The Unique Technical Benefits of an Engineered Solution for GPFS

December 2015

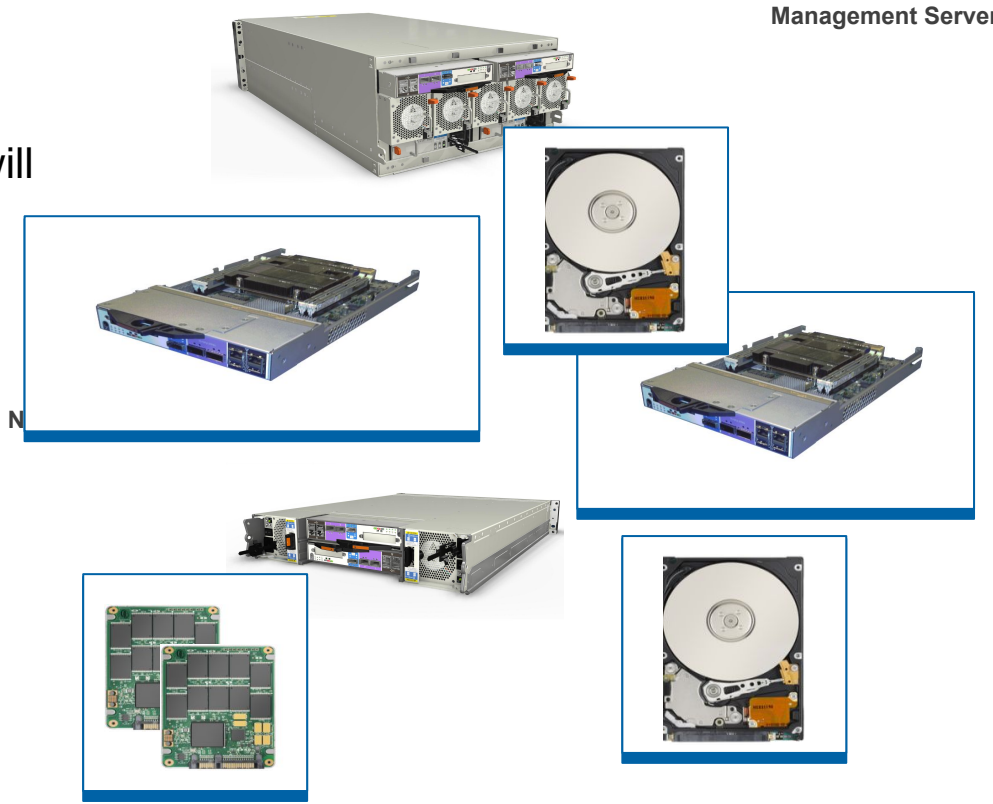
Marc Roskow

Non NDA

Why an Engineered Solution?

or Why not just a bag of parts?

1. Easier to acquire with confidence it will work
2. Faster time to operation
3. Easy to keep software current
4. Better support



Introducing ClusterStor's G200

Engineered Solution with IBM Spectrum Scale

Highest Performance De-clustered RAID solution for Spectrum Scale

- **Up to 63 GB/s per rack throughput**

ClusterStor HPC Drive Support

- **The industry's highest performance 3.5" drive that enhance the ability of GPFS to address a broad range of sequential, random and mixed customer workloads**

Pre-tested and configured integrated solution

- **Created from converged storage building blocks assuring fast, accurate installation and easy, modular expansion**

Introducing ClusterStor G200 Spectrum Scale

The **Enterprise Features and Reliability** of IBM's Spectrum Scale File System

- Proven at Scale across global organizations
- Enterprise ready: data protection, management, security and more
- Automated tiering and ILM from HDDs to Flash to Tape
- Easy to Use, get up and running in a few hours

+

The **Power** of the ClusterStor Architecture

- Industry's Fastest Converged Scale-Out Platform
- Industry's highest quality disk drives with lowest disk failure rate
- Highest performance throughput per hard disk drive
- Robust Management and Support

