

ALICE DP status

Mihaela Gheata

DPHEP Collaboration Workshop

8-9 June 2015

Work directions

- Making realistic analysis available on the open-access portal
- Deployment of AOD data and of some AOD-based analysis
- Preservation of analysis procedures

Releasing realistic analysis

- Currently the available ALICE open access analysis examples are meant for education only
 - Physics masterclasses and ALICE analysis framework tutorial
- Planning to deploy published analysis, requiring reduced statistics
 - Such as charge particle multiplicity measurements
 - Just planning for now, few steps to do before:
 - Deploy ALICE standard QA analysis train in the public VM
 - Summer student project, understand problems, limitations
 - Improve user interface to allow configuring the analysis
 - Provide software and document all analysis steps

AOD-based analysis

- Deployed ~8TBytes of ESD data so far from 2010 (~10% of data)
 - 2TB of AOD data coming soon
- Most analysis can run on both formats
 - ALICE AOD is a primary format with no filtering, suitable for >90% of analysis
 - No performance studies possible on AOD, just analysis
- No filtered data releases for specific analysis planned so far, but possible in future
 - E.g. muons, rare signals, ...

Preservation of analysis procedures

- All ALICE analysis production history preserved in the **LEGO framework**
 - ALICE analysis train system + job deployment & monitoring + bookkeeping in Monalisa DB
 - Data available for the last 3-4 years
 - No clear analysis metadata before
- Contains detailed metadata on the analysis procedures
 - Software versions, used datasets, processed statistics, ...
 - No analysis documentation in the train system
 - Some effort needed to put it together in a uniform format
- To add a public query API to access this metadata
 - Push data into the analysis preservation system
 - Trying to understand the requirements

Summary

- Working integrated solution
 - VM(CernVM-FS) + OpenAccess portal + dataset access (starting from json metadata) + user interface + documentation
- Data released as planned in the ALICE policy
- Summer student project to be starting soon
- Investigating several issues
 - Improvement of the open access solution
 - Deployment of realistic analysis
 - MC datasets (?) – probably first to deploy MC data directly related to the deployed analysis.