Lab Validation Report

Panasas ActiveStor

High Performance HPC Storage for the Enterprise

By Tony Palmer, Senior Lab Analyst

January 2013

Contents

ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about emerging technologies and products in the storage, data management and information security industries. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by Panasas.

Introduction

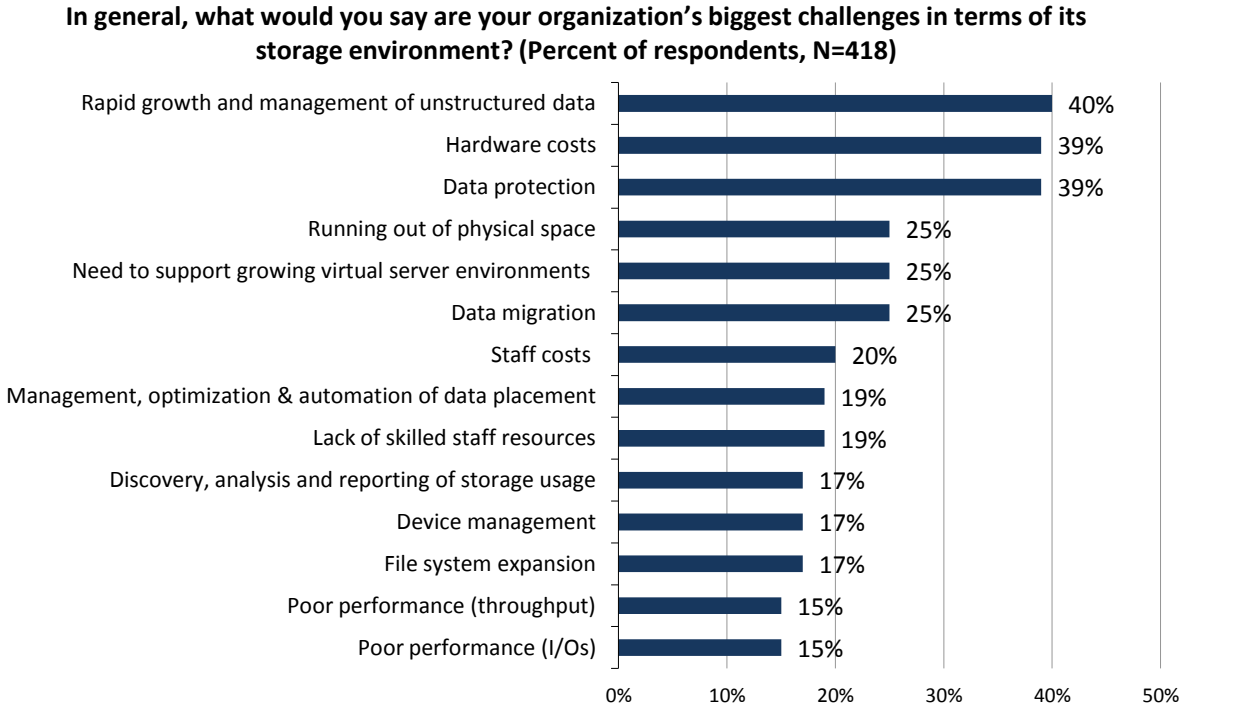
This report documents the results of ESG Lab hands-on testing of the performance scalability of the Panasas ActiveStor 11 and ActiveStor 14 high performance computing (HPC) storage systems. ActiveStor is a plug-and-play, scale-out NAS solution designed with enterprise manageability in mind. ESG Lab evaluated the solution with an eye on validating the claims of appliance-like simplicity in support of large scale storage environments with demanding HPC performance requirements.

Background

IT managers within organizations of all sizes are challenged by the cost and complexity associated with managing growing volumes of digital data. More than half of IT managers within enterprises (1,000 employees or more) surveyed by ESG (54%) reported managing 250TB or more of data[1] and they're struggling to keep up with data growth. Nearly half of the respondents to a separate ESG research survey (46%) reported that data is growing at an annual rate of 20% or more.[2] For those organizations, data volume is doubling every two to five years.

Faced with these challenges, it's no surprise that the top enterprise storage challenges reported by IT managers would include rapid growth and management of unstructured data, hardware and staff costs, and even simply running out of physical space, as shown in Figure 1.[3] It should also be noted that the need to support growing virtual server environments and the management, automation, and placement of data made the list of top ten challenges as well.

Figure 1. Storage Challenges



Source: Enterprise Strategy Group, 2012.

But traditional file storage is only part of the challenge. Large enterprises have been using HPC to garner a competitive advantage for a long time. From technology development, to financial modeling, to seismic analysis and bio/pharmaceutical applications, HPC is critical in the development of the products and services that drive

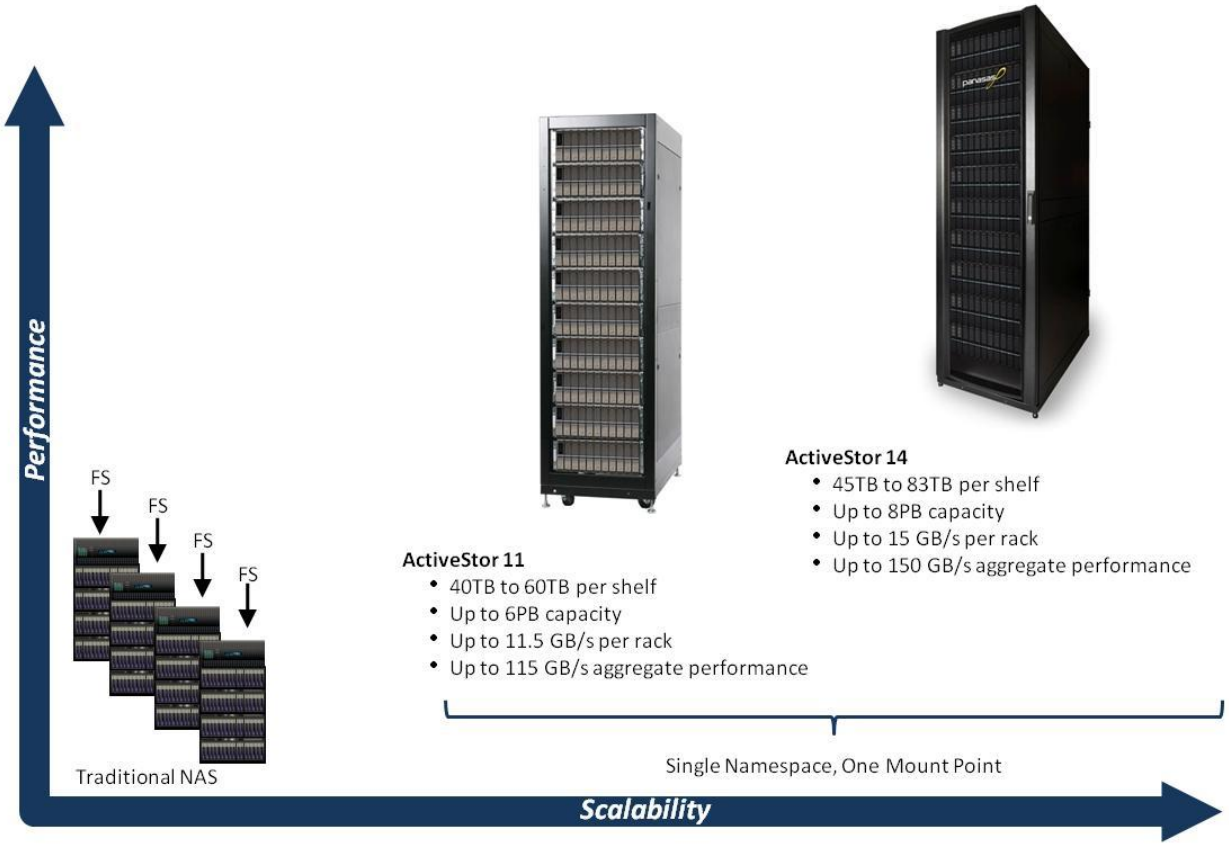
[1] Source: ESG Research Brief, *Disk-based Storage Capacity Trends*, September 2012.
[2] Source: ESG Research Report, *Trends in Data Protection Modernization*, August 2012.
[3] Source: ESG Research Report, *2012 Storage Market Survey*, November 2012.

revenue in the enterprise. HPC has a special set of requirements that traditional file storage solutions cannot address, including support for very large file systems and the ability for a large number of systems to read from and write to not only the same *file share*, but the same *file* with high throughput, to name two. While storage performance is important but not a top-ten challenge for enterprises surveyed by ESG, storage performance in an HPC context is paramount. This is because storage performance has a direct impact on HPC application performance: if compute nodes in an HPC cluster are waiting to access data needed for processing, the job takes longer to complete, leading to delays in time-to-market and potentially lost revenue.

Panasas ActiveStor

Panasas ActiveStor is a networked, parallel storage system designed to support large scale storage deployments with manageability and ease-of-use geared for the enterprise user. ActiveStor 14 is the fifth generation of Panasas's storage blade architecture, which runs the Panasas PanFS parallel file system and provides simultaneous, multi-protocol file services over DirectFlow, NFS (with Parallel NFS support planned), and CIFS. ActiveStor 14 is the highest performing member of the ActiveStor family; Panasas also offers ActiveStor 11 as a high-value alternative. ActiveStor architecture is comprised of two primary building blocks: director blades manage metadata and provide access via traditional NAS protocols, while storage blades each serve their content directly to clients using the Panasas DirectFlow client. These two independent types of blades enable ActiveStor and PanFS to aggregate capacity and performance in a linearly-scalable system. ActiveStor is engineered for highly demanding HPC applications in both core research and commercial sectors. The Panasas ActiveStor portfolio is shown in Figure 2.

