

CASE STUDY

FACILITATING CUTTING-EDGE RESEARCH WITH A HIGH-PERFORMANCE CLOUD THAT IS MANAGEABLE AT SCALE

Quobyte software storage accelerates the time to scientific breakthroughs in life science research by simplifying storage management, delivering HPC-grade parallel file system performance, and facilitating data ingest and processing with its hybrid file-object-storage capabilities.



Summary

Needs/Challenges: The University of Tübingen handles raw scientific data sets on a petabyte scale; hence, efficient data processing and analysis are paramount – as is security and protection of the sensitive data researchers are working on.

Solution: Quobyte's Data Center File System

Platform:

- Quobyte Data Center File System
- OpenStack
- 40GbE and InfiniBand-based network
- Servers with SSDs and HDDs
- 13 PB of data and growing
- Solution design, deployment, and support by Hamburgnet

Use Case: Life Sciences, Bioinformatics, Genomics, Scientific Computing

Key Benefits

- High-performance parallel file system and native S3 access speed up data ingest and processing
- Scale-out storage eliminates the need for tedious capacity planning
- Seamless integration with OpenStack
- File, block, and object storage packaged into one system
- Multi-purpose volumes are possible thanks to multi-tenancy and ACLs
- Multi-tenancy provides strong separation of data and greatly facilitates operations
- No storage or data silos enable barrier-free collaborative research
- Automated management enables low-touch operations with a small team of admins

“ A high-performance and reliable IT infrastructure is crucial to perform analyses with continuously growing data sets and Quobyte plays a key role in future-proofing our Infrastructure. ”

Dr. Jens Krüger

Group Leader High Performance and Cloud Computing

University of Tübingen and de.NBI – Trailblazing the Cloud Infrastructure for Life Sciences Research in Germany

The University of Tübingen provides one leg out of five for de.NBI Cloud, the German Network for Bioinformatics Infrastructure. With the de.NBI Cloud, some of Germany's leading universities join efforts in biomedical research to provide researchers with the computational resources required to help them achieve faster scientific breakthroughs in the life sciences. It's a collaborative solution that enables integrated analyses for the entire life sciences community in Germany ensuring the efficient use of data in cutting-edge research and its applications.

Today's life science research is not only about compute resources for data analysis, but storage for enormous amounts of data also poses a key challenge. New genome sequencing and imaging technologies result in the generation of massive amounts of data which, in many cases, are highly sensitive and require solid protection.

University of Tübingen's part of the research cloud provides biomedical researchers with the infrastructure for their data-intensive research workflows. The genomics data are exploding and require a huge amount of storage capacity and – to efficiently run analytics on that data – high-performance read and write access.

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



Meeting the Demands for a Competitive Research Cloud

“A high-performance and reliable IT infrastructure is crucial to perform analyses with continuously growing data sets and Quobyte plays a key role in future-proofing our infrastructure”, says Jens Krüger, Leader of the High Performance and Cloud Computing Group at the University of Tübingen.

The call for tenders for the new storage infrastructure had three major requirements: First, the storage system needed to integrate seamlessly and painlessly with OpenStack which is the basis for the entire de.NBI research cloud. Second, it had to be multi-purpose and able to handle both high throughput and high IOPS workloads. And third, a geo-redundant replication concept had to be set up that secured the data at a second location.

Integrating Seamlessly with OpenStack

Originally looking at Ceph for their OpenStack storage foundation, Krüger and his team found that it added another level of complexity to the already complex OpenStack infrastructure. Looking around for alternatives, they heard about how Quobyte eliminates storage complexity. Testing Quobyte software storage, they learned that the installation was a breeze and the storage system would also save them maintenance costs thanks to its auto-monitoring and self-healing features.

One Storage System for File, Block, and Object Storage

Quobyte fit right in with University of Tübingen’s needs for the triad of file, block, and object storage. The native S3 interface makes data ingest easy and passes data to the parallel file system so that researchers can run their analytics and number crunching directly on a file system – without the need for further migrations. The data analysis jobs researchers run require a high throughput, meaning access to the file system needs to be fast which guarantees that the jobs yield results quickly.

Making Capacity Planning Obsolete

Quobyte also helps the University of Tübingen get rid of tedious capacity planning. Thanks to Quobyte’s ability to work with any standard server hardware, the storage team can just add disks (HDD, SSD, NVMe, or a mixture of them) and servers as required – depending solely on their needs. Quobyte’s powerful policy engine takes care of the rest and can be configured on the fly, giving the team full flexibility over the storage infrastructure.

Multi-Purpose Infrastructure thanks to Multi-Tenancy

The key requirement of working collaboratively in a cloud environment is multi-tenancy. OpenStack provides strong, secure multi-tenant capabilities that are crucial to having highly sensitive data and proprietary research on a shared infrastructure.

Quobyte storage fully supports multi-tenancy, letting the Tübingen team define isolated namespaces for their storage. Tenant group members can only see the volumes assigned to them, while tenant admins can manage resources of their own group but are prevented from seeing the entire cluster. Quotas can be shared by tenant members, and administrators can further isolate tenants by controlling which physical hardware they have access to.

The key to effortlessly administer and maintain such collaborative research is Quobyte’s fine-grained access control to multi-purpose volumes. On top of ACLs, access to sensitive patient data can also be secured by using certificate-based authentication protocols.

Disaster Recovery through Geo-Replication

As a major plus, Quobyte is able to provide the University of Tübingen’s de.NBI installation with geo-replication so that a second cluster could be set up as an asynchronous mirror of the scientific data. So, in addition to an already secure storage infrastructure, this second cluster even shields the data against major physical failures and disaster at site one.

The Impact: A Solid Foundation for a Modern Research Cloud

With Quobyte in place and running in production, the University of Tübingen and its part of the de.NBI cloud are future-proof in crucial ways: first, the storage infrastructure is easily expandable when the need arises and scales out seamlessly. Second, the system delivers the performance today’s research workloads require. Third, the fact that one storage system runs file, block, and object workloads makes things simpler and more efficient for the storage admins. And fourth and fifth, ACLs and X.509 certificates allow for the full-fledged security that’s required for even the most sensitive research data while at the same time allowing for the silo-free flexibility that collaborative research needs.

The design and implementation for this project was realized in cooperation with Quobyte’s trusted partner **Hamburgnet**, a company with excellent expertise in the field of software-defined data center solutions.



 [google.com/+Quobyte](https://plus.google.com/+Quobyte)

 [linkedin.com/company/quobyte](https://www.linkedin.com/company/quobyte)

 twitter.com/quobyte

Quobyte[®]

Data Center File System.™ Fast and Reliable Software Storage.

Quobyte Inc.

4633 Old Ironsides Drive
Suite 150
Santa Clara, CA 95054
info@quobyte.com
+1 650-564-3111

Quobyte GmbH.

Hardenbergplatz 2
10623 Berlin
info@quobyte.com
+49 30-814 591 800