Enterprise Strategy Group | Getting to the bigger truth.™

Lab Validation Report

SGI Modular InfiniteStorage Server with Scality RING Organic Storage

Object and file storage with enterprise-class performance and resiliency

By Mike Leone, ESG Lab Analyst, Aviv Kaufmann, ESG Lab Analyst, and Terri McClure, Senior Analyst

October 2013

Contents

ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about data center technology products for companies of all types and sizes. 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 SGI.

Introduction

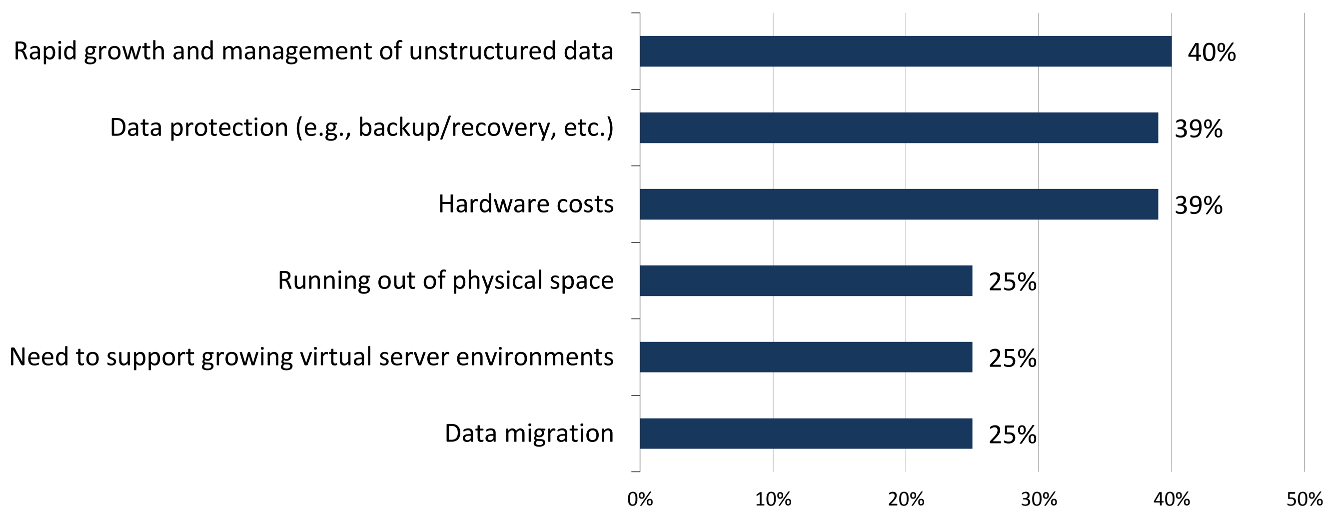
This ESG Lab Validation report documents hands-on evaluation and testing of SGI Modular InfiniteStorage (SGI MIS) servers with Scality RING Organic Storage software. Testing focused primarily on the performance of both object-based and file-based storage, each with two data protection methods, replication, and erasure coding. ESG Lab also tested the resiliency and high availability of the solution by measuring rebuild times for simulated drive failures.

Background

Today, organizations are faced with a number of unique challenges that may justify a rethinking of their storage strategies. As shown in Figure 1, respondents to ESG research reported that their primary storage challenges are the rapid growth of unstructured data, data protection, hardware costs, the increasing scarcity of physical space and data migration.[1]

Figure 1. Organization's Top-five Storage Challenges

In general, what would you say are your organization's biggest challenges in terms of its storage environment? (Percent of respondents, N=418)



Source: Enterprise Strategy Group, 2013.

While these challenges are not new to the storage industry, traditional block storage (SAN) and file storage (NAS) designs are not suited to address all of these challenges within a single architecture. A system designed around object storage technology would help address the need to deal with unstructured data, data protection, and data migration, while a cost-effective, highly dense hardware design would help to minimize cost and maximize the use of floor space. Providing this in a flexible, scale-out, distributed design would minimize the burden of planning for future growth.

[1] Source: ESG Research Brief, *Key Challenges in Data Storage*, December 2012.

SGI Modular InfiniteStorage

SGI Modular InfiniteStorage (SGI MIS) is a high-density integrated storage and compute platform built into a 4U chassis. Available in either single or dual server versions, SGI MIS delivers industry-leading density with the maximum of flexibility. Key features of the SGI MIS platform include:

- **Storage Density:** A single SGI MIS 4U chassis can be configured with up to 72 3.5” HDDs or 144 2.5” SSDs for up to 288TB of storage capacity per chassis.
- **Flexible Configurations:** The SGI MIS server(s) can be configured with a mix of SSD and SAS drives along with one or two Intel Xeon E5 dual socket motherboards and up to 256GB of memory per server. Up to 8 PCIe Gen 2 expansion slots allow for further flexibility in customizing the networking and storage configuration.

Scality RING Organic Storage

Scality RING Organic Storage is a software-based storage solution that offers highly scalable object and file data access using industry-standard hardware components. The RING utilizes a distributed and decentralized peer-to-peer (P2P) “shared nothing” architecture. Key design features of the RING design include:

