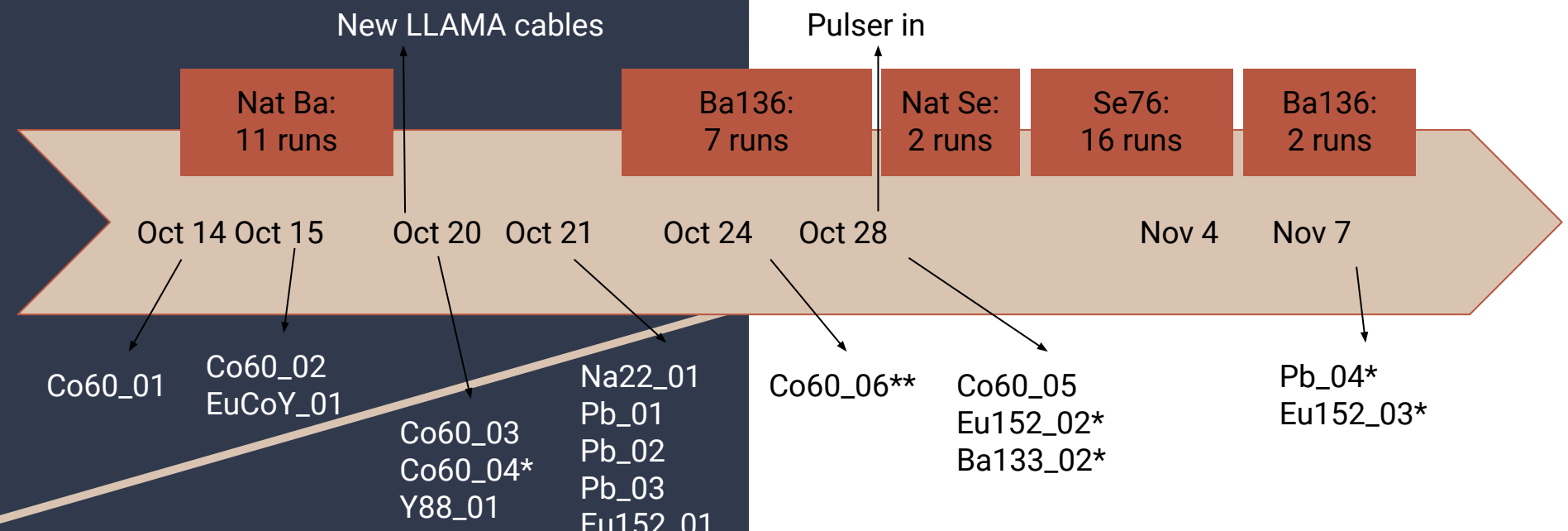
Overview of LLAMA data structure and analysis

Elisabetta Bossio (TUM)
OMC4DBD Collaboration Meeting Online
26th April 2022

Data timeline

Runs do not have same length!



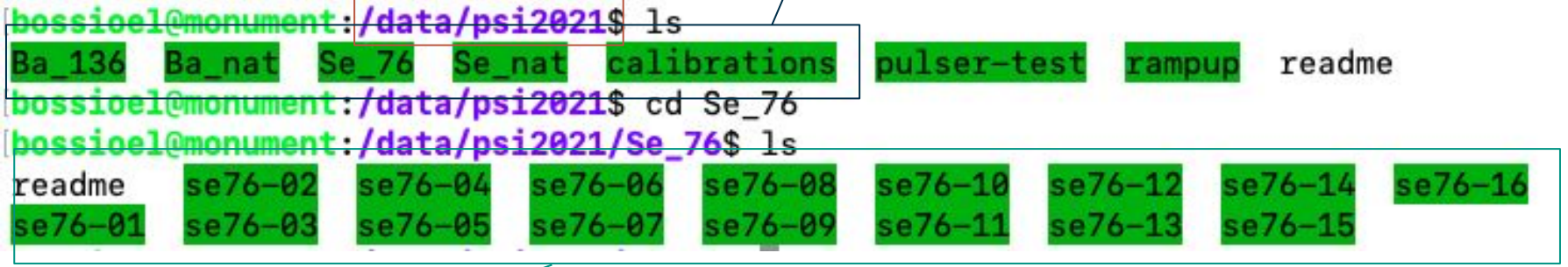
*TUM pulser used

**calibration source directly below detector 9

Data @ TUM-server

Data location for beam time 2021

Data divided by target/calibration source



Multiple runs for each target/calibration source: run length highly non-uniform :)

