Quantum®

CASE STUDY

Embedding Archive into the DNA of the Australian Genome Research Facility with Quantum Multi-Tier Storage

The Australian Genome Research Facility (AGRF) amasses large amounts of genome sequencing data to support academic, industry, and government customers. The team at AGRF adopted a multi-tier Quantum storage solution to accommodate rapid data growth and better protect their valuable data.

agrf
Empowering Australian Genomics

FEATURED PRODUCTS

StorNext® 5

"Previously, we couldn't back up our data in a reasonable window. By continuously archiving data using the Quantum solution, we can better preserve and protect data."

Rowan Gronlund
Head of ICT, AGRF

SOLUTION OVERVIEW

- Quantum StorNext® Scale-out Storage
- Quantum Artico™ NAS Appliance
- Quantum Scalar® Tape Library

KEY BENEFITS

- **Supports tremendous data growth**
  while enabling users to easily access old data
- **Helps protect valuable sequencing data**
  with a multi-tier storage environment
- **Simplifies storage management**
  and provides readily available support
- **Reduces hardware costs**
  by capitalizing on cost-effective tape archives

The Australian Genome Research Facility (AGRF) provides genomic sequencing for a broad range of customers, including academic researchers, healthcare organizations, commercial businesses, government agencies, and more. "We support clinical research and DNA analysis," says Rowan Gronlund, head of information and communications technology (ICT) at AGRF. "We also work with agricultural firms, helping identify crop strains for maximum yields."

After receiving biological samples, AGRF teams perform quality control checks, and then conduct genome sequencing on a range of machines, from Sanger sequencers to next-generation sequencers from Illumina. They can then run bioinformatics or other analyses before returning the data to customers.

Running approximately 20 Sanger sequences and 30 next-generation sequences along with other processes every month produces tremendous data volumes. "Our goal is to process a large amount of data and hand it back to customers within a reasonably fast timeframe," says Gronlund.

FACING RAPID GROWTH OF GENOMIC DATA

For Gronlund's ICT group, storing that sequencing data presents serious challenges. "The amount of data we store has tripled in the last two and a half years," says Gronlund. "Every time we add a new sequencer or there is a change in chemistry, data grows significantly. For example, we recently had a chemistry change on our next-generation sequencers that meant we could do twice the runs in half the time—so effectively, we multiplied the amount of data we had to store by four times."

Data growth had limited the amount of time that AGRF could hold data. "In the past, we could only hold onto data for 200 to 300 days before we had to remove it from our systems to make room for the next data sets," says Gronlund. "To keep the genomic data longer, we needed a new approach."

RECOGNIZING THE CRITICAL IMPORTANCE OF PROTECTING SEQUENCING DATA

The AGRF IT group also lacked a way to protect the valuable genomics data produced by researchers. "We stored sequenced data in a single array—without any data protection," says Gronlund. "If the array died, we would have had to re-sequence the biological sample, which is not always possible."

If re-sequencing had been required, it would have been difficult to keep existing projects on track. "Sequencing can take four to six weeks," says Gronlund. "If we had to re-sequence data, it would seriously impact the current processing load and force us to delay existing projects."

Just as important, data loss could have significant negative effects on AGRF's reputation. "We're a service organization. If we lose data, our reputation takes a serious hit," says Gronlund. "Our clients wouldn't trust us. There's a great potential for loss of customers and income."

IMPLEMENTING A QUANTUM MULTI-TIER SOLUTION TO STORE AND ARCHIVE DATA

After considering a number of options for expanding its storage infrastructure, the ICT group selected a multi-tier storage solution from Quantum. "We had a good understanding of how multi-tier environments work, and we had heard very positive feedback from other Quantum customers," says Gronlund.

The ICT group selected a Quantum Artico NAS appliance, which is powered by StorNext data management software. Artico provides more than 30TB of high-performance disk-based storage as the front end for a multi-tier storage and archive environment. The ICT group has written custom scripts that move data to the Artico environment as soon as it comes off the sequencer.

The ICT group also implemented a Quantum Scalar tape library behind the Artico appliance. The 80-slot device has two LTO-6 drives. "We're up to 140 tapes, some in the vault," says Gronlund. "Between Artico and the Scalar tape library, we have approximately 230TB of native storage capacity."