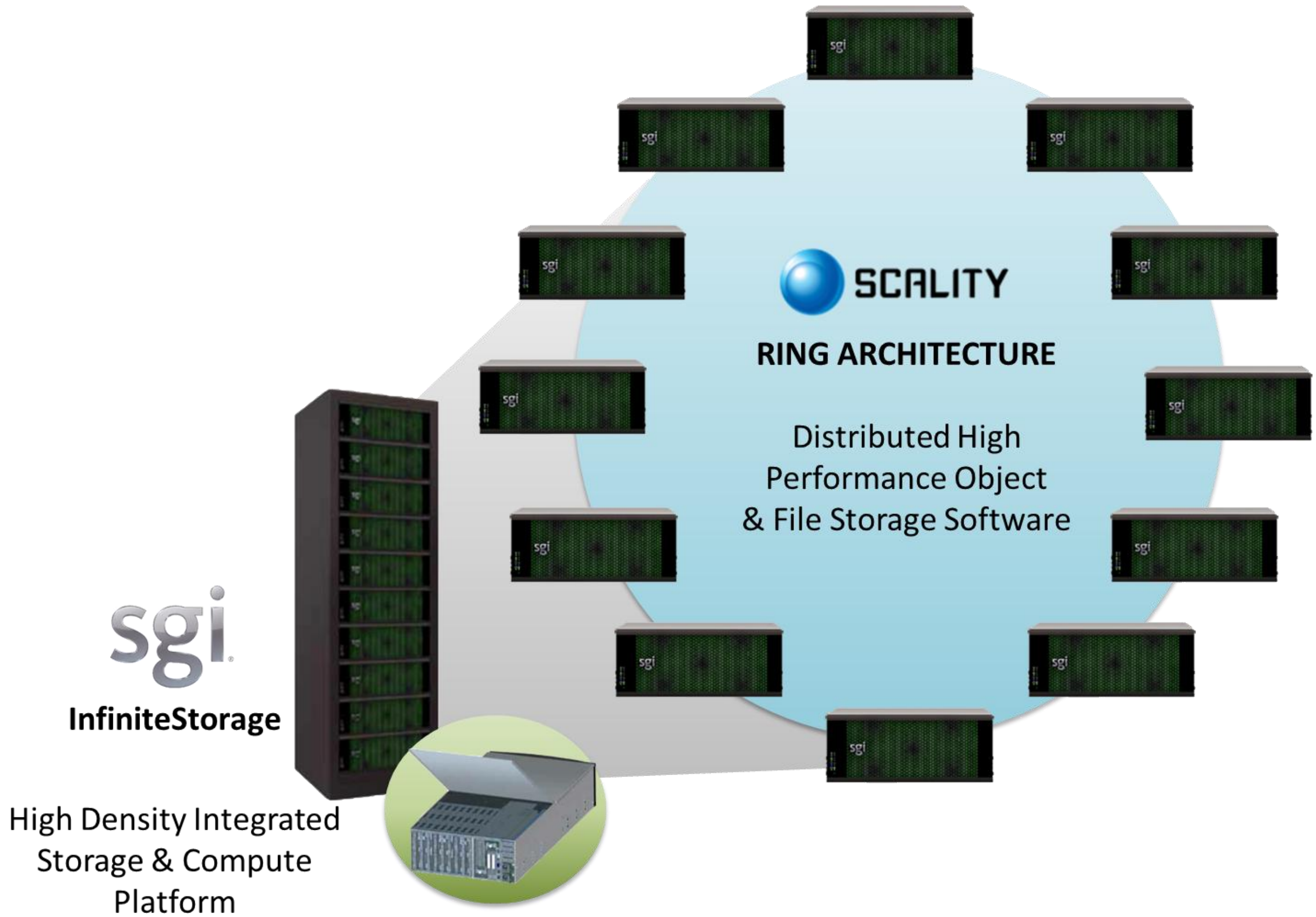
- **High Performance:** Every node (server + storage + networking) in the RING is fully independent of all other nodes. Every node contains its own data store and all nodes own their own CPU, memory, and disk resources. Because there is no single point of contention across the entire system, performance continues to scale linearly as nodes are added. High performance is maintained through automatic rebalancing and automatic tiering of data.
- **High Scalability:** Because the Scality RING architecture is based on a “shared nothing” design, the RING does not have limitations in the number of nodes that it can scale or on the addressable size of the data used for object storage. Each node must only be aware of the existence of two other nodes for the system to function, avoiding heavy chat between nodes, which is the negative side of some peer-to-peer technology.
- **High Availability:** The RING architecture contains no single point of failure and data is always protected against component failure, whether it is a hard-drive or a server or even a full rack in larger configurations. Scality supports two modes of data protection: replication, which offers higher read performance and is recommended for smaller files, up to 1 MB, and erasure code. Scality’s proprietary erasure code technology, ARC, offers flexible data protection scheme, and delivers high data durability and high write throughput, with lower disk overhead, and minimal read penalty. ARC requires a lower storage overhead than RAID6—while offering much better resiliency—including protection against server loss.
- **Flexible Data Access:** Although its inner technology is object storage, Scality RING can be natively accessed through different storage methods. Scality RING supports posix-file system through its SOFS (Scality Scale-Out File System) implementation, with NFS v3, CIFS, FTP, FUSE, and AFP compatibility. Scality RING also supports multiple REST-based API, including their proprietary API, an API compatible with Amazon S3, and support for CDMI, the SNIA standard for cloud storage. Furthermore, all data stored via either the file system or the REST API can be processed, in place by Hadoop task. This allows the user and/or application to choose the best way to access data. SOFS is a distributed file system and can scale far beyond typical NAS deployments.

SGI Modular InfiniteStorage with Scality RING Organic Storage

SGI's Modular InfiniteStorage represents an ideal platform on which to deploy the Scality RING software to create a cost-effective, scalable, high-density, highly available, high-performing object and file storage system.

The SGI MIS platform gives the customer the flexibility to decide the optimal hardware balance to power Scality RING nodes. CPU, memory, networking, and storage resources can be carefully chosen to meet customer performance, capacity, availability, and cost requirements. A view of the combined solution is shown in Figure 2.

Figure 2. SGI InfiniteStorage with Scality RING Organic Storage



ESG Lab Validation

ESG Lab audited hands-on evaluation and testing of SGI and Scality's joint solution at a remote facility. Testing was designed to demonstrate object- and file-based performance capabilities, with both replication and erasure coding data protection method, and storage high availability during potential drive failures.

Object storage operates differently from standard file system storage. With a standard storage infrastructure, content is managed through a hierarchical file system using an index table that points to the physical storage location of each file, and tracks only simple metadata. This approach limits the number of files that can be managed in a single directory. Object storage data is organized into containers of flexible sizes ("objects"). Each object has a unique ID (instead of a file name) with metadata that can include detailed attributes. This metadata can be used to set up automatic storage policies such as the migration of aging data from high-performance to more cost-efficient capacity-based disk, or the deletion of data when it expires. Object storage offers a simpler design and greater scalability, easily managing billions of individual objects. The disadvantage of object storage in the past has been the lack of performance and the necessity to integrate with proprietary REST-based API. The goal of this study was to measure how well Scality RING running on the SGI MIS hardware has alleviated these concerns.