Data scheme

```
├── ba136-07
│   ├── log
│   ├── raw
│   ├── tier1
│   └── tier2
├── ba136-08
│   ├── log
│   ├── raw
│   ├── tier1
│   └── tier2
├── ba136-09
│   ├── log
│   ├── raw
│   ├── tier1
│   └── tier2
```

*Size and length (in time) of files highly non-uniform :)

- **Raw**: binary files*

Example:

ba136-20211022-000600.bin

Data type

Timestamp

File key: same at all levels

- **Tier1**: root files

Example:

ba136-20211022-000600.tier1.ge.root
ba136-20211022-000600.tier1.pmt.root

Detector type

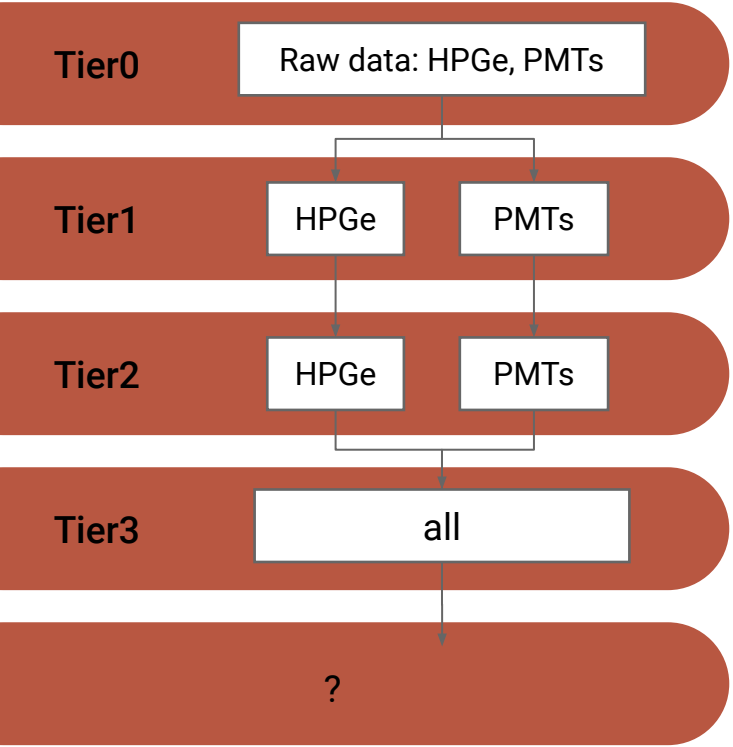
Tier level

- **Tier2**: root files

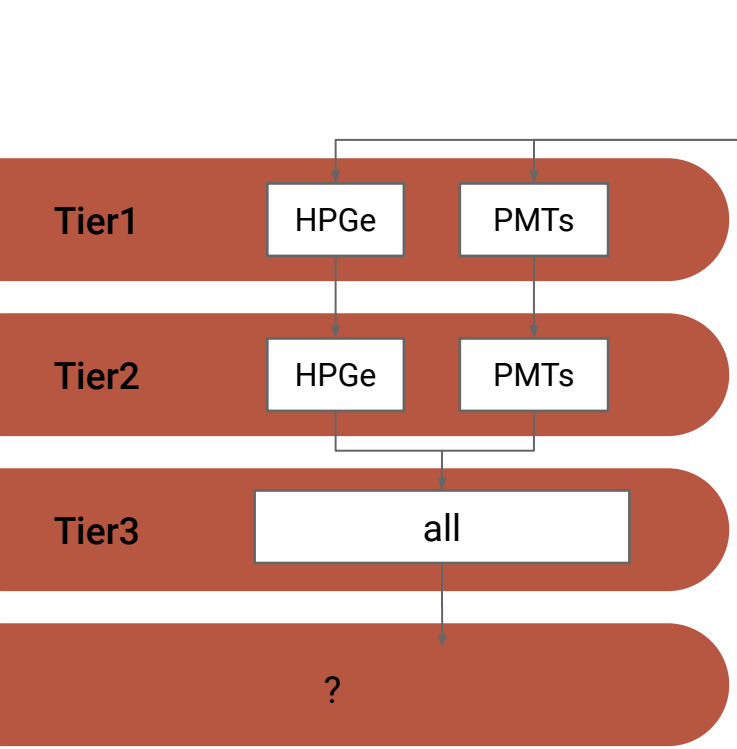
Example:

