

---

---

# **Skimming the data, the filtered trees**

**Federica Sozzi (FIAS/GSI)**  
**Marian Ivanov (GSI)**  
**Jacek Otwinowski (IFJ PAN)**

**Offline week**  
**November 2015**

---

---

---

# Filtered Trees

---

**FilterEvents\_Trees.root** files produced during

- AOD processing, physics passes
- LEGO processing (AdHocCalibration train)
- **New**: AOD processing, for cpass1 from LHC15k onwards
- **New**: for MC in next incoming productions (*ATO-289*,  
*C. Ristea*)

Recent improvement is also the merging of the files in the case of AOD processing: one or few files per run, easier to handle (*PWGPP-166 C.Grigoiras*)

---

# Content of Filtered Trees

---

Files contain different root trees with full objects and derived variables:

- **HighPt** ESD tracks, propagated tracks, reference points, MC info
- **V0s** selected V0s (dedicated cuts to reduce combinatorial background), corresponding ESD tracks, reference points
- **CosmicPairs** Cosmic track candidates with reference point
- **dEdx** high dEdx tracks
- **Laser** Laser tracks with reference points
- **MCEffTree** information on all MC generated tracks (e.g. reconstructed or not) to compute efficiency + other MC info

Full information for the interesting events in a reduced size data set :

LHC15f pass2 ESDs size 33 TB

Trees size 35 GB

They allow for **fast analysis** and are the **ideal benchmark** for new implementations and developments

---

# Code

---

## Task

*\$ALICE\_PHYSICS/PWGPP/AliAnalysisTaskFilteredTree.h[.cxx]*

## Cuts definition :

*\$ALICE\_PHYSICS/PWGPP/AliFilteredTreeEventCuts.h[.cxx]*

*\$ALICE\_PHYSICS/PWGPP/AliFilteredTreeAcceptanceCuts.h[.cxx]*

## Configuration

*\$ALICE\_PHYSICS/macros/AddTaskFilteredTree.C*

---

# Example of usage of Filtered Trees

---

**“ExpertQA”** Tracking QA, QA (resolution, efficiency), pt resolution ... already used in the past, at the moment being automated (ATO-102 service task of J. Gronefeld)

**“Splines”** calibration: use V0s to calibrate the Bethe Bloch, used for PID in the actual framework

**“Transfer function”** calibration: dEdx calibration using V0s and cosmics (later it will use also highPt): will be used for PID in future

**RAW event streaming** list of interesting events created from trees, merge corresponding raw data and create ESD. Faster feedback loop O(1 day): reconstruction/new implementation can be tested and compared much faster

**Analysis level:** parameterization of resolution, efficiency, acceptance

---

# Usage at GSI

---

**Download** : Cron jobs at GSI downloading the FilteredTrees once produced, using the script: `$ALIPHYSICS/PWGPP/QA/scripts/alienSync.sh` (by M. Krzewicki)

Input for the script is a query, finding the correct files : e.g.

*`alien_find /alice/data/2015/ SpacePointCalibrationMerge/*/FilterEvents_Trees.root`*

for all 2015 data, AOD processing.

(it would be possible also to use xml files with the lists to avoid the query, if these xml files can be created, e.g. after merging)

**Post-processing:** scripts on the new downloaded files (example next slide)

**Storage:** (input and output) follows naming conventions described in

<https://twiki.cern.ch/twiki/bin/view/ALICE/TPCAlienAccount>

- on the [GSI farm](#)
- on [alien catalog](#) (tpcdrop or pgwpp account). This part is not automated yet

---

# Post processing - PID

---

Script for the post processing:

*`$NOTES/reconstruction/dEdxRegionCalibration/code/processFilteredTree  
s.sh`*

**Input: list of files from different sources.** Files can belong to different paths (LEGO or AOD, chunks or merged runs) and to different periods/productions.

**1- SplitList** : input list is split in different lists for the different productions (according to period, pass, AOD or LEGO type) .

