Contribution ID: **10**   Type: **Oral Presentation**

Reinforcement Learning for Smart Caching in the CMS experiment

Tuesday, 23 March 2021 14:00 (20 minutes)

In the near future, High Energy Physics experiments' storage and computing needs will go far above what can be achieved by only scaling current computing models or current infrastructures. Considering the LHC case, for 10 years a federated infrastructure (Worldwide LHC Computing Grid, WLCG) has been successfully developed. Nevertheless, the High Luminosity (HL-LHC) scenario is forcing the WLCG community to dig for innovative solutions. In this landscape, one of the initiatives is the exploitation of Data Lakes as a solution to improve the Data and Storage management. The current Data Lake model foresees data caching to play a central role as a technical solution to reduce the impact of latency and network load. Moreover, even higher efficiency can be achieved through a smart caching algorithm: this motivates the development of an AI-based approach to the caching problem.

In this work, a Reinforcement Learning-based cache model (named QCACHE) is applied in the CMS experiment context. More specifically, we focused our attention on the optimization of both cache performances and cache management costs.

The QCACHE system is based on two distinct Q-Learning (or Deep Q-Learning) agents seeking to find the best action to take given the current state. More explicitly, they try to learn a policy that maximizes the total reward (i.e. hit or miss occurring in a given time span). While the Addition Agent is taking care of all the cache writing requests, clearly the Eviction Agent deals with the decision to keep or to delete files in the cache. We will present an overview of the QCACHE framework with results comparing standard replacement policies using "Real-world" data (i.e. historical data requests aggregation used to predict dataset popularity filtered for Italy and US regions). Moreover, we will show the planned subsequent evolutions of the framework.

Primary authors: TRACOLLI, Mirco (University of Perugia, INFN Perugia and University of Florence); TEDESCHI, Tommaso (University and INFN, Perugia (Italy))

Co-authors: CIANGOTTINI, Diego (INFN Perugia); STORCHI, Loriano (INFN Perugia); Dr SPIGA, daniele (INFN-PG)

Presenter: TEDESCHI, Tommaso (University and INFN, Perugia (Italy))

Session Classification: Data Management & Big Data Session

Track Classification: Data Management & Big Data