

A V E R E

Mount Sinai School of Medicine

Mount Sinai School of Medicine is internationally recognized as a leader in groundbreaking clinical and basic science research and is known for its innovative approach to medical education. The Mount Sinai community is highly regarded for its cutting-edge biomedical research, including genetics and proteomics. Recent breakthroughs by Mount Sinai genomics researchers include identifying a gene in the brain that may be the cause in the development of schizophrenia, identifying the common gene variant linked to autism and identification of the gene that is defective in most cases of prostate cancer.



MOUNT SINAI
SCHOOL OF
MEDICINE

Mount Sinai Case Study

Modern experimental technologies for genetic and genome sequencing projects involve massive datasets that require researchers to not only think about the biology of their projects but the technologies needed to handle the analysis of hundreds of millions of genomic sequences generated.

At Mount Sinai, gene sequencers and genomics analyzers leverage a compute farm with 500 processor cores, all of which have at least 64GB of RAM, to meet the demands of processes that can take up to 12 hours for a single sample. The complexity is driven by the computing challenge of processing a single genetic sample that requires 100 million sequences to recognize any anomalies.

The genomic sequencing process of RNA or DNA material is also a very disk-intensive operation stressing the I/O capabilities of the storage system with hundreds of thousands of reads and writes.

But the powerful compute capabilities did nothing to mitigate Mount Sinai's storage I/O bottleneck that was extending processing time to up to 12 hours for researchers. "The process of looking for fragments of RNA and DNA within genomic sequencing and comparing these against a 100 million sequencing reference genome is a very disk-intensive process with lots of reads and writes," said Ravi Sachidanandam, Assistant Professor of Genetics and Genomic Sciences at Mount Sinai.

The conventional approach to solve the NAS I/O problem is to add large numbers of high-performance and expensive Fibre Channel or Serial Attached SCSI hard disk drives to generate marginal performance improvement. The small performance boost comes at the cost of over-provisioning storage capacity, low disk utilization, inefficient use of limited data center space and wasted power consumption.

Mount Sinai Medical School turned to a state-of-the-art solution from Avere Systems to match its state-of-the-art research capabilities. Mount Sinai added Avere FTX 2300 Tiered NAS Appliances that sit between the compute farm and the storage to radically improve data I/O by dynamically placing heavily accessed data on tiers of Solid State Disk and fast-access Serial Attached SCSI disk drives in the appliance.

The Avere FXT 2300 appliance provides the high-performance storage media that enables it to accelerate the performance and reduce the cost of NAS environments. The FXT 2300 appliance incorporates 64GB of DRAM,

“The process of looking for fragments of RNA and DNA within genomic sequencing and comparing these against a 100 million sequencing reference genome is a very disk-intensive process with lots of reads and writes. Avere has helped minimize these read/write bottlenecks, dramatically reducing the amount of time needed to process each job.”

Ravi Sachidanandam
Assistant Professor of Genetics
and Genomic Sciences

1GB of NVRAM and eight 146GB 15k SAS drives, supporting an application working set of 1.2TB raw capacity per appliance and up to 29TB raw per cluster.

It leverages Avere’s Demand-Driven Storage™ technology to intelligently and dynamically locate hot data to the optimal storage tier based on business demand. The most active files are stored on FXT 2300 appliance while inactive data is placed on the NAS cluster.

The intelligent algorithms within the Avere OS constantly monitor data access patterns and actively manage data placement to increase performance, distribute workload in the cluster, and minimize accesses to the mass storage system, dramatically improving data I/O with no expenditure to upgrade the backend NAS cluster or the compute farm.

The effect of installing the Avere appliances was immediate and dramatic. Mount Sinai slashed its processing time from 10-12 hours to under 6 hours. “Avere has helped minimize these read/write bottlenecks, dramatically reducing the amount of time needed to process each job,” said Sachidanandam. “With our hardware tuned to the needs of our research and development, we will be able to handle even more data generated as we upgrade our sequencing solution in the near future.”

With its decision to go with Avere, Mount Sinai has removed current I/O bottlenecks in its computing infrastructure. In addition, the Avere solution has given Mount Sinai the option for future I/O scalability with the ability to add up to 25 FXT appliances in a single cluster to support any expansion in its genomics research processing capabilities.

A V E R E