Figure 2.The Panasas Parallel Storage Portfolio



ActiveStor 14 includes Solid State Disk (SSD)/SATA storage tiers designed to be an intelligent, unified, and cost-effective way to address both large file throughput and small file IOPS workloads. By storing all file system metadata and files up to 60KB in size on SSD, small file IO, directory listings and other metadata-heavy tasks are fast and preserve the ability of ActiveStor 14 to deliver the best possible streaming throughput performance at the same time. Because the use of SSD is an automatic, integral part of the file system, no additional systems

management is required to set policies or migrate data to and from SSD storage. The heightened performance characteristics of ActiveStor 14 also benefit high reliability and availability – storage blades with 4TB hard drives can be rebuilt in the same time or faster than on ActiveStor 11 or ActiveStor 12 with 3TB drives.

Panasas ActiveStor products are hardware compatible across families, protecting user investment while helping to solve complex computing problems and accelerate innovation.

- **Extreme Performance** – Up to 150GB/sec throughput or 1.4 million IOPS, with simultaneous multi-protocol access to a single file system.
- **High Scalability** – Modular hardware architecture scales seamlessly and online to meet customer storage needs.
- **Data Integrity** – Object-based RAID protection inside the file system delivers extremely fast RAID rebuilds and eliminates the implementation headaches of third-party RAID controllers.
- **Enterprise-class Manageability** – Panasas provides a single point of management over a distributed and highly scalable architecture to eliminate data and performance "islands."

ESG Lab Validation

ESG Lab performed hands-on evaluation and testing of the ActiveStor 11 and ActiveStor 14 high performance storage systems at Panasas headquarters in Sunnyvale, CA. Testing was designed to demonstrate the scalability and performance of ActiveStor systems using industry standard tools and methodologies. Ease-of-deployment and system management were also examined.

Getting Started

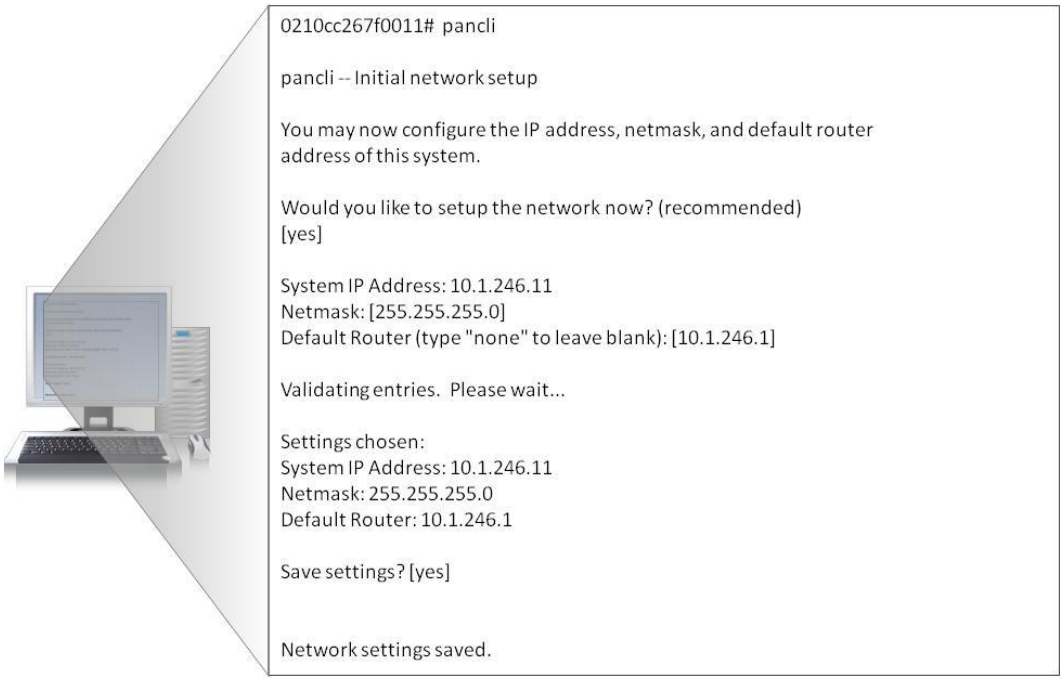
ESG Lab testing began with an ActiveStor 14 system configured with one director blade and ten storage blades connected with two aggregated 10Gb Ethernet links to a Dell Force10 S4810 switch, as were 24 Linux clients (see Figure 3). The Linux clients accessed the storage using both DirectFlow and standard NFS protocols. An ActiveStor 11 system was also attached for additional testing.

Figure 3. The ESG Lab Test Bed

ESG Lab Testing

ESG Lab examined the ease of deploying a new ActiveStor 14 storage solution by walking through the steps required to install, configure, and connect it to the system. The hardware was pre-installed in the rack with one ActiveStor shelf connected to the network switch. ESG Lab began testing by powering up the shelf and opening a CLI session with a terminal connection to the director blade as shown in Figure 4.

Figure 4. IP Address Assignment

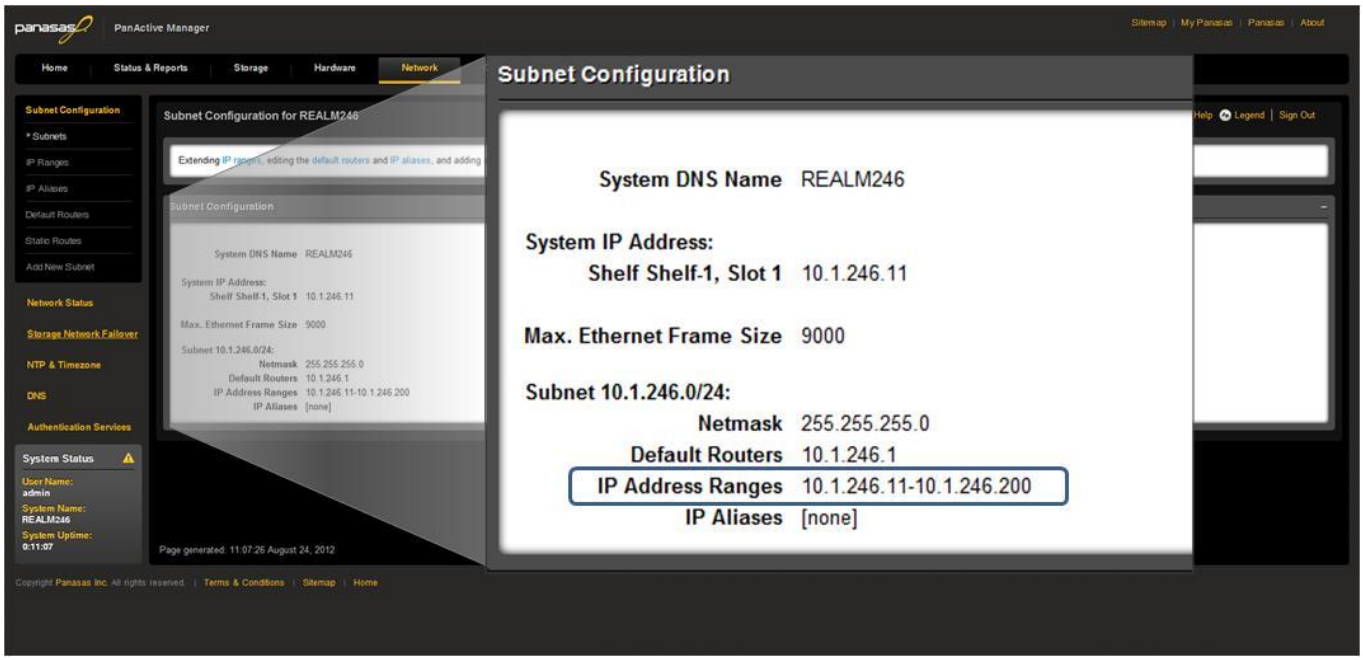


The only step required in the CLI session is the assignment of an IP address for the director blade to allow a web-based connection. ESG Lab assigned an IP address, netmask, and default router and closed the CLI session.

ESG Lab completed the remainder of the configuration steps via the install wizard which is accessed using PanActive Manager, a browser-based connection to the director blade. The first step is to install a license, which ESG Lab accomplished by entering a customer ID provided by Panasas on the license page. ESG Lab was then able to assign a password to the administrator account on the "Administrator" page of the wizard.

As shown in Figure 5, the next step required for setup is a configuration of the subnet to be used for the storage system. This allows an administrator to assign a range of IP addresses that are automatically assigned to storage blades as they come online and to assign a unique name for the storage system that can be used in client mounts. ESG Lab created a DNS name called "REALM246" and a subnet for the blade system and assigned a range of IP addresses for the system blades.

