

21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015)



Contribution ID: 297

Type: **oral presentation**

Integrating CEPH in EOS

Thursday, 16 April 2015 12:00 (15 minutes)

The EOS storage software was designed to cover CERN disk-only storage use cases in the medium-term trading scalability against latency. To cover and prepare for long-term requirements the CERN IT data and storage services group (DSS) is actively conducting R&D and open source contributions to experiment with a next generation storage software based on CEPH.

CEPH provides a scale-out object storage system (RADOS) and additionally various optional high-level services like S3 gateway, RADOS block devices and a POSIX compliant file system (CephFS). The acquisition of CEPH by Redhat underlines the promising role of CEPH as the open source storage platform of the future.

CERN IT is running a CEPH service in the context of OpenStack on a moderate scale of 1 PB replicated storage. Building a 100+PB storage system based on CEPH will require software and hardware tuning. It is of capital importance to demonstrate the feasibility and possibly iron out bottlenecks and blocking issues beforehand.

Main idea behind this R&D is to leverage and contribute to existing building blocks in the CEPH storage stack and implement a few CERN specific requirements in a thin customizable storage layer.

The presentation will introduce various open source developments & contributions, their applicability and first performance figures of a next generation storage platform aka EOS Diamond:

- radosFS - a RADOS based client API providing a scale-out meta data catalog and pseudo-hierarchical storage view with optimized(POSIX-light) directory and file access, parallel meta data queries, modification time propagation, file striping ...
- Erasure Codes - local reconstruction and reed solomon codes for storage cost reduction and improved reliability
- a multi-protocol CEPH Overlay Service based on the XRootD framework providing data localization, multi-user policies, strong authentication, WebDAV/XRootD access ...
- a read-write Federated Storage Cloud platform based on storage virtualization and infrastructure aware scheduling

Primary author: Mr PETERS, Andreas Joachim (CERN)

Co-authors: VAN DER STER, Dan (CERN); PEREIRA ROCHA, Joaquim Manuel (CERN)

Presenter: Mr PETERS, Andreas Joachim (CERN)

Session Classification: Track 3 Session

Track Classification: Track3: Data store and access