

Large Hadron Collider pushing computing to the limits

March 1 2019



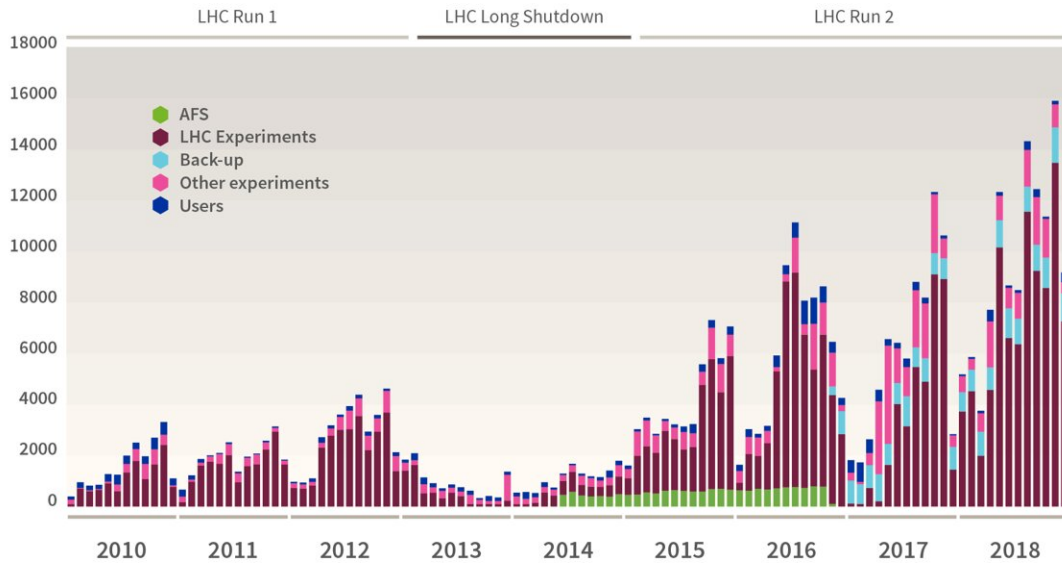
Racks of computers in CERN's computing centre are just a fraction of the hardware needed to store and process the data from the LHC. Credit: Anthony Grossir/CERN

At the end of 2018, the [Large Hadron Collider \(LHC\)](#) completed its

second multi-year run ("Run 2") that saw the machine reach a proton–proton collision energy of 13 TeV, the highest ever reached by a particle accelerator. During this run, from 2015 to 2018, LHC experiments produced unprecedented volumes of data with the machine's performance exceeding all expectations.

This meant exceptional use of computing, with many records broken in terms of data acquisition, data rates and data volumes. The [CERN Advanced Storage system \(CASTOR\)](#), which relies on a tape-based backend for permanent data archiving, reached 330 PB of data (equivalent to 330 million gigabytes) stored on tape, an equivalent of over 2000 years of 24/7 HD video recording. In November 2018 alone, a record-breaking 15.8 PB of data were recorded on tape, a remarkable achievement given that it corresponds to more than what was recorded during the first year of the LHC's Run 1.

The distributed storage system for the LHC experiments exceeded 200 PB of raw storage with about 600 million files. [This system \(EOS\)](#) is disk-based and open-source, and was developed at CERN for the extreme LHC computing requirements. As well as this, 830 PB of data and 1.1 billion files were transferred all over the world by [file transfer service](#). To face these computing challenges and to better support the CERN experiments during Run 2, the entire computing infrastructure, and notably the [storage systems](#), went through major upgrades and consolidation over the past few years.



Data (in terabytes) recorded on tape at CERN month-by-month. This plot shows the amount of data recorded on tape generated by the LHC experiments, other experiments, various back-ups and users. In 2018, over 115 PB of data in total (including about 88 PB of LHC data) were recorded on tape, with a record peak of 15.8 PB in November. Credit: Esma Mobs/CERN

New IT research-and-development activities have already begun in preparation for the LHC's Run 3 (foreseen for 2021 to 2023). "Our new software, named CERN Tape Archive (CTA), is the new tape storage system for the custodial copy of the physics data and a replacement for its predecessor, CASTOR. The main goal of CTA is to make more efficient use of the tape drives, to handle the higher data rate anticipated during Run 3 and Run 4 of the LHC," explains German Cancio, who leads the Tape, Archive & Backups storage section in CERN's IT department. CTA will be deployed during the ongoing second long shutdown of the LHC (LS2), replacing CASTOR. Compared to the last year of Run 2, data archival is expected to be two-times higher during Run 3 and five-times higher or more during Run 4 (foreseen for 2026 to

2029).

The LHC's computing will continue to evolve. Most of the data collected in CERN's data centre is highly valuable and needs to be preserved and stored for future generations of physicists. CERN's IT department will therefore be taking advantage of LS2, the current maintenance and upgrade of the accelerator complex, to perform the required consolidation of the computing infrastructure. They will be upgrading the [storage](#) infrastructure and software to face the likely scalability and performance challenges when the LHC restarts in 2021 for Run 3.

Provided by CERN

Citation: Large Hadron Collider pushing computing to the limits (2019, March 1) retrieved 17 July 2024 from <https://phys.org/news/2019-03-large-hadron-collider-limits.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.