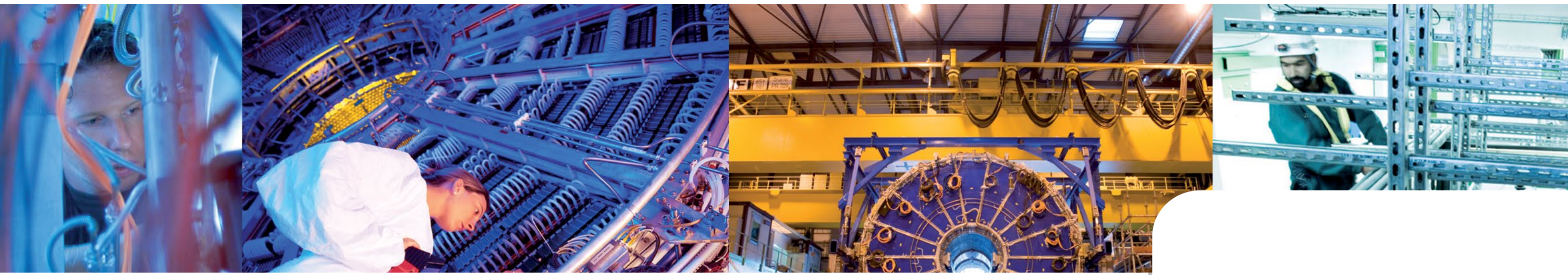
Quantum.

CASE STUDY



StorNext Software

Quantum StorNext Helps CERN Accelerate Research and Discovery

When it comes to exploring the origins of the universe, managing the massive volumes of data can be a galactic task. CERN, the world's leading particle physics laboratory, uses StorNext software to enable data from the ALICE experiment to be accessed and shared in a high-performance computing environment for scientists to quickly understand and exploit new ideas and discoveries.

MANAGING BILLIONS OF BITS OF DATA GENERATED EVERY SECOND

CERN—the European organization for nuclear research—is the world's leading laboratory for particle physics. Founded in 1954 and headquartered in Geneva, Switzerland, the laboratory was one of Europe's first joint ventures and now includes 20 Member States. In CERN's Large Hadron Collider (LHC) particle accelerator very high-energy protons collide against protons, and heavy ions (like the nuclei of lead) collide against heavy ions. This allows scientists to penetrate further into the structure of matter and recreate the conditions prevailing in the universe just a few millionths of a second after the "Big Bang."

One of the experiments assigned to the LHC is ALICE (A Large Ion Collider Experiment). This is one of the largest experiments in the world devoted to researching the physics of matter at an infinitely small scale. Hosted at CERN, it involves an international collaboration of more than 1,000 physicists, engineers and technicians from 30 countries. Together they contribute to resolving one of the earliest challenges in fundamental physics: recounting the birth of matter.

Using detectors, ALICE collects massive amounts of data that is generated by particle collisions. Pierre Vande Vyvre, Project Leader for ALICE Data Acquisition, was tasked with designing an information management system with rock-solid data acquisition, selection, transfer, storage and handling to manage the billions of bits of data generated every second. This cutting-edge system needed to handle the continued flow of data at 1.5 GB/second which is sustained day and night for at least one month. This can amount to more than one petabyte (PB) of data accumulated in a month. To further complicate requirements, the ALICE experiment takes place three kilometers away from CERN's main computer center where all of this data resides on mass storage systems.

"The buffering of data between the experiment and the mass storage site is essential," says Vande Vyvre. "The ALICE Data Acquisition system requires enough storage equipment for several hours of autonomous data acquisition before data is migrated to the computing center. After selection, we store the data to disk and subsequently transfer it off line to the main computer center."

Another challenge in the data acquisition process was that the Linux file system in place on several tens of cluster nodes used at ALICE was not sufficient to share between nodes.

"We needed to be able to access data quickly and easily on all hosts without having to pass the data over slow network connections," says Ulrich Fuchs, Data Acquisition (DAQ) Systems Manager. "Linux NFS was a possibility but it displayed limitations in terms of aggregated bandwidth and scalability of the number of client nodes."

