

Considerations on using CernVM-FS for dataset sharing within various research communities

Friday, 18 March 2016 09:40 (20 minutes)

Firmly established as a method of software distribution for the Large Hadron Collider (LHC) experiments and many other research communities at Grid sites, the CernVM File System (CernVM-FS) is reaching now a new stage, its advantages starting to be acknowledged by communities activating within and making use of various Cloud computing environments.

As the manipulation of research data within various grid and cloud infrastructures (EGI FedCloud, TWGrid, OpenScience Grid) becomes more important for many communities, their members started to look into the CernVM-FS as a technology that could bring expected benefits.

Also the developers have optimized the technology for access to conditions data and other auxiliary data, and because of the use of standard technologies (http, squid) the CernVM-FS can now be used everywhere (local clusters, grid and cloud environments) and for more than software distribution.

The presentation will give an overview on the status of the EGI CernVM-FS infrastructure developed for the benefit of the non-LHC communities and its integration with other similar infrastructures for consolidated and better software and data access across the globe.

We will explain when CernVM-FS can be used for dataset sharing without losing the main benefits of the technology and then information on how to properly use it will be given.

Pros and cons will be discussed and available use cases from different High Energy Physics (HEP) and non-HEP research communities (i.e. Space, Natural and Life Sciences) will be analysed.

Primary author: Mr CONDURACHE, Catalin (STFC Rutherford Appleton Laboratory)

Presenter: Mr CONDURACHE, Catalin (STFC Rutherford Appleton Laboratory)

Session Classification: Physics & Engineering Session

Track Classification: Physics (including HEP) and Engineering Applications