

Providing large-scale disk storage at CERN

Tuesday, 10 July 2018 14:45 (15 minutes)

The CERN IT Storage group operates multiple distributed storage systems and is responsible for the support of the infrastructure to accommodate all CERN storage requirements, from the physics data generated by LHC and non-LHC experiments to the personnel users' files.

EOS is now the key component of the CERN Storage strategy. It allows to operate at high incoming throughput for experiment data-taking while running concurrent complex production work-loads. This high-performance distributed storage provides now more than 250PB of raw disks and it is the key component behind the success of CERNBox, the CERN cloud synchronisation service which allows syncing and sharing files on all major mobile and desktop platforms to provide offline availability to any data stored in the EOS infrastructure.

CERNBox recorded an exponential growth in the last couple of year in terms of files and data stored thanks to its increasing popularity inside CERN users community and thanks to its integration with a multitude of other CERN services (Batch, SWAN, Microsoft Office).

In parallel CASTOR is being simplified and transitioning from an HSM into an archival system, focusing mainly in the long-term data recording of the primary data from the detectors, preparing the road to the next-generation tape archival system, CTA.

The storage services at CERN cover as well the needs of the rest of our community: Ceph as data back-end for the CERN OpenStack infrastructure, NFS services and S3 functionality; AFS for legacy home directory filesystem services and its ongoing phase-out and CVMFS for software distribution.

In this paper we will summarise our experience in supporting all our distributed storage system and the ongoing work in evolving our infrastructure, testing very-dense storage building block (nodes with more than 1PB of raw space) for the challenges waiting ahead.

Primary authors: ROUSSEAU, Herve (CERN); LAMANNA, Massimo (CERN); IVEN, Jan (CERN); ESPINAL CURULL, Xavier (CERN); MASCETTI, Luca (CERN); LO PRESTI, Giuseppe (CERN); MOSCICKI, Jakub (CERN); CHAN KWOK CHEONG, Belinda (CERN); VAN DER STER, Dan (CERN); CONTESCU, Cristian (CERN)

Presenter: ROUSSEAU, Herve (CERN)

Session Classification: T4 - Data handling

Track Classification: Track 4 - Data Handling