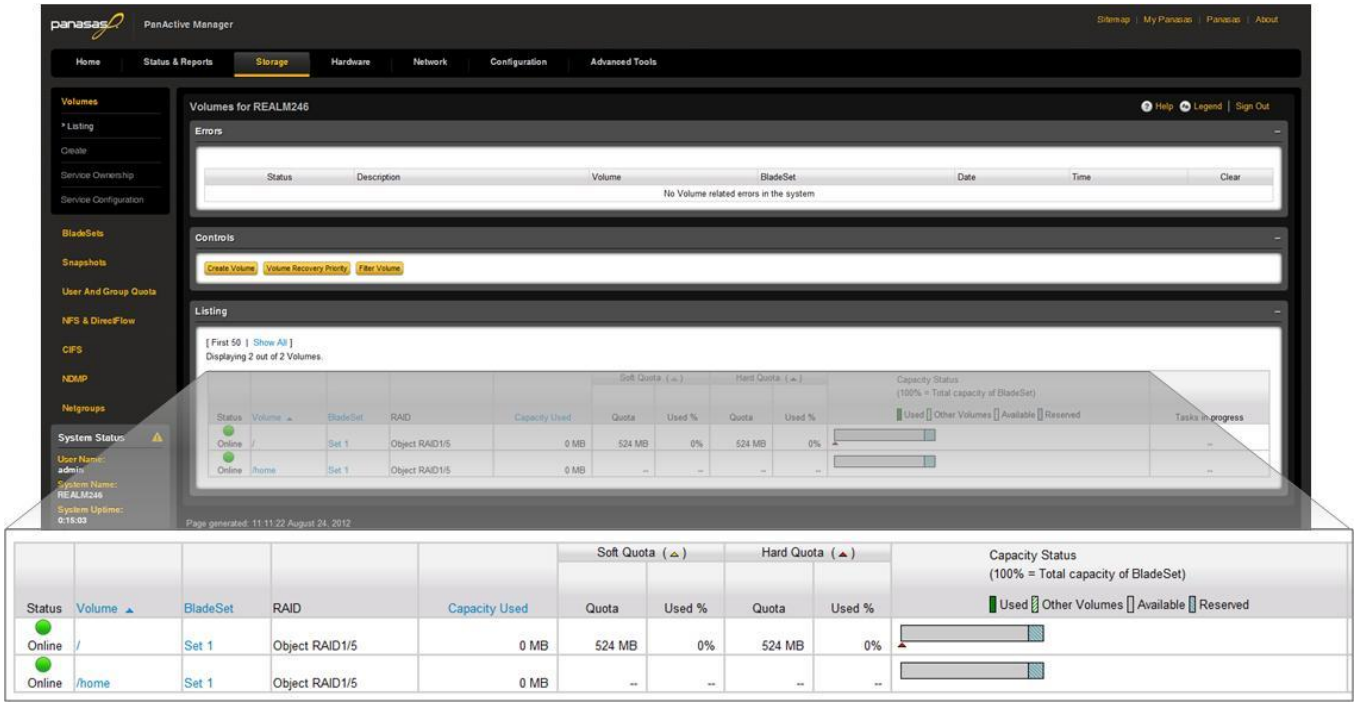
Figure 5. Configuring the Network Subnet



The next required step was to configure network services such as DNS and NIS domain names, which are used for web services, NTP, and SMTP. ESG Lab used the same domain name, REALM246, and accepted the default addresses for NTP. CIFS and NFS setup was the final step. The CIFS option could be activated to add the storage system into a Windows domain. ESG Lab activated NFS to allow for mounts from the Linux clients.

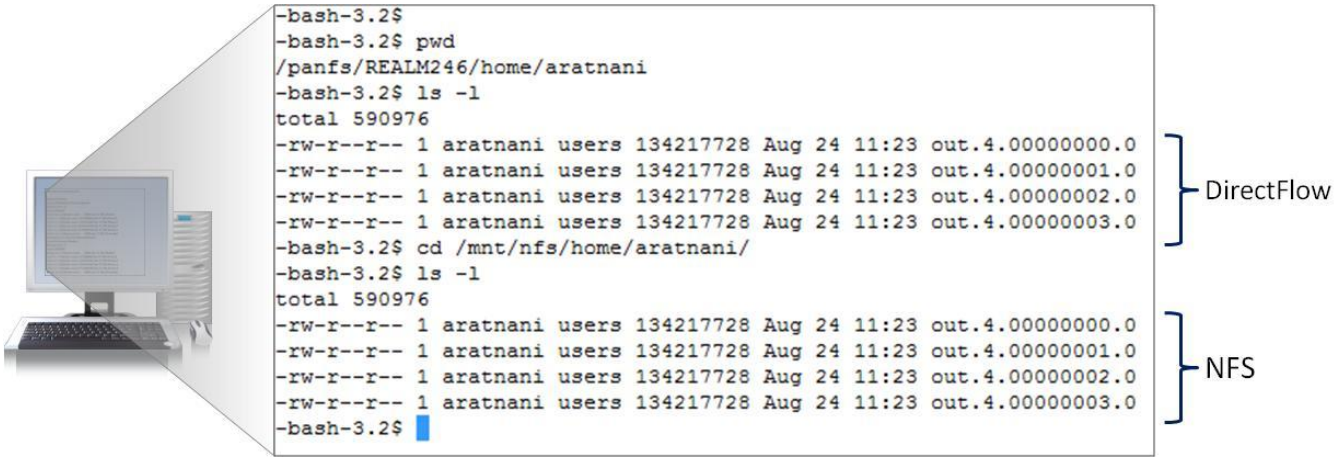
After confirming the configuration, ESG Lab observed that the system initialized the ten storage blades and added them to a blade set labeled "Set 1." In addition, ESG Lab examined the volumes configuration page, as seen in Figure 6, and noted a volume labeled "/home" that was automatically created and mounted to allow immediate connectivity tests from a client.

Figure 6. The Home Volume Mounted and Exported



ESG Lab used one of the 24 Linux clients to test connectivity to the new storage with both DirectFlow and NFS. The Panasas PanFS storage operating system supports NFS, CIFS, and DirectFlow connections simultaneously. DirectFlow is a Panasas-developed network protocol that allows high performance compute clusters to access storage in parallel. DirectFlow predates Parallel NFS which is part of the NFS 4.1 standard and many features of DirectFlow were incorporated into the standard.

Figure 7. Accessing the Home Volume via DirectFlow and NFS



The DirectFlow client is a kernel loadable module that was configured before testing by installing a standard RPM package. ESG Lab accessed a Linux client with a terminal session and created a mount for both DirectFlow and NFS to the /home directory. As Figure 7 shows, ESG Lab was able to list the contents of the /home volume with both DirectFlow and NFS access. In addition, ESG Lab created a new file in the /home volume using both DirectFlow and NFS.

From start to finish, ESG Lab was able to install, configure, and begin accessing an ActiveStor storage system in under ten minutes with 24 Linux clients accessing the /home volume and ready to process work requests.

Why This Matters

While companies that require a high performance computing environment have long dealt with the cumbersome administrative overhead of managing that environment, the tradeoff is worth it if the performance benefits are there. Ideally, though, these organizations would like to have the best of both worlds: HPC results delivered with tools that make managing the environment simpler and automated.

ActiveStor provides an answer to both challenges with a system that extends the HPC benefits established in earlier Panasas storage solutions along with a simple and intuitive administrative tool to manage hardware, storage, high availability, and reporting functions that have traditionally been handled manually through burdensome command line and scripting tasks.

ESG Lab was able to set up a fully functioning, high performance ActiveStor storage system quickly and easily and was accessing volumes and files with a Linux client in under ten minutes. In addition, ESG Lab was able to access the functions required to manage and maintain the ActiveStor system via a simple web-based administrative tool.

Scalable Performance

Panasas PanFS File System

The Panasas PanFS file system is the foundation of the Panasas family of scale-out NAS solutions. PanFS creates a single pool of storage with either single or multiple mount points under a global namespace that provides customers with the flexibility to support multiple applications and workflows from a single point of management. A single pool of storage enables extreme performance scaling for complex technical applications. System reliability and availability are protected by PanFS through dual parity data protection and active capacity balancing. PanFS supports Panasas DirectFlow, NFS, and CIFS protocols simultaneously.

The Panasas PanFS File System is designed to scale capacity and performance linearly leveraging the concept of object-based storage, which is fundamental to the PanFS architecture. In PanFS, individual elements of the system can serve up their file content directly to clients as needed. This eliminates the centralized points of service typically found in traditional (NAS) systems, which can limit performance in large scale deployments. One of the key functions that makes this possible is the decoupling of the control path (file system metadata) from the data path (reads and writes). These functions are serviced by specific types of blades—director blades manage and control the system, storage blades address file content.

Performance Methodology

The IOR and mdtest performance utilities were used to generate workloads and measure performance during each of the tests presented in this report. The IOR benchmark was developed as an HPC parallel IO benchmark by the Department of Energy at Lawrence Livermore National Laboratory. IOR is an open-source project and available for download at Sourceforge.net.[4] The IOR version used was the current released version of 2.10.3. ESG Lab used POSIX for I/O operations and MPI for process synchronization. The mdtest utility is also an open-source project available for download at Sourceforge.net. ESG Lab used mdtest version 1.8.4, not yet available from Sourceforge, which is capable of using user-specified file sizes rather than all zero-length files – this version was selected to produce a workload that is more in keeping with real-world environments.