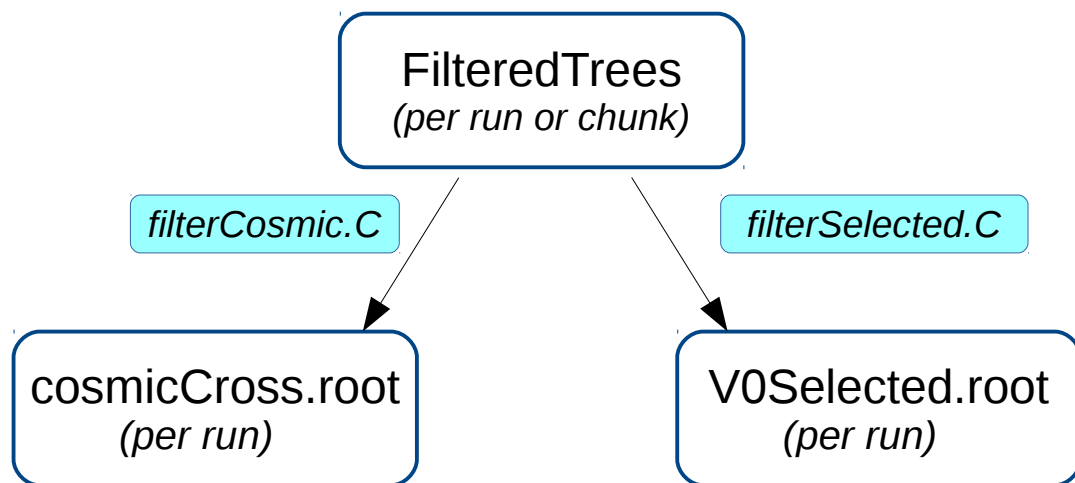
## **2- Processing for every list**

For each list retrieve all the run numbers, check if the file is a merged run or a chunk (from the path) and start the post processing for each run:

In case of a **merged run**, just launch the job on this

In case of **chunks**, consider all the chunks already existing on the GSI farm and launch one job for run

# Post processing - PID



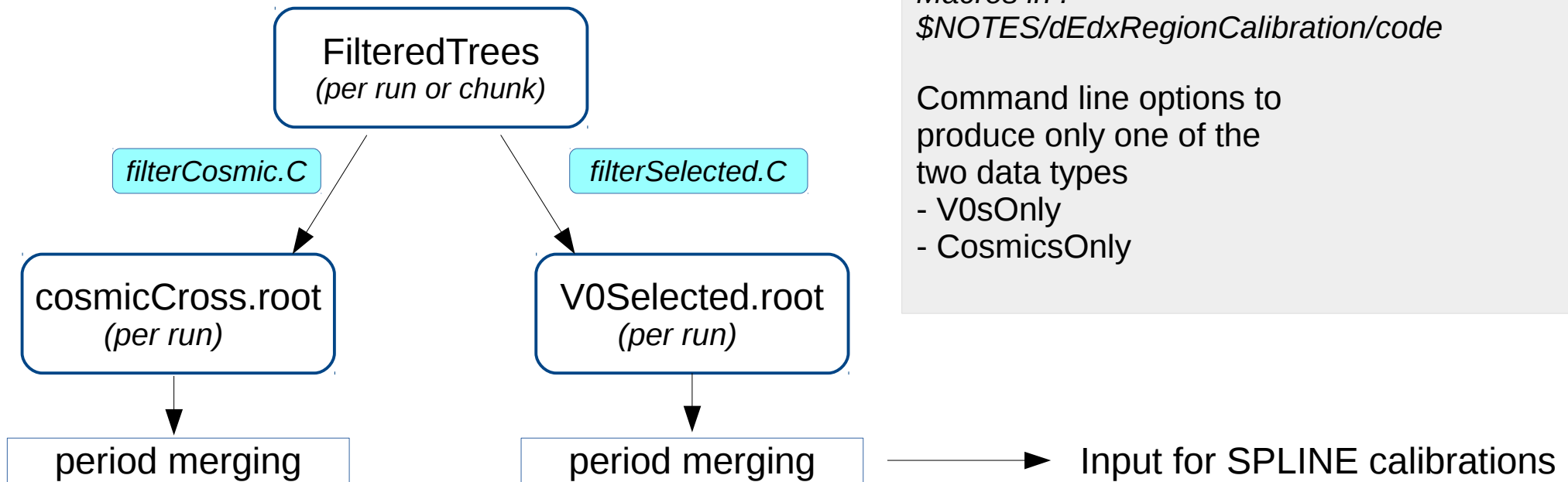
Macros in :  
\$NOTES/dEdxRegionCalibration/code

