

How Fair is my Fair-Sharing? Exposing Some Hidden Behavior Through Workload Analysis

Wednesday, March 27, 2019 2:25 PM (25 minutes)

Monitoring and analyzing how a workload is processed by a job and resource management system is at the core of the operation of data centers. It allows operators to verify that the operational objectives are satisfied, detect any unexpected and unwanted behavior, and react accordingly to such events. However, the scale and complexity of large workloads composed of millions of jobs executed each month on several thousands of cores, often limit the depth of such analysis. This may lead to overlook some phenomena that, while they are not harmful at the global scale of the system, can be detrimental to a specific class of users.

In this talk, we illustrate such a situation by analyzing the large High Throughput Computing (HTC) workload trace coming from the Computing Center of the National Institute of Nuclear Physics and Particle Physics (CC-IN2P3) which is one of the largest academic computing centers in France. The batch scheduler implements the classical Fair-Share algorithm which ensures that all user groups are fairly provided with an amount of computing resources commensurate to their expressed needs for the year. However, the deeper we analyze this workload's scheduling, especially the waiting times of jobs, the clearer we see a certain degree of unfairness between user groups. We identify some of the root causes of this unfairness and propose a drastic reconfiguration of the quotas and scheduling queues managed by the job and resource management system. This modification aims at being more suited to the characteristics of the workload and at improving the balance across user groups in terms of waiting. We evaluate the impact of this modification through detailed simulations. The obtained results show that it still guarantees the satisfaction of the main operational objectives while significantly improving the quality of service experienced by the formerly unfavored users.

Primary authors: AZEVEDO, Frédéric (CC-IN2P3); KLUSÁČEK, Dalibor (CESNET); SUTER, Frederic (CNRS / CC-IN2P3)

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

Session Classification: Computing & Batch Systems

Track Classification: Computing & Batch Services