



Contribution ID: 95

Type: **Presentation**

## The Cern VM File System as global software delivery fabric

*Wednesday 29 January 2020 15:40 (20 minutes)*

The Cern VM File System (CVMFS) is a service for fast and reliable software distribution on a global scale. It is capable of delivering scientific software onto physical nodes, virtual machines, and HPC clusters by providing POSIX read-only file system access. Files and metadata are downloaded on demand by means of HTTP requests and take advantage of aggressive caching on the client and at intermediate caches. The choice of the HTTP protocol enables the exploitation of standard web servers and web caches, including commercially-provided content delivery networks.

CVMFS was developed to assist the High Energy Physics (HEP) community to run data processing applications on the Worldwide LHC Computing Grid (WLCG). The scale of the deployment for HEP counts more than 1 billion files accessed by 100,000 nodes and cached on 5 replica servers and 400 web caches.

Potential applications of CVMFS, however, are not confined to the HEP world. The recent addition of S3 as data storage backend for CVMFS makes it readily deployable on Amazon Web Services and compatible with the Ceph-provided S3 API. In addition, the specialized DUCC (Daemon that Unpacks Container images into CVMFS) component supports the publication of container images in their extracted form into CVMFS. Such functionality replaces and goes beyond the service provided by container registries (e.g., Docker Hub) as published images are usable by container daemons (e.g., Docker, Singularity) without the need of pulling and unpacking them first.

**Author:** BOCCHI, Enrico (CERN)

**Presenter:** BOCCHI, Enrico (CERN)

**Session Classification:** Fabric and platforms for Global Science

**Track Classification:** User Voice: Novel Applications, Data Science Environments & Open Data