Designed for the world's most data intensive workflows

Pre-integrated, tested, tuned, ready to deploy

Removes metadata bottlenecks with SSDs

x86 based De-clustered RAID Solution

Drastically lowers Total Cost of Ownership

ClusterStor G200 Overview

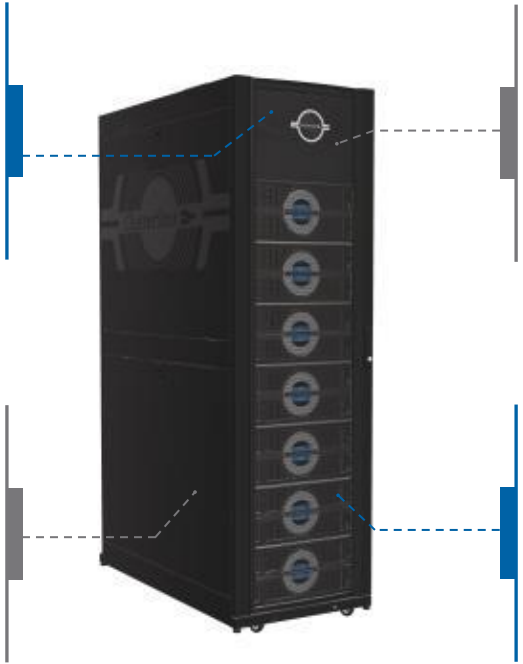
Designed for HPC, Big Data and Cloud

Connectivity

- IB – FDR, QDR and 40 GbE
- Exportable via CNFS, CIFS, Object storage, HDFS connectors
- Linux and Windows Clients

Robust Feature Set

- Spectrum Scale / GPFS v4.1 Standard
- Global Shared Access with Single Namespace across cluster/file systems
- Snapshot and Rollback
- Integrated Lifecycle Management
- Backup to Tape Options
- Non Disruptive Scaling, Restriping, Rebalancing
- Synchronously replicated data and metadata



Management and Support

- Clusterstor CLI Based Single Point of Management
- RAS/Phone home
- SNMP integration with Business Operation Systems
- Low level Hardware Monitoring & Diagnostics
- Embedded monitoring,
- Proactive alerts

Hardware Platform

- Building Block approach with Scalable performance and capacity
- GridRAID De-clustered RAID Support
- Embedded High Availability NSD Servers Integrated into Storage Enclosure
- Fastest Available IO per Rack Unit
- Extremely Dense Storage Enclosures with 84 drives in 5U

ClusterStor Spectrum Scale Performance Density Rack Configuration

Key components:

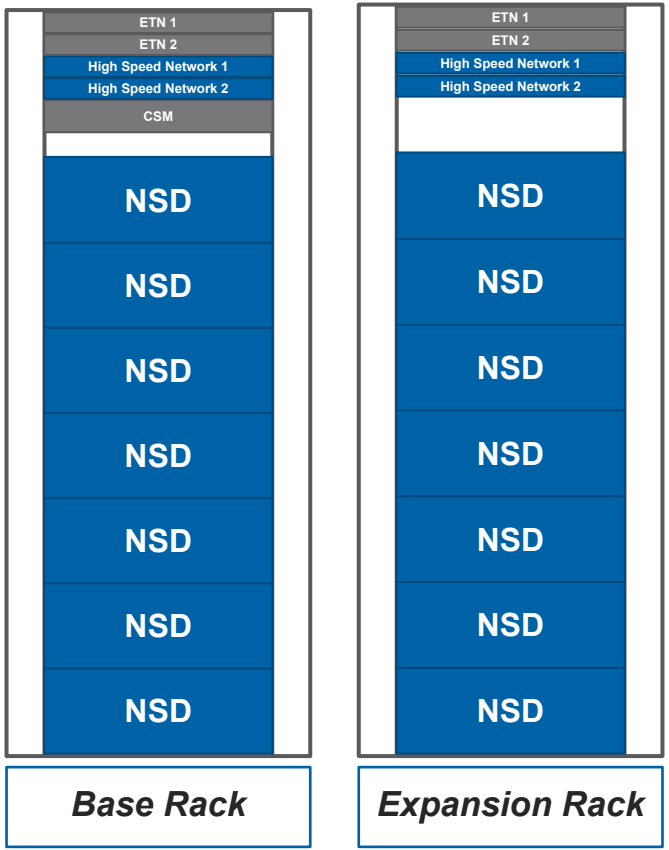
- **ClusterStor Manager Node (2U enclosure)**
  - 2 HA management servers
- **2 Management switches (1Gbe)**
- **2 Data Network switches (IB or Ethernet)**