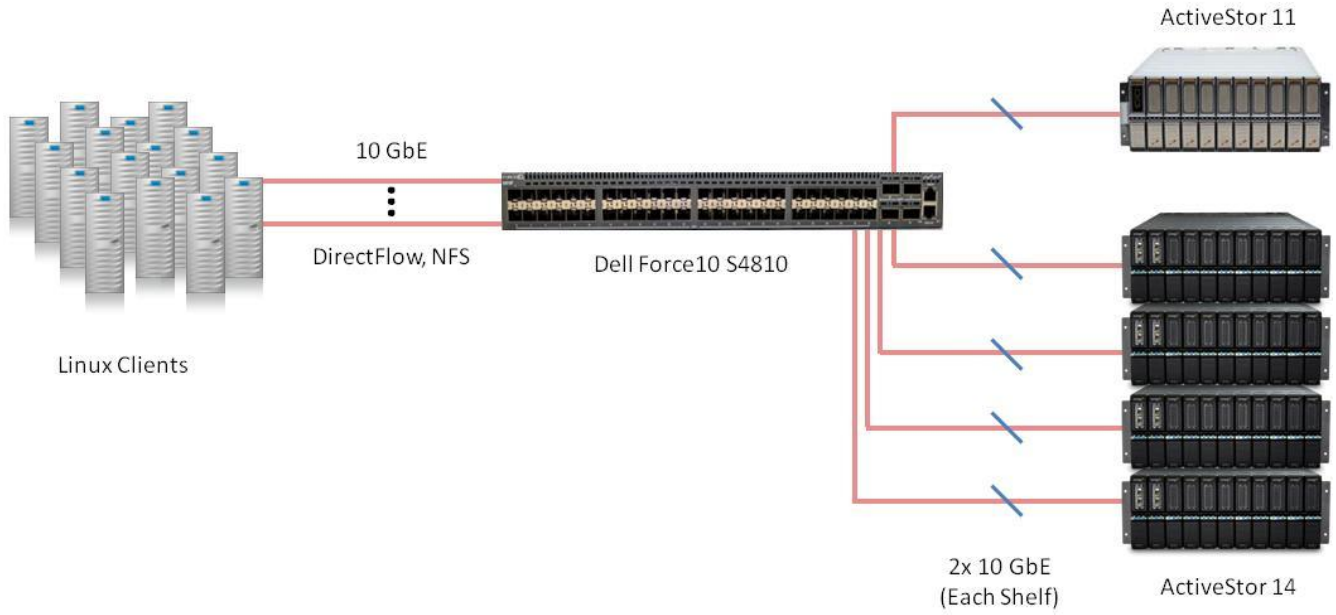
ESG Lab audited the source code of both utilities and verified that they use standard file system calls to open, write, read, and close files. Performance data was observed independently at the accessing client to verify the authenticity of performance results reported by the tools.

ESG Lab Testing: Throughput Performance

Performance testing was executed first with one shelf of both ActiveStor models and then again with four shelves of ActiveStor 14 to demonstrate scalability, as shown in Figure 8. Each ActiveStor shelf contained one director blade and ten storage blades. The ActiveStor 11 shelf contained twenty 3TB SATA drives, for a total raw capacity of 60TB. The four ActiveStor 14 shelves contained twenty 2TB SATA drives and ten 480GB SSD drives each, for a total capacity of 179.2TB.

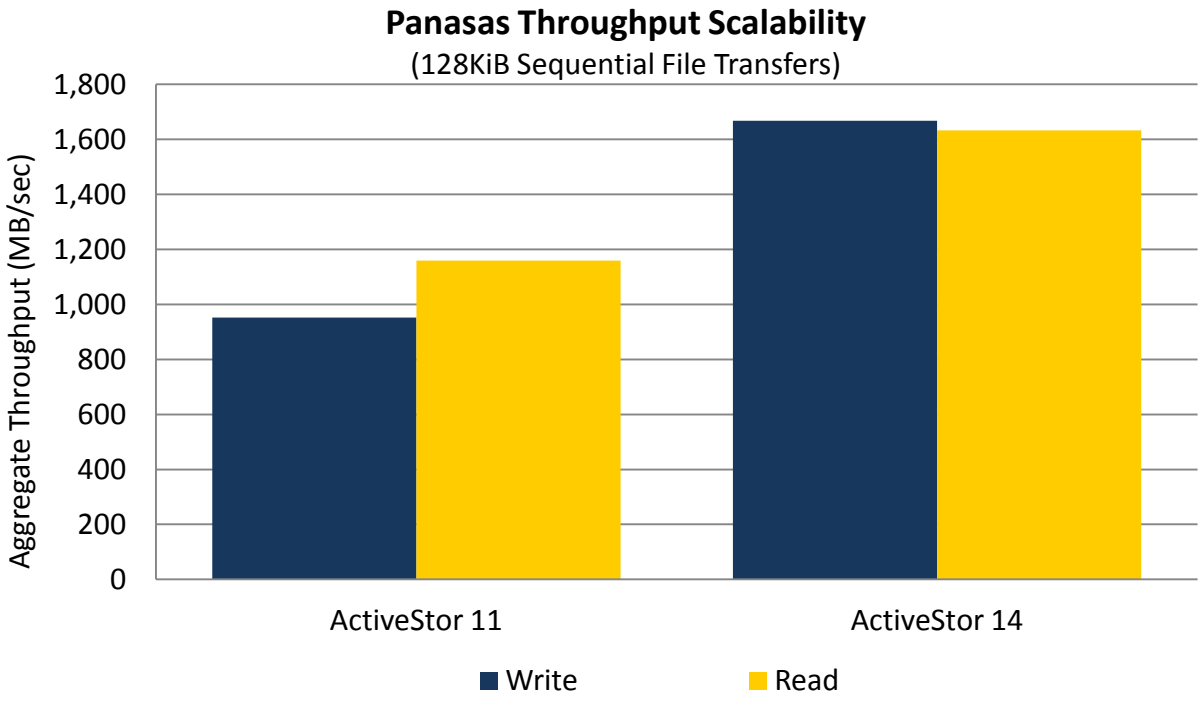
[4] http://sourceforge.net/projects/ior-sio/

Figure 8. The Performance Scalability Test Bed



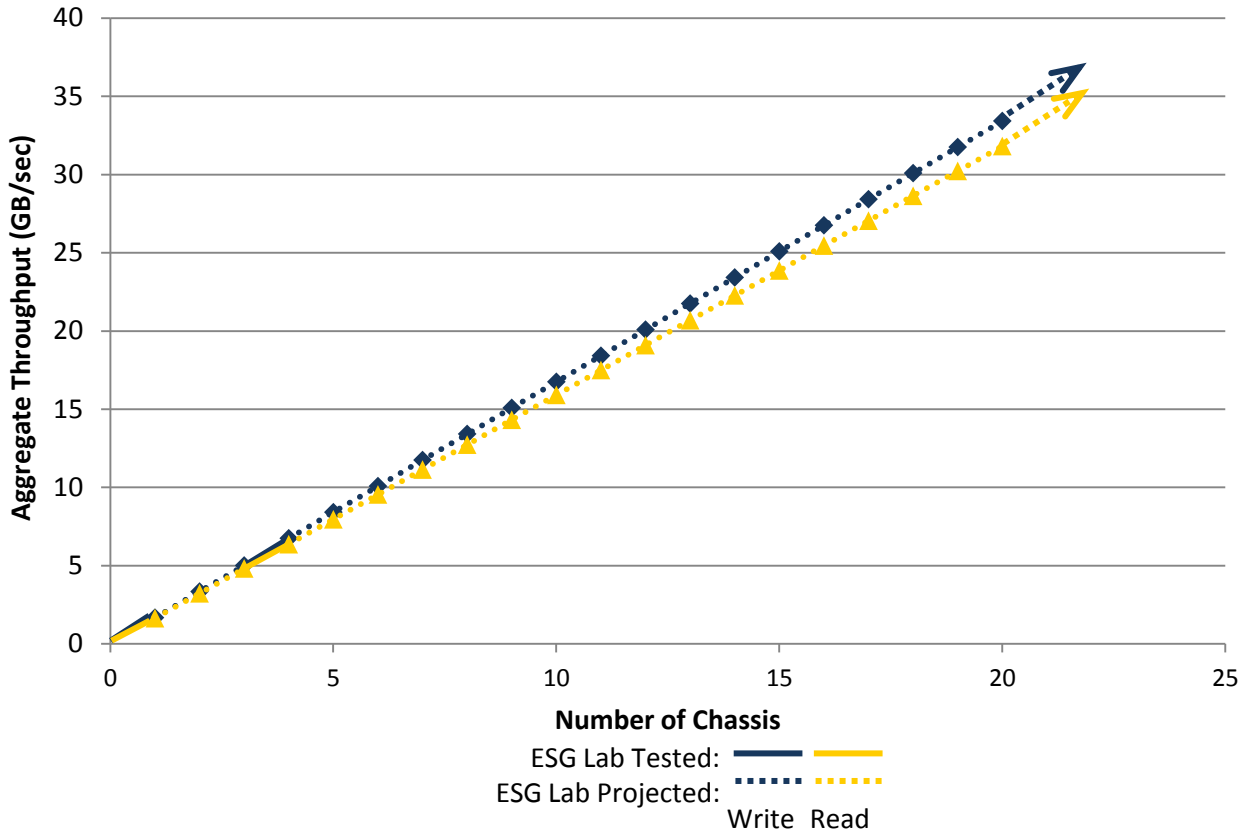
First, tests were run against a single shelf of each product using one 10GbE connection to the LAN for ActiveStor 11 and two 10GbE connections for ActiveStor 14. Then three additional shelves were added to the ActiveStor 14 blade set and the tests were repeated.[5] Figure 9 shows the results of IOR benchmark testing of one ActiveStor 11 and one ActiveStor 14 shelf.

