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



Contribution ID: 448

Type: Oral presentation to parallel session

Using the CVMFS for Distributing Data Analysis Applications for the Fermilab Intensity Frontier

Tuesday 15 October 2013 17:47 (22 minutes)

The Cern Virtual File System (CVMFS) provides a technology for efficiently distributing code and application files to large and varied collections of computing resources. The CVMFS model and infrastructure has been used to provide a new, scalable solution to the previously difficult task of application and code distribution for grid computing.

At Fermilab, a new CVMFS based application and code distribution system has been deployed as an alternative for previous central storage systems. This centralized server system has been successfully adopted by the Intensity Frontier experiments (NOvA, g-2, Mu2e, Minerva) to distribute their application code efficiently to both Fermilab grid resources running thousands of concurrent jobs and to individual experiments who develop and run analysis code on their laptops and desktop system. The new system removes many of the issues related to system stability and to I/O bottlenecks previously encountered with other central storage and code distribution systems during large-scale production running.

The Fermilab solution simplifies the process of code and release management for the experiments while simultaneously addressing security concerns related to code integrity and application consistency by providing a single centrally managed and secure server for the experiments to push their updates to. The system has made the previously unthinkable possibility of experiments being able to work transparently both on their laptops and on the grid a reality.

Author: NORMAN, Andrew (Fermilab)

Co-author: Dr LYON, Adam (Fermilab)

Presenter: NORMAN, Andrew (Fermilab)

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

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