Performance:

- **Up to 63GB /sec per rack**

Key components:

- **5U84 Enclosure Configured as NSDs + Disk**
  - 2 HA Embedded NSD Servers
  - 76 or 80 7.2K RPM HDDs
  - 4 or 8 SSDs
- **42U reinforced Rack**



ClusterStor Spectrum Scale
Capacity Optimized Rack Configuration

Key components:

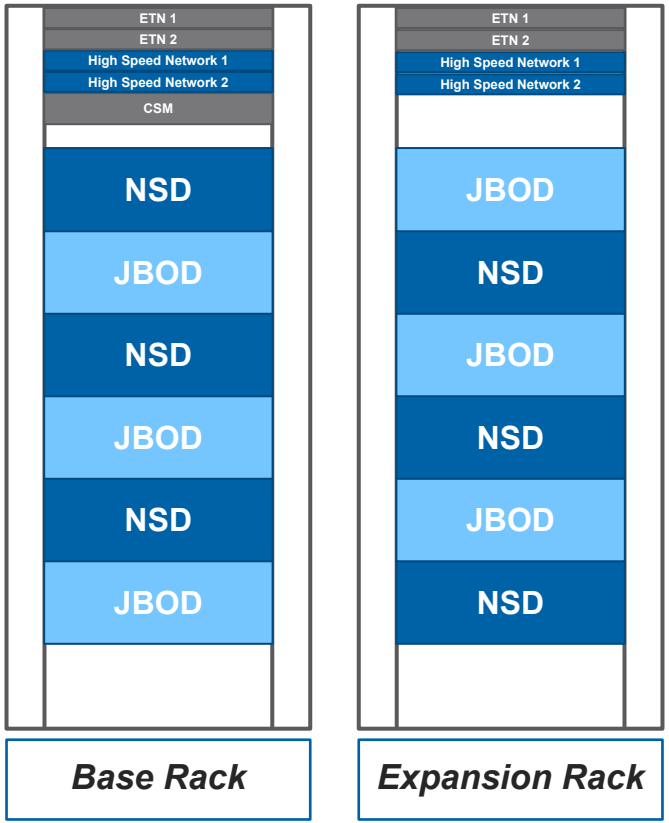
- **ClusterStor Manager Node (2U enclosure)**
  - 2 HA management servers
- **2 Management switches (1Gbe)**
- **2 Data Network switches (IB or Ethernet)**

Performance:

- **Up to 27GB /sec per rack**

Key components:

- **5U84 Enclosure Configured as NSDs + Disk**
  - 2 HA Embedded NSD Servers
  - 76 or 80 7.2K RPM HDDs
  - 4 or 8 SSDs
- **5U84 Enclosure Configured as JBODs**
  - 84 7.2K RPM HDDs
  - SAS connected to NSD servers, 1 to 1 ratio
- **42U reinforced Rack**



ClusterStor Spectrum Scale – Standard Configuration

SSU - NSD (MD) Server x 2 (SATI)

Large File Sequential Performance

- 9GB/sec per 5U84 (Clustered)

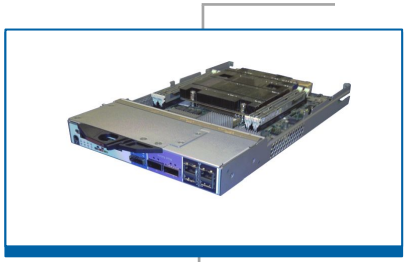
- 5GB/sec per 5U84 (Scatter)

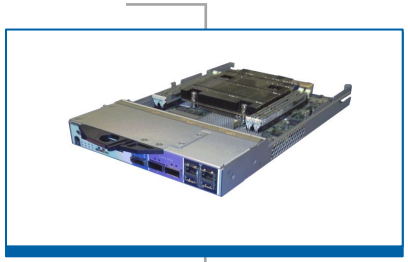
Meta Data Performance

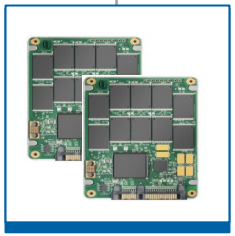
26K File Creates per Second Average

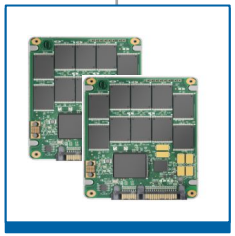
2 Billion Files Capacity per 5U84

Management Server









Metadata SSD Pool
~13K File Creates / sec
~ 1Billion files,
800 GB SSD x 2

User Data Pool
~4GB/sec
HDD x qty (40)

