

Towards a responsive CernVM-FS architecture

Tuesday, July 10, 2018 4:40 PM (20 minutes)

The CernVM File System (CernVM-FS) provides a scalable and reliable software distribution service implemented as a POSIX read-only filesystem in user space (FUSE). It was originally developed at CERN to assist High Energy Physics (HEP) collaborations in deploying software on the worldwide distributed computing infrastructure for data processing applications. Files are stored remotely as content-addressed blocks on standard web servers and are retrieved and cached on-demand through outgoing HTTP connections only. Repository metadata is recorded in SQLite catalogs, which represent implicit Merkle tree encodings of the repository state. For writing, CernVM-FS follows a publish-subscribe pattern with a single source of new content that is propagated to a large number of readers.

This talk focuses on the work to move the CernVM-FS architecture in the direction of a responsive data distribution system. A new distributed publication backend allows scaling out large publication tasks across multiple machines, reducing the time to publish. For the faster propagation of new published content, the addition of a notification system allows clients to subscribe to messages about changes in the repository and to request new root catalogs as soon as they become available. These developments make CernVM-FS more responsive and are particularly relevant for use cases where a short propagation delay from repository down to individual clients is important, such as using CernVM-FS as an AFS replacement for distributing software stacks. Additionally, they permit the implementation of more complex workflows, with producer-consumer pipelines, as for example in the ALICE analysis trains system.

Primary author: POPESCU, Radu (CERN)

Co-authors: BLOMER, Jakob (CERN); GANIS, Gerardo (CERN)

Presenter: POPESCU, Radu (CERN)

Session Classification: Posters

Track Classification: Track 3 –Distributed computing