Object-based Performance with Replication Data Protection

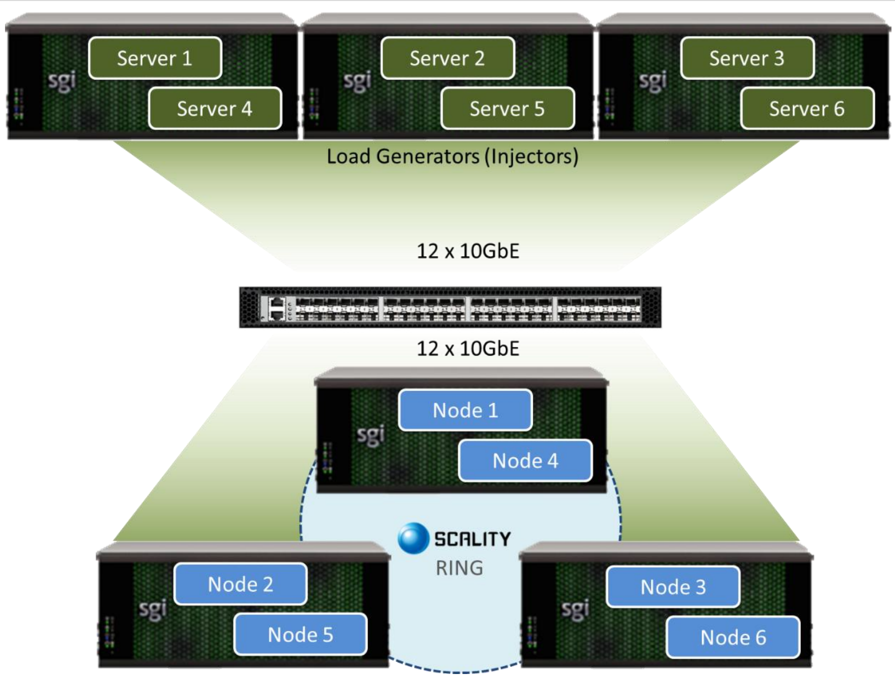
The first phase of testing focused on object-based performance using 3x replication for data protection.

ESG Lab Testing

The test bed used for this testing is shown in Figure 3. A six node Scality RING in replication mode was configured using three SGI MIS servers with dual motherboards. Each node contained 32 4TB SAS HDDs, two 400GB SSDs, 128GB of RAM, and two 10GbE ports, providing a total of 768TB of raw storage. The Scality RING was connected to an Edgecore 10GbE switch that was then connected to six injector nodes. Each injector node contained two 400GB SAS HDDs, 24GB of RAM, and two 10GbE port. Moving the workload generation off the Scality RING nodes and over to injector nodes allowed for the maximum thresholds to be hit in major server performance categories including CPU, disk, memory, and network. The complete configuration was designed with a goal of ensuring sufficient workload generation to saturate the Scality RING from both a network and operations-per-second standpoint.

The numbers are then extrapolated to a full rack of SGI MIS servers, composed of 10 SGI MIS equipped with 4TB drives, which represents just a little over 2.5 PB of raw storage.

Figure 3. Test Bed for Object-based Performance Measurements



Using a Scality-developed test harness, ESG Lab audited performance results when using object-based REST interface calls under the replicated object model.

Using "replication" data protection means that multiple copies of the same data are spread across the storage nodes within the Scality RING. In the test, three copies of the data were made, which makes the system completely resilient over two component failures, whether these are hard drive or server failures. The test was performed with two object sizes, 4KB and 1MB, representing a typical use case range. When objects are larger, erasure coding data protection method is recommended. For both tests, ESG Lab increased the simulated client base workload from 480 clients to 2,880 clients. ESG Lab measured the number of objects/second, which is meaningful for small objects, the throughput, which is meaningful for large objects, and the Time To Full Upload (TTFU) or Time To Full Download (TTFD). Note that these performance metrics are meaningful for an object store, but are not directly comparable with traditional SAN performance, which measures performance metrics like IOPS and Time To First Byte (TTFB). A correlation can be made from an object read or an object write typically requiring multiple input/output operations on a traditional SAN.

Figure 4. Object-based Performance Results – Objects/second in Replication Mode

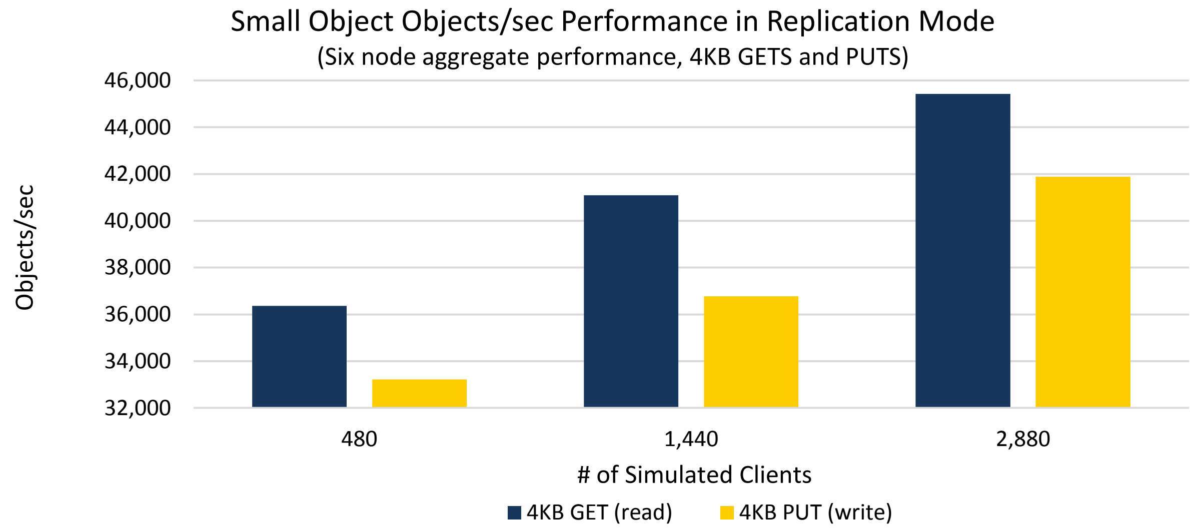


Table 1. Object-based Performance Results – Objects/sec in Replication Mode

Number of Clients	4KB Object Size – Objects/second	
	GET (read)	PUT (write)
480	36,362	33,219
1,440	41,099	36,776
2,880	45,420	41,891

What the Numbers Mean

- Using the REST interface, a 4KB object-size workload was run to measure the number of objects per second produced by a six node Scality RING with powerful SGI MIS servers in 3x replication mode.
- Both GETS (reads) and PUTS (writes) were measured as the simulated number of clients increased from 480 to 2,880.
- ESG Lab witnessed impressive performance, ranging from 36,362 to 45,420 objects/sec for GETS and from 33,219 to 41,891 objects/sec for PUTS.
- With a slight increase in the number of objects processed per second, a manageable increase in the time it took to upload or download an object was witnessed. Even under heavy load, it remained below one-tenth of a second and would seem instantaneous to any end-user.
- These numbers can be extrapolated to deliver 151,996 objects/second for GETS and 139,636 objects/second for PUTS with a full rack of storage and 9,600 clients.

Next, the object size was increased to 1MB and the total achievable throughput was measured for GETS and PUTS from the same six node Scality RING configuration. The results are shown in Figure 5 and Table 2.

