

ALICE Data Preservation

M. Gheata

Program objectives

- Preserve data, software and know-how inside ALICE Collaboration
 - Foundation for the long term DP strategy
 - Continuous effort, DP is taken into account alongside SW evolution (e.g data formats)
- Sharing data and associated software with larger scientific community
 - Sharing additional resources
 - Requires additional storage, distributed computing, person-power
 - Formalizing/simplifying data formats and analysis procedure, accessibility issues, intellectual property issues arising from data/SW usage
- Open access to reduced data sets and SW/documentation to general public
 - Educational and outreach purposes
 - Continuous effort aiming to provide meaningful examples/demonstrators to approach ALICE physics

Strategy and scope approved by the Collaboration, implementation under discussion with the ALICE PB

Analysis reproducibility

- Currently storing the builds of every AliRoot Git tag in CVMFS
- Reproducibility requires an extension of the VM concept
 - To include OS, software framework, condition data, analysis macros and documentation
- Published analysis metadata coming with all versions required to reproduce the original environment
 - Analysis reproducible at any moment of time on any platform
 - Still possible to change analysis parameters or input data

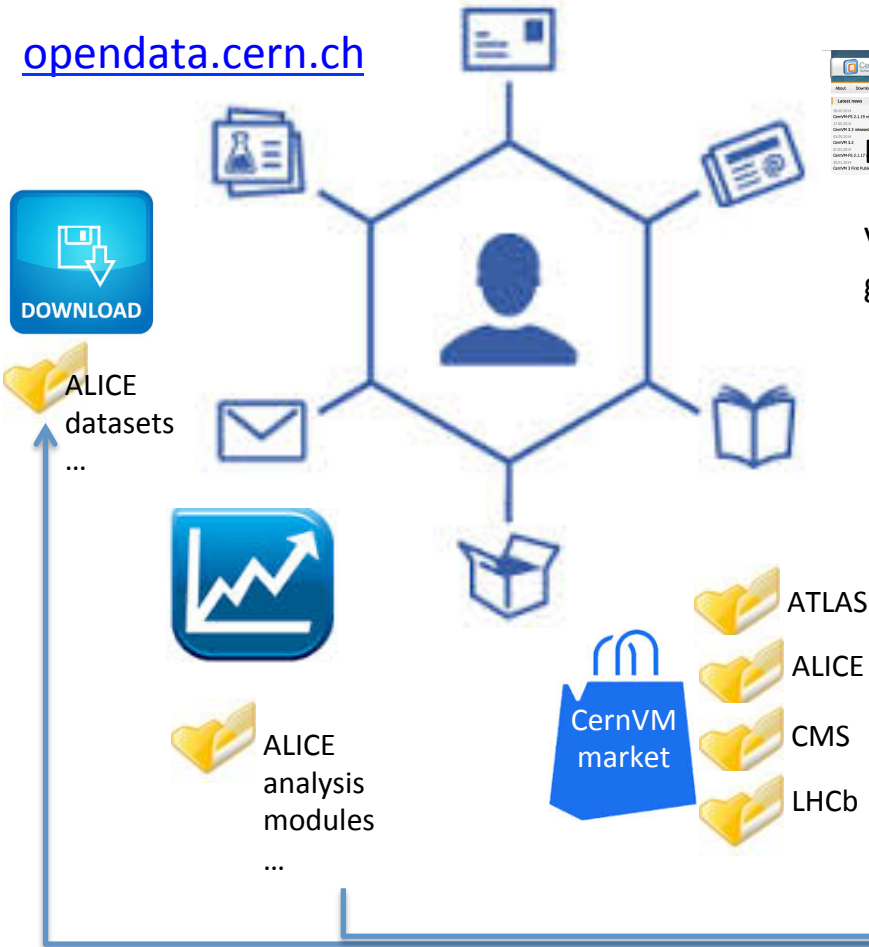
To be exercised first by ALICE collaborators, then gradually exposed via Open Data sharing platforms

Education and Outreach

- First effort to expose the ALICE examples from the CERN Master Class program
 - Simple demonstrator embedded in a CernVM, featuring few simple ALICE physics examples and a basic event display
 - Out of the box basic procedure: no compilation, no environment setting
- Provides access to data, SW tools and needed documentation
 - Very limited data adequately selected from p-p and Pb-Pb collisions in Run1
 - Via OpenAccess portal at CERN
- To be extended to more complex analysis
 - More data to be made available
 - Analysis still limited by the “single box” approach until the system will be parallel computing enabled

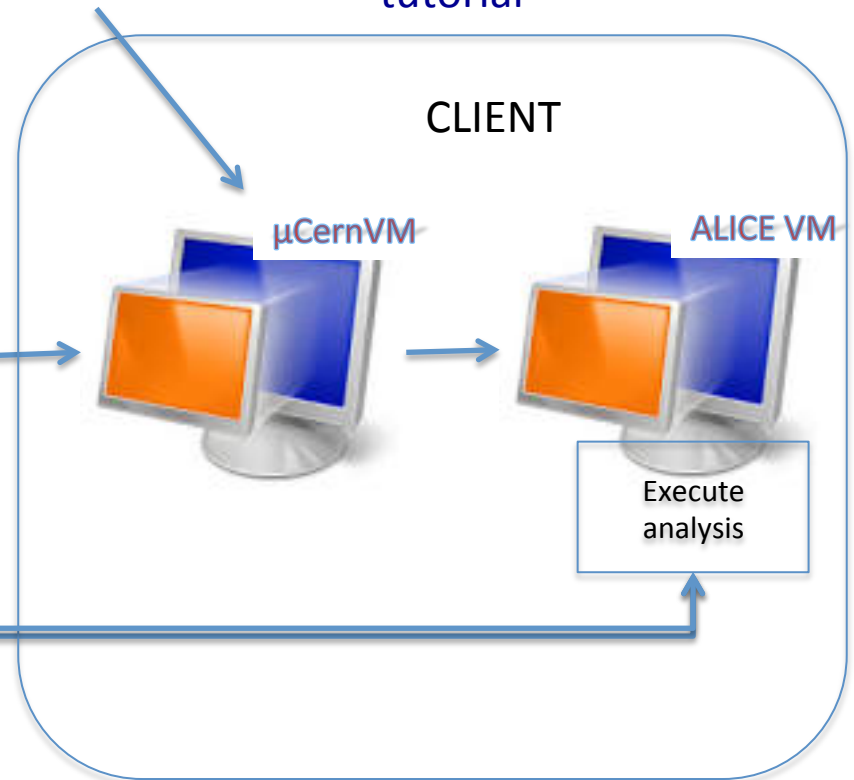
ALICE open access

opendata.cern.ch



VM installation guide

- Provide a custom VM running masterclass examples
- Upgrade in future by adding simple analysis examples
 - Currently the Pt analysis tutorial



Tools required by ALICE SW

- ROOT, AliRoot
- CernVM, CernVMFS
- PYTHIA6, PYTHIA8, DPMJET, PHOJET, ...
- Geant3, Geant4