

Application Characteristics that Best Fit the CernVM FileSystem and the Frontier Distributed Database Caching System

Tuesday, 20 May 2014 14:10 (20 minutes)

The CernVM FileSystem (CVMFS) and the Frontier Distributed Database Caching System (Frontier) are both based on using local http proxy caches provided by grid sites. This approach enables satisfying a high volume of simultaneous read requests from similar grid jobs with a small number of servers. CVMFS is optimized for distributing software, and Frontier is optimized for distributing data stored in databases. Both systems work great for the applications they're designed for, and could be put to good use in more applications, but they're also both subject to limitations. This presentation discusses the application characteristics that work best with these two systems and the characteristics that don't work well. Recent and planned features for both systems are included.

URL(s) for further info

<http://frontier.cern.ch>
<http://cernvm.cern.ch/portal/filesystem>

Wider impact and conclusions

There are many more potential users and applications of CVMFS and Frontier. This presentation will help users in the EGI community decide whether or not these systems will work for their applications, and how best to use them.

Description of work

An http proxy (squid) infrastructure was first deployed on the Worldwide LHC Computing Grid (WLCG) to support CMS and ATLAS loading Conditions/Calibrations data from databases to grid worker nodes with Frontier. Later CVMFS was deployed using the same proxy infrastructure to manage distribution of software. The primary author led the Frontier development & deployment, and has been supporting the deployment of CVMFS. The co-author is a CVMFS developer.

Primary author: DYKSTRA, Dave (Fermi National Accelerator Laboratory)

Co-author: MEUSEL, Rene (CERN)

Presenter: DYKSTRA, Dave (Fermi National Accelerator Laboratory)

Session Classification: EGI services for global software and common data distribution

Track Classification: Integrated e-Infrastructure services (Track Leaders: K. Kurowski, A. Parodi)