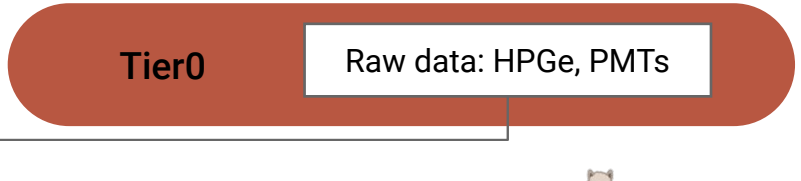
ba136-20211022-000600.ge.tier2.root
ba136-20211022-000600.pmt.tier2.root

Multi-level data structure *á la GERDA*



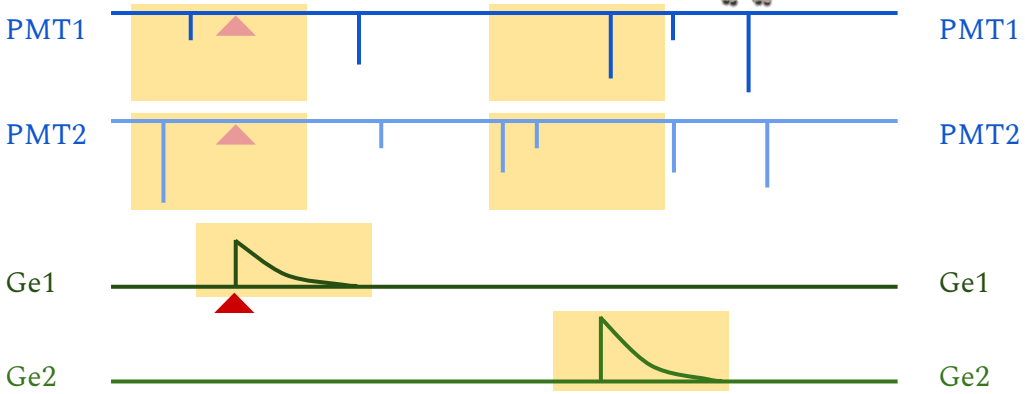
Multi-level data structure *á la GERDA*



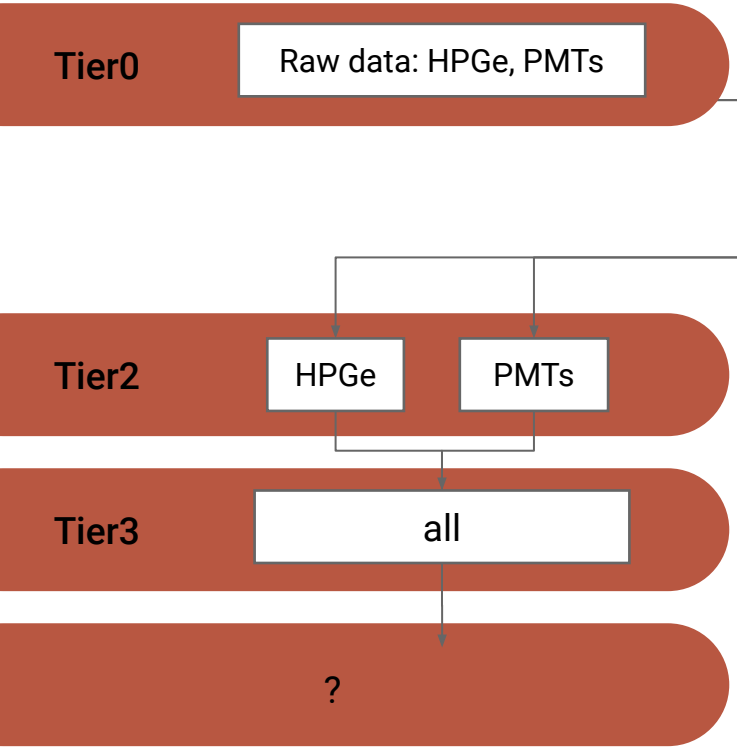


- Raw data from LLAMA-DAQ

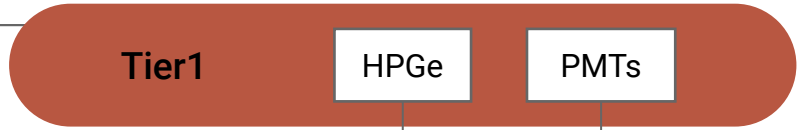
Mario's talk!



