

Garvan Institute Drives Genomics Revolution in Medicine

Panasas ActiveStor Scale-Out NAS Storage Anchors Infrastructure Upgrade That Increases Genome Sequencing Capacity by 50 Times

SUMMARY

CUSTOMER

Garvan Institute of Medical Research

SIZE

600 employees

LOCATION

Sydney, Australia

INDUSTRY

Life sciences

CHALLENGE

- Boost performance level to handle flood of genomics data
- Minimize installation and maintenance time requirements to avoid hiring staff
- Avoid disrupting work of 80 researchers using the institute's high-performance infrastructure

SOLUTION

- Upgraded performance to enhance user experience for researchers
- Installed state-of-the-art genome sequencers
- Developed innovative workflow to improve throughput

RESULTS

- Increased production capacity by 50 times without adding staff
- Boosted productivity of 80 internal researchers
- Delivered results quickly to researchers around the world



“Panasas lives up to its promise of terrific performance with negligible maintenance and administration time.”

Dr. Warren Kaplan

Chief of Informatics
Garvan Institute of Medical Research

The sequencing of the first human genome in 2003 took 10 years and cost US\$3 billion; the same task today takes just three days and costs just over \$1,000. Garvan Institute of Medical Research aims to take advantage of this astonishing technological advance to improve clinical practices such as assessing cancer risk and diagnosing children with intellectual disabilities.

Garvan is widely recognized for its genomics expertise, thanks in part to its willingness to adopt new technology. When Illumina, a leading manufacturer of genomics sequencing instruments, introduced the HiSeq X Ten sequencing platform in 2014, Garvan immediately decided to upgrade to this breakthrough technology—one of only three organizations in the world to do so at product introduction.

However, building a genomics production line presents daunting technical challenges. The X Ten system produces up to 5TB of data per day. To keep the line rolling, downstream analyses and archiving operations must be able to handle this torrent of

information. “Meeting our commitments to researches requires extremely high computational power that is available 24/7,” says Dr. Warren Kaplan, chief of informatics at Garvan.

A second challenge involves funding. Many of the institute's high-performance computing and data infrastructure improvements rely on grants, which only cover hardware, not personnel. Therefore, ease of installation and maintenance are vital to avoid diverting existing resources or hiring additional staff.

Finally, ongoing research could not be compromised. The X Ten system had to be integrated into the current high-performance infrastructure in a way that would scale to support growth and, ideally, enhance the user experience.

Performance Bottlenecks Vanish Thanks to Panasas, Greatly Enhancing Researcher Productivity

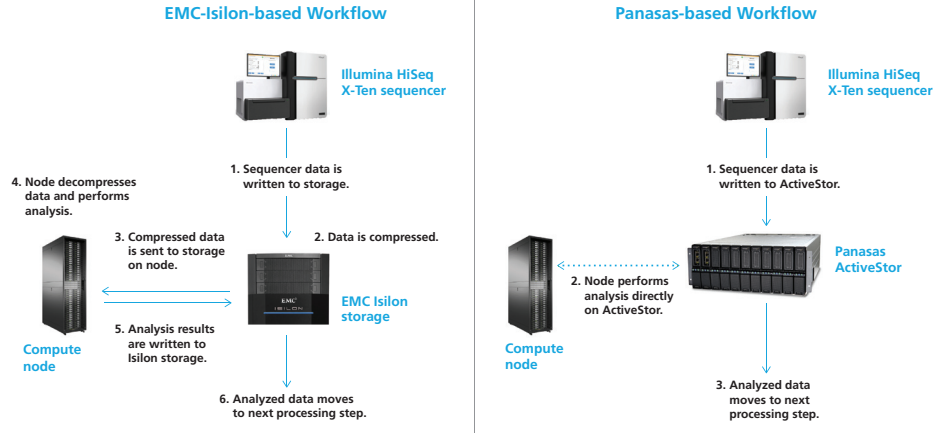
To run the Illumina system at full capacity, Garvan needed to make changes to the existing infrastructure, most notably, implementing parallel processing. Lacking funding for more staff, Kaplan decided against labor-intensive open source software such as Lustre. Instead, he opted for Panasas® storage with the Panasas PanFS® parallel file system.

Timely migration to Panasas boosts application response and user satisfaction

Five Panasas ActiveStor® network-attached storage (NAS) appliances arrived just in time to quell a user mutiny over painfully slow application response. The root cause was an expansion of the research staff from 10 to 80, which had overloaded the existing storage system. Kaplan’s group migrated the researchers to the Panasas storage as soon as it was installed. “Their response? ‘Problem solved!’” says Kaplan. “Thanks to Panasas, they were back to full productivity.” The team later installed one more ActiveStor storage system, bringing total Panasas storage to 400TB.

Inventive workflow bests EMC Isilon-based approach

When it came to designing the workflow, Panasas again saved the day. In Illumina’s recommended workflow, sequencer data moves back and forth between the EMC Isilon central storage and local storage on the compute nodes. “Thanks to Panasas’ exceptional performance, our sequencing data stays in the central repository throughout the analysis,” Kaplan says. “This streamlined workflow saves time and bandwidth, enabling us to deliver results quickly to researchers around the world.” The new Illumina sequencers and the high-performance platform with Panasas storage have increased Garvan’s sequencing capacity to 50 genomes per day on average—a fiftyfold improvement.



Comparison shows how Panasas approach saves time by reducing unnecessary steps and streamlining Garvan’s data processing workflows.

As temperatures soar, Panasas takes the heat

A cooling system failure caused Kaplan to worry about equipment damage—needlessly, as it turned out. “The Panasas system automatically shut down, avoiding damage that could corrupt data and reduce IT and researcher productivity,” says Kaplan. “Panasas lives up to its promise of terrific performance with negligible maintenance and administration time.”

Garvan Drives Toward Goal of Better Clinical Practices Through Genomics

Garvan has an ambitious goal, nothing less than transforming the practice of medicine through genome sequencing. The

institute envisions itself as an enabler that can rapidly prototype and evaluate specific analyses. Once verified, those analyses will be made available to downstream research institutions as well as businesses working to commercialize genomics technology. “Fundamental to our work is maintaining an extraordinary infrastructure that makes it all possible,” says Kaplan, “and Panasas is a key part of that.”

To learn more about Panasas ActiveStor platforms that bring plug-and-play simplicity to large-scale storage deployments, visit www.panasas.com/products.