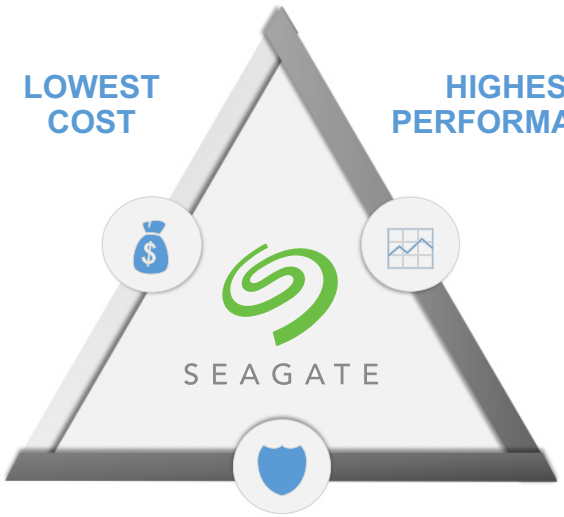
ClusterStor A200

Delivers Industry-best cost, performance, and availability metrics



The ClusterStor A200 is a tiered active archive object store for the ClusterStor product line. The Clusterstor A200 solution includes active archiving and hierarchical storage management components that provides an automatic policy-driven archive storage tier for ClusterStor systems.

The ClusterStor A200 solution is engineered for productivity and provides an optimal combination of cost, performance, and availability for the most demanding HPC & commercial workloads. The A200 tiered archive delivers up to 50% lower cost than HPC primary storage and 2-3x more performance and higher availability than other archive solutions in the marketplace, while providing near limitless scale-out capabilities in terms of capacity, number of objects and throughput.

The ClusterStor A200 utilizes network erasure coding that is tuned to the storage platform to deliver unmatched availability, scalability and performance for active archiving applications. The A200 is capable of providing extremely high levels of durability by reconstructing data from a failed 8TB drive in less than 1 hour in the reference configuration. The A200 utilizes massive parallelization to deliver linear performance and capacity scalability to meet the active archiving needs of the most demanding high performance computing environments. The ClusterStor A200 has been designed with extreme system resiliency, with no single points of failure and the ability to replace any failed component without downtime.

Benefits

Customers benefit from industry-leading economics, not only in terms of storage and performance costs but also from reduced data center floor space, power, cooling and administrative costs; up to 60% better than competitive active archive solutions. Customers using the A200 can scale the system indefinitely and save up to 50% in storage and operational costs by migrating static data from Lustre or Spectrum Scale primary storage to the A200 active archive object store.

- Storage Value-Utilizes cost effective SMR drives & efficient erasure coding.
- Limitless Scale - 3.4 Undecillion (2^128) objects, unlimited object size, unlimited storage capacity.
- Linear performance scaling - Up to 10GB/s per rack, add additional racks to increase performance linearly (up to the network limit).
- 4 "Nines" of Availability. Less than 55 min of downtime per year for base configuration.

Features

- Archive object storage tier for ClusterStor platform.
- Combine with ClusterStor HSM to provide automatic policy-driven data migration & retrieval.
- Optimized for a balance of cost, performance & density.
- Utilizes network erasure coding to provide high levels of data availability and data durability.
- No single points of failure, replace any failed component without downtime.
- Dual 10Gb Ethernet node connectivity.

Economics

- Up to 50% $/TB storage cost savings vs HPC primary storage
- Utilize most cost effective drives - 8TB+ SMR SATA HDDs
- Cost efficient 8+2 erasure coding

