Contribution ID: **62**   Type: **not specified**

CVMFS service evolution and infrastructure improvements

Wednesday, 14 October 2020 09:40 (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 intermediate caches and clients. The choice of the HTTP protocol also enables the exploitation of standard web servers and web caches, including commercially-provided content delivery networks.

CVMFS is widely adopted in the HEP community for the distribution of production software, integration builds, auxiliary datasets, and has recently introduced new capabilities to broaden its scope of application. As a prime example, it implements extensive support for container images with DUCC (Daemon that Unpacks Container images into CVMFS), a specialized component that unpacks container images and publishes their extracted form on a repository, and tight integration with container runtimes, making published container images usable by widely-adopted container platforms (e.g., Singularity, Docker, Kubernetes). Such functionality provides an alternative to traditional container registries (e.g., Docker Hub, GitLab Container Registry) and makes the distribution of container images more efficient by leveraging on file-based deduplication and on-demand caching provided by CVMFS.

CVMFS at CERN has also been subject to several infrastructural updates. The repository storage for Stratum Zero servers is now hosted on the Ceph-based S3 service, which provides a relevant performance improvement with respect to block storage provided via Cinder volumes. Also, content distribution to clients has been made more resilient by deploying dedicated caches for sets of repositories, which greatly reduces the problem of interference across repositories and cache thrashing phenomena.

Primary authors: BOCCHI, Enrico (CERN); FURANO, Fabrizio (CERN); VAN DER STER, Dan (CERN); MOSCI-CKI, Jakub (CERN)

Presenter: BOCCHI, Enrico (CERN)

Session Classification: Wednesday morning

Track Classification: Storage & Filesystems