Multi-level data structure *á la GERDA*

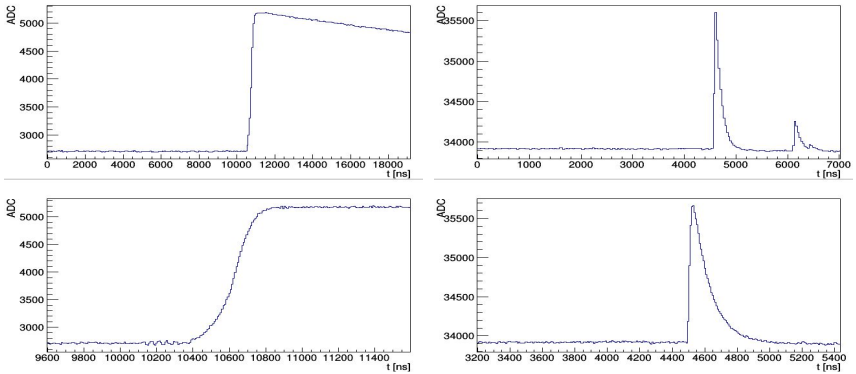


"Rootification"

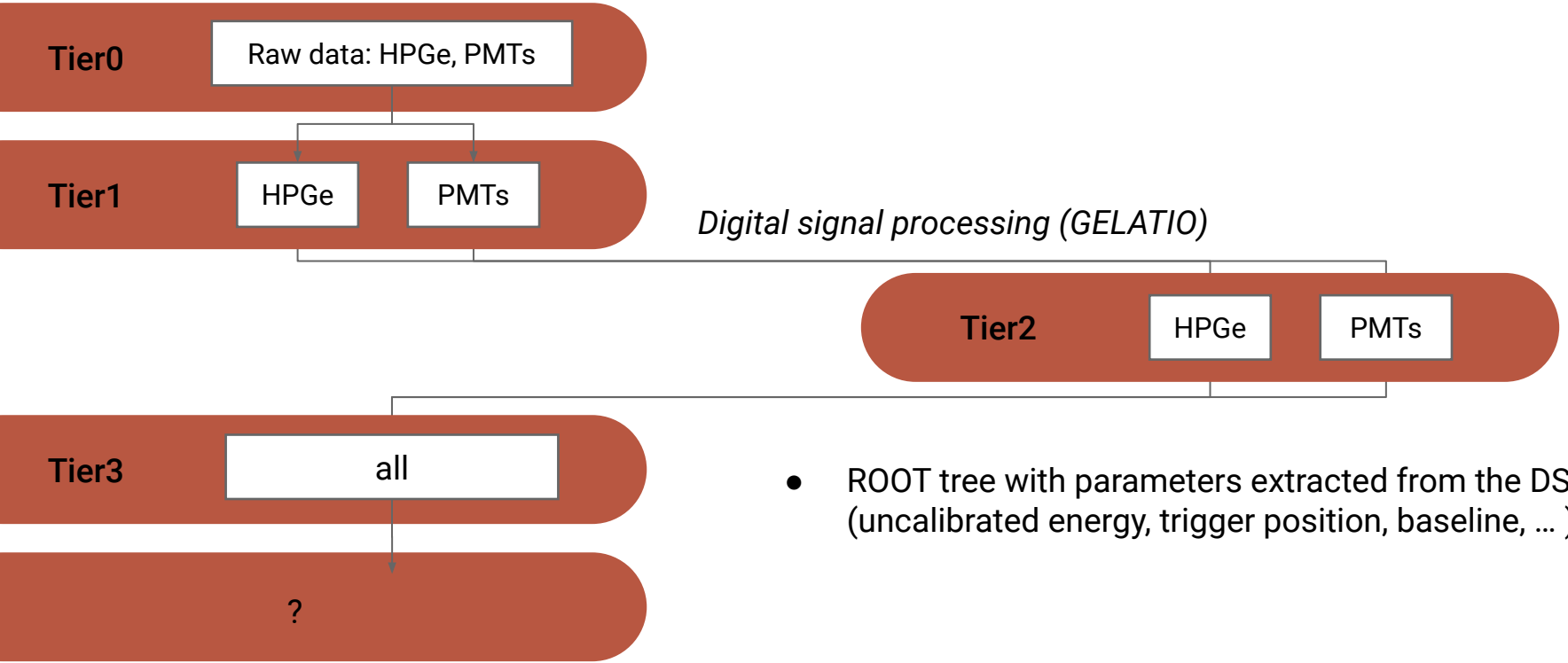


[J.Phys.Conf.Ser. 375 (2012) 042027]

