Contribution ID: **105**   Type: **presentation**

Using a dynamic data federation for running Belle-II simulation applications in a distributed cloud environment

Thursday 12 July 2018 15:00 (15 minutes)

The dynamic data federation software (Dynafed), developed by CERN IT, provides a federated storage cluster on demand using the HTTP protocol with WebDAV extensions. Traditional storage sites which support an experiment can be added to Dynafed without requiring any changes to the site. Dynafed also supports direct access to cloud storage such as S3 and Azure. We report on the usage of Dynafed to support Belle-II production jobs running on a distributed cloud system utilizing clouds across North America. Cloudscheduler, developed by the UVic HEP Research Computing group, federates Openstack, OpenNebula, Amazon, and Microsoft cloud compute resources and provides them as a unified grid site which on average runs about 3500 Belle-II production jobs in parallel. The input data for those jobs is accessible through a single endpoint, our Dynafed instance. This Dynafed instance unifies storage resources provided by Amazon S3, Ceph, and minio object stores as endpoints, as well as storage provided by traditional DPM and dCache sites. We report on our long term experience with this setup, the implementation of a grid-mapfile based X509 authentication/authorization for Belle-II access, and we show how a federated cluster can be used by Belle-II through gfalFS. We also report on the advantages and challenges we encountered in using it to support the Belle-II experiment in a distributed cloud environment.

Author: Dr EBERT, Marcus (University of Victoria)

Co-authors: LEAVETT-BROWN, Colin Roy (University of Victoria (CA)); DRIEMEL, Colson (University of Victoria); FERNANDEZ GALINDO, Fernando (TRIUMF (CA)); BERGHAUS, Frank (University of Victoria (CA)); CASTEELS, Kevin (University of Victoria (CA)); PATERSON, Michael (University of Victoria (CA)); SOBIE, Randy (University of Victoria (CA)); SEUSTER, Rolf (University of Victoria (CA)); TAYLOR, Ryan (University of Victoria (CA))

Presenter: Dr EBERT, Marcus (University of Victoria)

Session Classification: T4 - Data handling

Track Classification: Track 4 - Data Handling