Contribution ID: **362** Type: **presentation**

# Ceph File System for the CERN HPC Infrastructure

*Tuesday 10 July 2018 15:45 (15 minutes)*

The Ceph File System (CephFS) is a software-defined network filesystem built upon the RADOS object store. In the Jewel and Luminous releases, CephFS was labeled as production ready with horizontally scalable metadata performance. This paper seeks to evaluate that statement in relation to both the HPC and general IT infrastructure needs at CERN. We highlights the key metrics required by four users, including: POSIX compliance, single-user small-file latency, multi-user metadata throughput, and metadata horizontal scalability and failure tolerance. We will report about the experience so far and future evolution of the service.

In addition, we describe a new suite of micro-benchmark which measure the small latencies of buffered and synchronous filesystem operations and can be used to quantify the evolving quality of a storage cluster over time. We also introduce a simple ping-like latency tool—fsping—which evaluates the time needed for two clients to notice file modifications in a shared filesystem. These tests, in combination with several classical HPC benchmarks run in single- and multi-user scenarios, paint a picture of CephFS which is emerging as a viable option for HPC storage and NFS-appliance replacements.

**Primary authors:** VAN DER STER, Dan (CERN); LAMANNA, Massimo (CERN); MOURATIDIS, Theofilos (National and Kapodistrian University of Athens (GR)); ROUSSEAU, Herve (CERN); Dr WIEBALCK, Arne (CERN); LLOPIS SANMILLAN, Pablo (CERN)

**Presenter:** ROUSSEAU, Herve (CERN)

**Session Classification:** T4 - Data handling

**Track Classification:** Track 4 - Data Handling