



Contribution ID: 319

Type: Oral

Distributing User Code with the CernVM FileSystem

Thursday, November 7, 2019 11:45 AM (15 minutes)

The CernVM FileSystem (CVMFS) is widely used in High Throughput Computing to efficiently distributed experiment code. However, the standard CVMFS publishing tools are designed for a small group of people from each experiment to maintain common software, and the tools don't work well for the majority of users that submit jobs related to each experiment. As a result, most user code, such as code to do specific physics analyses, is still sent with jobs. That process is relatively inefficient, especially when the user code is large. To overcome these limitations, we have built a CVMFS user code publication system. Users submit code as a tarball with their jobs, and the job submission client sends the tarball to one of a pair of user code publishing servers. The code is published in one of a small number of shared CVMFS repositories, and updates are expedited through the standard CVMFS infrastructure to be available on worker nodes worldwide within several minutes. A job wrapper automatically passes the location of the code to the jobs. The user code is automatically deleted from CVMFS after a period of no use. Most of the software for the system is available as a single self-contained open source rpm called `cvdfs-user-pub`, and is available for other deployments.

Consider for promotion

No

Primary author: DYKSTRA, Dave (Fermi National Accelerator Lab. (US))

Co-authors: BOX, Dennis (F); KIM, Hyunwoo (Fermilab); LEVSHINA, Tanya

Presenter: HERNER, Kenneth Richard (Fermi National Accelerator Laboratory (US))

Session Classification: Track 3 – Middleware and Distributed Computing

Track Classification: Track 3 – Middleware and Distributed Computing