Performance

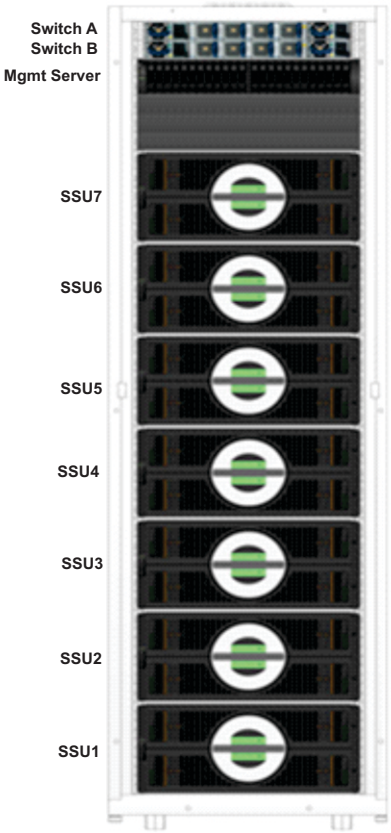
- Parallel read/write access paths
- Linear multi-session scaling increases research productivity
- Read/write throughput of up to10GB/s per rack (7SSUs)

Availability/Durability

- 4 “nines” availability for base configuration
- No single points of failure, replace any failed component without downtime.
- Improved data availability & durability via network erasure coding

ClusterStor A200

Delivers Industry-best cost, performance, and availability metrics



A200 Base Rack

ClusterStor A200 Tiered Archive

The ClusterStor A200 is a pre-configured rack-based storage solution that delivers predictable performance and eliminates single points of failure. The A200 consists of 5-7 scalable storage units (SSU), a dedicated ClusterStor management unit (CMU) and redundant 10/40GbE switches.

A200 Scalable Storage unit

The Scalable Storage Unit (SSU) is the storage sub-system for the A200 active archive. Each 5U84 A200 SSU contains an embedded storage controller, eighty two 8TB SMR drives, two metadata drives, dual 10GigE network ports, and provides 524 usable TB per tray (over 3.6 usable PB per rack). The result is an easy to deploy, easy to use and easy to manage solution. There is no need to guess at how to scale performance for multi-session workloads; each SSU is a balanced performance building block delivering a predictable level of performance and system can be scaled by simply adding SSUs and expansion racks as needed.

Additional SSU's and expansion racks can non-disruptively be added as needed to provide linear capacity and performance scalability. Overall system performance is directly proportional to the number of SSUs to satisfy performance and/or data capacity needs.

Seagate Capacity Optimized 8TB SMR Hard Drives

Shingled Magnetic Recording (SMR) drives are optimized for cost, efficiency and capacity as data is written wide and the next adjacent track write “trims” the previous track. This makes SMR drives a perfect choice for archival applications, where capacity and cost are paramount. Readable tracks are narrower than originally written tracks, which results in 30-40% more tracks being stored on the platter than standard Perpendicular Magnetic Recording (current SAS technology) drives.

A200 ClusterStor Manager

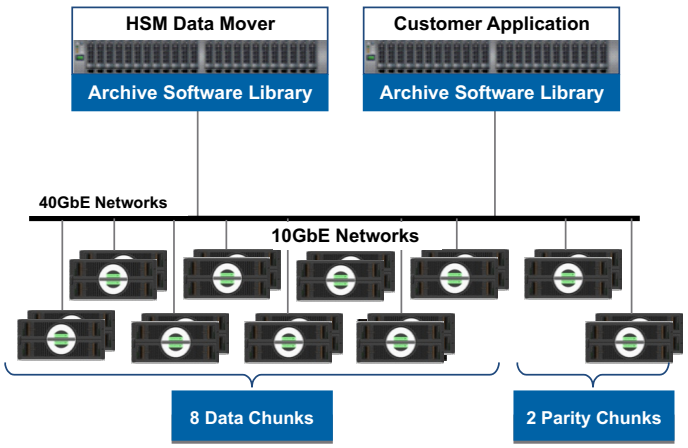
The A200 ClusterStor Manager is a comprehensive system management application that is part of the CS A200's distributed management framework and is responsible for consolidating health and performance monitoring and administration in a single managed system view.

ClusterStor Manager consolidates management of the CS A200 storage infrastructure, network erasure coding data protection layer, and HSM management and monitoring into a single, easy-to-use administrator interface.

