STARBOARD
STORAGE SYSTEMS

CASE STUDY

The Marine Biological Laboratory

The Marine Biological Laboratory (MBL) is an international center for research, education, and training in biology, biomedicine, and ecology. MBL is dedicated to scientific discovery and improving the human condition through research and education in biology, biomedicine and environmental science. The MBL is a private nonprofit corporation founded in Woods Hole, Massachusetts in 1888.

Background

The MBL installed the Starboard AC72 Storage System™ to store and manage their modern application workloads consisting of massive amounts of unstructured, virtualized and structured data associated with their biological sequencing systems. The system has been utilized specifically for the Josephine Bay Paul Center for Comparative Molecular Biology and Evolution (JBPC, a division of MBL), affiliated with Brown University.

The JBPC's mission is to explore the evolution and interaction of genomes in response to environmental changes. The JBCP chooses those genomes, which have the greatest impact on human health and environmental biology. In addition, the JBCP shares laboratory resources with the Mobile Genetic Element Cluster, which studies mobile DNA. Mobile DNA is a specific sequence present in variable locations (commonly referred to as jumping genes). These genes change position via replication and are therefore a common culprit for the mutations that lead to cancerous cell growth.

The analysis of these large-scale genomic transcriptomic and metagenomic projects present a unique set of unstructured, virtualized and structured data that must be stored, analyzed, and archived according to usage patterns. The MBL currently manages between 16 and 24 active experiments per year. Each experiment generates approximately 3.25TB of data in a mixed-workload environment. The MBL chose Starboard Storage for its support of modern application workloads and data that is stored in multiple protocols. MBL also chose the Starboard Storage AC72 based upon its state-of-the-art autonomic tiering capabilities, built-in SSD performance and affordable scalability. The Starboard Storage AC72 delivers all these features at a much lower cost than comparable hybrid and unified storage systems. Implementing the Starboard Storage AC72 allowed the MBL to reduce cost and complexity when compared to their legacy NAS systems.

Starboard Storage Solutions for the Scientific Research and Life Sciences Market

Scientific Research and Life Sciences companies generate a tremendous amount of data. Bioinformatics and biostatistical analysis have a pressing need for storage solutions that are purpose built to support the accessibility of modern application workloads produced by application intensive analytics tools.

Organization
The Marine Biological Laboratory

Industry
Scientific research and discovery concerning education in biology, biomedicine and environmental science

Challenges

- Store research data produced by next generation biological sequencing systems
- Increase performance to meet growing access demands
- Make capacity easily scalable in line with actual business requirements
- Simplify provisioning and management of mixed workloads
- Archive and retrieve data according to access requirements

Solution
Starboard AC72 Storage System

Benefits
High performance, high availability storage system to support business critical applications

Scalable and cost-effective, protecting current and future IT investments

Robust management with modern technologies that ease provisioning and pools storage resources

> *Compared to other storage solutions, we found Starboard Storage to be the best fit for our managing our data growth. As an administrator, I appreciate how Starboard Storage sets the bar for consolidating and simplified management of modern application workload environments and delivering the industry's best price for performance. We are thrilled with the results we have seen so far and will be adding on additional Starboard Storage capacity in the future.*
>
> - Rich Fox, System Administrator

Performance

Bioinformatics companies such as the MBL, and its subsidiary branches, need high performance storage for application intensive studies such as gene profiling and sequencing. Often the entire genome is mapped for research and in other cases large panels are constructed using nanostring technology[1]. This is an especially relevant issue, as digital RNA sequencing will likely overtake microarray hybridization[2] in the near future. RNA sequencing creates much more data than its predecessor technology and will require much higher availability to respond to application demands.

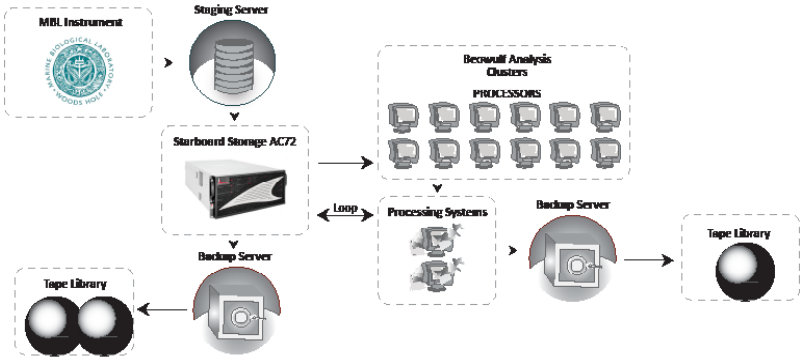
Capacity

The MBL also requires high capacity storage to house the gene profiles of archival samples. Large gene expression profiling studies can contain upwards of 1,000 samples[3] translating to hundreds of terabytes of profile information. (Waldron L., 2012) These samples may be resurrected at a later date for current studies in order to broaden the sample base resulting in a lower margin of error in statistical analysis. Additionally, this repository of information is invaluable in providing the basis for stringent quality control initiatives common to the industry (tracking the causes and occurrences of contaminated samples for example).

To meet the capacity needs of the MBL now and in the future, the Starboard Storage AC72 leverages its thin-provisioning feature to virtually allocate blocks of storage as demanded instead assigning physical disk space. This eliminates whitespace and increases overall utilization efficiency in line with actual usage patterns without adding administrative overhead. Dynamic Pooling allows the user to mix drive types and add them directly to the storage pool. The Starboard Storage AC72 stripes for redundancy across the entire drive pool (unlike the MBL's previous solution that offered striping across the RAID group only) automatically load balancing the data across newly added drives. SSD Tiering assures high-performance across the AC72 and maximizes response-time for the genomic sequencing system. The AC72 eliminates the need for IT administrators to set user created policies. Instead, the system intelligently performs workload balancing across tiers by analyzing application demands with built-in I/O detection technology. Data is placed on the desired class of disk storage autonomically and nondisruptively. When physical capacity must be expanded, the Starboard Storage AC72 supports 16-bay and 45-bay expansion shelves, totaling up to 576 TBs behind each AC72 node. Jeanne Wilson of Condor Storage states, "As a reseller and trusted advisor we are delighted that MBL successfully implemented the Starboard Storage AC72 to address the intense demands of their modern application workload environment, which includes storing large amounts of fixed content unstructured data."

Results

MBL has experienced significant performance improvement over their legacy DAS and NAS systems, and reduced operational overhead. In addition, the AC72's state-of-the-art storage features and technologies allow MBL to scale performance and capacity as their storage needs intensify. Starboard Storage Systems is an ideal fit for managing the MBL's modern application environment. ■



[1]Digital technology based on direct multiplexed measurement of gene expression that is capable of high level precision and sensitivity at less than 1 transcript copy per cell. Uses molecular "barcodes" and single-molecule imaging to detect and count hundreds of unique transcripts in a single reaction. Daniel E. Zak, A. A. (2009). A systems view of host defense. Nature and Biotechnology #27.

[2]Identification of organisms in an ecological sample by hybridizing DNA extracted from the sample to a set of known DNA specimens arrayed on a solid support, usually a microscope slide Welch, J. M. Microarray Hybridization. JBPc, MBL, http://jbpc.mbl.edu/wheelbase/ToolsTechniques/Microarray_Hybridization.html.

[3]Waldron L., S. P. (2012, March 29). Report on emerging technologies for translational bioinformatics: a symposium on gene expression profiling for archival tissues. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22458912