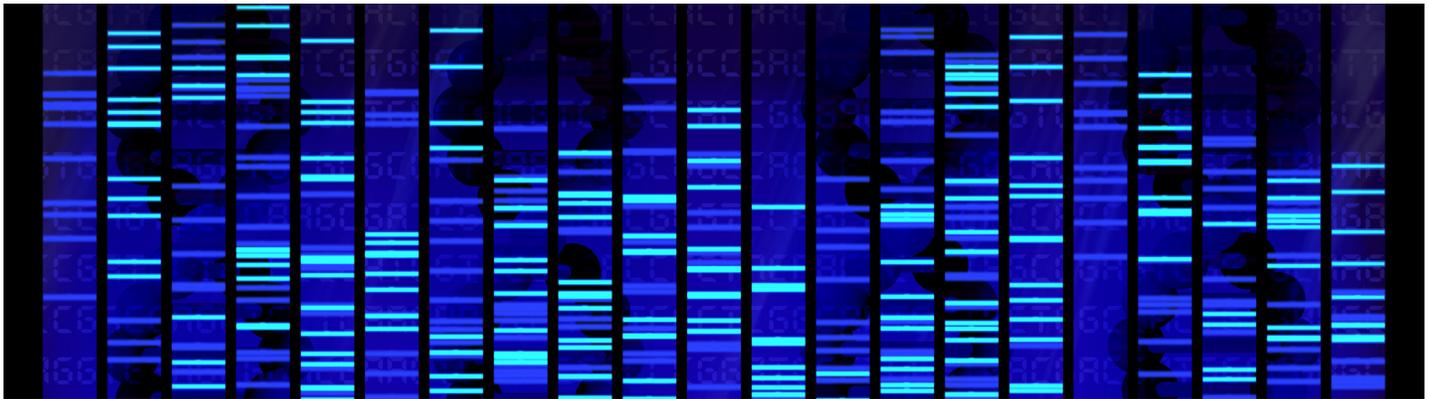


SOLUTION BRIEF

Meeting the Storage Challenge of Exponential Data Growth in Life Sciences



Life Science Storage Needs:

- Scalability
- Global namespace
- Support for all data types and access patterns
- Concurrent access
- Data integrity and durability
- Extremely low TCO
- Easily adapt to changing storage needs



OPPORTUNITY / CHALLENGE

The pace of change in life sciences is staggering as new technology allows researchers to see and do things never before possible. Laboratory instruments such as multi-dimensional microscopy, functional MRI, and DNA sequencers generate terabytes of data each, adding up to petabyte-scale storage requirements. DNA sequencing has seen dramatic proliferation due to a rapid reduction in cost. Thousands of sequencers sit in labs and hospitals around the world as they begin to sequence patients' genomes to customize care. A single DNA sequencer can easily generate multiple terabytes of data per day.

The challenge in today's life sciences world is how to efficiently store and manage the massive amount of diverse data and access patterns with concurrent usage for local and remote users and applications. The overwhelming requirement for researchers is shared access to common storage so lab instruments, compute clusters and workstations see all the same data. They also want access to all data whether derived or primary. The answer unfortunately often leads to implementing SAN solutions at the lab level that creates silos of data and are expensive to scale.

Traditional Storage Not Suited for Today's Life Science Workloads

Much of the data would eventually be stored on tape – that is not possible today due to the requirement for global online access and collaboration. Traditional RAID-based SAN systems are expensive and complex, and by themselves, cannot guarantee data will never be lost. This requires a backup solution that increases operational cost. All of this adds unnecessary complexity and overhead that strains IT budgets.

SOLUTION

Amplidata has addressed these and other storage challenges with Himalaya™, an object-based software defined storage platform that delivers unbreakable durability, infinite, scalability, and extreme efficiency. It is optimized for Intel-based commercial off-the-shelf hardware from a variety of vendors and provide flexible, cost effective exabyte scalability. Designed to take full advantage of the latest Intel* Xeon* Processors, performance scales linearly with each controller added to the system.

