Cure Life Sciences **Storage & Data** Management Ills

Highlights

- **Increase data accessibility and flexibility** by aggregating disparate storage into a single pool in a global namespace

- **Easily integrates** into existing GPFS environments

- **High performance** flash storage to accelerate CPU and I/O intensive workloads

- **Reduce cost** by automatically moving unused or infrequently accessed data to an ActiveScale™ storage system

- **Extreme data durability and integrity** at petabyte scale helps to ensure valuable data is protected long-term

- **Improve IT agility** by quickly provisioning and redeploying storage resources to support new business needs

Rapid advances in life sciences research and development are making it possible for medical professionals to make more-informed and faster decisions, bringing a whole new dimension to personalized patient care. Imagine a gravely ill infant with a life expectancy of 60 days. Sequencing their DNA to determine the best course of treatment currently takes days, which can represent a significant percentage of their life. Accelerating genomic analysis to make critical treatment decisions can be essential to their survival.

As life science organizations pursue the objective of precision medicine, their storage and data management challenges remain constant. Limitations with massive scalability, storage performance and the accommodation of a wide variety of file types, sizes and access patterns are the same. The challenge of managing and sharing millions to billions of files over a long period of time seems unsurmountable.

Challenge

- **Costly to scale high performance storage systems** to keep up with data growth

- **Need for higher performance** to meet the need for faster turnaround time for analysis

- **High management overhead** to enable collaboration and sharing of data among global, distributed teams

- **Cost effectively keep data protected and accessible** over increasingly longer retention periods

Solution

- **The ActiveScale™ object storage system and General Atomics® EasyHSM™** hierarchical storage management solution for GPFS. The combined solution enables policy-based movement of files from GPFS to ActiveScale and back.

ActiveScale + EasyHSM for Life Sciences
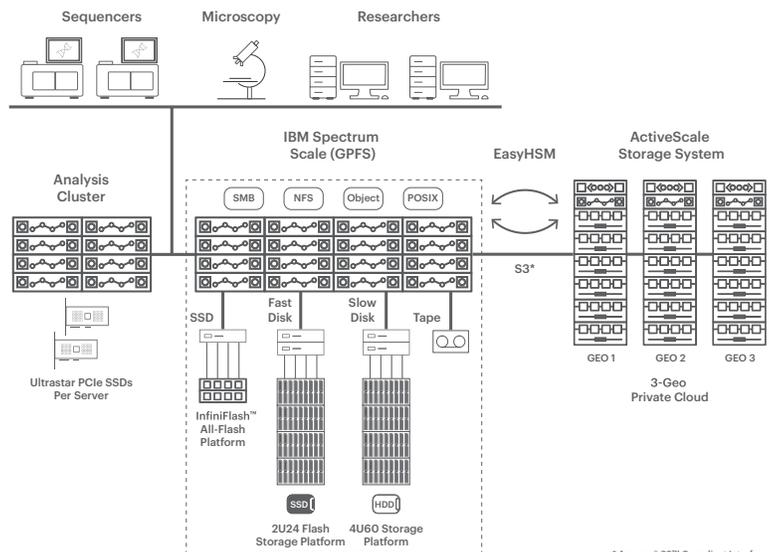
Life sciences data management architecture example

Accelerating Life Sciences Workloads

To support clinicians' ability to identify and personalize treatment options for their patients, the IT infrastructure they rely on must be up to the task. As genomic data sets continue to grow, life science organizations are rethinking their storage and data management needs. What is needed is a highly scalable distributed storage solution that can deliver various levels of data throughput, is affordable, scalable and globally accessible. In addition, the use of flash storage can have a profound impact on accelerating discovery and ultimately better patient care. The diagram on the previous page shows where flash storage fits.

Primary Storage Overload

Life Sciences organizations are at the epicenter of data growth. Each new generation of instruments catapult data volumes higher, driving the cost to maintain and operate traditional storage systems to unsustainable levels. Some organizations are dealing with storage sprawl, while others keep all the data on expensive high performance Tier1 storage. It's easy to understand "if it's not broke, don't fix it" when managing valuable data; however, there is a better way.

Data Management Made Easy

HGST and General Atomics are working together to dramatically reduce the cost of storage. The combined solution includes HGST's ActiveScale object storage system, and EasyHSM from General Atomics. ActiveScale is designed to cost-effectively protect and preserve petabyte-scale data over long periods of time. EasyHSM is a high-performance data management software for IBM SpectrumScale Storage (GPFS) environments that automatically and transparently moves data from expensive Tier 1 storage to ActiveScale. EasyHSM keeps everything the same for users and applications. User files remain accessible from high speed GPFS storage when needed. Application access remains unchanged with the same standard POSIX interface (SMB, NFS, GPFS) no matter where the data is stored.

Storage Simplicity at Petabyte-scale

Using next generation object storage technology, ActiveScale systems can help affordably protect valuable life science data over the long-term.

Easy to Install and manage, simply add power and network connections and ActiveScale is ready to go. The system self-protects and heals, by automatically addressing hardware and data integrity issues in the background. Unlike legacy systems that require an immediate response, there is no need to immediately swap-out the failed component. This can be done more efficiently during scheduled maintenance.

Advanced data protection helps ensure valuable data is protected and available. ActiveScale delivers up to 17 nines durability and site-level fault tolerance in a multi-geo configuration. Robust data integrity checks occur automatically and transparently, protecting data over the long-term; each object can tolerate up to 1000 bit-errors without data loss.

Capacity and performance to keep up with fast growing life science data sets. Whether it's the next generation sequencers or super resolution bio-imaging, ActiveScale has the capacity to easily scale to over 50 petabytes in a multi-site configuration. Aggregate throughput for a single rack can reach up to 8GB per second.

Advanced tools to gain insights and enable predictive analysis. ActiveScale CM is a cloud-based management tool. It analyzes single or multi-site deployments and provides analytics on capacity trends and performance bottlenecks, to help avoid issues and to keep things running smoothly.

Conclusion

Researchers and clinicians need access to that data for smarter patient-care decisions. ActiveScale and EasyHSM deliver the scale and efficiency with fewer resources and at a lower total cost than traditional approaches.

To learn more about ActiveScale visit:
http://www.hgst.com/products/systems

Contact Sales

North America: salesna@hgst.com | Europe/Middle East/Africa (EMEA): salesemea@hgst.com | Asia Pacific: salesap@hgst.com | Japan: salesjp@hgst.com

Western Digital Corporation

5601 Great Oaks Parkway
San Jose, CA 95119 USA

Information & Technical Support
www.hgst.com
www.hgst.com/support

Partners First Program
channelpartners@hgst.com
www.hgst.com/partners