SGI® OmniStor™

Scale-out Storage to Keep Ahead of the Needs of Today's Big Data Environments.
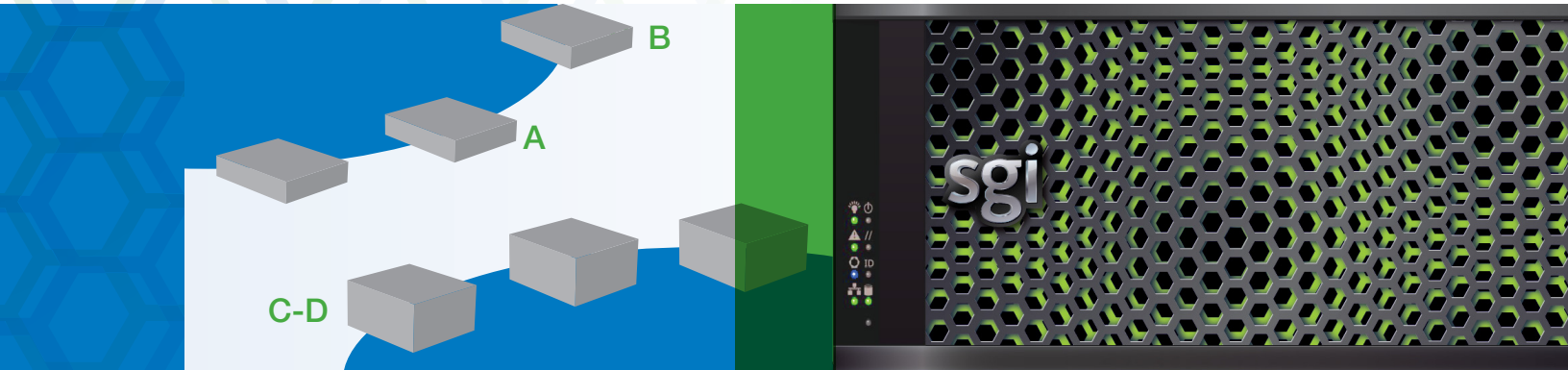
Scale-out Storage for all workflows

New IT and business challenges abound as cloud-scale computing and storage become a reality. These challenges must be addressed by innovative, cost-effective products and solutions. The SGI® OmniStor™ platform is just such a solution: a high performance scale-out storage solution for cloud-type applications as well as conventional file system access, that accommodates millions of users, tens of millions of transactions and multiple billions of files, while providing the tools to deliver 'always on' services and ubiquitous access to data.

What is SGI OmniStor?

SGI OmniStor is a complete scale-out storage solution to address petabyte-scale data problems, but with the kind of flexibility and performance that is normally associated with conventional RAID platforms. Unlike many object storage solutions with performance only suitable to archives, the SGI OmniStor solution incorporates the best of both worlds, such that one system can now provide disk-based scale-out storage with public and private cloud architectures while also enabling high performance NFS/CIFS and block storage access.

And with support for SGI® DMF™ tiered storage virtualization, SGI OmniStor solutions enable IT managers the cost and data protection advantages of a cloud architecture while also bridging into existing conventional file system infrastructures. And this is possible at a lower cost than public cloud providers like Amazon S3.