PROVIDING LONGER-TERM STORAGE FOR SEQUENCING DATA

With the new solution, AGRF can now provide longer-term storage for data produced by sequencers. "The Quantum solution allows us to take sequencer data and put it in an archive for an extended period of time," says Gronlund. "Instead of having to remove data within 300 days to make room for new data sets, we can now hold onto it for up to three years."

Keeping that information longer enables customers to access and analyze data again without having to migrate it to and from their own IT infrastructure. "Customers can now request older data, and we can produce it quickly and easily," says Gronlund.

REDUCING COSTS WITH ACTIVE VAULT

The AGRF ICT group uses Active Vault—a feature that allows infrequently used tapes to be stored in a separate partition of the library—to avoid excessive hardware purchases. "Active Vault lets us keep costs under control," says Gronlund. "By using the Artico solution plus the Scalar tape library with Active Vault, we are able to bring the total cost down to half the cost of a competing solution."
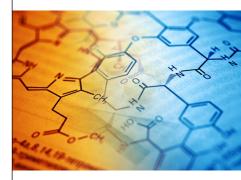
STREAMLINING IT ADMINISTRATION

The ICT group selected the Quantum solution in part because it offered an easier way to manage storage than solutions from other vendors. "Competing solutions were overly complicated," says Gronlund. "The Quantum solution doesn't require a high level of expertise for management, and that's perfect for us."

The Quantum solution, which includes an easy-to-use web interface, helps save administration time. "The amount of time we have to spend administering the Quantum environment is minimal," says Gronlund. "Since we completed the initial setup and policy configurations, we really haven't touched the environment except for purchasing more tapes."

If the ICT team needs assistance understanding administrative functions, Quantum support is readily available.

"With the Quantum solution, we can avoid having to re-sequence data and—more importantly—we can minimize the risk of losses that would affect our organization's reputation."

Rowan Gronlund,
Head of ICT, AGRF

ABOUT AGRF

A not-for-profit organization, the Australian Genome Research Facility (AGRF) is dedicated to improving the quality of life through exceptional life science. AGRF provides genome sequencing services and expertise that are easily accessible throughout Australia. By exchanging knowledge, fostering collaboration, and offering state-of-the-art technologies, AGRF enables Australian universities, industry organizations, and government agencies to lead genomic research internationally. AGRF relies on a Quantum StorNext solution to protect and maintain long-term access to data generated by the organization's genomic sequencers.

"As a manager, I want great support for whatever we buy," says Gronlund. "The Quantum team has been quick to respond whenever we've had questions."

PROTECTING SEQUENCING DATA WITH THE MULTI-TIER QUANTUM ENVIRONMENT

The Quantum solution helps the ICT group better protect valuable sequencing data by enabling more robust backup and disaster recovery (DR) strategies. "Previously, we couldn't back up our data in a reasonable window," says Gronlund. "By continuously archiving data using the Quantum solution, we can better preserve and protect data."

FACILITATING DISASTER RECOVERY WITH MULTIPLE COPIES, ON-SITE & OFF-SITE

The Quantum solution also facilitates disaster recovery. "Artico allows us to keep an untainted version of the data produced by sequencers," says Gronlund. "If anything goes wrong with post-sequencing analysis, we can recover original data easily."

The tape archives provide additional layers of DR protection. "We keep two copies of data on tape—one on-site and another off-site," says Gronlund. "If there is some disaster in our primary data center, data stays protected. With the Quantum solution, we can avoid

having to re-sequence data and—more importantly—we can minimize the risk of losses that would affect our organization's reputation."

PREPARING FOR FUTURE DATA GROWTH

With a new, scalable solution in place, the ICT group at AGRF can address the continued growth of genomics data well into the future. "Data growth was one of the primary drivers for acquiring the Quantum solution," says Gronlund. "We can reduce costs and extend the front-end disk system by moving data into a cost-effective archive. As a result, we can much more easily accommodate our customers' growing data requirements."

To contact your local sales office, please visit www.quantum.com

Quantum.