HIMALAYA™

HIMALAYA™

Reduced CapEx

- COTS hardware optimized for Intel* Xeon* processors
- Reduced storage overhead
- Lower software cost
- No backup required

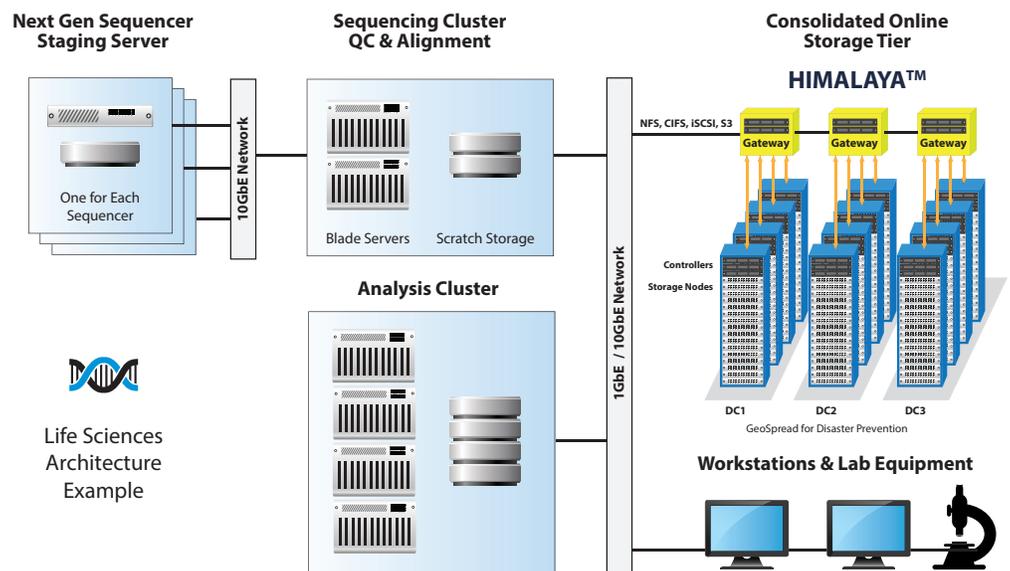
Reduced OpEx

- High availability
- Single system view
- Longer system useful life
- Fail in place model
- Automation
- Remote monitoring
- High density storage
- Low power consumption

A locality aware global namespace allows data to be accessed by instruments, users and applications regardless of location. Together with certified gateway partners, Himalaya easily integrates with existing environments with support for standard access protocols including NFS, CIFS, iSCSI*, HTTP/REST, S3* and iRODS*.

Himalaya eliminates the need for standard backup processes using its patented BitSpread® technology delivers fifteen 9s of data durability while tolerating up to 19 simultaneous drive failures. Patented BitDynamics® performs continuous data integrity checks and self heals in the background with little to no impact to system performance. Drive repair happens in parallel and out of band with all affected nodes participating. Failed drives can remain in place with their replacement shifting to a scheduled maintenance routine vs. the all too familiar RAID fire drill. With no additional software, GeoSpread offers disaster prevention by spreading data across multiple sites using 65% less capacity than RAID 6 and replication. Data remains protected and accessible even if an entire site becomes unavailable.

Performance and capacity can be scaled independently by simply adding Intel* Xeon* processor-based controllers for more performance and storage nodes for capacity. Changing the durability policy or system configuration can be done on-the-fly. The system will dynamically add the resources without having to migrate any data. Below is an example of a common life sciences system architecture.




Life Sciences
Architecture
Example

Contact Amplidata

North America: 408-433-1600

EMEA: +32 9 324 25 90

Email: info@amplidata.com

CONCLUSION

With Himalaya, life science businesses can dramatically reduce their total cost of storage ownership and keep more of their large data sets on-line and accessible. The modular architecture can enable significant operational efficiencies and workflow improvements while superior data durability and scalability allow organizations to move beyond traditional RAID-based storage and tape infrastructures. Himalaya is the ideal solution for today's life science workloads that require elastic scalability, high data durability and performance at a lower cost.

AmpliStor, BitSpread and BitDynamics are registered trademarks of Amplidata. *Other Brands and Names are the property of their respective owners.