Contribution ID: **6**   Type: **not specified**

OPERA-P: An adaptive scheduler for dynamically provisioning Big Data frameworks on-demand

Thursday 8 June 2017 12:15 (25 minutes)

Purpose:

The objective of this project is to optimize Big Data (BD) workload scheduling, using a hybrid framework (dedicated and non-dedicated) that blends the best of both Hadoop YARN and HTCondor worlds in a single analytical environment.

Method:

The proposed OPERA-P, short for OPportunistically, Elastically Resource Allocation and Provisioning scheduler, is a new hybrid BD platform that combines High-Throughput and High-Performance Computing, i.e., HTCondor and Yarn (see Figure 1). By utilizing OPERA-P, an HTCondor opportunistic pool and an Apache Yarn dedicated cluster can collaborate, and we can achieve an enhanced tasks throughput, for the benefits of BD applications, with minimal cost of deployment. This model is very similar to how multiple applications run concurrently on a laptop or smartphone. In that, new threads are spawned, and more resources are asked as they are needed; consequently, the OS arbitrates among all of the requests. In comparison, OPERA-P will represent the OS, by keep spawn new Docker containers among the idle HTCondor workstations (creating an opportunistic container-based cluster on the HTCondor pool) and ensures efficiently provisioning for the Hadoop dedicated cluster on-demand.

Conclusion:

OPERA-P is an enabling technology that can be used to take advantage of leveraging all of the resources within an enterprise or cloud as a single pool of resources, to achieve full flexibility, scalability, and elasticity provisioning on-demand. OPERA-P provides a seamless bridge from the pool of resources available in HTCondor to the YARN tasks that want those resources. In the presentation, we will discuss further our project and the ongoing efforts behind it. Also, we will discuss OPERA-P design, challenges, and the prototype opportunities.

Authors: Mr AWAYSHEH, Feras (USC - CiTIUS); Mr CADERNO, Pablo (USC -CiTIUS)

Co-authors: Dr PENA, Tomas (USC - CiTIUS); Dr CABALERIO, Carlos (USC - CiTIUS)

Presenter: Mr AWAYSHEH, Feras (USC - CiTIUS)

Session Classification: Workshop

Track Classification: HTCondor presentations and tutorials