Figure 5. Object-based Performance Results – Throughput in Replication Mode

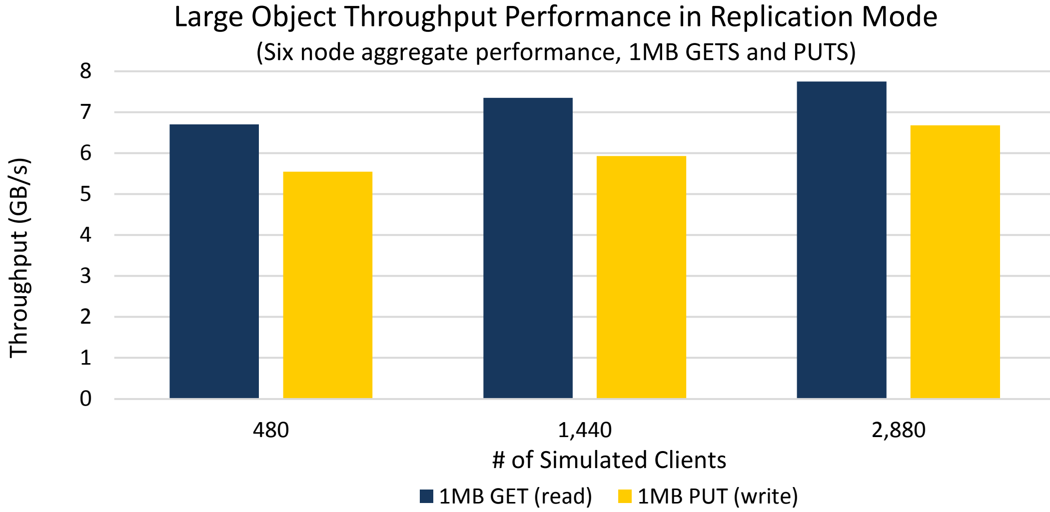


Table 2. Object-based Performance Results – Throughput in Replication Mode

Number of Clients	1MB Object Size – GB/s	
	GET (read)	PUT (write)
480	6.7	5.55
1,440	7.35	5.93
2,880	7.75	6.68

What the Numbers Mean

- Using the REST interface, a 1MB object-size workload was run to measure the throughput produced by a six node Scality RING with powerful SGI MIS servers in replication mode.
- Both GETS (reads) and PUTS (writes) were measured as the simulated number of clients increased from 480 to 2,880.
- ESG Lab witnessed impressive performance, ranging from 6.7 GB/s to 7.7 GB/s for GETS and from 5.5 GB/s to 6.7 GB/s for PUTS.
- Again, as the number of simulated user count increased, latency remained manageably low. At the sweet spot for the three server system, with 1,440 clients, the test showed 7.3 GB/s for GET with an average time to full download (TTFD) of an object of 0.2s. At the highest tested load of 2,880 simulated clients, the TTFD had roughly doubled to 0.4s.
- These numbers can be extrapolated to deliver over 25 GB/s of read throughput and 22 GB/s of write throughput with full rack of storage and 9,600 clients.

Object-based Performance with Erasure Code (ARC) Data Protection

Scality's proprietary implementation of erasure coding, ARC, helps address the long-term storage requirements for larger amounts of data. Based on the Information Dispersal Algorithm (IDA), this erasure-code technology operates within a Scality RING to intelligently protect data from potential disk, server, rack, or site failures. ARC eliminates the process of storing duplicate copies of the same data. Instead, chunks of data are dispersed throughout the Scality RING along with checksums of the same data. The checksums are mathematical combinations of the original chunks of data, but are calculated in a way that allows for the data to be reconstructed despite potential data losses.

ESG Lab Testing

Up to this point, performance testing had been completed using the replication data protection method. The next phase of testing revisited object storage performance, but now using ARC erasure-code technology. Using the same six node Scality RING test bed with SGI MIS servers, ESG Lab measured the achievable throughput of 10MB objects. ARC can be configured to various protection scheme. ARC (x,y) means that the data is sliced into x chunks and that y additional checksums are created so that the system can survive y component (whether it is hard drive or server) failures. For the test, we used ARC (4,2), where for only 50% overhead (as opposed to 200% overhead for 3x replication), the system can withhold two failures (exactly the same as 3x replication). The REST interface was used to PUSH and GET 10MB objects with a goal of achieving high levels of performance that compared well with the replication mode results. The results are shown in Figure 9 and Table 4.

Figure 6. Object-based Performance Results – Throughput in ARC data protection mode

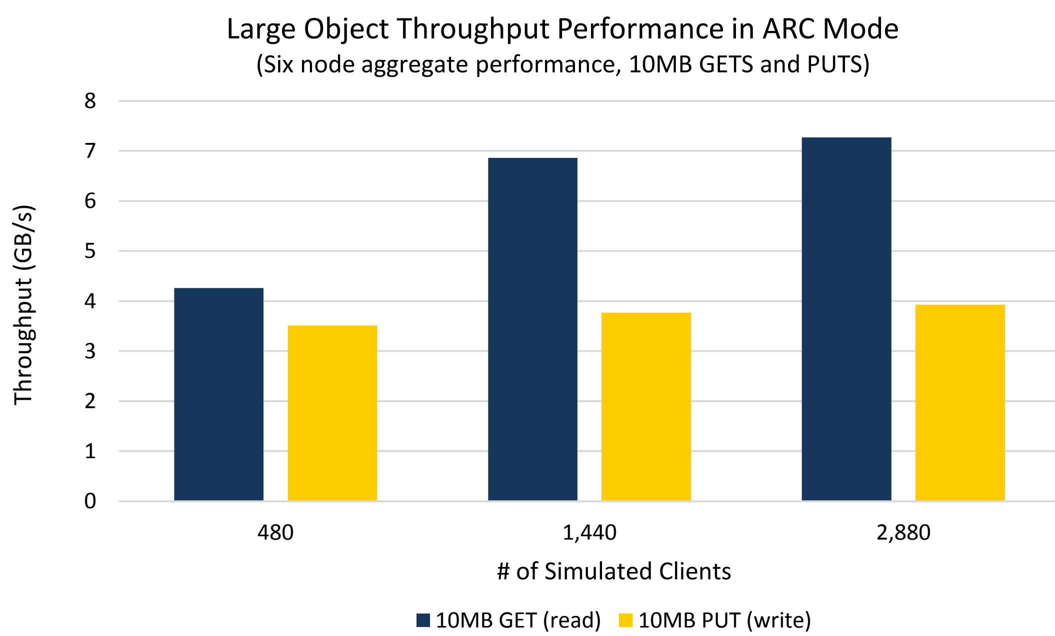


Table 3. Object-based Performance Results –Throughput in ARC data protection mode

