Seagate Cloud Systems and Solutions
Delivering next-generation workloads with Intelligent Information Infrastructure

OEM
- 2+ million enclosures
- 17+ Petabytes shipped
- Drive Variety (HDD, SAS, SATA, SSD, hybrid)
- Enclosures, controllers
- Customer-driven partnership
- Services: Logistics, fulfillment, warranty, design, supply chain

HPC
- Engineered to optimize capacity and performance
- 40% fewer racks required
- >1TB/sec file system performance

Scale-Out Systems
- Engineered solutions for object storage
- Validated architectures for open source and software-defined storage
- Private cloud appliances for backup and recovery
- Modular, scalable components for DIY customers

Cloud Services
- Backup as a service
- Disaster recovery as a service
- Archive as a service
- Endpoint backup
- Managed services
ClusterStor Product Line Overview

ClusterStor 1500
1GB/sec increments
1 to 110GB/sec
Up to 7PB
Standard Rack

ClusterStor 6000
6GB/sec increments
Up to 1TB/sec
Up to 25PB
Custom Rack

ClusterStor 9000
9GB/sec increments
Up to 1TB/sec+
Up to 25PB+
Custom Rack

Performance
ClusterStor Engineered End to End Solution

Creating and supporting your own Lustre solution with high availability and productivity is challenging.

Seagate is the ONLY Vendor to Provide a complete End to End Lustre Solution

- High density disk enclosures
- Storage Bay Bridge Object Storage Servers
- Dedicated Lustre software development, integration and support team
- System Administration Management CLI & GUI
- Manufacturing integration and test validation
- Support and Professional services
- Storage Media (HDD, SSD, PCIe)
Vertical Markets

- Weather
- Healthcare
- Energy
- Finance
- Pharmacology
- Engineering
- Academic
- Defense
Top 5 Misconceptions about Seagate ClusterStor
Dispelling some myths

1. ClusterStor uses an ancient version of Lustre
2. ClusterStor does not use an open source compatible version of Lustre
3. ClusterStor is a proprietary, closed solution
4. Seagate cannot support Lustre
5. Seagate does not contribute to Lustre
Lustre 2.5 Distributed Name Space (DNE)

DNE is a hardware OPTION with ClusterStor v2.0 or greater

- Same Base Rack as ClusterStor v1.x with active/passive MDS for MDT0
- Additional Directory Units (ADUs) are configured with active / active MDT pairs

Scale up to 8 ADUs

- Up to 16 additional MDTs
- 1.885 billion additional files per ADU

Achieve up to 280K file creates per second and 15.984 billion files

Integrated management, monitoring, and serviceability
ClusterStor Lustre HSM Partners

Cray TAS and SGI DMF

ClusterStor v2.0 or greater with Lustre 2.5 is HSM Ready

Partners supply additional HSM Gateway hardware and back end Archive engine
Seagate Lustre

Seagate Lustre 2.5.1 is focused on production quality

- Seagate's Lustre distribution incorporates hundreds of stability and performance improvements
- Contributions from Intel, Seagate, and the community are curated to improve the quality of Lustre beyond the original OpenSFS release baseline
- Seagate's Lustre 2.5 distribution is engineered to operate reliability at scale in production environments

Contributions

- Seagate developers
- Other community developers
- Intel developers

3541 hours of Lustre testing per month!
### Seagate Roles in the Lustre Community

<table>
<thead>
<tr>
<th>OpenSFS Board Member and EOFS Member</th>
<th>Active Contributor to Lustre</th>
<th>Integration into Lustre ClusterStor™</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct funding of and contributions to Lustre tree and roadmap development</td>
<td>Developed and provided over 292 patches</td>
<td>Based from community Lustre tree; stabilization efforts contributed</td>
</tr>
<tr>
<td>Leadership and Participation in working groups and PACs</td>
<td>Liberator of Lustre trademark and IP</td>
<td>Proven at scale on BlueWaters and sites around the world</td>
</tr>
</tbody>
</table>
Seagate Financial Investment in Lustre

Seagate has been an OpenSFS Promoter member for 4 consecutive years.

Seagate has invested $2.3M in Lustre development and tree maintenance through contributions to OpenSFS.

This is above and beyond Seagate’s considerable investment (in millions) to maintain its own internal Lustre development and support capability and capacity.
Lustre.org will be jointly managed by the Open Scalable File Systems, Inc. (OpenSFS) and European Open File System (EOFS) as trusted stewards of the community.

“...this move is a great indication that Seagate intends to be an active, contributing member of the Lustre community...”, Rich Brueckner, Inside HPC

“We wanted to send out a special thank you to Seagate for showing such a strong commitment to Lustre...They continue to show their excitement and commitment to Lustre with concrete moves like this...”, OpenSFS
Seagate Contributes Hadoop on Lustre Connector

Improves IO performance and eliminates HDFS pre-staging overhead

Improves workflow efficiency for users of both Lustre and Hadoop by eliminating the need to copy data into HDFS prior to running Apache Hadoop jobs

Provides an alternative to Hadoop’s reliance on the HDFS file system

Enables Hadoop eco-system tools such as Mahout, Hive and Pig to take advantage of the Lustre file system

TeraSort Benchmark
HDFS vs. Hadoop Workflow Accelerator

- 63% Faster Overall Data Transfer not Needed
- 38% Faster Analysis Only
- 53% Faster Overall Data Transfer not Needed
- 28% Faster Analysis Only

On the 25 Node, 216 Hadoop Cores, 160 Lustre HDDs, 216 HDFS HDDs cluster. Connector is:
- 53-63% faster by eliminating import into HDFS!
- 28-38% on analytics alone!
Seagate Open Sources Xperior
Automated Build and Test Platform for Lustre

Used by Seagate Engineering and System Test groups for Lustre test automation
• Test management
• Execution control
• Test result reporting

Flexibility for extensions
• Supports the addition of new test models for new Lustre release via wrapper functions

Next steps
• Connect directly to community Jenkins repository
• Scale testing on on Amazon Web Services (AWS)
Support for Lustre and Beyond
The World’s Leading Data Services Company

<table>
<thead>
<tr>
<th>Company</th>
<th>Years of Services</th>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seagate</td>
<td>Over 37 Years</td>
<td>With a satisfaction rating above 95%, we have proven that we are the place to go for data recovery</td>
</tr>
<tr>
<td>EVault</td>
<td>Over 18 Years</td>
<td>Founded in 1997 as the original Cloud Backup and Archive Services company</td>
</tr>
<tr>
<td>LaCie</td>
<td>Over 23 Years</td>
<td>Founded in 1997 as the original Cloud Backup and Archive Services company</td>
</tr>
<tr>
<td>Xyratex</td>
<td>Over 21 Years</td>
<td>With expertise in both global data storage and networking technologies</td>
</tr>
</tbody>
</table>
Seagate Powers 4 of 5 1+ TB/sec file systems
Supported 100% by Seagate & our partners
Seagate Collaborative Research Efforts

Results for collective I/O performance optimization through E10 work in the DEEP-ER project.

New E10 hints infrastructure built on top of MPI-IO.

Funded initiative which will target E10 work a "Marie Curie Initial Training Network" program in Europe which funds PhD students.