Summer Student Programme - Project description

Student	Mr. Nils Leif FISCHER
Main supervisor	Dr. Dirk DUELLMANN
Main supervisor e-mail	Dirk.Duellmann@cern.ch
Main supervisor div	IT-ST-AD
2nd supervisor	Mr. Valentin Y KUZNETSOV
2nd supervisor e-mail	vkuznet@gmail.com
2nd supervisor div	EP-UCM
Project ID	16152
Project content (%)	10/0/90 (Physics/Engineering/Computing)

Project proposal

This project will be executed as collaboration between CMS and IT:

The candidate will work on the aggregation and pre-analysis of logs for the CMS WMArchive project using the IT analytics services. The project includes the need to develop regular snapshots of various aggregated metrics coming from CMS Framework job reports (e.g. job efficiency of CMS production agents). The developed analysis system should provide access to number of jobs, events yield and other job related quantities within certain time interval and will enable the joint optimisation of experiment efficiency between experiment and service providers.

Training value

The student will learn about the CMS workflow system and its performance metrics and will get first hand experience in using the hadoop environment for a concrete parallel processing task. The student will be supervised jointly by the IT and CMS responsible (D. Duellmann and V. Kuznetsov).

Computing knowledge

JAVA:Other

Computing knowledge (other)

Hadoop (Map-Reduce), Python