



Contribution ID: 93

Type: **Parallel**

Status and Future Perspectives of CernVM-FS

Tuesday, May 22, 2012 2:20 PM (25 minutes)

The CernVM File System (CernVM-FS) is a read-only file system used to access HEP experiment software and conditions data. Files and directories are hosted on standard web servers and mounted in a universal namespace. File data and meta-data are downloaded on demand and locally cached. CernVM-FS has been originally developed to decouple the experiment software from virtual machine hard disk images and to be used as a replacement of the shared software area at Grid sites. Here it allows for the provision of an essentially zero-maintenance software service. CernVM-FS solves the scalability issues of network file systems such as AFS, NFS, or Lustre, which are traditionally used for shared software areas.

Currently, CernVM-FS distributes around 30 million files and directories. It is installed on a large portion of the Worldwide LHC Computing Grid (WLCG) worker nodes supporting the ATLAS and LHCb experiments. In order to scale to the order of 10^5 worker nodes, CernVM-FS uses replicated repository servers and a hierarchy of web caches. Repository replica servers are operated at CERN, BNL, RAL, and ASGC Tier 1 sites. We will report on the lessons learned from the HEP community feedback and the experience from large-scale deployment.

For the server side, we present a new, streamlined and improved toolset to maintain repositories. The new toolset is supposed to reduce the delay for distributing new software releases to less than an hour. It provides parallel preprocessing of files and it introduces “push replication” of updates by means of a replication manager. The simplified repository maintenance also lowers the bar for small collaborations to distribute their software on the Grid.

Finally, we present the roadmap for the further development of CernVM-FS. The roadmap includes Mac OS X support, variable algorithms for file compression and content hashing, as well as a distributed shared memory cache for diskless server farms.

Summary

The CernVM File System (CernVM-FS) provides a scalable, reliable and essentially zero-maintenance software distribution service. It was developed to assist HEP collaborations to deploy their software on the worldwide-distributed computing infrastructure used to run their data processing applications. CernVM-FS is deployed on a wide range of computing resources, ranging from powerful worker nodes at Tier 1 grid sites to simple virtual appliances running on volunteer computers. The key contribution is a new approach to stage updates and changes into the file system, which aims to reduce the delay in distributing a software release to less than an hour. In addition, it significantly reduces the complexity with respect to both required capabilities of the master storage as well as installation and maintenance. We will report on key scalability figures gathered from normal operational in production use cases. Furthermore, we will discuss new requirements for additional features that have been arisen from HEP community feedback and present the road map for the future development of the file system.

Primary author: BLOMER, Jakob (Ludwig-Maximilians-Univ. Muenchen (DE))

Co-authors: HARUTYUNYAN, Artem (CERN); LARSEN, Dag (University of Bergen (NO)); BUNCIC, Predrag (CERN); MEUSEL, Rene (CERN)

Presenter: BLOMER, Jakob (Ludwig-Maximilians-Univ. Muenchen (DE))

Session Classification: Software Engineering, Data Stores and Databases

Track Classification: Software Engineering, Data Stores and Databases (track 5)