Metadata SSD Pool
~13K File Creates / sec
~ 1Billion files,
800 GB SSD x 2

User Data Pool
~4GB/sec
HDD x qty (40)

ClusterStor Spectrum Scale – Performance Configuration

SSU - NSD (MD) Server x 2 (SATI)

Large File Sequential Performance

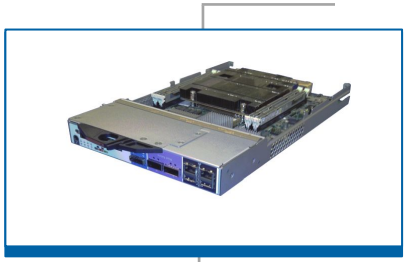
- *9GB/sec per 5U84 (Clustered)*
- *5GB/sec per 5U84 (Scatter)*

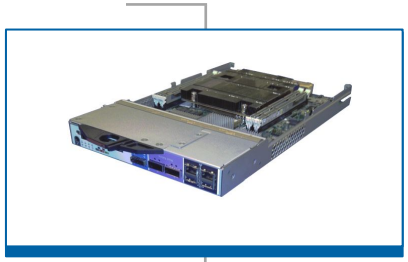
Meta Data Performance

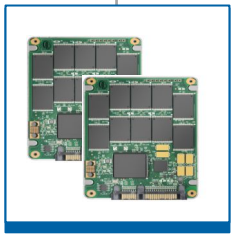
~40K File Creates per Second Average

4 Billion Files Capacity per 5U84

Management Server

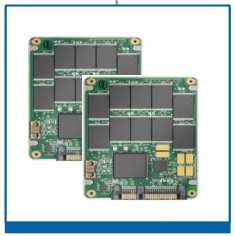






Metadata SSD Pool
~20K File Creates / sec
~ 1Billion files,
800 GB SSD x 4

User Data Pool
~4GB/sec
HDD x qty (36)



Metadata SSD Pool
~20K File Creates / sec
~ 1Billion files,
800 GB SSD x 4

User Data Pool
~4GB/sec
HDD x qty (36)

NSD Storage Servers, Chassis, and Disk Array

Management Servers - Seagate Ultra High Density - CS 2224

- 2U24 Enclosure
  - Dual-ported 2.5” SAS HDD Support
- Pair of H/A Embedded Management Servers

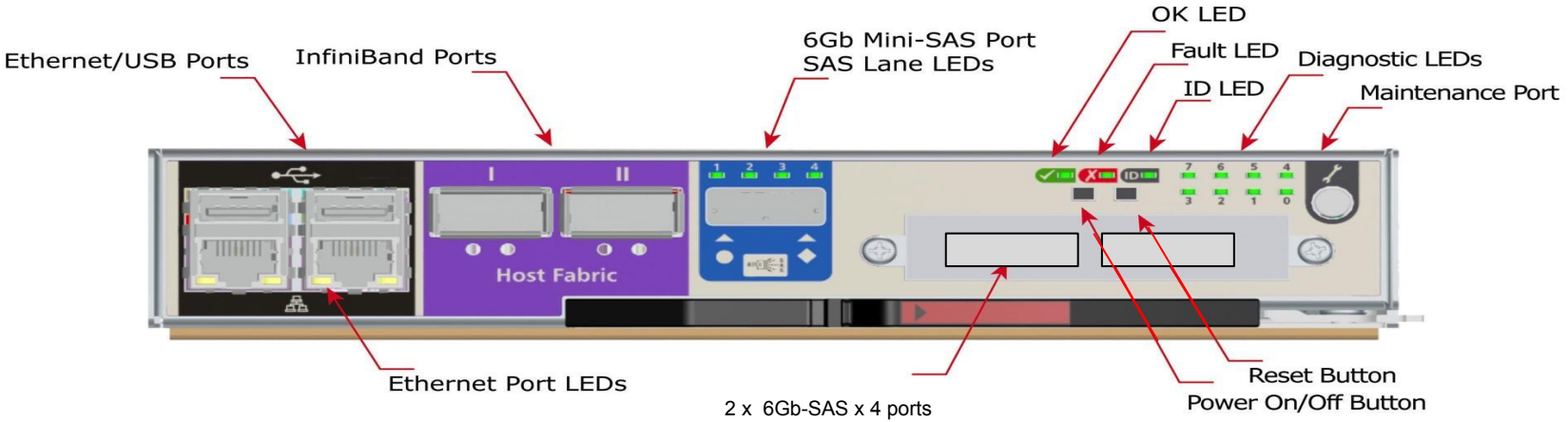
NSD & Meta Data Servers - Seagate Ultra High Density - CS-2584

