

CASE STUDY

How the Karlsruhe Institute of Technology (KIT) Migrated 100 PB from Oracle to Spectra



Image: Karlsruhe Institute of Technology

100 PB. No operational downtime. Here's how KIT achieved reliability, performance, and cost efficiency with the Spectra TFinity® Plus tape system.

THE CHALLENGE

Operating on aging Oracle tape libraries with diminished support introduced vulnerabilities to long-term reliability for KIT's GridKa, a Tier 1 center for CERN. To ensure continuous access to research data and maintain operational stability, **KIT needed to migrate 100 PB from legacy Oracle technology** to a scalable and fully supported storage infrastructure.

THE SOLUTION

KIT selected the Spectra TFinity Plus tape library for its unmatched capacity, reliability, and cost efficiency, along with its multi-media flexibility. Over two years, **KIT migrated 100 PB from its Oracle systems to the Spectra TFinity Plus** and TS1160 tapes and drives while keeping both libraries active to maintain daily operations. During migration, KIT optimized its environment and workflows, eliminating HDD buffers in favor of flash to maximize tape drive performance.

THE RESULT

KIT now operates on a future-ready storage foundation that can scale seamlessly as data demands grow. The modernized architecture delivers **improved tape library diagnostics, maximum drive performance, and world-class technological support**, enabling faster, more reliable access to critical datasets for researchers around the world.

THE SYSTEM



THE ENVIRONMENT SNAPSHOT

- Spectra TFinity Plus tape library, 17 frames
 - TS1160 tape drives and media
- LumOS library management software
- HPSS storage software

THE STORY

Located in the famously fan-shaped city of Karlsruhe is one of Germany's most prestigious research institutions: the Karlsruhe Institute of Technology (KIT).

Formed in 2009 through the merger of the University of Karlsruhe and the Karlsruhe Research Center of the Helmholtz Association, KIT educates the next generation of scientists while also operating large-scale national research infrastructure.

One of KIT's most important units is the Scientific Computing Center (SCC), which hosts GridKa — a German Tier 1 data center that stores, processes, and distributes scientific data from all four CERN Large Hadron Collider experiments (ATLAS, CMS, ALICE, and LHCb).

Using GridKa's 100-petabyte archive, scientists probe some of the deepest questions in physics, from the nature of antimatter to the fundamental structure and evolution of the universe.

Doris Ressmann, Deputy Head of Department for Scientific Data Management at KIT, and her team steward this vast and continuously growing dataset.

Unlike many scientific archives that primarily store infrequently accessed cold data, GridKa operates as a high-throughput, always-active system: continuously ingesting new experimental data while simultaneously retrieving stored datasets for the global scientific community. Data flows out almost as much as it flows in.

For years, this critical workload was supported by a trio of Oracle (StorageTek) STK8500 tape libraries. As these tape systems aged and support services waned, Ressmann and her team undertook a major modernization effort. **The goal: migrate 100 petabytes to a new tape infrastructure designed for greater affordability, flexibility, and long-term resilience.**

Here's how KIT successfully executed this large-scale migration, ensuring that decades of vital research data remains both accessible and enduring.

ABOUT KARLSRUHE INSTITUTE OF TECHNOLOGY (KIT)



Image: Magali Hauser/Karlsruhe Institute of Technology

KIT develops solutions for urgent challenges — from climate change, energy transition, and sustainable use of natural resources to artificial intelligence, sovereignty, and an aging population.

The Scientific Computing Center is the information technology center of KIT.

FROM STAGNANT INFRASTRUCTURE TO PLANNING FOR THE FUTURE

As Oracle support services and tape roadmap development diminished, KIT's legacy Oracle tape libraries no longer provided the reliability and performance required by its fast-moving Worldwide LHC Computing Grid (WLCG) community.

To assuage the risks of remaining on a fixed, maintenance-only platform with no clear upgrade path, **KIT initiated a migration from Oracle libraries, drives, and media to a Spectra TFinity Plus tape library with IBM TS1160 media and drives.**

This modern, fully supported archive solution is designed for long-term scalability and operational continuity. KIT selected the TFinity Plus tape library because it delivers:

- **Density & Scalability:** TFinity Plus is the highest-capacity tape library in the world, with an efficient footprint.
- **Future-Ready Compatibility:** TFinity Plus aligns with KIT's TS1160 strategy, while maintaining flexibility for potential integration of LTO technology in the future.
- **Cost Efficiency:** TFinity Plus combines high performance with long-term affordability at scale.
- **Intelligent Diagnostics:** Spectra LumOS library management software enables organizations to manage, monitor, and optimize tape storage with unmatched efficiency and reliability.

