Contribution ID: **379** Type: **presentation**

Improving efficiency of analysis jobs in CMS

Tuesday 10 July 2018 14:30 (15 minutes)

Hundreds of physicists analyse data collected by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) using the CMS Remote Analysis builder (CRAB) and the CMS GlideinWMS global pool to exploit the resources of the World LHC Computing Grid. Efficient use of such an extensive and expensive resource is crucial. At the same time the CMS collaboration is committed on minimizing time to insight for every scientist, by pushing for the fewer possible access restrictions to the full data sample and for freedom of choosing the application to run. Supporting such varied workflows while preserving efficient resource usage poses special challenges, like: scheduling of jobs in a multicore/pilot model where several single core jobs with an undefined runtime run inside pilot jobs with a fixed lifetime; balancing usage of every available CPU vs. use of CPU close to the data; avoiding that too many concurrent reads from same storage push jobs into I/O wait mode making CPU cycles go idle; watching over user activity to detect low efficiency workflows and prod them into smarter usage of the resources.

In this paper we report on two complementary approaches adopted in CMS to improve the scheduling efficiency of user analysis jobs: job automatic splitting, and job automatic estimated running time tuning. They both aim at finding an appropriate value for the scheduling runtime, a number that tells how much walltime the user job needs, and it is used during scheduling to fit user's jobs into pilots that have enough lifetime. With the automatic splitting mechanism, an estimation of the runtime of the jobs is performed upfront so that an appropriate value can be estimated for the scheduling runtime. With the automatic time tuning mechanism instead, the scheduling runtime is dynamically modified by analyzing the real runtime of jobs after they finish. We also report on how we used the flexibility of the global computing pool to tune the amount, kind and running locations of jobs allowed to run exploiting remote access to the input data.

We discuss the strategies concepts, details, and operational experiences, highlighting the pros and cons, and we show how such efforts helped improving the computing efficiency in CMS.

Primary authors: BELFORTE, Stefano (Universita e INFN Trieste (IT)); WOLF, Matthias (University of Notre Dame (US)); IVANOV, Todor Trendafilov (University of Sofia (BG)); MASCHERONI, Marco (Univ. of California San Diego (US)); PEREZ-CALERO YZQUIERDO, Antonio (Centro de Investigaciones Energéti cas Medioambientales y Tecno); LETTS, James (Univ. of California San Diego (US)); BALCAS, Justas (California Institute of Technology (US)); WOODARD, Anna Elizabeth (University of Notre Dame (US)); BOCKELMAN, Brian Paul (University of Nebraska Lincoln (US)); DAVILA FOYO, Diego (Autonomous University of Puebla (MX)); CIANGOTTINI, Diego (Universita e INFN, Perugia (IT))

Presenters: IVANOV, Todor Trendafilov (University of Sofia (BG)); HERNANDEZ, Jose (CIEMAT)

Session Classification: T3 - Distributed computing

Track Classification: Track 3 –Distributed computing