



# Agilysys...

High Performance Storage  
for Grid Applications

For Financial Services Organizations

## Business Drivers

Financial Services companies gain a business advantage by executing trades more quickly than their competition. The efficiency of the algorithms your business implements, which analyze large amounts of real-time trade data as well as the speed of that analysis, are key to executing trade orders moments ahead of the competition. In response to growth in financial services data, which boast some of the heaviest transaction loads anywhere, Agilysys and Sun have architected a blended software and hardware storage solution to address throughput demands. Designed with proven, open technologies, this storage solution takes grid application performance to the next level.

## Technical Drivers

The grid computing architecture is well-suited for applying trading algorithms to real-time financial trade data. The large size of the financial trade data-set can be divided and submitted to many CPUs for trade execution processing and achieve compute parallelism cost-effectively. Grid computing is, in fact, the preferred architecture that most Financial Services companies use for this type of data processing.

**...this storage solution takes grid application performance to the next level.**

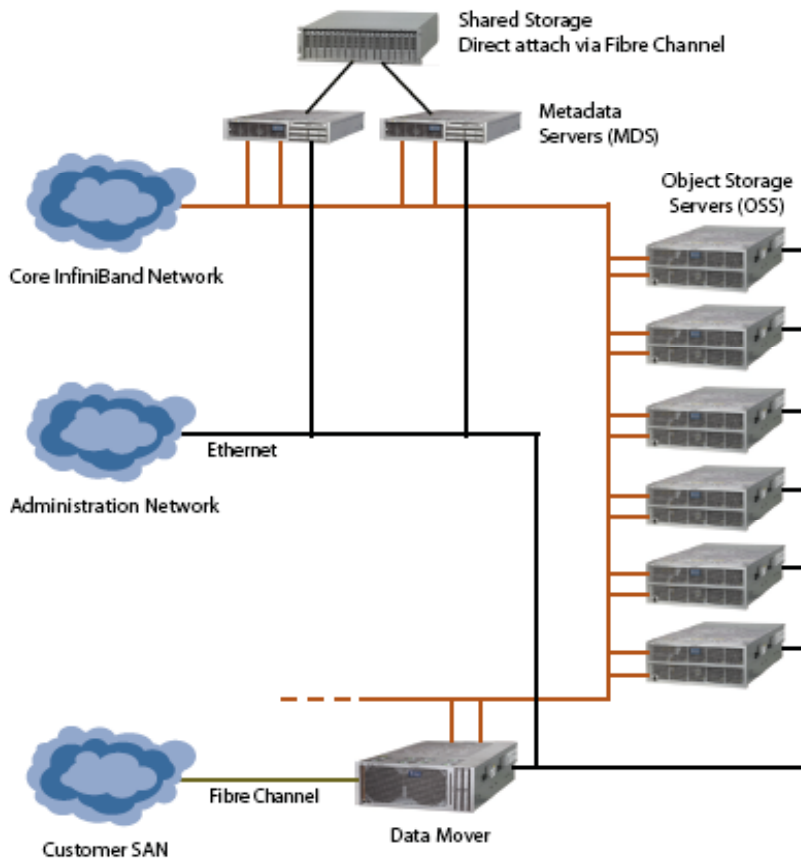
While most initial grid deployments start out their life-cycle with fairly well-balanced performance, as the demands on the grid grow and greater numbers of increasingly faster Execution Nodes are added, the storage response issues become more evident. Performance bottlenecks for this analysis are most often found to be in the storage sub-system. The ability of the storage sub-system to deliver data at a rate that the Execution Nodes can compute is a challenging design point. Many factors throttle the delivery of data to the execution nodes, including:

- Storage area network (SAN) bandwidth and speed of the fibre channel (FC) switch and network interface cards (NICs)
- Layout of logical unit numbers (LUNs) and raid groups
- Disk drive type, density and revolutions per minute (RPMs)
- Network file system (NFS) protocol limitations - NFS version, operating system (OS) version, transfer control protocol/user datagram protocol (TCP/UDP) stack

It is a common approach to continue to add more hardware to the storage sub-system to meet the increased performance demand. This is often in the form of more (NAS) heads, disk storage and network infrastructure. Unfortunately, this approach comes at the expense of increased cost and complexity. Also, at some point the addition of more hardware reaches a point of diminishing returns.

## The Solution

The Sun™ Customer Ready Scalable Storage Cluster provides a high performance storage solution for high performance computing (HPC) environments. This powerful solution features high-performance SunFire™ servers, the Lustre scalable cluster file system, and high-speed/low-latency InfiniBand interconnects, fully tested and integrated into one system. The expandable design scales from 48 terabytes (TB) of data to multiple petabytes of raw storage capacity.



In the configuration shown above, two Sun Fire X4200 servers are employed as Metadata Servers (MDS). These servers are connected to a single Sun StorageTek™ 2540 array for shared metadata storage between the two servers. Six Sun Fire X4500 servers act as Object Storage Servers (OSS) for the Lustre File System, and use captive storage of six SATA channels each with eight drives, for a total of 48 TB raw capacity per server.

