Contribution ID: **64** Type: **Oral Presentation**

Simulating a network delivery content solution for the CMS experiment in the Spanish WLCG Tiers

Friday, 25 March 2022 13:30 (20 minutes)

The LHC experiments at CERN, the world's largest particle collider, have produced an unprecedented volume of data in the history of modern science since it started operations in 2009. Up to now, more than 1 Exabyte of simulated and real data have been produced, being stored on disk and magnetic support and processed in a worldwide distributed computing infrastructure, comprising 170 centers in 35 countries, known as WLCG (World-wide LHC Computing Grid). LHC operates in yearly periods, characterized by incremental steps in the number of particles that collide, gradually increasing the amount of experimental data to be stored and analyzed. By 2026, the experiment will face the High-Luminosity LHC (HL-LHC) era, where the produced data will increase a factor 10 as compared to today's values. The compute budget is not expected to increase substantially, so the LHC community is exploring novel ideas to integrate into experiment compute models in order to alleviate the expected compute and storage demands in that period. In terms of data management and access one of the strategic directions is to integrate storage caches as network delivery content solutions, and consolidating the main storage systems of WLCG into fewer sites. One of the benefits would be to be able to run the sites with less computing contribution without having its own storage system deployed. This technology would also have an impact over the latency hiding for remote reads and accelerate data delivery to opportunistic compute clusters and Cloud resources. In this contribution we simulate different behaviors and configurations of Least Recently Accessed (LRU) data caches for the CMS experiment in the Spanish CMS region, using real data accesses from both PIC Tier-1 and CIEMAT Tier-2. We expose and discuss the most efficient features and configurations to optimize caches and executed job performances in terms of the most relevant identified metrics.

Primary authors: Dr SIKORA, Anna (Universitat Autònoma de Barcelona (UAB)); PEREZ DENGRA, Carlos (CIEMAT); FLIX, Josep (PIC / CIEMAT)

Presenter: PEREZ DENGRA, Carlos (CIEMAT)

Session Classification: Data Management & Big Data

Track Classification: Track 6: Data Management & Big Data