth, limiting its ability to expand client service offerings and complicating IT management. With the EMC Isilon solution, the company has been able to seamlessly unify and scale its storage infrastructure in line with business growth, improving its competitive edge and simplifying IT management.

## **FINAL THOUGHTS AND ESSENTIAL GUIDANCE**

IT managers face several challenges as data growth and data management demands strain storage environments. Scale-out storage promises to mitigate those challenges, offering a data growth and management solution that is cost-effective in Big Data environments.

Implementation decisions are critical for IT managers. Costly reconfigurations and migrations can be avoided with the right solution. When evaluating suppliers, organizations must consider factors beyond simply price and performance. IT managers and storage executives need to seek solutions that scale and support the storage and business growth that strengthens a company's competitive position while reducing management and cost overheads. From this study, three critical areas came to the fore:

- ☒ Increase in storage efficiency
- ☒ Cost reductions in storage, facilities, and power
- ☒ Gains in IT productivity

With Big Data, the opportunity arises to now liberate data from any single application and infrastructure. Big Data enables enterprises to create new and innovative models for serving their customers. At the foundation of Big Data technologies is a reliable, flexible, and virtualized infrastructure. At the base of that infrastructure is an efficient, easy-to-manage, and scalable storage system.

---

## **Copyright Notice**

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2011 IDC. Reproduction without written permission is completely forbidden.