



Contribution ID: 24

Type: **not specified**

Learning-based Approaches to Estimate Job Wait Time in HTC Datacenters

Tuesday, 13 October 2020 11:40 (20 minutes)

High Throughput Computing (HTC) datacenters are a cornerstone of scientific discoveries in the fields of High Energy Physics and Astroparticles Physics. These datacenters provide thousands of users from dozens scientific collaborations with tens of thousands computing cores and Petabytes of storage. The scheduling algorithm used in such datacenters to handle the millions of (mostly single-core) jobs submitted every month ensures a fair sharing of the computing resources among user groups, but may also cause unpredictably long job wait times for some users. The time a job will wait can be caused by many entangled factors and configuration parameters and is thus very hard to predict. Moreover, batch systems implementing a fair-share scheduling algorithm cannot provide users with any estimation of the job wait time at submission time.

Therefore, in this talk we investigate how learning-based techniques applied to the logs of the batch scheduling system of a large HTC datacenter can be used to get such an estimation of job wait time. After having illustrated the need for users to get an estimation of the time their jobs will wait, we identify some intuitive causes of this wait time based on the analysis of the information found in the batch system logs. Then, we formally analyze the correlation between these intuitive causes and job wait time and propose learning-based estimators of both job wait time and job wait time ranges. We conclude by presenting the obtained preliminary results and thoughts about how to deploy the proposed estimators in production.

Primary authors: Mr GOMBERT, Luc (CC-IN2P3 / CNRS); Mr SUTER, Frederic (CC-IN2P3 / CNRS)

Presenter: Mr SUTER, Frederic (CC-IN2P3 / CNRS)

Session Classification: Computing and Batch Services

Track Classification: Computing & Batch Services