

DUG Unlocks Faster RNA Sequencing with VAST Data Platform & Solidigm™ Storage

From seismic exploration to RNA sequencing, DUG harnesses VAST Data Platform, Solidigm QLC SSDs to redefine HPCaaS

As a leader in high performance computing (HPC) solutions, DUG Technology Ltd (DUG) is driving innovation, with applications ranging from seismic analysis for the resources sector to decoding genomes for healthcare. With its high-powered HPC-as-a-service (HPCaaS) offerings, DUG provides organizations with the computational resources to tackle complex challenges, offering scalability and flexibility without the need for on-prem infrastructure.

Founded in 2003 as DownUnder GeoSolutions, DUG began as a geoscientific service provider that operated out of a shed in co-founder Matthew Lamont's backyard. Fast forward 22 years, DUG now has offices in Perth, London, Houston, Kuala Lumpur and Abu Dhabi, supporting a diverse industrial client base. The company also designs, owns and operates a network of some of the largest supercomputers on the planet.

"Our HPC workloads include seismic data processing and imaging, bioinformatics, artificial intelligence, remote Earth observation, computational fluid dynamics and radio astronomy," said DUG Chief Information Officer (CIO) Harry McHugh. "The key takeaway here is that HPC is the enabler of modern scientific discovery – indeed it is now a necessity." A critical enabler of DUG's technology lies in its advanced storage architecture, which plays a pivotal role in accelerating data-intensive workloads, McHugh said.

DUG initially relied on HDD-based storage to support its HPC applications, but the company encountered performance bottlenecks, reliability concerns, and scalability limitations, particularly with its Lustre-based file system. As demand for data processing grew—especially with seismic workloads that required copying and writing data multiple times—DUG needed a more versatile and scalable solution to ensure consistent, high-speed input/output (I/O) performance. The company sought a storage architecture capable of handling exascale requirements while maintaining efficiency and cost-effectiveness.

"Compared to traditional storage systems, running these workloads on the VAST and Solidigm architecture yielded a **1.7x speed increase and a **40% reduction in runtime costs**," McHugh said. "The speedups we've achieved with VAST and Solidigm are remarkable."¹**

The Computational Challenges of Big Data

While most of DUG's compute power is used to process and image seismic data, the company's HPCaaS clients also include bioinformaticians – whose workloads include the computationally demanding challenges of RNA sequencing. Researchers rely on RNA sequencing to analyze genomic data, but this process generates an enormous volume of small files, requiring a storage system that can handle high bandwidth with minimal latency.

RNA sequencing workflows typically involve input data sizes of around 100 gigabytes, but can produce output datasets as large as five terabytes, McHugh said.

"That's a huge jump from input to output, with millions of small files," he added.

This explosion of data places immense stress on storage and computing infrastructure. The challenge is well-suited for the purpose-built VAST Data Platform, which utilizes Solidigm™ SSD technology. DUG chose VAST for its high-performance, scalable data solutions designed to handle big data workloads. The VAST Data Platform's storage layer combines Solidigm SSD storage with advanced data reduction and efficiency technologies, enabling organizations to manage massive datasets – like those in RNA sequencing – quickly and efficiently.

“A distributed file system is essential for that (bioinformatics) workload to happen,” McHugh said. “It needs to be high-performance; it needs to be reliable. It needs to have high IOPS. In other words, it needs to have all the things that VAST offers.” The VAST Data Platform speeds up HPC by providing the performance and scale of parallel file systems with the simplicity of NAS (network-attached storage).

Benefits of RNA Sequencing

RNA sequencing is revolutionizing how disease is understood and treated at the transcriptional level — giving a real-time snapshot of gene activity across cells and tissues. By providing a dynamic view of the genome in action, RNA sequencing offers deeper insight than DNA alone — and that's helping clinicians deliver more accurate diagnoses, better-targeted treatments, and improved patient outcomes. Here are some of the ways that RNA sequencing is transforming medicine:

Precision Diagnosis: RNA-sequencing reveals which genes are actively being expressed in a patient's cells, helping clinicians differentiate between disease subtypes, detect rare or misdiagnosed diseases, and guide biomarker-based diagnostics.

Personalized Treatment: By profiling how a patient's genes are turned on or off, RNA-sequencing enables selection of targeted therapies, monitoring of treatment response over time, and identification of drug resistance mechanisms.