- Waveforms in ROOT format (use the MGDO libraries)
- "Almost" same content as tier0 (data cleaning)

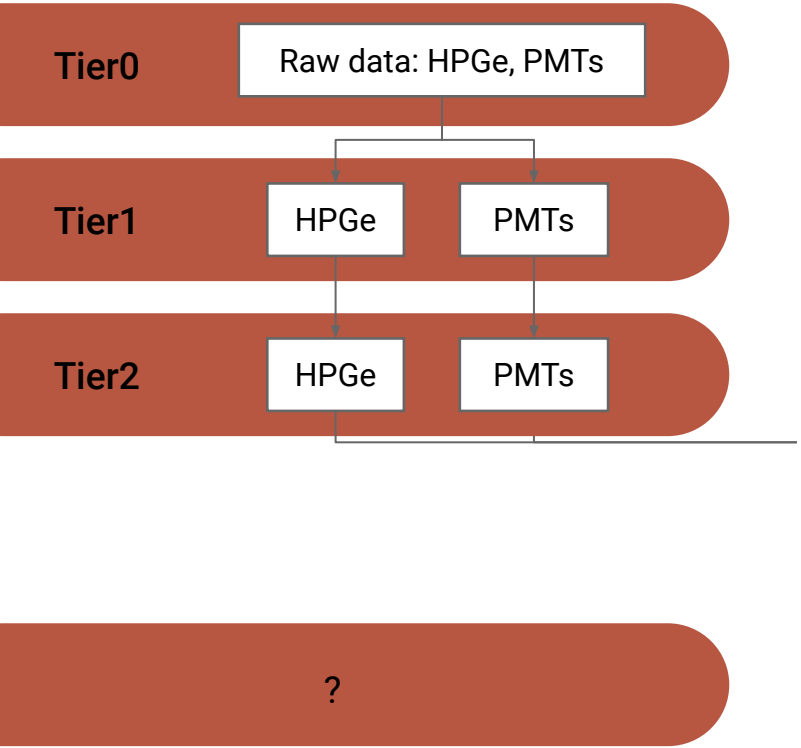


Multi-level data structure *á la GERDA*



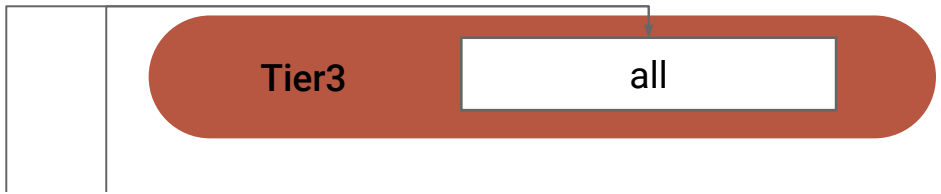
- ROOT tree with parameters extracted from the DSP (uncalibrated energy, trigger position, baseline, …)

Multi-level data structure *á la GERDA*



Dhanurdhar and Eli's talks!

Energy calibration and Quality cuts



- ROOT tree with subset of parameters (calibrated energy, quality cut flags, …)
- Integration of information from different HPGe and PMTs (multiplicity, HPGe-PMT time difference, …)

Data cleaning

- First selection of data for analysis based on run logs
- Main reason for data exclusion:
  - Detectors refilling
  - Beam tuning
  - Beam off/ beam down
  - Changes in the set up
- No tier1/tier2 are produced for excluded data (raw data still available)
- Lists of files to be used for analysis accessible via meta-data
- All documented on GitLab
  https://git.e15.ph.tum.de/monumentum/meta-data/-/merge_requests?scope=all&state=merged