Figure 9. ActiveStor IOR Benchmark Throughput Comparison-One Chassis



[5] See the Appendix for more configuration details.

Figure 10. ActiveStor 14 Throughput Scalability Projection



As seen in Table 1, performance scaled nearly linearly as the ActiveStor 14 system was scaled from one to four shelves. Using data collected from these tests, previous performance tests run by Panasas, and reports from multiple Panasas customers running large clusters, ESG Lab verified that Panasas ActiveStor scales nearly linearly as shelves are added while delivering performance at near-wire speed from each shelf.

Table 1. IOR Benchmark Results – Detail

Model	Number of Shelves	Number of Drives	Connectivity (10GbE)	Read (MB/sec)	Write (MB/sec)
ActiveStor 11	1	20	1	1,159	953
ActiveStor 14	1	30	2	1,632	1,667
ActiveStor 14	4	120	2	6,363	6,741

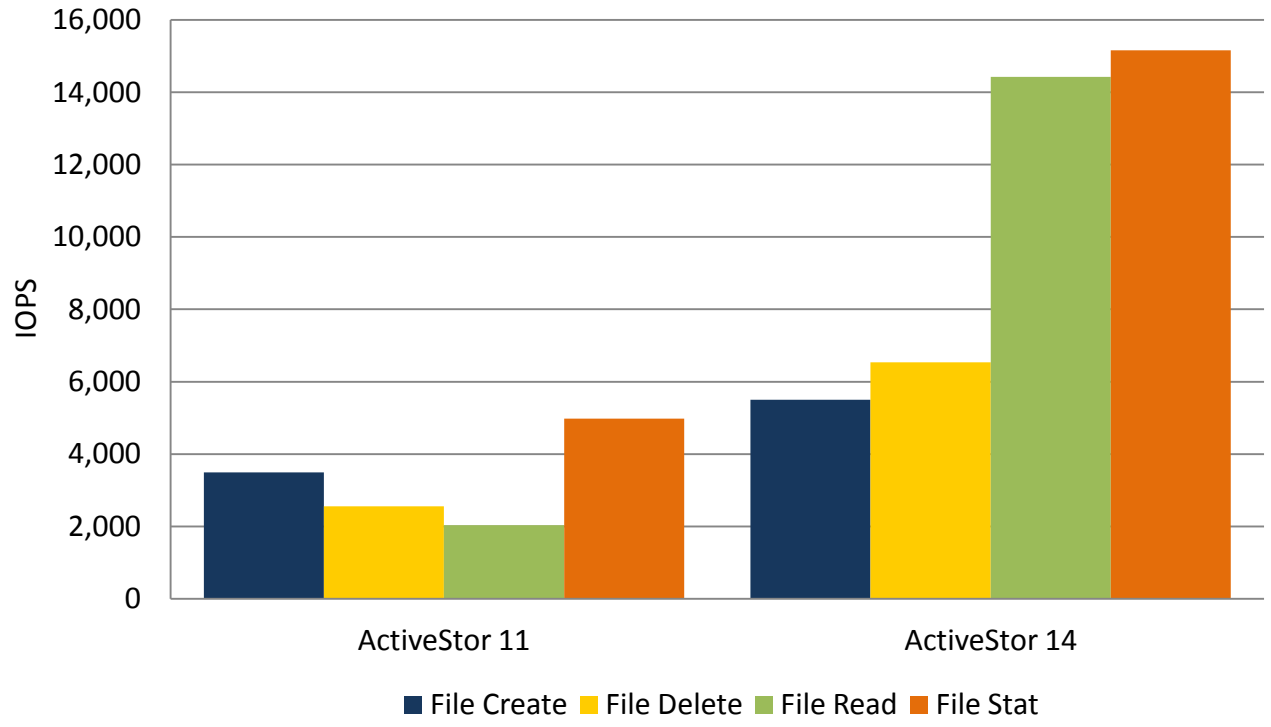
It is important to note that cache was purged and sufficiently large file sizes were used for the read tests so the results reflect the performance of reading data from ActiveStor without significant benefit of cache. In addition, volume failover to maximize file system availability was enabled for the tests, matching real-world deployment best practices.

ESG Lab Testing: Metadata and Small File Performance

ESG Lab also evaluated the ability of ActiveStor 11 and ActiveStor 14 to deliver metadata and small file performance using the mdtest performance utility. The ActiveStor configurations used were identical to those used for the throughput performance tests except that each shelf was configured with two director blades and nine storage blades. These performance tests measured the performance of a single PanFS volume on a single shelf of each product. To follow best practices, volume failover was again set up, using a second director blade in a different shelf (it was available to support a potential director blade failover but did not participate in the performance of the volume). Disabling volume failover would have improved write performance from these results but at the cost of giving up protection against a director blade failure. As a result, volume failover is always recommended except in certain scratch space implementations where performance may be more important than availability.

For this set of tests, mdtest was first used to create 100,000 4KB files, read back all of the files in the directory, perform a 'stat' command against all of the files, measuring the metadata performance of the file system, and finally delete all the files. Figure 11 shows the results of mdtest benchmarking of one ActiveStor 11 and one ActiveStor 14 shelf.

Figure 11. ActiveStor mdtest IOPS Comparison-One Chassis



Next, ESG Lab added three shelves to the ActiveStor 14 test configuration and ran the same set of tests. As seen in Figure 12, small-block IOPS performance scaled nearly linearly as the ActiveStor system was scaled from one to four shelves. Using data collected from these tests, previous performance tests run by Panasas, and reports from multiple Panasas customers running large clusters, ESG Lab verified that Panasas ActiveStor scales nearly linearly as shelves are added while delivering high IOPS performance from each shelf.

Figure 12. ActiveStor 14 IOPS Scalability Projection

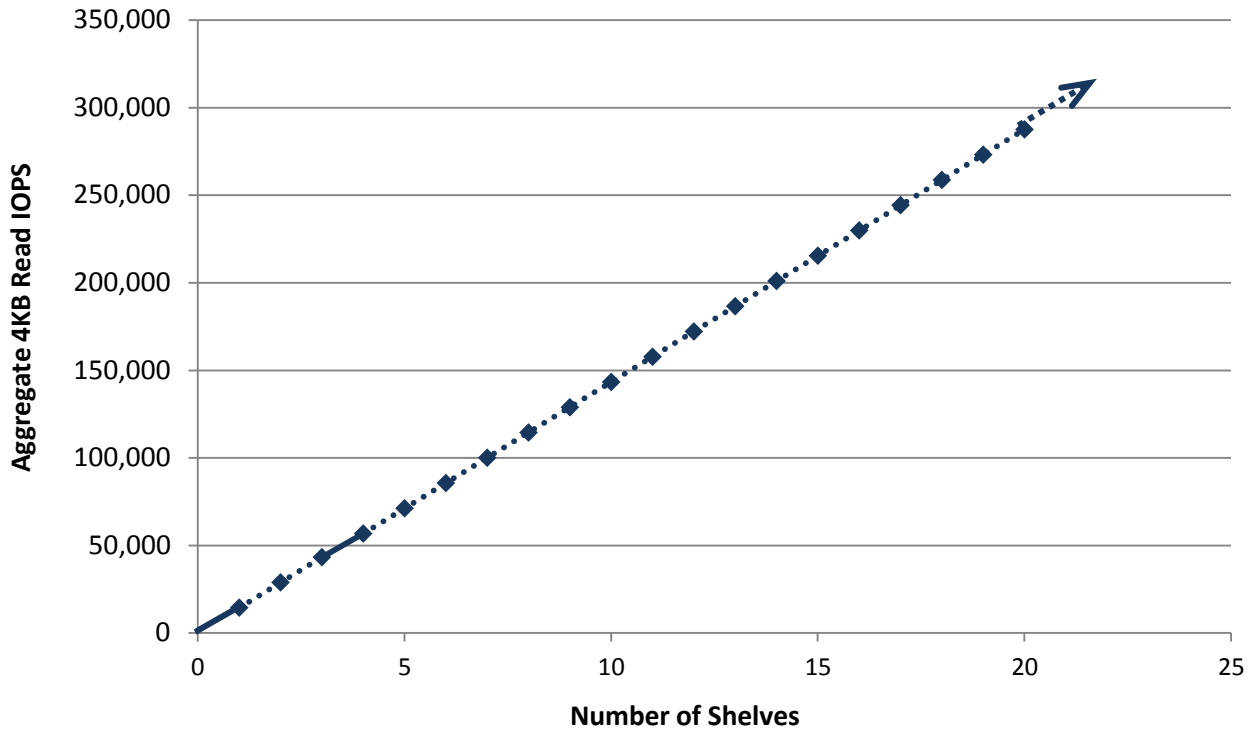


Table 2. mdtest IOPS Benchmark Results – Detail

Model	Number of Shelves	Number of Drives	Connectivity (10GbE)	Create (ops/sec)	Delete (ops/sec)	Read (ops/sec)	Stat (ops/sec)
ActiveStor 11	1	20	1	3,495	2,555	2,031	4,982
ActiveStor 14	1	30	2	5,498	6,536	14,424	15,160
ActiveStor 14	4	120	2	19,041	24,709	56,744	56,853

What the Numbers Mean

- The aggregate read and write throughput of the storage repository is one of the key factors to consider when architecting and sizing storage for an HPC environment.
- A typical HPC application (seismic analysis, for example) could require dozens, hundreds, or even thousands of nodes to access data sets of hundreds of terabytes. The ability to read in raw data and write out processed data quickly is essential.
- As an example, a 100-shelf Panasas ActiveStor 14 system could effectively enable an HPC compute cluster to read a 100TB data set into memory for processing in about ten minutes. When processing is complete, the compute nodes can write output data to disk at more than 10TB per minute.
- Metadata actions like file stats and directory listings as well as small file IO performance benefit strongly from the use of SSD storage in ActiveStor 14.
- Real-world workflows normally involve mixed file sizes, with both large file throughput, small file IOPS, and metadata access all occurring concurrently. The ability of ActiveStor 14 to accelerate small file and metadata performance makes it an excellent choice for private clouds and other mixed workload environments.
- Applications that benefit substantially from read cache may experience significantly better performance than observed here.

