



Contribution ID: 429

Type: Oral

Using DODAS as deployment manager for smart caching of CMS data management system

Monday, 11 March 2019 18:20 (20 minutes)

DODAS stands for Dynamic On Demand Analysis Service and is a Platform as a Service toolkit built around several EOSC-hub services designed to instantiate and configure on-demand container-based clusters over public or private Cloud resources. It automates the whole workflow from service provisioning to the configuration and setup of software applications. Therefore, such solution allows to use “any cloud provider”, with almost zero effort. In this talk, we demonstrate how DODAS can be adopted as deployment manager to set up and manage compute resources and services, required to develop an AI solution for smart data caching. The smart caching layer may reduce the operational cost and increase flexibility with respect to regular centrally managed storage of the current CMS computing model. The cache space should be dynamically populated with the most requested data. In addition, clustering such caching systems will allow to operate them as Content Delivery System between data providers and end-users. Moreover a geographically distributed caching layer will be functional also to a data-lake based model, where many satellite computing centers might appear and disappear dynamically. In this context, our strategy is to develop a flexible and automated AI environment for smart management of the content of such clustered cache system. In this contribution we will describe the identified computational phases required for the AI environment implementation, as well as the related DODAS integration. Therefore we will start with the overview of the architecture for the pre-processing step, based on Spark, which has the role to prepare data for a Machine Learning technique. A focus will be given on the automation implemented through DODAS. Then, we will show how to train an AI-based smart cache and how we implemented a training facility managed through DODAS. Finally we provide an overview of the inference system, based on the CMS-TensorFlow as a Service and also deployed as a DODAS service.

Primary authors: SPIGA, Daniele (Universita e INFN, Perugia (IT)); BOCCALI, Tommaso (INFN Sezione di Pisa, Universita' e Scuola Normale Superiore, P); BONACORSI, Daniele (University of Bologna); DONVITO, Giacinto (INFN-Bari); SALOMONI, Davide (Universita e INFN, Bologna (IT)); GAIDO, Luciano (Universita e INFN (IT)); CIANGOTTINI, Diego (INFN, Perugia (IT)); KUZNETSOV, Valentin Y (Cornell University (US)); ANTONACCI, Marica (INFN); TRACOLLI, Mirco (INFN Perugia); DUMA, Doina Cristina (INFN)

Presenter: TRACOLLI, Mirco (INFN Perugia)

Session Classification: Track 1: Computing Technology for Physics Research

Track Classification: Track 1: Computing Technology for Physics Research