Contribution ID: **162** Type: **Oral**

CEPHFS: a new generation storage platform for Australian high energy physics

Monday, October 10, 2016 3:15 PM (15 minutes)

CEPH is a cutting edge, open source, self-healing distributed data storage technology which is exciting both the enterprise and academic worlds. CEPH delivers an object storage layer (RADOS), block storage layer, and file system storage in a single unified system. CEPH object and block storage implementations are widely used in a broad spectrum of enterprise contexts, from dynamic provision of bare block storage to object storage backends of virtual machines images in cloud platforms. The High Energy Particle Physics community has also recognized its potential by deploying CEPH object storage clusters both at the Tier-0 (CERN) and in some Tier-1s, and by developing support for the GRIDFTP and XROOTD (a bespoke HEP) transfer and access protocols. However, the CEPH filesystem (CEPHFS) has not been subject to the same level of interest. CEPHFS layers a distributed POSIX file system over CEPH's RADOS using a cluster of metadata servers dynamically partitioning responsibility for the file system namespace and distributing the metadata workload based on client accesses. It is the less mature CEPH product and has been waiting to be tagged as a production-like product for a long time.

In this paper we present a CEPHFS use case implementation at the Center of Excellence for Particle Physics at the TeraScale (CoEPP). CoEPP operates the Australia Tier-2 for ATLAS and joins experimental and theoretical researchers from the Universities of Adelaide, Melbourne, Sydney and Monash. CEPHFS is used to provide a unique object storage system, deployed on commodity hardware and without single points of failure, used by Australian HEP researchers in the different CoEPP locations to store, process and share data, independent of their geographical location. CEPHFS is also working in combination with a SRM and XROOTD implementation, integrated in ATLAS Data Management operations, and used by HEP researchers for XROOTD or/and POSIX-like access to ATLAS Tier-2 user areas. We will provide details on the architecture, its implementation and tuning, and report performance I/O metrics as experienced by different clients deployed over WAN. We will also explain our plan to collaborate with Red Hat Inc. on extending our current model so that the metadata cluster distribution becomes multi-site aware, such that regions of the namespace can be tied or migrated to metadata servers in different data centers.

In its current status, CoEPP's CEPHFS has already been in operation for almost a year (at the time of the conference). It has proven to be a service that follows the best industry standards at a significantly lower cost and fundamental to promote data sharing and collaboration between Australian HEP researchers.

Tertiary Keyword (Optional)

Secondary Keyword (Optional)

Distributed data handling

Primary Keyword (Mandatory)

Storage systems

Primary authors: BORGES, Goncalo (University of Sydney (AU)); BOLAND, Lucien Philip (University of Melbourne (AU)); Mr CROSBY, Sean (University of Melbourne (AU))

Presenter: BORGES, Goncalo (University of Sydney (AU))

Session Classification: Track 4: Data Handling

Track Classification: Track 4: Data Handling