Why This Matters

High performance computing and big data applications share important requirements – to store, access, update and share vast amounts of file data. Application performance and productivity suffers when the storage system can't keep up with the volume of data. High performance storage, leveraging parallel I/O, provides faster time-to-results, increased productivity, increased revenue opportunities and a competitive advantage.

Enterprise applications across a broad range of industries can benefit from high performance file storage. Examples of these workloads include manufacturing design and simulation, seismic research and reservoir simulation, next generation genome sequencing, financial analytics, computer graphics and video editing, as well as academic and government research. A scale-out NAS appliance with linear scalability enables organizations to minimize their initial storage investments and reduce complexity while offering the ability to expand to meet future demands affordably.

Panasas ActiveStor is widely recognized for its high throughput performance. The metadata and small file performance results observed by ESG Lab demonstrate that ActiveStor 14 is not only relevant for large file throughput applications but also mixed workloads and private cloud implementations.

ESG Lab has confirmed that a single ActiveStor 14 shelf achieves impressive aggregate read and write throughput rates as well as IOPS-intensive small file and metadata performance; performance in both areas scales in a nearly perfect linear fashion as shelves are added to the PanFS file system. The single ActiveStor 14 shelf tested by ESG delivered 1,667MB/sec of aggregate write throughput and 1,632MB/sec of aggregate read throughput while a four-shelf ActiveStor 14 system delivered 6,741MB/sec of aggregate write throughput and 6,363MB/sec of aggregate read throughput. ActiveStor 14 scaled equally well with small file and metadata IO scaling to nearly 57,000 ops/sec with four shelves. This represents a more than 50% overall increase in throughput over ActiveStor 11 and an impressive 700% increase in small file read performance.

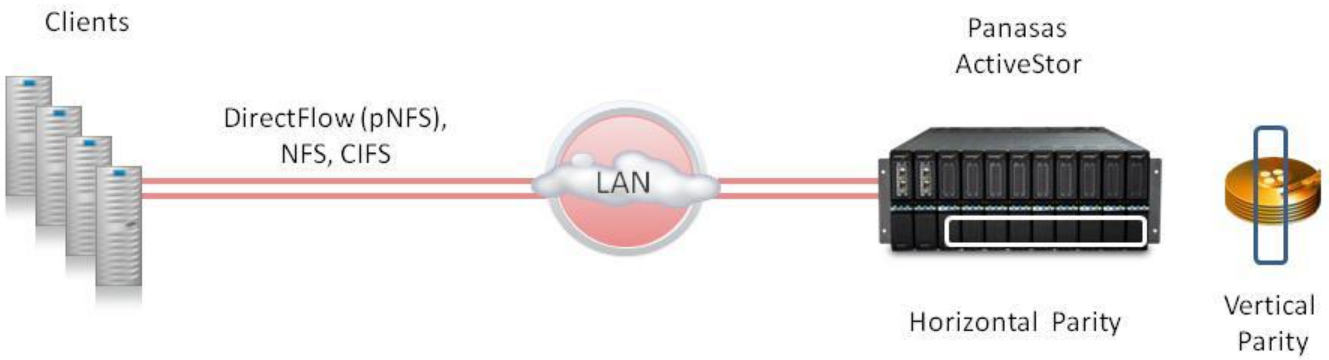
Enterprise Manageability and Reliability

Panasas is managed either from a full-featured PanActive Manager web-based interface or from its command-line equivalent. Both of these tools allow administrators to perform all important functions including:

- Managing physical hardware, logical storage, and user access
- Viewing system status and generating reports
- Managing system configuration such as security, networking, and file systems
- Performing upgrades and system diagnostics
- Managing high availability features

System reliability and availability are protected by PanFS through dual parity data protection, shown in Figure 13. In the PanFS implementation, horizontal parity protects data across blades while vertical parity protects data within each individual drive from uncorrectable sector errors.

Figure 13. Panasas Dual Parity RAID



ESG Lab Testing

ESG Lab examined several features that ease ActiveStor management, including adding storage, creating snapshots, and recovering from hardware failures. Testing began with ESG Lab adding another shelf to the storage system to examine the dynamic functionality and ease-of-use in increasing storage capacity.

Figure 14. Blade Set with One ActiveStor Shelf

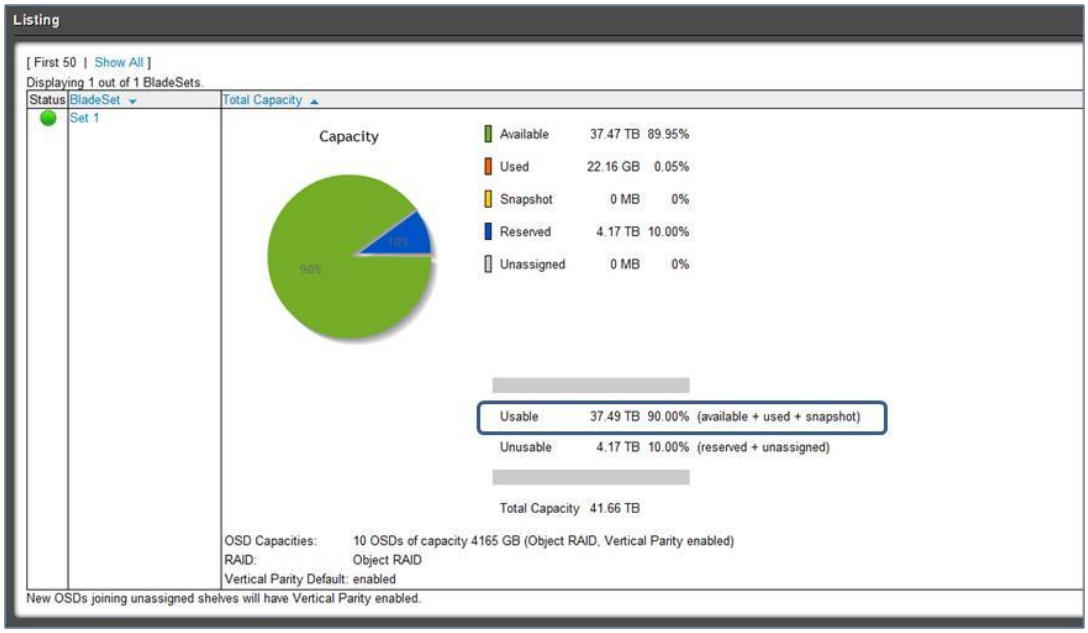
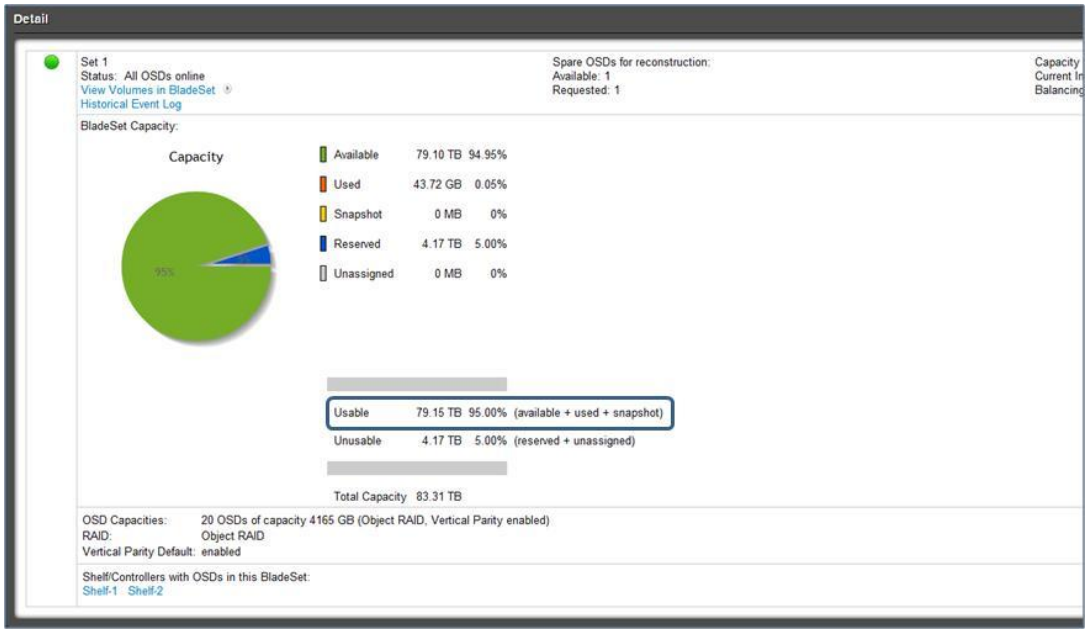


Figure 14 shows a blade set labeled “Set 1” with one shelf configured. Total capacity of the blade set is 38.96TB. ESG Lab connected a second storage shelf to the Dell Force 10 switch and monitored the results. The new shelf took approximately five minutes to initialize and was automatically added to the existing realm, REALM246. After the shelf initialized, it appeared on the hardware page of PanActive Manager unassigned to a blade set. ESG Lab was able to add the second shelf into the blade set “Set 1” in one step. In addition, the system automatically striped and load balanced the data from the first shelf.

