



Contribution ID: 15

Type: **not specified**

# EOS namespace on top of a key-value store

*Thursday 2 February 2017 13:55 (20 minutes)*

## EOS namespace on top of a key-value store

CERN has been developing and operating **EOS** as a disk storage solution successfully for over 6 years. The CERN deployment provides 140 PB of storage and more than 1.4 billion replicas distributed over two computer centres. Deployment includes four LHC instances, a shared instance for smaller experiments and since a couple of years ago an instance for individual user data as well. The user instance represents the backbone of the CERNBOX service for file sharing. New use cases like synchronisation and sharing, the planned migration to reduce AFS usage at CERN and the continuous growth has brought EOS to new challenges.

Recent developments include the integration and evaluation of various technologies to do the transition from a single active in-memory namespace to a scale-out implementation distributed over many meta-data servers. The new architecture aims to separate the data from the application logic and user interface code, thus providing flexibility and scalability to the namespace component. The aim of the new design is to address the two main challenges of the current in-memory namespace: reducing the boot-up time of the namespace and removing the dependency between the namespace size and the amount of RAM required to accommodate it. In order to achieve all this, we've developed an in-house solution for the back-end that combines interesting concepts from already existing projects like Redis, RocksDB, XRootD as well as state of the art consensus protocol like Raft.

This presentation details the basic concepts and the technology used to develop the key-value backend and the namespace interface modifications required to accommodate both the old and the new implementations. Furthermore, we explain the necessary steps to configure the new setup and the implications it has on the service availability and deployment process. Last but not least, we present some preliminary performance metrics of the new system that represent the basis for comparison with the request rates that we currently observe in production.

**Authors:** SINDRILARU, Elvin Alin (CERN); BITZES, Georgios (CERN)

**Presenter:** SINDRILARU, Elvin Alin (CERN)

**Session Classification:** EOS Workshop