Number of Clients	10MB Object Size – Throughput (GB/s)	
	GET (read)	PUT (write)
96	4.26	3.51
480	6.86	3.77
960	7.27	3.93

What the Numbers Mean

- Using the REST interface, a 10MB object-size workload was run to measure the throughput produced by a six node Scality RING with powerful SGI MIS servers in ARC mode.
- Both GETS (reads) and PUTS (writes) were measured as the simulated number of clients increased from 96 to 960.
- ESG Lab witnessed impressive performance, ranging from 4.3 GB/s to 7.3 GB/s for GETS and from 3.5 GB/s to 3.9 GB/s for PUTS.
- When compared with the throughput obtained with replication, ARC delivers similar read throughput, but slightly slower write throughput.
- These numbers can be extrapolated to deliver nearly 25 GB/s of read throughput with a full rack of storage and 3,200 clients.

Why This Matters

Object storage systems have become a viable alternative to file- and block-based systems—especially for public and private cloud infrastructures. The indexed object structure is ideal for unstructured data and provides far greater scaling, while the enhanced metadata offers superior data protection, capacity optimization, and security. Object storage allows files to be dispersed over wide geographic areas with a minimized effect on performance.

ESG Lab validated that by leveraging the power of SGI MIS servers with Scality RING organic storage in both replication and erasure coding data protection modes, organizations can expect high, predictable levels of object-based performance for both small and large-sized objects. Utilizing a six node Scality RING on three SGI MIS servers, ESG Lab measured impressive performance for both GETS and PUTS, with peaks of 45,420 objects/s for small objects GETS, 7.7 GB/s for large objects with replication data protection, and 7.3 GB/s with ARC. These numbers can be extrapolated to 150,000 object/s GET or nearly 25 GB/s in read throughput for a full rack of storage populated with ten SGI MIS servers. Furthermore, as ESG Lab increased the simulated client count, latency remained manageably low.

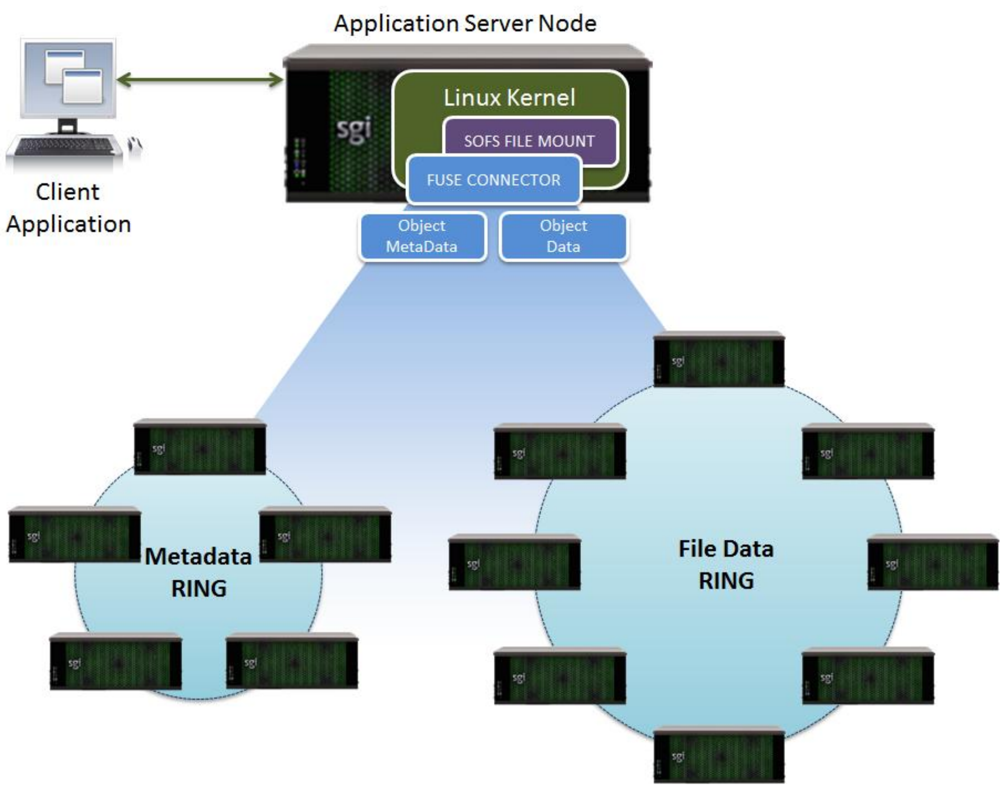
File-based Performance

Many organizations retain the need for file system data access in some of their core applications. File access serves the needs of a broader range of applications and does not require modifications to existing software. Traditional file systems are limited, however, in the size of files, size of the file systems themselves, and the number of file systems supported.

In order to implement a file system that works with the Scality RING architecture and provides higher efficiency than traditional file systems, Scality created a full POSIX-compliant file system known as the Scality Scale-Out File System (SOFS). The connector interacts with a metadata ring (MD Ring), which translates the file request into an object request that is satisfied by the Data Ring.

Because every file in the file system creates a single metadata object in the MD Ring, access times are predictable and do not degrade with the size of the file system. All metadata is fully protected through 5X replication. The data and metadata rings can be housed in the same physical ring nodes using logical separation, but can also be physically separated in larger configurations. For enhanced file system performance, the metadata ring can be stored on nodes with SSDs, while the data nodes can be populated with cost-effective SATA drives. Scality RING supports multiple access methods on top of its SOFS, ranging from NFS v3 to Apple Filing Protocol. The access method tested by ESG was FUSE, File System in User Space, which provides the best performance for Linux environments. The SOFS architecture with FUSE is shown in Figure 7.

Figure 7. SOFS Architecture with FUSE for File-based Workloads



ESG Lab Testing

The next phase of testing focused on file-based performance using the SOFS architecture with FUSE. The emphasis was put on large files, which corresponds to typical use case for video, backup, archive, or high performance computing (HPC) environments. With a goal of achieving linear performance scalability, ESG Lab audited results of a six node Scality RING running on three SGI MIS servers. Using a Scality-developed test harness, performance was measured as separate read and write workloads were generated from one to six injector nodes. Since there were four clients per injector node, this corresponds to four to twenty-four clients. The read and write workload consisted of 1GB files. Both replication and erasure coding (ARC) data protection mode were tested.