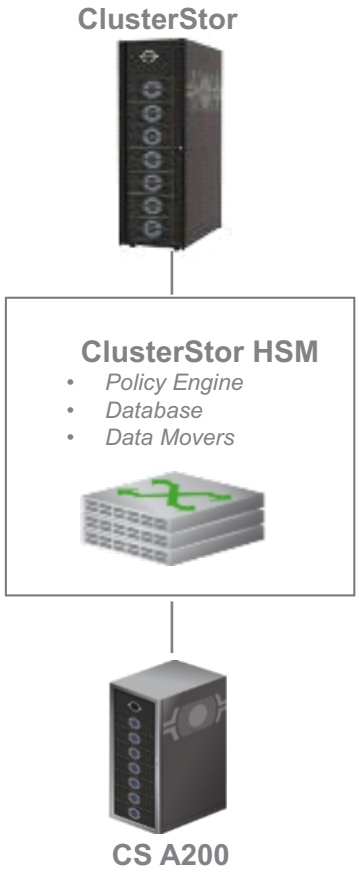
Initial installation, creation of data layouts, ongoing monitoring and management, expansion and software upgrades are accessed via a single pane of glass. This brings unprecedented system visibility to the storage management system administrator.

ClusterStor A200

Delivers Industry-best cost, performance, and availability metrics



ClusterStor A200 Data Layouts



ClusterStor HSM

CS A200 Data Protection

The CS A200 system supports network erasure coding to provide data and disaster protection. Multiple network erasure coding schemas (called “Data Layouts”) can be concurrently defined and utilized in a single A200 system to deliver separation of data across SSUs if desired.

Seagate’s network erasure coding distributes the data+parity shards across multiple SSUs which enables data to be available even if multiple SSUs or individual drives fail concurrently. The default supported CS A200 erasure coding schema is 8+2, meaning that objects are sharded into eight data shards and 2 parity chunks and written across 10 networked SSU’s. An A200 configuration containing ten or more SSUs maintains full data read availability even in the event of two concurrent SSU failures or two concurrent drive failures. In addition, network erasure coding extends the number of drives that can participate in data reconstruction across SSUs, allowing for massively increased parallelism and reconstruction speed. Increased reconstruction speed in turn shortens the failed drive data reconstruction time, leading to an increase in data durability as the system scales.

ClusterStor HSM

Key Storage Tiering (HSM) Capabilities

ClusterStor HSM is designed to offload data from a higher level storage tier to a less expensive ClusterStor A200 Active Archive storage tier based on customer defined policies to meet specific cost, performance and quality of service requirements while still maintaining transparent access to data. The initial ClusterStor HSM release is designed to manage Lustre ClusterStor configurations, and support for IBM Spectrum Scale (GPFS) is planned for a follow-on release.

ClusterStor HSM Benefits

- Storage costs – Up to 50% cost reduction by moving data off of expensive tier 1 storage.
- Data availability & durability – Improve availability & durability of archived data for long-term retention.
- Performance – Reduce primary storage fragmentation & latency / seek times.
- Automation / User productivity – Automation reduces effort & setup time to run a job.

ClusterStor Lustre Active Archive Use Case

ClusterStor HSM is a gateway appliance that utilizes data movers that act as Lustre clients to transfer files between ClusterStor primary storage Object Storage Targets(OSTs) and the ClusterStor A200 storage target based on customer-defined policies. Customer-specified policies determine when data is copied (migrated) to the secondary storage target, and when migrated files are stubbed in primary storage. Files are typically stubbed based on last access time, size, creation date, and/or some combination of other file properties. Migrated files can be also stubbed on demand when primary storage free space falls below a specific threshold. When a Lustre client/user attempts to use this file, it is automatically and transparently retrieved from the secondary storage tier before proceeding. Once a file is re-hydrated, it can again be stubbed per policy once it meets policy criteria.

