

# Evolution of CERNBox Sync&Share platform

*Monday, 30 January 2017 16:50 (20 minutes)*

We started CERNBox in 2013 as a small prototype based on a simple NFS storage and one of the initial versions of the Owncloud server. Some 3 months and 300 users later we have had enough of enthusiastic feedback to consider to open the sync&share service at CERN. Since then we witnessed a rapidly growing service in terms of number of accounts, files, transfers and daily accesses. At the same time we have been evolving the architecture of CERNBox in order to cope with new requirements, well beyond traditional sync&share services which are usually focused on office documents. This included not only the increasing performance expectations but also integration of the sync&share capabilities into diverse daily workflows of CERN users: from desktop applications and home directories to scientific data analysis.

Current CERNBox architecture integrates very closely with the EOS backend storage with a built-in support for HTTP-based synchronisation protocol used by Owncloud synchronisation clients. This allows to harmoniously integrate the native EOS storage access, such as filesystem, with the synchronisation layer. In this model the storage is exposed to the end users for direct access and thus it is not solely controlled by the sync and share layer. We have also been evolving the Owncloud web server to take into account such architectural changes.

In this presentation we will describe further evolution of CERNBox. Implementation of sharing directly on the storage, using EOS native access control mechanisms and metadata propagation features, is the next logical step to provide improved user experience. For the internal architecture we investigate a model based on micro-services to get more flexibility to evolve and improve individual functional subsystems on a longer run. To name just few examples, we consider evolution of the synchronisation protocol including file metadata synchronisation and efficiency improvements, especially for high-latency, low bandwidth and unreliable network connections. For web fronted we are revisiting the handling of metadata and pre-processed files, such as image preview, in a large-scale storage environment. The growing scale of operations is calling for efficient methods to detect and debug user problems remotely.

**Primary author:** MOSCICKI, Jakub (CERN)

**Presenter:** MOSCICKI, Jakub (CERN)

**Session Classification:** Technology