Results of the read and write tests for replication mode, including throughput and Time To Full Download/Upload, are shown in Figure 8 and Table 3.

Figure 8. File-based Performance Results – Read Throughput in Replication Mode

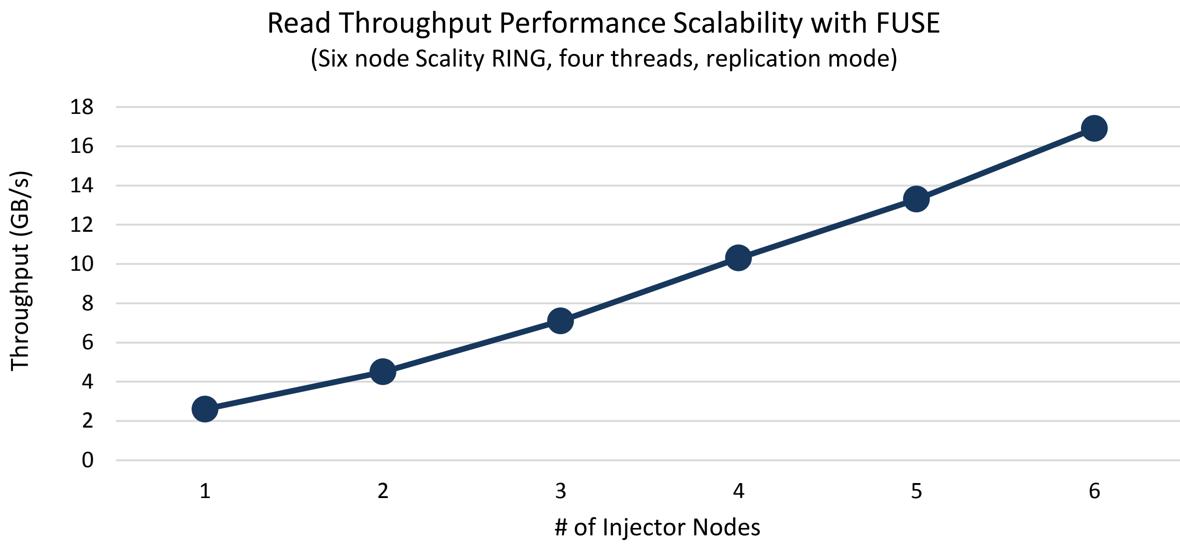


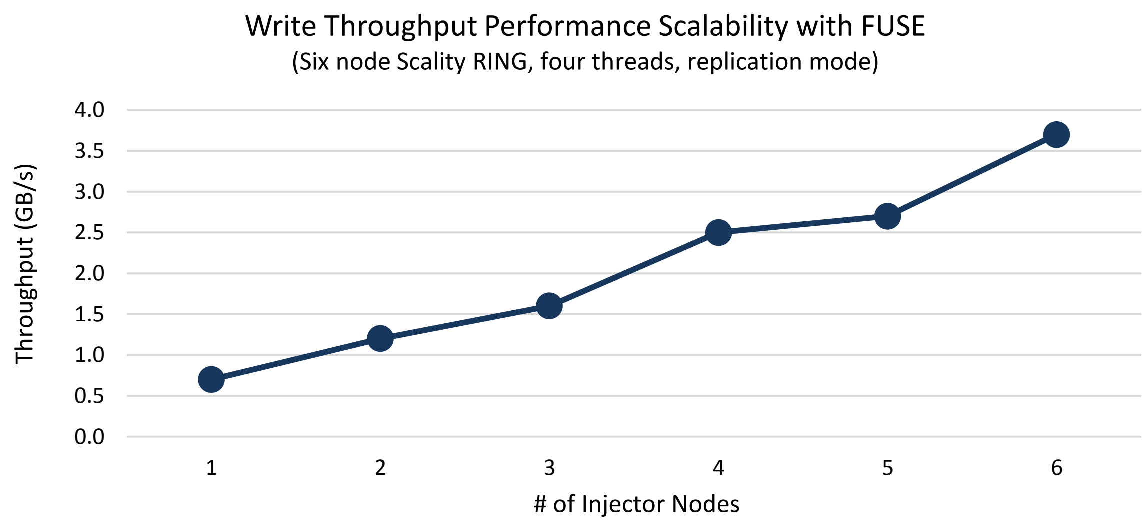
Table 4. File-based Performance Results –Throughput in replication data protection mode

Number of Clients	Throughput (GB/s)	
	1GB File Reads	1GB File Writes
4	2.6	0.7
8	4.5	1.2
12	7.1	1.6
16	10.3	2.5
20	13.3	2.7
24	16.9	3.7

What the Numbers Mean

- Using the SOFS architecture with FUSE, a read and write file-based workload was run to measure the throughput and Time To Full Download/Upload produced by a six node Scality RING with powerful SGI MIS servers in replication mode.
- 1GB file reads and 1GB file writes were measured as the workload increased from one to six injector nodes.
- For read testing, throughput performance scaled linearly. Four clients produced 2.6 GB/s while 24 clients were able to show an impressive 16.9 GB/s.
- The total time to download a full 1GB file remained extremely low as the workload being sent by the injector nodes increased. With a maximum of .3 seconds, TTFD appeared unaffected by the increased workload.
- As shown in Figure 9, the file write tests scaled in near-linear fashion as the workload increased with one to six injector nodes. ESG Lab witnessed 0.7 GB/s with one injector and a maximum of 3.7 GB/s with all six injectors.
- Also shown in Figure 8, latency continued to improve as the workload increased. It took only 1.36 seconds to upload a complete 1GB file with one injector, and a greatly improved average of 0.33 seconds with all six injectors.
- The numbers can be extrapolated for a 1GB file read to achieve 56 GB/s and under 0.3s for a time to full file download with a full rack of storage and 80 clients.

Figure 9. File-based Performance Results – Write Throughput in Replication Mode



As with object-based testing earlier, ESG Lab then went on to test the file interface with ARC erasure-code for data protection.

