Contribution ID: **368**                                                                      Type: **Talk**

# Enhancing CMS XCache efficiency: A comparative study of Machine Learning techniques and LRU mechanisms

*Thursday 24 October 2024 13:48 (18 minutes)*

The Large Hadron Collider (LHC) at CERN in Geneva is preparing for a major upgrade that will improve both its accelerator and particle detectors. This strategic move comes in anticipation of a tenfold increase in proton-proton collisions, expected to kick off by 2029 in the upcoming high-luminosity phase. The backbone of this evolution is the World-Wide LHC Computing Grid, crucial for handling the flood of data from these collisions. Therefore, expanding and adapting it is vital to meet the demands of the new phase, all while working within a tight budget. Many research and development projects are in progress to keep future resources manageable and cost-effective in managing the growing data. One area of focus is Content Delivery Network (CDN) techniques, which promise data access and resource use optimization, improving task performance by caching input data close to users. A comprehensive study has been conducted to assess how beneficial it would be to implement data caching for the Compact Muon Solenoid (CMS) experiment. This study, with a focus on Spanish computing facilities, shows that user analysis tasks are the ones that can benefit the most from CDN techniques. As a result, a data cache has been introduced in the region to understand these benefits better. In this contribution, we analyze remote data access from users in Spanish CMS sites to figure out the best size and network connectivity requirements for a data cache serving the whole Spanish region. Exploration of machine learning techniques, along with comparisons to traditional LRU mechanisms, allow for the identification and preservation of frequently accessed datasets within the cache. This approach aims to optimize storage usage efficiently, while prioritizing accessibility to the most popular data.

**Author:**    FLIX MOLINA, Jose (CIEMAT - Centro de Investigaciones Energéticas Medioambientales y Tec. (ES))

**Co-authors:**    Dr SIKORA, Anna (UAB);   DELGADO PERIS, Antonio (CIEMAT - Centro de Investigaciones Energéticas Medioambientales y Tec. (ES));   PEREZ-CALERO YZQUIERDO, Antonio (Centro de Investigaciones Energéticas Medioambientales y Tecnológicas);   RODRIGUEZ-CALONGE, Francisco Javie;   HERNANDEZ, Jose (CIEMAT);   Ms SERRANO SIERRA, Paula (UAB)

**Presenter:**    FLIX MOLINA, Jose (CIEMAT - Centro de Investigaciones Energéticas Medioambientales y Tec. (ES))

**Session Classification:**   Parallel (Track 1)

**Track Classification:**   Track 1 - Data and Metadata Organization, Management and Access