- 5U84 Enclosure – completely H/A
  - Two (2) trays x (42) 3.5” drive slots
  - Dual-ported 3.5” ClusterStor HPC, Nearline SAS & SSD Drive Support
- Pair of H/A Embedded NSD Storage Servers
- QSFP port Supports IB QDR/FDR or 10/40 GbE Network Link

ClusterStor NSD Server

NSD Storage Server

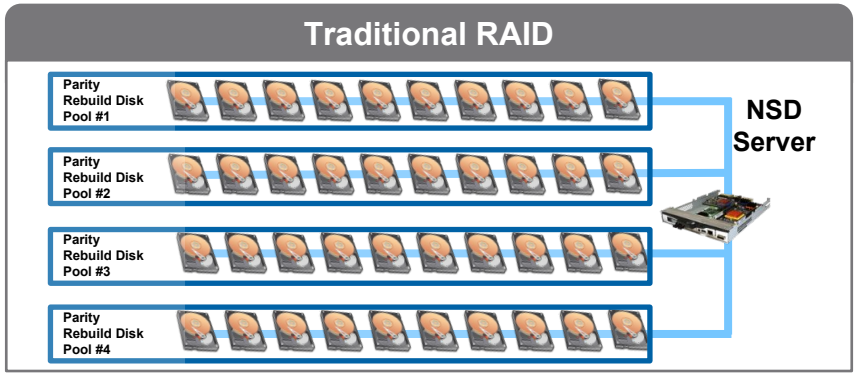
- QSFP Ports for Data Network Connectivity
- Embedded network support on the NSD Server motherboard
  - Supports both Infiniband QDR/FDR and 40Gb Ethernet data network connectivity

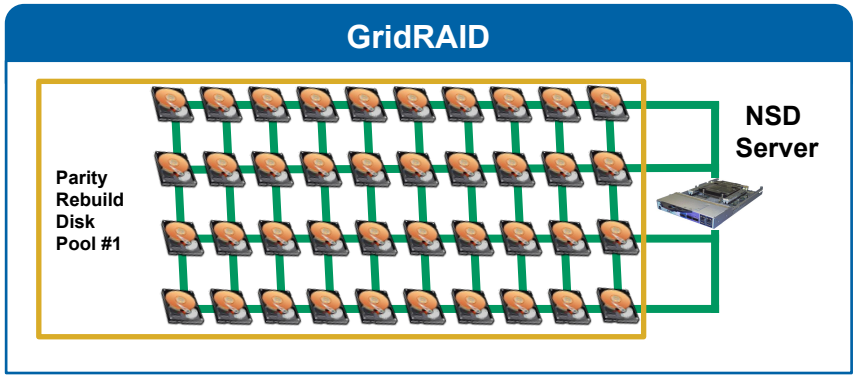


ClusterStor GridRAID

Single Line Definition of GridRAID with Benefit

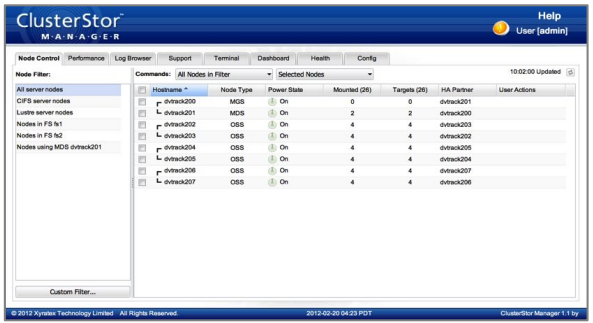
Feature	Benefit
De-clustered RAID 6: Up to 400% faster to repair **Rebuild of 6TB drive** – MD RAID ~ 33.3 hours, GridRAID ~ 9.5 hours	Recover from a disk failure and return to full data protection faster
Repeal Amdahl's Law: speed of a parallel system is gated by the performance of the slowest component	Minimizes application impact to widely striped file performance
Minimize file system fragmentation	Improved allocation and layout maximizes sequential data placement
ClusterStor Integrated Management	CLI and GUI configuration, monitoring and management reduces Opex