Figure 10. File-based Performance Results – Throughput with ARC data protection mode

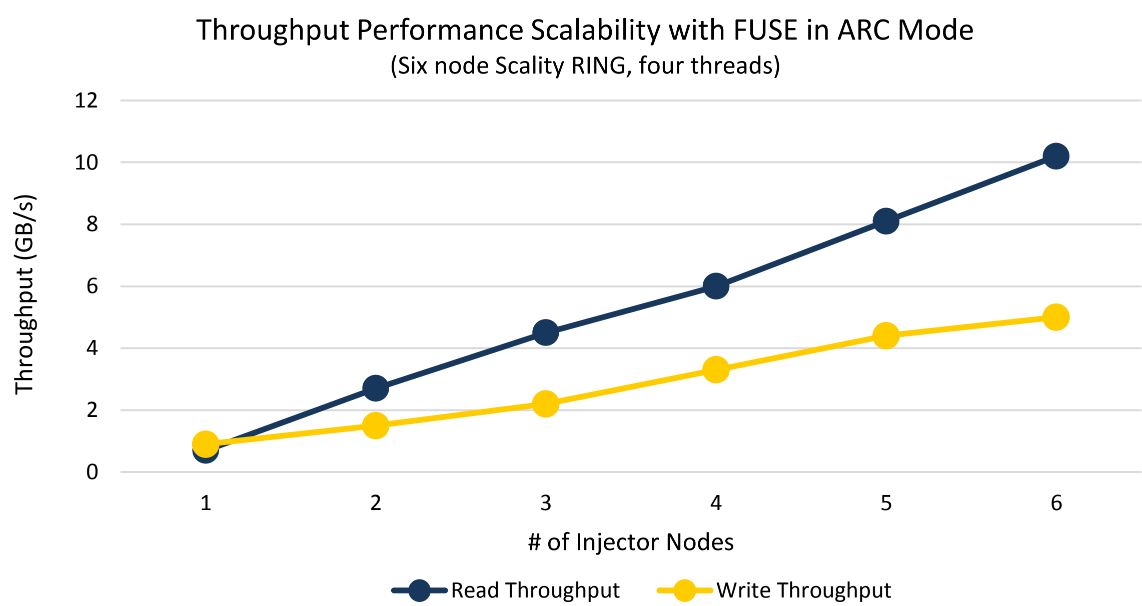


Table 5. File-based Performance Results –Throughput with ARC data protection mode

Number of clients	Throughput (GB/s)	
	1GB File Reads	1GB File Writes
4	0.7	0.9
8	2.7	1.5
12	4.5	2.2
16	6.0	3.3
20	8.1	4.4
24	10.2	5.0

What the Numbers Mean

- Using the SOFS architecture with FUSE, a read- and write file-based workload was run to measure the throughput produced by a six node Scality RING with powerful SGI MIS servers in ARC mode.
- 1GB file reads and writes were measured as the workload was increased from four to 24 clients.
- In both test cases, ESG Lab witnessed linear performance scalability, reaching peak performance of 10.2 GB/s for reads and 5.0 GB/s for writes.
- These results can be extrapolated to deliver 16.6 GB/s of write throughput with a full rack of storage and 80 clients.
- When comparing the read and write throughput performance of replication mode with ARC mode, ESG Lab noted that reads were significantly faster with replication, while write were faster with erasure code.

Why This Matters

Backup to disk, archival, content delivery, and big data analytic applications have strict performance demands for large data sets. A performance bottleneck in any of the systems between the application and the data can lead to lost revenue and dissatisfied customers. Meeting the bandwidth-intensive performance demands of these workloads using traditional storage architectures often leads to over-provisioning, wasted capital costs, increased complexity, and excessive demands on data center infrastructure.

ESG Lab validated that a Scality RING can meet the demands of customers asking for a high-performing file system with the SOFS architecture. ESG Lab witnessed impressive linear performance scalability as a file-based workload increased across a six node Scality RING in replication mode and in erasure code mode with three SGI MIS servers, which offered a high level of flexibility by allowing for the potential use of SSDs to service constant metadata requests. From a performance metric standpoint, as read and write throughput increased linearly, latency continued to improve, with both the read workloads remaining well under .4 seconds under the heaviest workload. These numbers can be extrapolated to 56 GB/s of read throughput using replication mode and 16.6 GB/s of write throughput using erasure coding with a full rack of storage and ten SGI MIS servers.

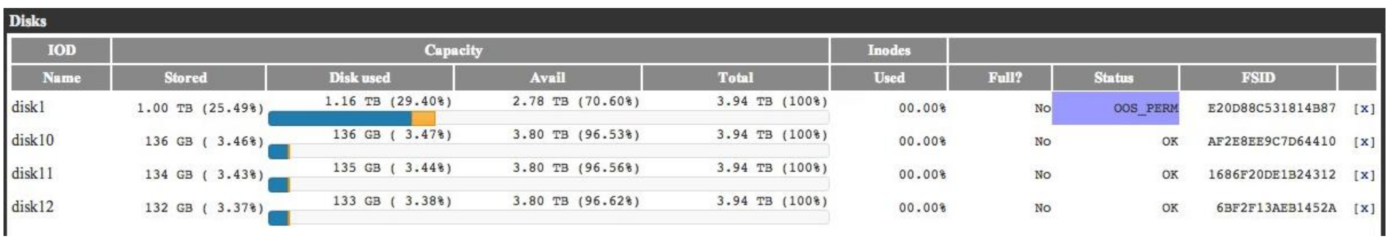
High Availability

From a hardware perspective, SGI MIS servers provide high levels of availability by using hot pluggable and redundant components. With easy access to servers from the front or back, drives or other components can non-disruptively be replaced when needed. From a software perspective, the fault tolerant design of a Scality RING ensures that storage services remain available during any type of hardware failure. When equipment fails or data corruption is detected, the organic self-healing design of the RING ensures that storage services remain available without user intervention.

ESG Lab Testing

The final phase of ESG Lab testing focused on the self-healing capabilities of a Scality RING in replication mode through an HDD failure. Two separate failure tests were conducted—one on a 750GB data set and the other on a 1TB data set. Figure 11 shows one of the drives ESG Lab tested with a status of Out Of Service (OOS).

Figure 11. Disk Out Of Service in Management Interface



Disks

IOD	Capacity				Inodes				
Name	Stored	Disk used	Avail	Total	Used	Full?	Status	FSID	
disk1	1.00 TB (25.49%)	1.16 TB (29.40%)	2.78 TB (70.60%)	3.94 TB (100%)	00.00%	No	OOS_PERM	E20D88C531814B87	[x]
disk10	136 GB (3.46%)	136 GB (3.47%)	3.80 TB (96.53%)	3.94 TB (100%)	00.00%	No	OK	AF2E8EE9C7D64410	[x]
disk11	134 GB (3.43%)	135 GB (3.44%)	3.80 TB (96.56%)	3.94 TB (100%)	00.00%	No	OK	1686F20DE1B24312	[x]
disk12	132 GB (3.37%)	133 GB (3.38%)	3.80 TB (96.62%)	3.94 TB (100%)	00.00%	No	OK	6BF2F13AEB1452A	[x]

After each disk failed, ESG Lab could query all the nodes in the Scality RING for any running tasks, which showed the status of the repair process running on each node. The software instance of the Scality RING on the node with the failed disk was responsible for reconstructing its own portion of data because the physical drives are shared and not normally bound to a single node. This helped to distribute the load and ensure system-wide scalability. ESG Lab ran an average workload and measured the time to reconstruct the lost data through the analysis of reconstructing logs and monitoring the reconstruction task through the Scality web interface. The measured results and extrapolated results are shown in Table 6.

