20th International Conference on Computing in High Energy and Nuclear Physics (CHEP2013)

Contribution ID: **308** Type: **Oral presentation to parallel session**

PROOF as a Service on the Cloud: a Virtual Analysis Facility based on the CernVM ecosystem

Tuesday, 15 October 2013 16:51 (22 minutes)

PROOF, the Parallel ROOT Facility, is a ROOT-based framework which enables interactive parallelism for event-based tasks on a cluster of computing nodes.

Although PROOF can be used simply from within a ROOT session with no additional requirements, deploying and configuring a PROOF cluster used to be not as straightforward. Recently great efforts have been spent to make the provisioning of generic PROOF analysis facilities with zero configuration, with the added advantages of positively affecting both stability and scalability, making the deployment operations feasible even for the end user.

Since a growing amount of large-scale computing resources are nowadays made available by Cloud providers in a virtualized form, we have developed the Virtual PROOF-based Analysis Facility: a cluster appliance combining the solid CernVM ecosystem and PoD (PROOF on Demand), ready to be deployed on the Cloud and leveraging some peculiar Cloud features such as elasticity.

We will show how this approach is effective both for sysadmins, who will have little or no configuration to do to run it on their Clouds, and for the end users, who are ultimately in full control of their PROOF cluster and can even easily restart it by themselves in the unfortunate event of a major failure. We will also show how elasticity leads to a more optimal and uniform usage of Cloud resources.

Primary author: BERZANO, Dario (CERN)

Co-authors: LESTARIS, Georgios (CERN); GANIS, Gerardo (CERN); CHARALAMPIDIS, Ioannis (Aristotle Univ. of Thessaloniki (GR)); BLOMER, Jakob (CERN); BUNCIC, Predrag (CERN); MEUSEL, Rene (CERN)

Presenter: BERZANO, Dario (CERN)

Session Classification: Distributed Processing and Data Handling A: Infrastructure, Sites, and Virtualization

Track Classification: Distributed Processing and Data Handling A: Infrastructure, Sites, and Virtualization