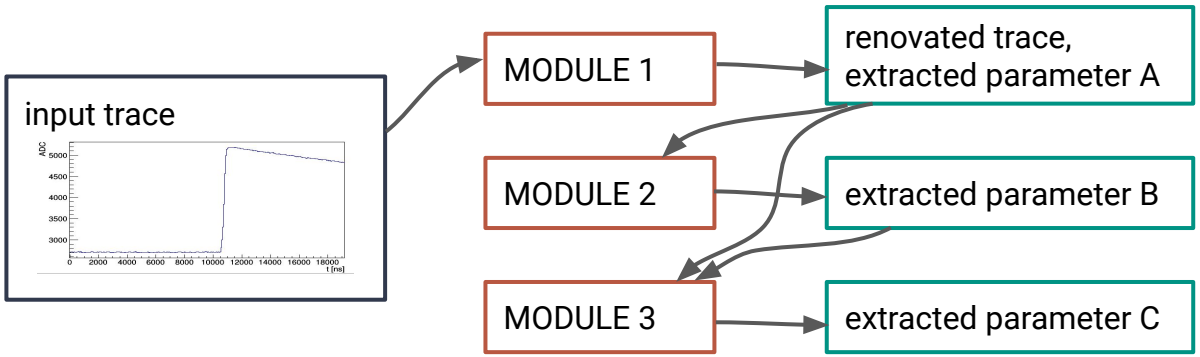
Lesson learnt for upcoming beam times!

Digital Signal Processing (DSP) with *GELATIO*

JINST 6 (2011) P08013

GELATIO (**GE**rda **LA**you**T** for **I**nput/**O**utput)

- Modular approach: each module implements a unique task of the DSP
- Output information can be used as input for other modules
- Chains of modules can be created, input parameters set via ASCII files (ini files)



Tier2 content

HPGe detectors

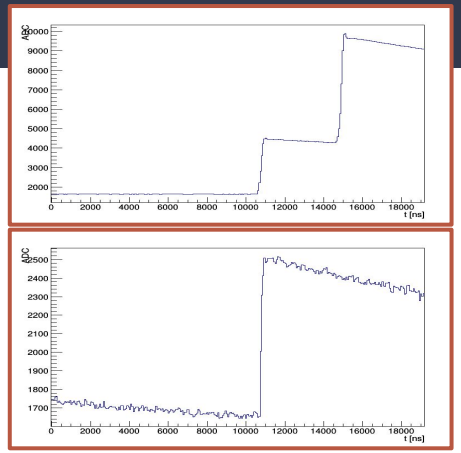
- **Trigger position**, number of triggers in the trace (**GEMDFTTrigger** and GEMDTrigger)
- Baseline position, baseline slope, baseline RMS (GEMDBaseline)
- Auxiliar baseline in different region of the trace to be used in quality cuts
- Minimum and maximum value of the trace (GEMDMinMaxFinder)
- **Energy**: Gaus filter and trapezoidal filter (GEMDEnergyGauss, **GEMDEnergyGast**), plus a shorter trapezoidal filter to be considered for pile-up events (GEMDEnergyGast_pileup)
- Rise time (GEMDRiseTime)
- Trigger and Rise Time for HF trace

PMTs

- **Trigger position**, number of triggers in the trace
- Amplitude of the signal: maximum amplitude and integral
- Since the PMTs signals are quite slow, the same variables are calculated also after shaping the trace with a MW deconvolution and a MW average: we expect to get better precision
- Same story for HF trace

https://docs.google.com/presentation/d/1kQcodmBLOilgSWmfPSSJ159s71B1BiTIYYESBOt8FS4/edit?usp=sharing

Towards tier3



- Main analysis issue: **pile-up** due to very high event rate (~50% of the data)
- We defined a strategy: separate between two data sets:

Golden data set

➔ Stringent requirement on the slope of the baseline to cut all pile-up events
➔ About 50% of the statistic
➔ Optimal performances: energy resolution and time resolution
➔ Ready to be analysed (ready to move to tier3 with calibration and quality cuts inputs)

Silver data set (Pile-up)

➔ Contains all the pile-up events cut from the GOLDEN data set
➔ Recover a lot of statistics
➔ Worse performances expected
➔ Not ready yet: need to find a proper pile-up correction method

Dhanurdhar and Eli's talks!