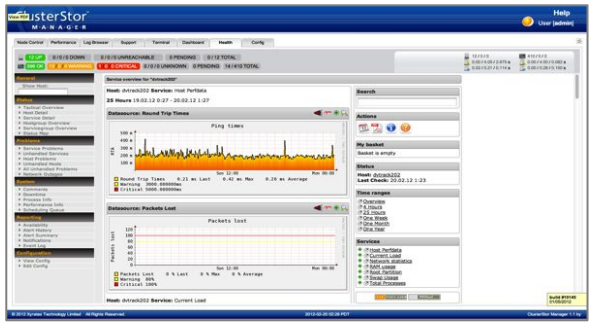




Management for ClusterStor

Based on ClusterStor Manager





Fully Integrated Solution Visibility and Management

Low level diagnostics, embedded monitoring, proactive alerts

Initial GA product:

- CLI-based File system management (as with Lustre)
- RAS/Phone home will be part of the solution
- SNMP integration with Business Operation Systems (Nagios, etc)
- Hardware monitoring (Health Alerts)

Release immediate following

- Updated GUI for performance reporting
- ReST API availability for RAS

ClusterStor HPC Drive: 4TB SAS HDD

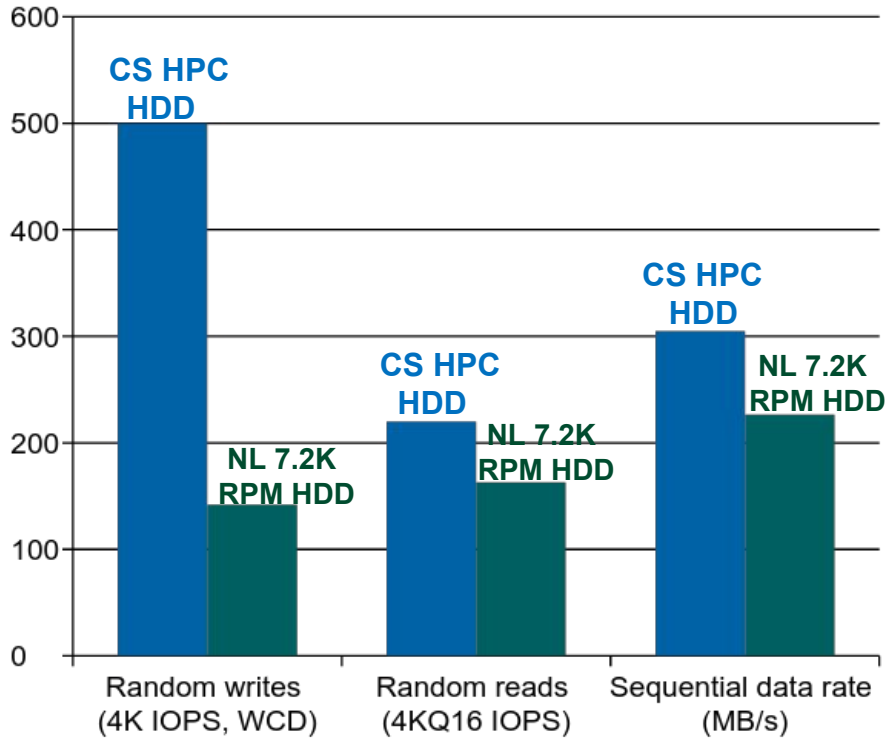
HPC Industry First; Best Mixed Application Workload Value

Performance Leader
World-beating performance over other 3.5in HDDs: *Speeding data ingest, extraction and access*

Capacity Strong
4TB of storage for big data applications

Reliable Workhorse
2M hour MTBF and 750TB/year ratings for reliability under the toughest workloads your users throw at it

Power Efficient
Seagate's PowerBalance feature provides significant power benefits for minimal performance tradeoffs