Table 6. Disk Reconstruction Times in Replication Mode

Size of Data Set	Time to Reconstruct	Number of Chunks Reconstruction	Type
750 GB	13 minutes	877,695	Verified
1 TB	16 minutes	1,455,384	Verified
2 TB	32 minutes	2,910,768	Extrapolated
3 TB	48 minutes	4,366,152	Extrapolated
4 TB	64 minutes	5,821,536	Extrapolated
10 TB	2.6 hours	14,553,840	Extrapolated

ESG Lab was impressed with the reconstruction speeds that the Scality RING achieved. It took roughly 16 minutes to reconstruct a 1TB data set with an average system load. Extrapolating larger capacity drive sizes, a 50% populated 4TB HDD will take approximately 32 minutes, while a fully populated 4TB HDD will take just over an hour. This means reconstruction speeds that achieve over 1GB/sec.

Why This Matters

When data is stored in any type of environment, it is expected to be available whenever it is needed. Unfortunately, as HDDs continue to increase in capacity, common data protection methods like RAID will mathematically introduce a higher probability of failure. For example, a 10TB HDD has a 55% probability of failure. As HDD capacity continues to grow with improvements in technology, so does the failure probability, reaching 99% for a 100TB HDD. Organizations require a solution that can address this potential problem while maintaining high service levels for storage platforms that manage billions of objects or files at a massive scale.

Many scale-out solutions today have limitations, whether at a drive level or controller level. By not using a solution with distributed resources like a Scality RING with SGI MIS servers, common drive failures can take longer than expected, especially for organizations with continuously growing data sets.

ESG Lab validated that with the high availability features of SGI MIS servers and the self-healing capabilities of Scality RING technology, data was always available during simulated disk drive failures. Disk reconstruction times were measured for two data sets and ESG Lab was impressed with the minimal time for recoverability. The Scality RING remained in degraded mode for just 13 minutes during a 750GB data set reconstruction, while 1TB of data took only 16 minutes to reconstruct.

ESG Lab Validation Highlights

- ☑ Using a six node Scality RING configured in replication mode with three SGI MIS servers, ESG Lab measured object-based performance using the REST interface to GET and PUT small (4KB) and large (1MB) objects. Peak performance was achieved with 2,220 simulated clients while latency remained manageably low in all the test scenarios; 4KB GETs reached 45,420 objects/second, 4KB PUTs reached 41,891 objects/second, 1MB GETs reached 7.75 GB/sec, and 1MB PUTs reached 6.68 GB/s.
- ☑ File-based read and write throughput performance was measured with Scality's SOFS architecture and ESG Lab witnessed linear performance scalability as the number of injector nodes driving the workload increased from one to six in a replicated Scality RING. Read throughput started at 2.6 GB/s with one injector and scaled up to 16.9 GB/s with six injectors. Write throughput started at 0.8 GB/s and scaled up to 3.7GB/s. Also, as the workload increased, latency continued to improve, creating a highly efficient SGI and Scality joint solution that met the performance and latency requirements of enterprise-class organizations.
- ☑ ESG Lab verified the ARC mode data protection method by comparing identical object- and file-based workloads to the measured replication mode results. In all cases, ARC mode performed well and in some cases, outperformed replication mode. With a 10MB object size, it took just 960 simulated clients to reach 7GB/s, while the file-based throughput simulation performed 30% faster than the replication mode result.
- ☑ ESG Lab witnessed impressive data recoverability rates to validate the solution's high-availability features. A 750GB and a 1TB data set were constantly accessible while data was quickly reconstructed. The 750GB data set recovered in just 13 minutes, while the 1TB data set took only 16 minutes.

Issues to Consider

- ☑ The 4.1 version of RING SOFS implementation provides highly performing file-based storage to organizations of all sizes. In the 4.2 RING release, the SOFS technology supports NFS v3 and CIFS as well. This includes concurrent access from multiple clients running NFS along with the first version of SOFS. The wide access and sharing facility will continue to simplify the integration and interoperability of applications.
- ☑ The test results presented in this report are based on simulated workload testing in a controlled lab environment. Due to the many variables in each production data center environment, capacity planning and testing in your own environment is recommended.

The Bigger Truth

The emergence of cloud computing has brought about an undeniable shift in the industry toward the creation of more unstructured data. It is difficult to predict and plan for the type and amount of data that people will create in an open cloud environment. Traditional block and file storage systems are designed around certain strengths, but are simply not designed to handle the scale and agility demanded by the cloud. Block systems perform well, but are too complex to scale and administer in a cloud environment. File-based systems integrate well into existing applications but are limited in scale.

Historically, one of the main deterrents to the use of object storage has been the concern that existing application performance would be negatively impacted. To address this concern and the increasing number of customer requests for a file-based approach, Scality has introduced a Scale Out File System (SOFS). This solution allows applications to perform as they always do without needing customizations to specifically work with Scality's object-based configuration. Now organizations can take advantage of object or file storage to meet their storage requirements.

ESG Lab validated the performance, scalability, resiliency, and high availability of a storage solution powered by SGI MIS servers with Scality RING software. Object storage reached high sustainable levels of performance, while file storage scaled linearly as read and write workloads increased across a six node Scality RING. With Scality's ARC erasure code, ESG Lab saw how different data protection methods could perform to meet different organizational storage requirements. Finally, high availability was witnessed through different sized data set failure simulations. The Scality RING remained online while data was quickly reconstructed across a Scality RING in less than 20 minutes for up to a 1TB data set.

SGI's Modular InfiniteStorage with Scality RING technology provides an ideal object and file storage solution to handle the volatility of unstructured data that is created by an organization's public and private cloud infrastructures. Together, SGI and Scality have created a flexible, scalable, highly available object and file storage system that can be configured to meet the needs of any organization in a cost-effective manner. SGI and Scality have shown and continue to show that object storage is no longer limited to archival functions. Now it is ready to meet the file storage demands of scale-out architectures in today's larger application and cloud deployments. The performance and data durability features of SGI's hardware and Scality's object store clearly demonstrate the maturing of object storage technology and its readiness to handle enterprise-class requirements.

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

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