SGI OmniStor is a software-defined storage solution that delivers scalability, flexibility, performance and cost efficiency. For optimal performance and density, the SGI OmniStor solution is deployed on SGI® InfiniteStorage™ products,
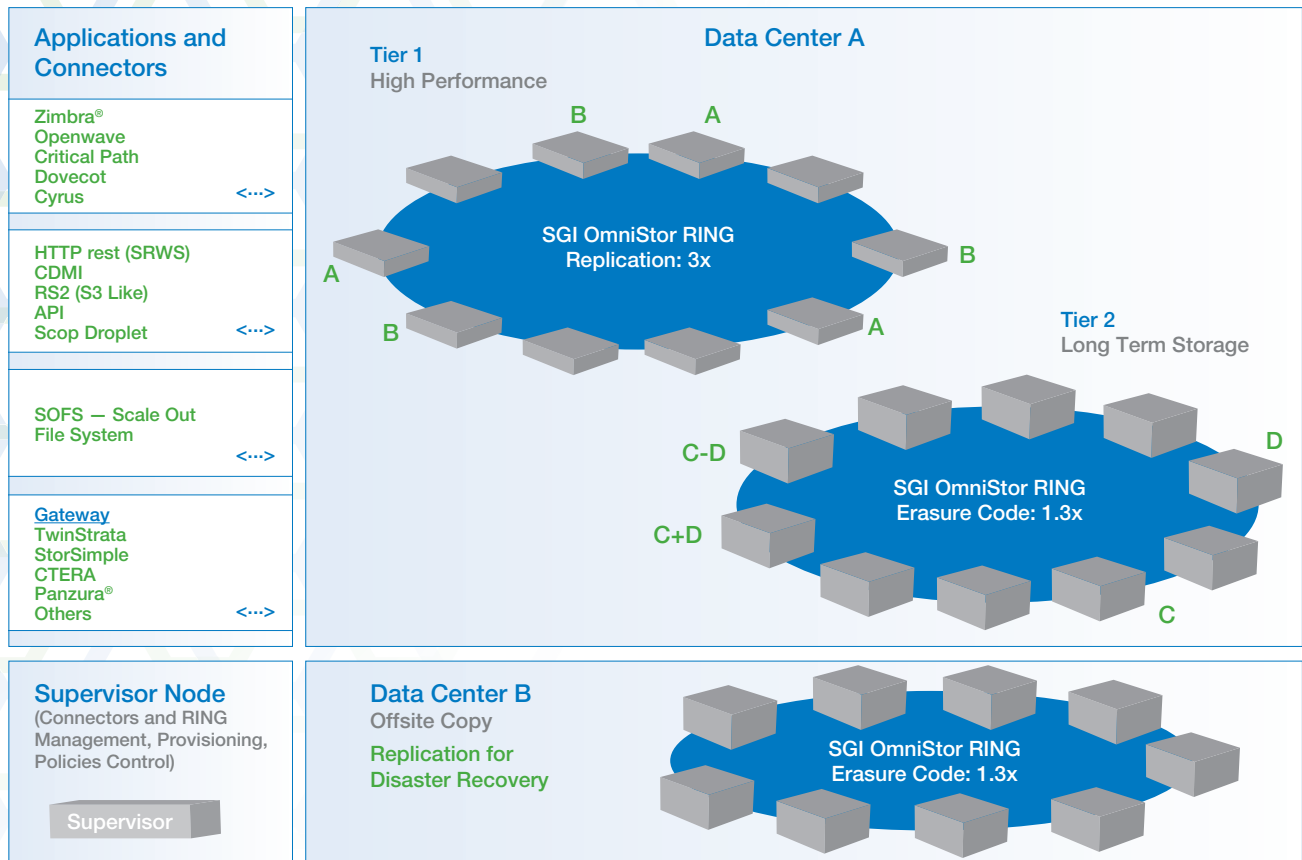
4-U storage servers that can house 72 HDDs and one or two dual-socket motherboards. SGI OmniStor can also be built with virtually any x86 servers, transforming a server farm into a large storage pool that can be integrated with enterprise and cloud applications to support a range of private and public cloud services.

For example, this solution is used by email service providers, the digital media industry and storage as a service or file storage services to provide durable and resilient file access to demanding users.

Applications can be connected to the SGI OmniStor solution using any of a variety of interfaces including the Amazon S3 API, a dedicated OmniStor API, and the SNIA CDMI standard. SGI OmniStor products can also provide NFS and CIFS/SMB access for direct NAS connectivity, enabling the solution to be used as scale-out network attached file system for on-premise usage.

SGI OmniStor provides many key features for new cloud-enabled data centers. It uses a peer-to-peer model and establishes end-to-end parallelism between client systems and storage nodes in the infrastructure. Data is protected in two ways: 1) via replication, both local or geographically distributed, and 2) also with proprietary erasure coding technology called ARC. ARC enables much greater storage efficiency while also providing high levels of protection. In this way, ARC can tolerate multiple simultaneous failures without losing data or access to data. The default configuration for ARC offers support for up to four failures with only an additional hardware overhead requirement of 30%, compared with 100% required for a RAID 10 or fully replicated environment.

Applications and Connectors

Zimbra®
Openwave
Critical Path
Dovecot
Cyrus   <···>

HTTP rest (SRWS)
CDMI
RS2 (S3 Like)
API
Scop Droplet   <···>

SOFS — Scale Out
File System
                                         <···>

Gateway
TwinStrata
StorSimple
CTERA
Panzura®
Others   <···>

Data Center A

Tier 1
High Performance

B   A

A   B

B   A

SGI OmniStor RING
Replication: 3x

Tier 2
Long Term Storage

C-D   D

C+D

SGI OmniStor RING
Erasure Code: 1.3x

C

Supervisor Node

(Connectors and RING
Management, Provisioning,
Policies Control)

Supervisor

Data Center B

Offsite Copy

**Replication for
Disaster Recovery**

SGI OmniStor RING
Erasure Code: 1.3x

Management Interface

The SGI OmniStor platform is managed through a supervisor console delivered as an intuitive Web GUI or a command line interface. The Web GUI admin tool provides easy access to reports and statistics via a dashboard, and controls configuration changes and maintenance tasks.

What is the Secret Sauce?

SGI OmniStor is a distributed system made of independent nodes that can be replaced over time as new technology becomes available.

• Object-based

• Distributed System

• Distributed metadata

• No Single point of failure

• Self healing

• Organic upgrades

Configurations for Capacity and Performance:

SGI OmniStor can be deployed in multiple different configurations, depending on the size of the environment, the growth rate of data, but also the performance level required. The software needs at least 6 nodes to create a Ring, which SGI can deliver in either 3 dual node chassis, or 6 single node chassis.

Standard configs include:

File System configuration:

The File System configuration is geared for the higher performance environments, including those with network-attached workflows. The File System configuration includes a mix of SSDs for fast metadata operations combined with SAS drives for reliability.