Simplified Installation – Hours vs. Days/Weeks

Out of the Box ClusterStor Solution

- **Sizing and Configuration optimization**
  - *Performance centric*
  - *Capacity centric*
- **Factory Integration & Staging**
  - *Rack Integration & cabling*
  - *Entire storage software stack factory pre-installed and pre-configured*
  - *System soak test and benchmark testing area at factory*

Spectrum Scale Version 4.1 Product Structure

Features	Express Edition	Standard Edition	Advanced Edition
Spectrum Scale v4.1 Includes 1 year subscription & support	NA (at this time)	GA Release	CY2016
Data sharing with a global namespace, simplified management at scale (quotas and snapshots) data integrity and availability	●	●	●
Create optimized tiered storage pools by grouping disks based on performance, locality, or cost characteristics		●	●
Simplify data management at scale with Information Lifecycle Management (ILM) tools that include policy based archiving to a low cost storage pool		●	●
Enable worldwide data access and empower global collaboration with Active File Manager (AFM)		●	●
Provide scalable file service with simultaneous access to a common set of data from multiple servers with Clustered NFS (cNFS)		●	●
Protect data at rest with native encryption and secure erase, NIST compliant and is FIPS certified			●

Seagate ClusterStor G200 Support Services

Seagate Provides All Levels of Support for Spectrum Scale

Professional Services

TIME TO VALUE

Strategic consulting, solution design, implementation and migration

Education Services

DEVELOP “IN HOUSE” EXPERTISE

Develop your team’s skills and expertise to optimize your investment

Support Services

EXPERT HELP

Support from trusted experts, with SLAs that meet your critical business needs

Expert Consulting, Services and Support from Seagate and Our Certified Partners

Award Winning ClusterStor Architecture

Powers the World's Fastest HPC Sites

Awards

Sample Customers

DOD HPC

"These new systems are a key component of our strategy of making sure the DOD's scientists and engineers have access to the most modern, capable, and usable computational tools available."

John West, director of the DOD's High Performance Computing (HPC) Modernization Program

DKRZ DEUTSCHES KLIMARECHENZENTRUM

"The ClusterStor solution provided the best performance density and, therefore, was the most efficient high-performance storage offering for our environment."

Professor Thomas Ludwig Director DKRZ and Research Team Leader

Met Office

"Improved observations, science and modeling, will deliver better forecasts and advice to support UK business, the public and government. It will help make the UK more resilient to high impact weather and other environmental risks."

Rob Varley
Met Office Chief Executive

KMA Korea Meteorological Administration

"We live in a weather sensitive environment, and people and businesses increasing rely on us for accurate environmental forecasting. Our new Cray supercomputers will be a valuable resource for us to meet our strategic, operational and research objectives."

Kyung Heoun Lee
Director of the National Center for Meteorological Supercomputing at KMA

ClusterStor Lustre Success Stories – *Powered by Seagate*

Seagate ClusterStor Lustre

5 of the 6 1TB/sec+ Storge File Systems are Seagate Clusterstor

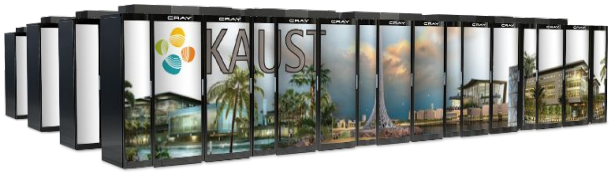
1TB/sec + ClusterStor Lustre Installations

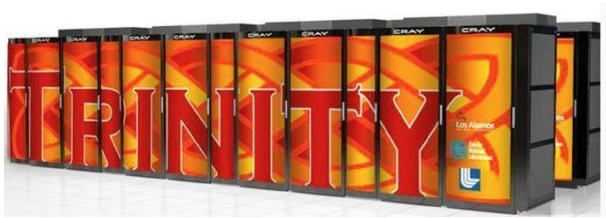
DKRZ - 1.5TB/sec 20PBs

"The ClusterStor solution provided the best performance density and, therefore, was the most efficient high-performance storage offering for our environment"

Thomas Ludwig, Director of DKRZ

Kaust - 1.0TB/s 17PBs





LANL - 1.7TB/s 77PBs

The Fastest File System in the World!

UK Met - 1.5TB/s 20PBs

Thank You