

GWDG Delivers Robust AI and Data Services with the VAST Data Platform



The Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) is at the forefront of artificial intelligence (AI) research in Germany. As a joint facility of the University of Göttingen and the Max Planck Society, GWDG provides critical IT services and infrastructure to support cutting-edge scientific research.

In 2022, GWDG faced a pivotal challenge: they needed a data platform to support their national AI service center while addressing their broader data needs. After an extensive evaluation, GWDG selected the VAST Data Platform as the foundation for their next-generation data platform infrastructure.

Prof. Dr. Julian Kunkel, Deputy Head of High-Performance Computing at GWDG, explains the decision: “We initially approached VAST because their platform is useful for AI workloads, particularly NVIDIA GPUDirect® support and read-intensive training. After our evaluation, we realized we could use it for many other purposes.”

Streamlined Data Management

One key benefit GWDG has experienced with VAST Data is simplified data management. Prof. Dr. Kunkel notes, “VAST is the entry point for managing multiple data platforms. Depending on the workload, we utilize different data systems.”

This centralized approach allows GWDG to manage diverse workloads and user requirements efficiently. “If a user has a high-performance workload, their data gets placed on the VAST Data Platform. If a user has a project with many large files that are not frequently accessed, we put them on traditional storage,” Prof. Dr. Kunkel explains.

The flexibility of the VAST Data Platform has enabled GWDG to consolidate multiple systems and workloads. Various workloads like home directories, software distribution, AI projects, and node deployment now leverage the VAST Data Platform.

¹NVIDIA Magnum IOTM GPUDirect® is a family of technologies that accelerate data movement and access for NVIDIA data center GPUs.

Why the VAST Data Platform?

The VAST Data Platform is a highly scalable and affordable all-flash data system that allows you to run AI-scale analytics at less than half of the cost of traditional all-flash solutions.

DASE Architecture

With its unique Disaggregated and Shared-Everything architecture, VAST breaks tradeoffs deliver significant savings and make flash affordable for all of your data.

VAST DataStore

The VAST DataStore is an enterprise all-flash NAS platform built to meet the needs of today's powerful AI computing architectures and beyond.

VAST DataBase

The VAST DataBase is a key strategic piece of the Data Platform targeting structured data. It provides transactional and analytical data warehouse functionality on top of the scalable, performant, all-flash base layer to ingest and process data at best-in-class speeds.

VAST Data Catalog

The VAST Data Catalog leverages the Database and SQL to provide a built-in metadata index that allows you to search and find data easily – structured or unstructured – at scale and in a fraction of the time of traditional methods.

Exceptional Performance and Reliability

For GWDG, the VAST Data Platform has delivered on its performance promises. “We get a 2:1 data reduction ratio, with zero downtime since we started using VAST in production,”

Prof. Dr. Kunkel reports. This reliability is crucial for GWDG’s role as a service provider to the research community.

The system’s ability to handle diverse workloads has been particularly impressive. Prof. Dr. Kunkel highlights, “Our Software Tree, a read-heavy workload, works like a charm on VAST.”

The ease of use and robust performance have had a tangible impact on GWDG’s operations. As Prof. Dr. Kunkel humorously puts it, one key benefit has been the ability to “sleep better,” knowing the data platform is stable and reliable.



“When we talk to the VAST team, they are very knowledgeable. What they say is not marketing speak; it’s true. They tell us ‘You can do that.’ We then try it and it works.”

Prof. Dr. Kunkel, GWDG

Powering AI Innovation

As a national AI service center, GWDG leverages the VAST Data Platform to support various AI initiatives. “We offer services like LLM fine-tuning and racking services,” Prof. Dr. Kunkel explains. “These are mostly hosted on the VAST Data Platform.”

The platform’s performance characteristics make it ideal for AI workloads. Researchers across Germany use GWDG’s infrastructure for model training and fine-tuning tasks. The organization is also exploring AI applications to improve its operations, including automated ticket processing and enhanced system monitoring.

Seamless Scalability and Future-Proofing

GWDG’s initial VAST deployment started at 600 terabytes, with plans to expand beyond one petabyte. The platform’s ability to scale seamlessly is crucial for GWDG’s evolving needs. Prof. Dr. Kunkel appreciates the value proposition: “Long-term infrastructure use is an exciting opportunity for us to minimize data migration efforts.”

The VAST Data Platform’s unified file and object storage approach has opened new possibilities for GWDG. Prof. Dr. Kunkel notes, “We found the unified namespace for S3 to be an outstanding concept. We want to explore and offer this to our users.”

This flexibility positions GWDG to tackle future data management challenges. Prof. Dr. Kunkel is particularly interested in emerging AI-driven data organization techniques: “There are new data management programs where they just pull data in, and then it organizes the directory structures for you.”

A Partnership Built on Trust

Throughout the deployment and ongoing use of the VAST Data Platform, GWDG has found a true technology partner. Prof. Dr. Kunkel praises the VAST team's expertise: "When we talk to the VAST team, they are very knowledgeable. What they say is not marketing speak; it's true. They tell us 'You can do that.' We then try it and it works."

This level of trust and technical competence has solidified GWDG's confidence in their choice. Prof. Dr. Kunkel summarizes his experience, "For me, this product comes with plenty of features and works as it should."



"VAST is the entry point for managing multiple data platforms. Depending on the workload, we utilize different data systems. We offer services like LLM fine-tuning and racking services, these are mostly hosted on the VAST Data Platform."

Prof. Dr. Kunkel, GWDG

Looking Ahead

As GWDG continues to expand its AI and other data services, the VAST Data Platform will play a central role in enabling new research capabilities and driving scientific discovery. With a robust, scalable, and flexible storage foundation, GWDG is well-positioned to meet the evolving data challenges of the scientific community it serves.

About VAST Data

VAST Data is the leading data platform software company bringing businesses into the AI era. The VAST Data Platform delivers scalable performance, simplifies data management, and enhances productivity by accelerating time-to-insight for workload-intensive applications. Launched in 2019, VAST has become history's fastest-growing data infrastructure startup.

Looking to simplify your data center and discover insights in all of your data? **Contact us at hello@vastdata.com**