



Contribution ID: 84

Type: Oral

## Analysis and modeling of data access patterns in ATLAS and CMS

*Tuesday, 5 November 2019 15:15 (15 minutes)*

Data movement between sites, replication and storage are very expensive operations, in terms of time and resources, for the LHC collaborations, and are expected to be even more so in the future. In this work we derived usage patterns based on traces and logs from the data and workflow management systems of CMS and ATLAS, and simulated the impact of different caching and data lifecycle management approaches. Data corresponding to one year of operation and covering all Grid sites have been the basis for the analysis. For selected sites, this data has been augmented by access logs from the local storage system to also include data accesses not managed via the standard experiments workflow management systems. We present the results of the studies, the tools developed and the experiences with the data analysis frameworks used, and assess the validity of both current and alternative approaches to data management from a cost perspective.

### Consider for promotion

Yes

**Authors:** CMS COLLABORATION AND CERN/IT; SCIABÀ, Andrea (CERN); SCHULZ, Markus (CERN); MATSKOVSKAYA, Victoria

**Presenter:** SCHULZ, Markus (CERN)

**Session Classification:** Track 4 –Data Organisation, Management and Access

**Track Classification:** Track 4 –Data Organisation, Management and Access