Command line options to produce only one of the two data types

- V0sOnly
- CosmicsOnly

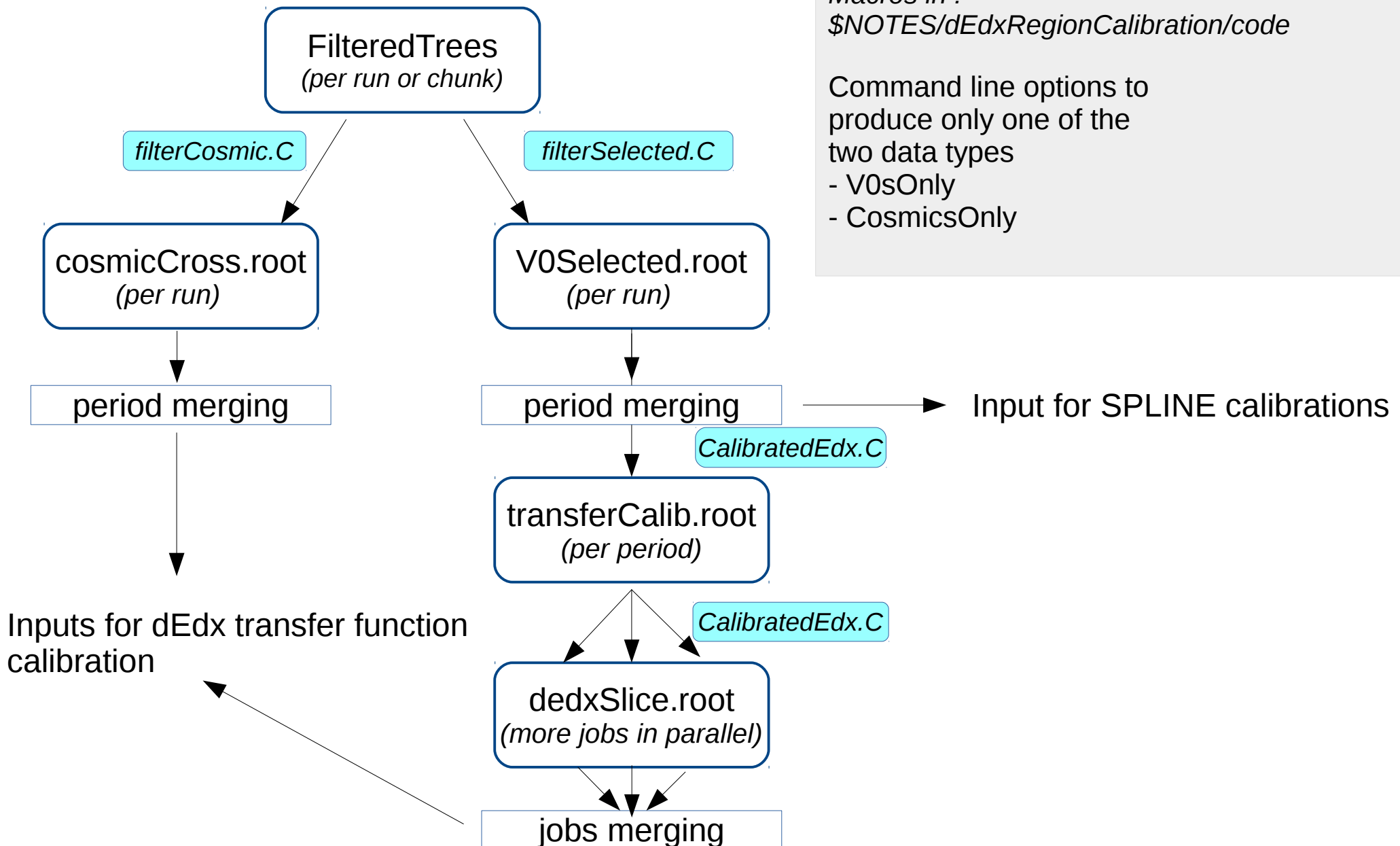


# Post processing - PID



*Merging jobs wait for all the single jobs on the runs to be finished.  
They take into account also the runs already processed  
in previous times*

# Post processing- PID



---

# Post processing- PID

---

- Script is already used
- Further developments going on :
  - Synchronization of data with alien accounts  
(tool to synchronize local dir with alien dir would be useful: further development of AlienSync with Mikolaj or usage of alisync by G.Eulisse )
  - Automatic check of failed jobs and resubmission
  - Clean log of the job outputs

---

# Summary

---

- Filtered trees already produced and used in the past for many usage cases
- They will be produced during AOD processing and also in MC and cpass1 from now onwards
- Post processing for different purposes is being automated (PID, expertQA)

# C.Zampolli, 9 Nov Computing board

## Calibration workflow for Run2 PbPb