As part of the migration, KIT also swapped its storage software from IBM Storage Protect to HPSS, which supports demanding requirements for large-scale, long-term scientific data preservation and access.

TWO-YEAR DATA MIGRATION PROCESS

"Migrating data at this scale was a massive undertaking," says Ressmann.

Compounding the challenge, the **migration had to occur with minimal disruption to daily operations.** Because the flow of critical research data does not pause, the team executed a dual-operation strategy — often referred to as a hybrid migration — running both the legacy and modern environments in parallel over a two-year period.

Both the Oracle and new Spectra tape libraries remained active throughout the entire migration. This approach preserved operational stability by decoupling the migration process from immediate decommissioning.

It afforded KIT the flexibility to transition data from Oracle technology to the Spectra TFinity Plus and TS1160 media and drives methodologically, one experiment at a time.



MIGRATION: AN OPPORTUNITY FOR OPTIMIZATION

The migration wasn't just a transfer of data from one system to another. It was also a rare opportunity to examine and improve KIT's archiving workflows for GridKa.

Ressmann and her team leveraged the migration project to modernize their entire environment. They identified and resolved performance bottlenecks, developed custom client software to optimize data write and recall operations, uncovered and resolved system bugs, and even re-architected their infrastructure to ensure tape drives could operate at maximum efficiency.

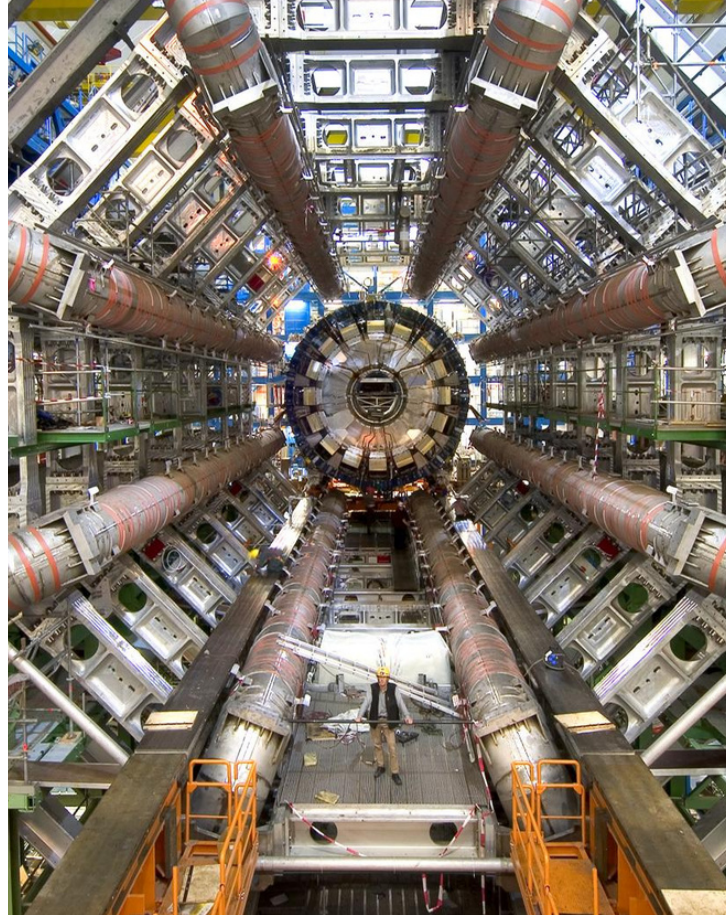
One key upgrade was the elimination of traditional HDD buffering in front of the tape library.

Unlike KIT's previous tape-based infrastructure, its new Spectra TFinity Plus tape library system relies solely on flash. "Hard drives didn't have enough IOPS to keep the tape drives running optimally," says Ressmann. "With flash, we're now able to keep the drives streaming at their maximum speed."

A TRUSTED STORAGE PARTNER

Through careful planning, continuous optimization, and a commitment to modernization, **KIT successfully migrated 100 petabytes of critical CERN research data** while also creating a supported, future-proof, and flexible storage environment.

Today, the GridKa archive stands on a more scalable, resilient, and cost-efficient foundation — one that is able to support the relentless pace of scientific discovery.



KIT's GridKa is a Tier 1 data center for CERN's Large Hadron Collider, which generates massive amounts of irreplaceable scientific data. Image: CERN

Contact Spectra Logic