ClusterStor A200

ClusterStor A200 Specifications - General Information (1) and (2)	
System Capacity (raw)	Up to 4.59PB Raw (3.63 PB usable) per rack Up to 2^128 objects
System Performance	Up to 10 GB/s bandwidth performance per rack (req's dual 40GbE connections to each rack TOR switch)
SSU Storage	Seagate 5U84 Storage enclosure, 84 drives, Integrated Storage Server, 10GbE Internal rack network connections
SSU's per rack	Up to 7 per rack (base or expansion)
Maximum Object Size	Unlimited (up to the size of the available storage)
Client Software	Seagate HSM Data Mover
Network Requirements	
External system management interface (base rack only)	2x 1GbE
External data Interface	4x 40GbE (2x from each rack TOR switch)
External switch infrastructure	40GbE switch pair in a HA configuration with LACP support, VLAN support
A200 Disk Drives	
Data Store	Seagate 8TB SMR SATA (Lamar) drives (82 per SSU)
Metadata Store	Seagate 6 Gb/s 6TB SAS drives (2 per SSU)
Dimensions	
Height	1,991 mm (78.4 in)
Width	600 mm (23.62 in)
Depth	1,200 mm (47.24 in)
Weight	1,194Kg (2,26 lbs) (preliminary data)
System Availability	
Data Availability	Four "nines" (99.99%) with 12 SSU configuration (req's HA switch infrastructure and dual connections to each TOR switch)
Data Durability	Eleven "nines" (99.999%) with 12 SSU configuration
Hot Swappable Components	Disk Drives, Power Supplies, Fans, Power Cooling Modules and Server Modules
Power	Redundant Power Supplies and Power Cooling Modules
Power Cooling Modules	Dual fans per SSU
Power Consumption	
Base Storage Rack (7 SSUs)	8.25 Kilowatts Peak (preliminary data)
Expansion Storage Rack (7 SSUs)	8.00 Kilowatts Peak (preliminary data)
Single SSU	1.10 Kilowatts Peak (preliminary data)
Heat Dissipation	
Base Storage Rack Configuration (7 SSUs)	28,150 BTU/hr
ClusterStor HSM Specifications - General Information	
Supported Target Systems	ClusterStor 9000 Lustre 2.5.1
Supported Archives	CS A200 archive
HSM PE/DB appliance	2U24 storage appliance including Seagate ClusterStor HSM SW (RobinHood)
HSM Data Mover software	Seagate-supplied Data mover 'copytool' software Requires Customer Supplied Linux Server[1] Min 2 servers required, maximum supported 16
Network Requirements	
HSM PE/DB appliance	2x 1GbE management interface 1x 40GbE/ FDR IB (connected to CS9000)
Customer Supplied Linux Server (for HSM Data Mover)	1x 40GbE (connected to A200) and 1x IB/40GbE (connected to CS9000)
Performance	
IO Throughput	Up to 1.5 GB/s per HSM Data Mover [2]
Update tracking method	Lustre RobinHood ChangeLog
Power Consumption	
HSM PE/DB appliance	700W peak (preliminary data)
Heat Dissipation	
HSM PE/DB appliance	2,388 BTU/hr

Note

1. HSM Data Mover Server Specifications

a. Intel CPU (2609 v2 or greater), 32GB DRAM, 100GB Storage (OS).
b. Centos 7.0 (7.0.1406) or compatible OS.
c. Lustre 2.5.1 or later client software.
d. 40GbE ethernet HBA connected to the A200 network infrastructure.
e. Network connection to Lustre network (IB/40GbE Ethernet).
f. Ability to add CS-A200 specific software RPMs, including userspace applications and kernel modules.

2. Target reference platform requirements

a. A200 archive configured with 12 SSU's and dual 40GbE connections to each A200 TOR switch.
b. 40GbE connection from data mover server to A200 network infrastructure.

seagate.com

AMERICAS
ASIA/PACIFIC
EUROPE, MIDDLE EAST AND AFRICA

Seagate Technology LLC 10200 South De Anza Boulevard, Cupertino, California 95014, United States, 408-658-1000
Seagate Singapore International Headquarters Pte. Ltd. 7000 Ang Mo Kio Avenue 5, Singapore 569877, 65-6485-3888
Seagate Technology SAS 16–18, rue du Dôme, 92100 Boulogne-Billancourt, France, 33 1-4186 10 00