Each enclosure in the File System configuration includes 64 4TB HDDs for 256TB/4U raw. In addition, each chassis also includes and 4.8TB of SSD as a performance cache. This architecture uses two RINGs: one RING with SSDs for metadata and a second RING with HDDs for data.

sgi

As with all SGI OmniStor configurations, after an initial 6-node Ring is established, additional nodes can be added at any time without limitation or system interruption. The system will incorporate the new nodes and capacity into the Ring, and load-balance all of the data across all nodes automatically. This is true whether it is 1PB, 100PB or more, in a single location or distributed globally. There is no limitation to the scale of this system.

Volume configuration:

The Volume configuration is geared for the optimizing the cost of large volumes of data which might have lower performance requirements. The Volume configuration eliminates SSDs, adding additional HDD capacity to each node.

Each enclosure in the Volume configuration includes 72 4TB HDDs for 288TB/4U raw. In this configuration, a single rack can house nearly 3PB of storage.

As with the other configuration one HTTP/REST connector runs on every node in the system, enabling massively parallel IO as the system grows.

In either configuration ARC can be configured to tolerate more or fewer failures depending on the requirements of the project. Doing so will change the effective storage space.

Both configurations examples run RedHat® Enterprise Linux® 6.2 installed on dedicated 64-bit internal boot drives. SGI recommends deploying these configurations with dedicated Top of Rack switches with a 10GB port. An additional server is included as a Supervisor Node for system management. This is typically deployed as an additional 1U server, but can also be a VM connected to the racks.

Conclusion

SGI OmniStor is the perfect architecture to address high demand cloud computing and storage challenges. This integrated solution offers one of the industry's highest density, reducing infrastructure costs. Coupled with the leading object storage software and the latest Intel® Xeon® processors, it provides one of the industry's most flexible and best performing scale-out storage solutions. Easy to integrate with your applications, this business and data center ready platform represents the state-of-the art in scale-out storage infrastructure today.
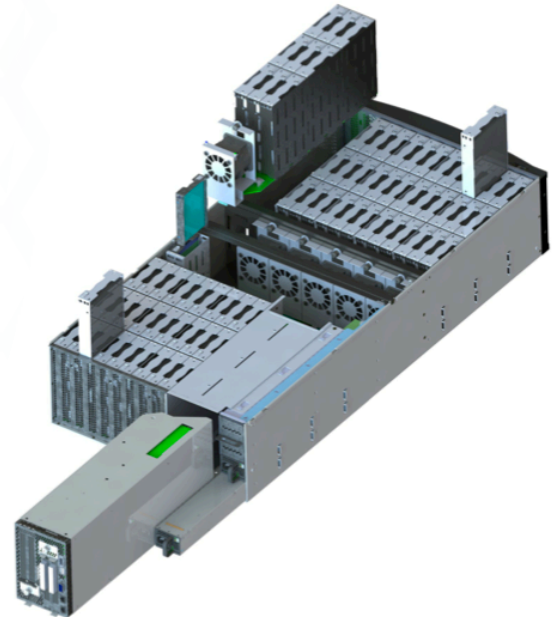
High Density: SGI Modular InfiniteStorage is the high-density hardware building block of SGI OmniStor. As a commodity storage server with 72 drives and one or two motherboards in a 4-U chassis, SGI MIS is optimized for packing the most data into the smallest space.

SGI OmniStor with SGI MIS hardware enables nearly 3PB of storage in a single rack.

SGI OmniStor Configurations

Type	Starter File System Interface	Starter Object Interface	Volume File System Interface	Volume Object Interface
Minimum # MIS Servers	• 3 dual server systems	• 3 dual server systems	• 6 single server systems	• 6 single server systems
# of SSD	• 8/system - 480GB/SSD	--	• 8/system - 480GB/SSD	--
# of SAS HDD	• 256TB (64 x 4TB)/system	• 288TB (72 x 4TB)/system	• 256TB (64 x 4TB)/system	• 288TB (72 x 4TB)/system
Total raw Storage	• 768TB	• 864TB	• 1.5 PB	• 1.7 PB
Total usable* Storage	~590TB	~650TB	~1.1 PB	~1.3 PB
Network	• One 10GB Ethernet Switch (1U)	• One 10GB Ethernet Switch (1U)	• One 10GB Ethernet Switch (1U)	• One 10GB Ethernet Switch (1U)
Software	• SGI OmniStor v.4.0 + STaaS Pack	• SGI OmniStor v.4.0 + STaaS Pack	• SGI OmniStor v.4.0 + STaaS Pack	• SGI OmniStor v.4.0 + STaaS Pack
OS	• Red Hat EL 6.2 64 bit	• Red Hat EL 6.2 64 bit	• Red Hat EL 6.2 64 bit	• Red Hat EL 6.2 64 bit
Supervisor node	• 1	• 1	• 1	• 1
Connector node (Minimum)	• 2	--	• 2	--

*based on OmniStor ARC ratio = 1.3