In addition, a Sun Fire X4600 server is used as the Data Mover in this configuration. The Data Mover component is a key feature of the Scalable Storage Cluster, moving data between the Lustre file system and the long term storage significantly faster than traditional network transfer methods.

QLogic InfiniBand edge switches and host channel adapters (HCAs) in the Sun Fire server are used for the high-speed, low-latency cluster interconnect. Additionally, Ethernet switches implement a separate management network, with connections to all cluster nodes. The Sun Fire X4600 server, used as the Data Mover for the Lustre File System, also contains Fibre Channel host bus adapters (HBAs) for connection to the site's existing SAN.

## Performance Results

In a recent benchmark study conducted at a Sun Microsystems facility in London, the following results were achieved:

The benchmarks exceeded the requirement to achieve 1-2GB/s read performance from 4-6 x4500 storage servers using the supplied test script. The script executed repeated dd (1M) commands of large file sizes.

The following servers were used to build the Lustre configuration for this test.

- 1 MDS - SunFire X2200-M2 connected to 1 Sun Storedge 3510 via Fibre Channel interface.
- 5 OSS - SunFire X4500 (48 500GB internal disk) Configured with 6 OSTs: 6 [6 disks volume / RAID5]
- 55 SunFire x2200-M2 2xOpteron 2220(RH Linux version 2.6, lustre.1.6.3)
- The clients were connected to the x4500s and x2200 MDS using a Silverstorm SDR IB fabric.

### Solution Highlights

- Lustre software architecture allows for different hardware storage implementations. The OSS storage can be internal or directly attached to a SAN fabric.
- Main components of the solution include: (Lustre filesystem, Lustre clients, MDS, Object Store Servers).
- The Lustre Client sits on execution nodes.
- The clients access the Lustre filesystem thru Infiniband connections.
- MetaData servers and Object Storage Servers implement the filesystem and communicate with the Lustre clients.
- The Lustre FS uses an object-based storage model treats files as objects, which are located by the MDS.
- MDS support all FS namespace operations, for example, file lookups, file creation, file and directory attribute manipulation.
- File data is stored in objects on the OSSs.
- The MDS directs actual file I/O requests from the clients to OSSs, which manages storage that is physically located on the underlying storage devices.
- Once the MDS identifies the location of a file, all subsequent file I/O is performed between the Lustre client and the OSSs.

### Results and Comments

In this configuration, Linux and Lustre were tested as an out-of-the-box installation, with no tuning to OS or file system parameters. The benchmark ran on 55 execution nodes and delivered slightly better than 2GB/s aggregate read performance at the larger thread counts.

It is anticipated that by extending the configuration with each additional x4500 OSS, throughput performance will increase by approximately 0.5 GB/second.

### The Future of Grid-Hosted Commercial Applications

This case study updates IT expectations for raw data throughput in business grids. The Sun Scaleable Storage Cluster solution uses a blended software and hardware approach to solve the problem of high-throughput and I/O scalability to the grid computing environment. This approach differentiates itself from the complexity and cost of Big Iron (hardware only) solutions.

The design point and supporting benchmark data suggests a wider role for the Sun Scaleable Storage Cluster solution in the HPC environment with no

ceiling on scaling in sight. Sun Microsystems technologies and the Agilysys implementation team stand ready to deliver this demonstrated robust, scalable solution for near real-time performance in business-critical commercial applications.

# Agilysys...

## References

- The SUN Customer Ready Scaleable Storage Cluster: High Performance Storage for High Performance Computing; Sean Cochrane, Sun Global Storage Practice. July 2007, <http://www.sun.com/servers/cr/scalablestorage/>
- Sun Fire X4500 Server, <http://www.sun.com/servers/x64/x4500/>
- Lustre File System, <http://www.sun.com/software/products/lustre/get.jsp>  
Case Study – Reference Architecture  
<http://wikis.sun.com/display/BluePrints/Tokyo+Tech+Tsubame+Grid+Storage+Implementation>

## Who We Are and What We Do

This white paper and the research survey behind it are sponsored by Agilysys, Inc., a leading provider of information technology (IT) solutions serving corporate and public-sector customers with special expertise in select markets, including retail and hospitality. Agilysys provides technology solutions—including hardware, software and services—to help customers resolve their most complicated IT needs. Our expertise includes enterprise architecture and high availability, infrastructure optimization, storage and resource management, virtualization, identity management and business continuity; along with software and services designed specifically for the retail and hospitality markets. We operate from locations throughout North America, and in the United Kingdom and China, with headquarters in Boca Raton, Florida.



*For more information about how Agilysys Sun Technology Solutions can help with your company's needs, call 732.692.1919 or visit [sun.agilysys.com](http://sun.agilysys.com).*