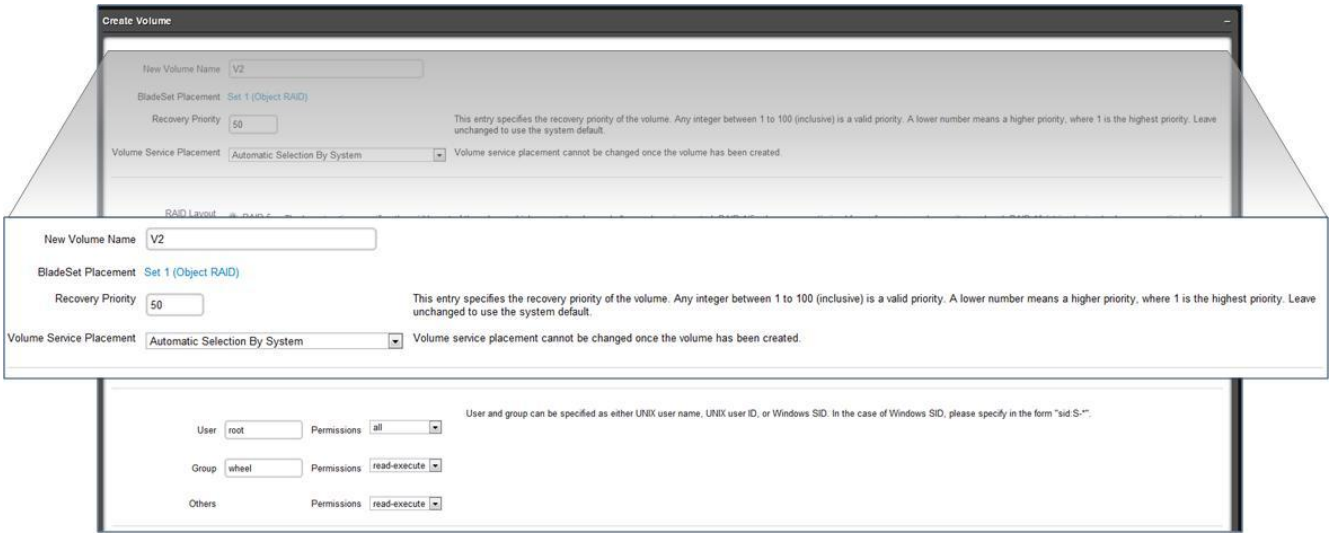
After the second shelf was added, ESG Lab examined the blade set “Set 1” and confirmed that the total capacity had increased from 37.49TB to 77.15TB.

Figure 15. Blade Set After Adding Second Shelf



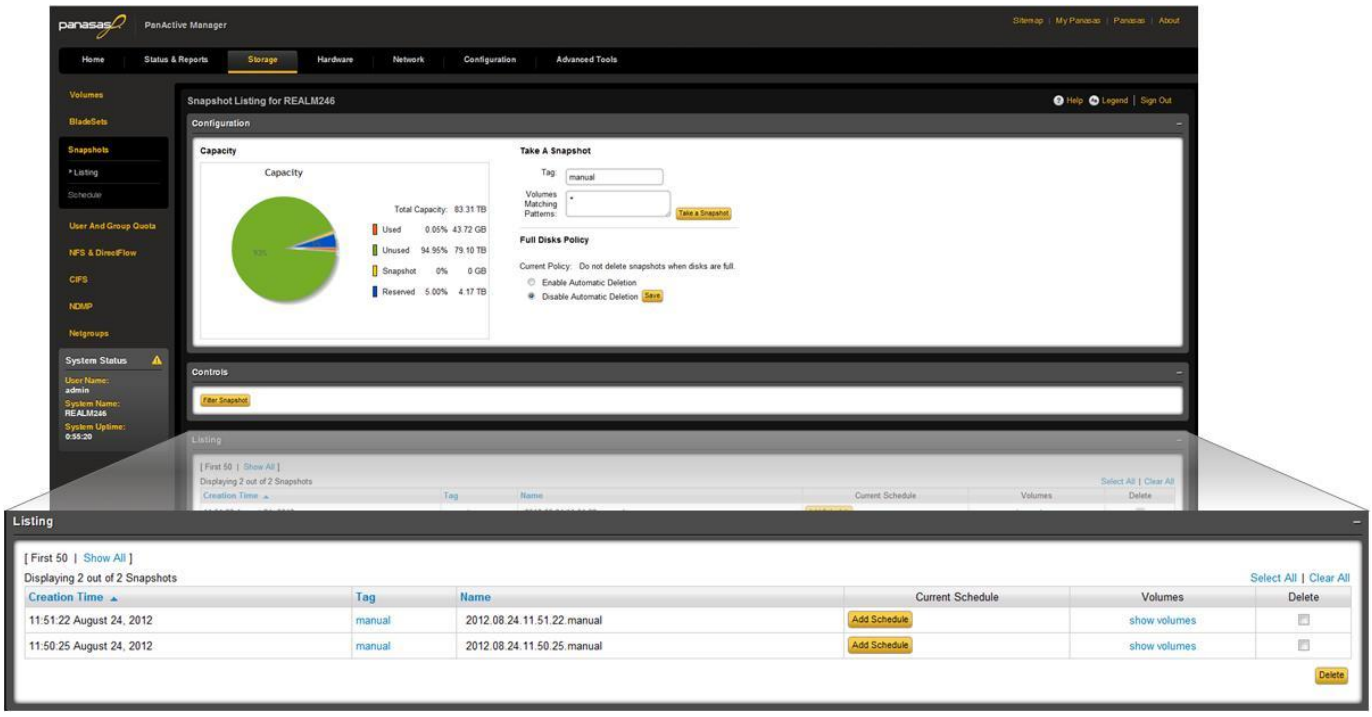
After the second shelf was added, ESG Lab tested the creation of an additional volume to the storage system. Volumes are managed by accessing the storage tab in the administrative tool and can be managed round robin by multiple director blades or assigned to a specific one. Figure 16 shows the setup page for a new volume. ESG Lab created a volume named “V2” and assigned the new volume to the blade set “Set 1.” Additional options for the volume include RAID layout, hard and soft quotas, and user permissions. ESG Lab set a soft quota of 1TB for the new volume. The volume was created instantly and ESG Lab was able to connect successfully to the volume with a Linux client.

Figure 16. Creating a Volume



Snapshot configuration with the PanActive Manager was straightforward and easy to use. ESG Lab accessed the snapshot page from the storage tab. Snapshots can be run manually or scheduled at regular intervals, and can contain an entire realm or just a volume. ESG Lab performed two manual snapshots, one of the entire realm and one of a single volume, /home.

Figure 17. Point-in-Time Snapshots

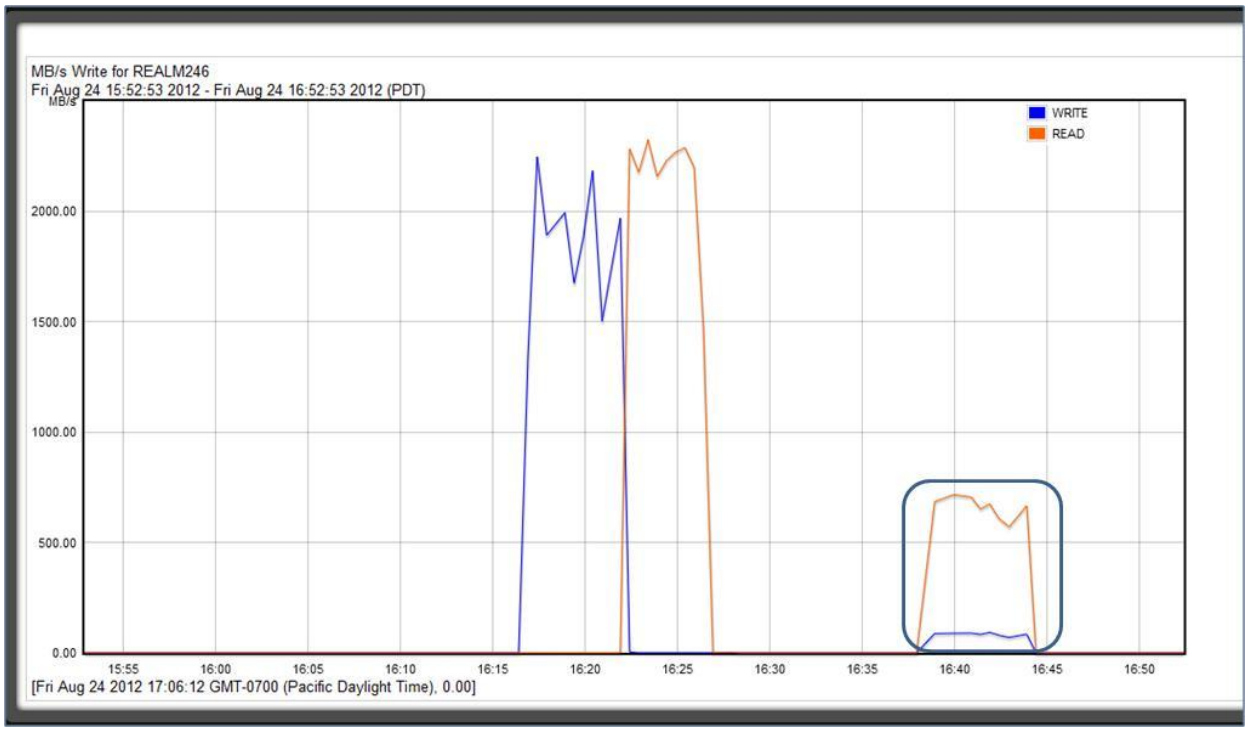


To test the snapshot functionality, ESG Lab deleted several files on volume /home using a Linux client. After the files were removed, ESG Lab accessed the snapshot directory located on the same volume. The deleted files were available and ESG Lab was able to restore the recently deleted files back to their original location. ESG Lab also examined the snapshot scheduler and found the process for scheduling snapshots to be very intuitive.

