

Gaining insight from large data volumes with ease

Tuesday, 10 July 2018 16:45 (15 minutes)

Efficient handling of large data-volumes becomes a necessity in today's world. It is driven by desire to get more insight from the data and to gain a better understanding of user trends which can be transformed into economic incentives (profits, cost-reduction and various optimization of data workflows and pipelines). In this talk we discuss how modern technologies are transforming a well established pattern to handle the data based on relational databases. This trend which was originally introduced in commercial world is adopting at a large pace in HEP. We present a few examples of such systems successfully designed and deployed in CMS to handle millions of Framework Job Report documents from distributed agents as well as efficiently aggregating large data volumes coming from a variety of data providers (Data-Bookkeeping System, Data location system, EOS, AAA, CRAB, WMArchive) and understanding the data trends based on build-in dashboards.

Primary author: KUZNETSOV, Valentin Y (Cornell University (US))

Presenter: KUZNETSOV, Valentin Y (Cornell University (US))

Session Classification: Posters

Track Classification: Track 4 - Data Handling