Meta-data

Collection of files to be used as input for different steps of the data processing and analysis:

Tier0 –> Tier1

- List of files:

  “analysis” list to be used for all analysis steps

```
bossioel@monument:~/meta-data/psi2021/Ba_136$ ls
ba136-01-allFiles.txt  ba136-04-allFiles.txt  ba13
ba136-01-analysis.txt  ba136-04-analysis.txt  ba13
ba136-02-allFiles.txt  ba136-05-allFiles.txt  ba13
ba136-02-analysis.txt  ba136-05-analysis.txt  ba13
ba136-03-allFiles.txt  ba136-06-allFiles.txt  ba13
ba136-03-analysis.txt  ba136-06-analysis.txt  ba13
```

Tier1 –> Tier2

- Ini files for tier2 production:

  GeDet.ini & PMT.ini

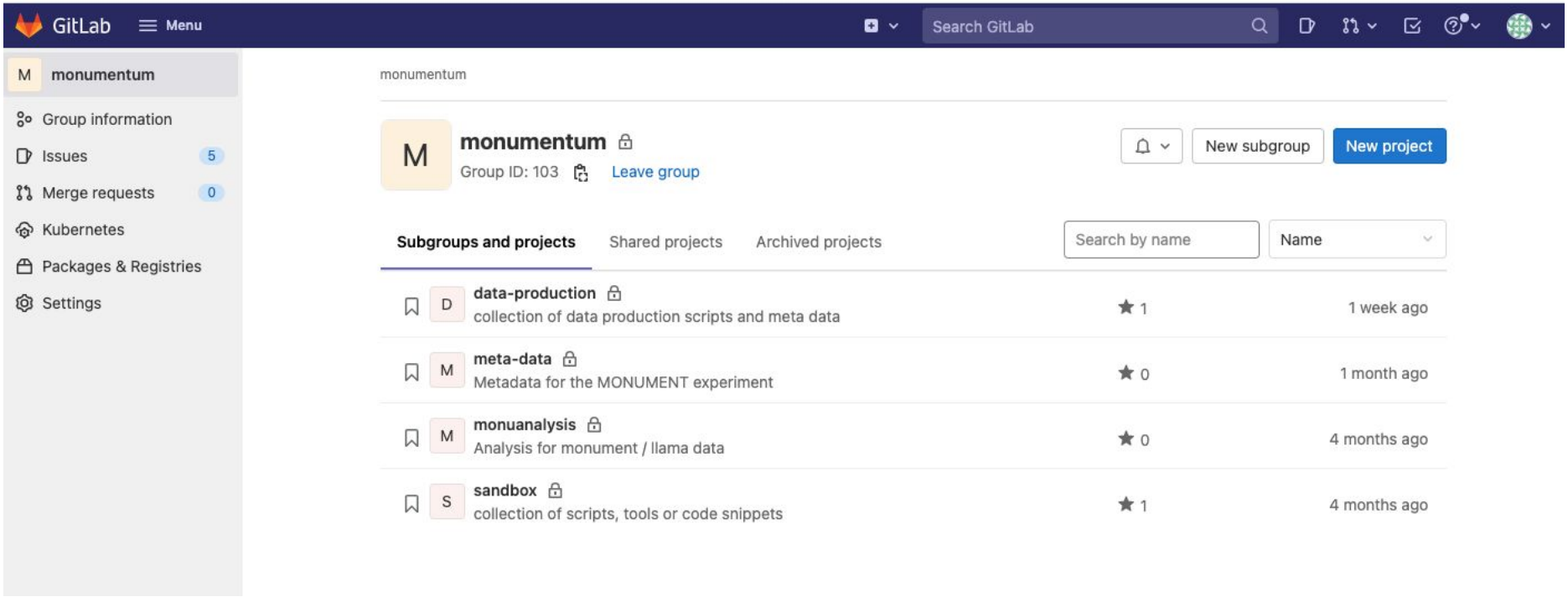
Tier2 –> Tier3

- Results of calibration / quality cut definition for tier3 production:

  To be implemented!

- …

GitLab project @ TUM



Summarizing...

- LLAMA data are available on the TUM server, organized in a multi-level data structure
- All data are available at the tier2 level: energy and time variables can be easily plotted
- With the inputs from calibration and quality cuts we are ready to move to the tier3 level for the GOLDEN data set
- The “new” LLAMA DAQ can count on well known tools and the previous successful experience with GERDA data
- We still need to think about the SILVER data set (in other words, how to handle pile-up)
- We have few lessons learnt for the upcoming beam times :)