Infectious Disease Management: RNA-sequencing helps detect and characterize viruses, bacteria, and other pathogens — even when they're hard to culture. It also enables real-time monitoring of pathogen evolution (e.g., COVID-19 variants) and detection of host immune responses to infection.

Understanding Complex Disorders: For diseases like neurodegenerative disorders, autoimmune conditions, and psychiatric illnesses, RNA-sequencing uncovers patterns of gene misregulation, insights into disease pathways and progression, and novel targets for treatment.

Cancer Research & Oncology: RNA-sequencing has become a foundational tool in oncology, helping identify gene fusions and mutations driving tumor growth, stratify patients for clinical trials and inform the development of RNA-based therapeutics.

How Solidigm and VAST Data Enable DUG's HPCaaS Offerings

Underlying the VAST Data Platform are the high-density QLC SSDs from Solidigm – a technology that has solved a litany of challenges for DUG.

VAST's data platform, enabled by Solidigm QLC SSD technologies, provided DUG with a breakthrough solution. By implementing Solidigm QLC SSDs, VAST offered a combination of all-flash storage with affordability that is more competitive than HDD-based systems.

By transitioning to the VAST Data Platform with Solidigm SSDs, DUG achieved exascale scalability with no single points of failure, enhanced data protection and compression capabilities, and a more efficient multi-tenant cloud environment.

Moreover, the remote appliance management capabilities provided by VAST Data significantly reduced operational overhead for DUG, allowing zero-downtime updates and proactive performance monitoring. The result was a highly reliable, cost-efficient, and future-ready storage infrastructure, enabling DUG to expand its compute capabilities while reducing complexity. With this solution, storage has transformed from a bottleneck into a strategic asset, empowering DUG to pursue new markets and drive innovation across multiple industries.

Over the years, VAST upgraded its underlying Solidigm storage technology from one generation to the next, enabling higher capacities.

“We've been with VAST from the start, and we've grown the amount of VAST we have installed over our sites over the years,” McHugh said. “So, I think that says a lot about that continuing relationship we have.”

McHugh noted that traditional storage systems are often optimized for large streaming reads and writes, which work well for applications like seismic processing, but struggle with small, fragmented file structures common in genomics research.

“Having Solidigm SSD storage in our infrastructure allows us to manage these high-random IOPS workloads effectively,” McHugh says. “It gives us the ability to scale while maintaining speed and efficiency.”

Comparing Performance of Traditional HDDs vs. Solidigm Storage Solutions

For DUG’s bioinformatics customers, the VAST and Solidigm combination translates into substantial operational benefits. The system’s ability to handle massive data throughput and rapid file access is key to accelerating genomic sequencing processes.

DUG conducted benchmarking that demonstrated that they were able to significantly improve the efficiency of bioinformatics workloads by leveraging Solidigm’s advanced storage solutions.



Faster Performance



Runtime Cost Reduction

“Compared to traditional storage systems, running these workloads on the VAST and Solidigm architecture yielded a 1.7x speed increase and a 40% reduction in runtime costs,” McHugh said. “The speedups we’ve achieved with VAST and Solidigm are remarkable.”¹

A Growing Collaboration with VAST Data

The reliability of VAST’s platform is another critical factor that keeps DUG investing in the technology – further enabled by the quality and reliability of the Solidigm storage devices.

“Reliability and performance often come at a trade-off, but with VAST and Solidigm, we see both,” McHugh explains. This balance ensures that DUG’s customers can trust the infrastructure supporting their most data-intensive workloads.

As industries continue to push the boundaries of data-driven science, the need for robust, high-performance computing solutions will only grow. DUG’s use of the VAST Data ecosystem enabled by Solidigm enterprise SSD storage is a great example of how technology can drive innovation across multiple verticals, from seismic processing to bioinformatics.

Appendix

1. Source: DUG

Nothing herein is intended to create any express or implied warranty, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, or any warranty arising from course of performance, course of dealing, or usage in trade.

The products described in this document may contain design defects or errors known as “errata,” which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Solidigm does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Contact your Solidigm representative or your distributor to obtain the latest specifications before placing your product order.

SOLIDIGM and the Solidigm “S” logo are trademarks of SK hynix NAND Product Solutions Corp. (d/b/a Solidigm), registered in the United States, People’s Republic of China, Japan, Singapore, the European Union, the United Kingdom, Mexico, and other countries.