ESG Lab examined the failover functionality of the ActiveStor 14 storage system, beginning with simulation of a blade failure. ESG Lab disabled a blade in the PanActive Manager and examined the event log to record the result. The affected volumes were promptly reconstructed across the remaining blades with no interruption of service.

ESG Lab then re-enabled the offline blade and monitored the results. As the system began to rebalance the blade set automatically, the rebuild of the storage blade generated around 40MB/sec of write traffic and 400MB/sec of read traffic. Figure 18 shows the traffic results of the blade rebuild, which took four minutes and caused no interruption of service.

Figure 18. Read and Write Traffic for Storage Blade Rebuild



ESG Lab also simulated failure of a network interface by removing one 10GbE link from the aggregate and disabling the link. Before the link was removed, ESG Lab began the IOR benchmark test to create IO against the storage. As the link was removed, throughput decreased as expected but the benchmark continued with no errors. ESG Lab then restored the original link to the aggregate and the throughput climbed back to original levels with no errors or interruption of service.

Why This Matters

In a high performance computing environment, there can be no doubt that interruptions of service are extremely costly. Panasas has designed its ActiveStor 14 storage to handle the HPC demands of bio/pharmaceutical, energy, finance, industrial manufacturing and media customers that demand not just performance but the assurance that applications experience no service interruptions in the case of hardware failure. In addition, providing protection of data and easy restoration in the case of data loss can save companies hours, if not days, in lost productivity. ActiveStor 14 delivers on those requirements by providing enterprise-level functions such as high availability and snapshot backups and restores to ensure that services and data are always available.

ESG Lab tested high availability and data protection and found that hardware failures caused no disruptions in service, and storage rebuilds took minutes to complete. Additionally, data was easy to back up and restore using snapshot functionality, and adding data on the fly was simple using the PanActive administrative tool.

ActiveStor Customer Experience

ESG Lab spoke with a high-profile Panasas customer using ActiveStor to support seismic data processing and analysis. Seismic applications typically require hundreds of computational nodes and hundreds of terabytes of storage. After transitioning to a new application design intended to improve performance and shorten processing time, this customer found that its traditional NAS infrastructure could not keep up with the I/O demands of 150 compute nodes reading or writing simultaneously.

"We evaluated several clustered, parallel file systems including home grown Linux-based systems as well as pre-packaged offerings from other vendors. We found that Panasas was able to service hundreds of nodes with performance to spare while providing greatly simplified management that well suits our needs." This customer has deployed more than 100 Panasas shelves and more than 6PB of storage and has found management to be remarkably easy, especially when compared to traditional NAS or home grown clustered file systems. Another area in which Panasas excels for this customer is support: "Panasas provides excellent support. Their support engineers are expert, methodical, and understand our applications and needs. On the infrequent occasions we have had to engage support, they've analyzed and resolved our issues quickly and effectively."

ESG Lab Validation Highlights

- ☑ Using a simple and intuitive installation wizard, ESG Lab was able to set up a high performance ActiveStor storage system quickly and easily and was accessing volumes and files with a Linux client in ten minutes.
- ☑ ESG Lab confirmed that the PanFS file system performs and scales extremely well in HPC environments.
- ☑ Large, sequential file access scaled in a nearly perfect linear fashion as ActiveStor 14 shelves were added to a single PanFS file system.
- ☑ A single ActiveStor 14 shelf processing large files achieved an aggregate throughput rate of 1,667MB/sec for writes and 1,632MB/sec for reads.
- ☑ A system with four ActiveStor 14 shelves scaled to deliver 6.7GB/sec of write throughput and 6.4GB/sec of read throughput.
- ☑ Metadata actions like file stats and small file IO performance benefitted strongly from the use of SSD storage in ActiveStor 14, scaling to nearly 57,000 ops per second with four ActiveStor 14 shelves
- ☑ ESG Lab tested high availability and data protection and found that hardware failures caused no disruptions in service.
- ☑ Snapshot functionality was robust and easy to use for both the administrator and clients.
- ☑ Adding capacity on the fly was fast and simple using the PanActive Manager administration tool.

Issues to Consider

- ☑ While industry-standard NFS and CIFS, which are built into most operating systems, are fully supported, maximum performance, as achieved in the tests in this report, was obtained using the Panasas DirectFlow protocol. While easy to deploy and manage, the DirectFlow client represents an additional component that needs to be installed in order to achieve the performance levels required by HPC environments. The planned future support of the open standard pNFS protocol will alleviate the need to deploy a proprietary client to achieve these high performance levels.
- ☑ The performance results presented in this report are based on benchmarks deployed in a controlled lab environment. Due to the many variables within each production data center, capacity planning and testing in your own environment is required to determine exactly how Panasas high performance parallel storage will perform with your applications.
- ☑ The ActiveStor 14 performance tests for this report were made using the ActiveStor 14T 45TB configuration with 16GB of memory cache per storage blade (vs. the 8GB per blade available in other ActiveStor 14 models). Panasas advises that while the 81 and 83TB shelf configurations of ActiveStor 14 perform similarly to the ActiveStor 14T for large file throughput they can be expected to deliver about 20% lower metadata and small file performance, still significantly higher than the ActiveStor 11 results we measured.

The Bigger Truth

Commercial enterprises are realizing the benefits of scalable clustered computing—from financial applications, such as stock market analysis, to biotech companies performing gene sequencing. High performance computing has a special set of requirements that traditional file storage solutions cannot address, including support for very large file systems and mount points and the ability for a large number of systems to read from and write to not only the same file share, but the same file, with high throughput. Storage performance has a significant impact on HPC application performance as well: if compute nodes in an HPC cluster are waiting to access data needed for processing, the job takes longer to complete, leading to delays in time to market and potentially lost revenue.

Many vendors over many years have tried to solve the problems of predictable, cost effective scale-out NAS performance. Scaling has turned out to be a very difficult problem to solve. Most solutions work well for large or small files, but not for both—leading many companies to deploy a scale-out NAS solution alongside a commodity, entry-level NAS solution to deal with metadata. The difficulty is further demonstrated by the fact that storage industry behemoths, while making progress, have not been able to get the job done either by acquisition or internal development.

Parallel, distributed file systems exist where administrators can build their own storage clusters from the ground up using commodity Linux servers with internal storage. These parallel file systems can provide speed and availability, but are complex to implement and manage and are missing enterprise-class storage features such as snapshots, NDMP, and automated failover which address the needs of the commercial enterprise customer.

ESG Lab was impressed not just with Panasas technology and near linear scalability, but Panasas's commitment to usability, manageability, and enterprise-class functionality. ESG Lab confirmed that the PanFS file system delivers outstanding levels of performance and scalability. An excellent large file throughput rate of 1.63GB/sec can be achieved by a single ActiveStor 14 shelf scaling in near-linear fashion as shelves are added to 150GB/s in a single file system. Small file performance was also outstanding as a single ActiveStor 14 created at a rate of 19.8 million 4KB files per hour. In ESG's experience, scale-out NAS systems are usually optimized for either large file throughput or small file IOPS; ActiveStor 14 showed strong performance for both types of workloads, making it a particularly versatile scale-out NAS solution.

The challenges that scale-out NAS solves are now in the mainstream and Panasas makes the ActiveStor solution consumable by storage administrators without a PhD. Organizations facing massive file system growth and complexity challenges would be wise to consider Panasas. PanFS includes commercial features like snapshots, user and group quotas, and replication, and lets pNFS, NFS, and CIFS work on the same file system. This scale-out architecture will grow with enterprise requirements and meet the needs of the business without interruption. Panasas brings the massive scalability of high performance parallel storage, once relegated to national laboratories and universities, to the commercial enterprise.

Appendix

Table 3. ESG Lab Test Bed

Storage	PanFS Version
1x ActiveStor 11 and 4x ActiveStor 14T Shelves: 1 director blade + 10 storage blades (IOR test) 2 director blades + 9 storage blades (Mdtest)	PanFS: 5.0.0-762195.128
Connectivity	**Configuration**
Dell Force10 S4810	48x 10GbE ports, MTU=9000
Clients	**OS**
24 Supermicro servers: Dual socket, 6-core Intel Xeon X5650 CPU @ 2.67 GHz (Westmere) clients with 24GB RAM	RHEL5 2.6.18-128 (CentOS 5.3)
Benchmark Software	**Version**
IOR Interface: POSIX for IO, MPI for process synchronization	2.10.3
Mdtest	1.8.4

Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel:508.482.0188 Fax: 508.482.0218 | www.enterprisestrategygroup.com