"Data is CERN's most precious commodity. Quantum StorNext is instrumental in collecting that data quickly and reliably, thereby enabling the scientific community to understand and exploit new ideas and discoveries."

Pierre Vande Vyvre
Project Leader

SOLUTION OVERVIEW

- Quantum StorNext File System and Distributed LAN Clients
- InfoTrend Fibre Channel Storage Arrays
- 4 GB/s disk arrays
- 75 servers running Linux
- QLogic Fibre Channel adapters and switches

KEY BENEFITS

- Provides continuous and simultaneous access to data while reducing storage costs
- Scales to achieve 4.5 GB/s write and 2.5 GB/s read speeds
- Supports LAN clients for cost-efficient data sharing expansion
- Preserves user choice and eliminates dependence on any single vendor
- Enables infrastructure to operate at peak performance using Affinity to direct data to specific primary disks

STORNEXT SCALES TO MEET CLUSTER FILE SYSTEM DEMANDS

Vande Vyvre, Fuchs and the rest of the CERN team investigated several different options available for a clustered file system (CFS). The main requirements for the CFS were maximum aggregate bandwidth performance; a minimal footprint for the hardware equipment (the ALICE Data Acquisition room was fairly small); scalability to manage the volume of clients (up to 100 user nodes); and—most importantly—independence between the CFS and the underlying hardware.

Among the half-dozen products initially considered, the team narrowed it down and performed extensive testing using identical hardware (one SAN fabric, six storage arrays and nine client nodes). The performance target was at least one fifth of the global performance required: i.e., 300 MB/s writing, 300 MB/s reading and 600 MB/s for combined read and write.

After several weeks of heavy testing, the team selected StorNext®.

During the initial stage of implementation, the ALICE Data Acquisition team began with one server, one client, and one disk array. The system now consists of 180 Fibre Channel 4G ports and 75 transient data storage arrays, with each divided into 3 volumes. There are 105 nodes accessing data over Fibre Channel. The ALICE storage architecture also has 90 StorNext distributed LAN clients accessing data over IP. All nodes communicate over a Gigabit Ethernet network.

ENABLING FAST, EFFECTIVE DATA ACQUISITION

One of StorNext's key benefits comes from its Affinity feature. This allows the team to direct data to specific primary disks, by writing to the affinities' associated relation point.

"If we put all the disk arrays into one global file system, we would have two machines connected with reading and writing taking place together in parallel," explains Fuchs. "Under such conditions, our tests have shown that performance would drop by 90 percent. Affinity allows us to direct data traffic to pre-determined disks. This means that all the machines operate at maximum performance at all times."

The ALICE system continues to scale and perform as planned. Over a seven month period, the ALICE experiment observed proton-to-proton collisions writing data at up to 500 MB/s sustained. A performance test was able to achieve 4.5 GB/s write speed and 2.5 GB/s read speed at a total load of 7 GB/s on the file system in preparation for a heavy-ion run. This performance headroom above the original specification has been instrumental: the heavy-ion run began in November 2010 and has produced data throughput of up to 2.5 GB/s sustained over one month. During that same time period, a total of 700TB of data has been collected. The total amount of data collected in 2010 was more than 2PB.

"StorNext is delivering the high speed, shared workflow operations and large-scale, multi-tier archiving required by ALICE," said Vande Vyvre. "It is important to appreciate the logistics involved in CERN's experiments, where a massive amount of data is required for analysis. Data is CERN's most precious commodity. Quantum StorNext is used to collect that data quickly and reliably, thereby enabling the scientific community to understand and exploit new ideas and discoveries."

"Affinity allows us to direct data traffic to pre-determined disks. This means that all the machines operate at maximum performance at all times."

Ulrich Fuchs
DAQ Systems Manager

ABOUT CERN

CERN is the European Organisation for Nuclear Research, the world's largest particle physics centre. Located near Geneva, CERN is a laboratory where scientists unite to study the building blocks of matter and the forces that hold them together. CERN exists primarily to provide them with the accelerators, which accelerate particles to almost the speed of light and contribute to the design and the construction of the detectors to make the particles visible. Founded in 1954, the laboratory was one of Europe's first joint ventures and includes now 20 Member States.