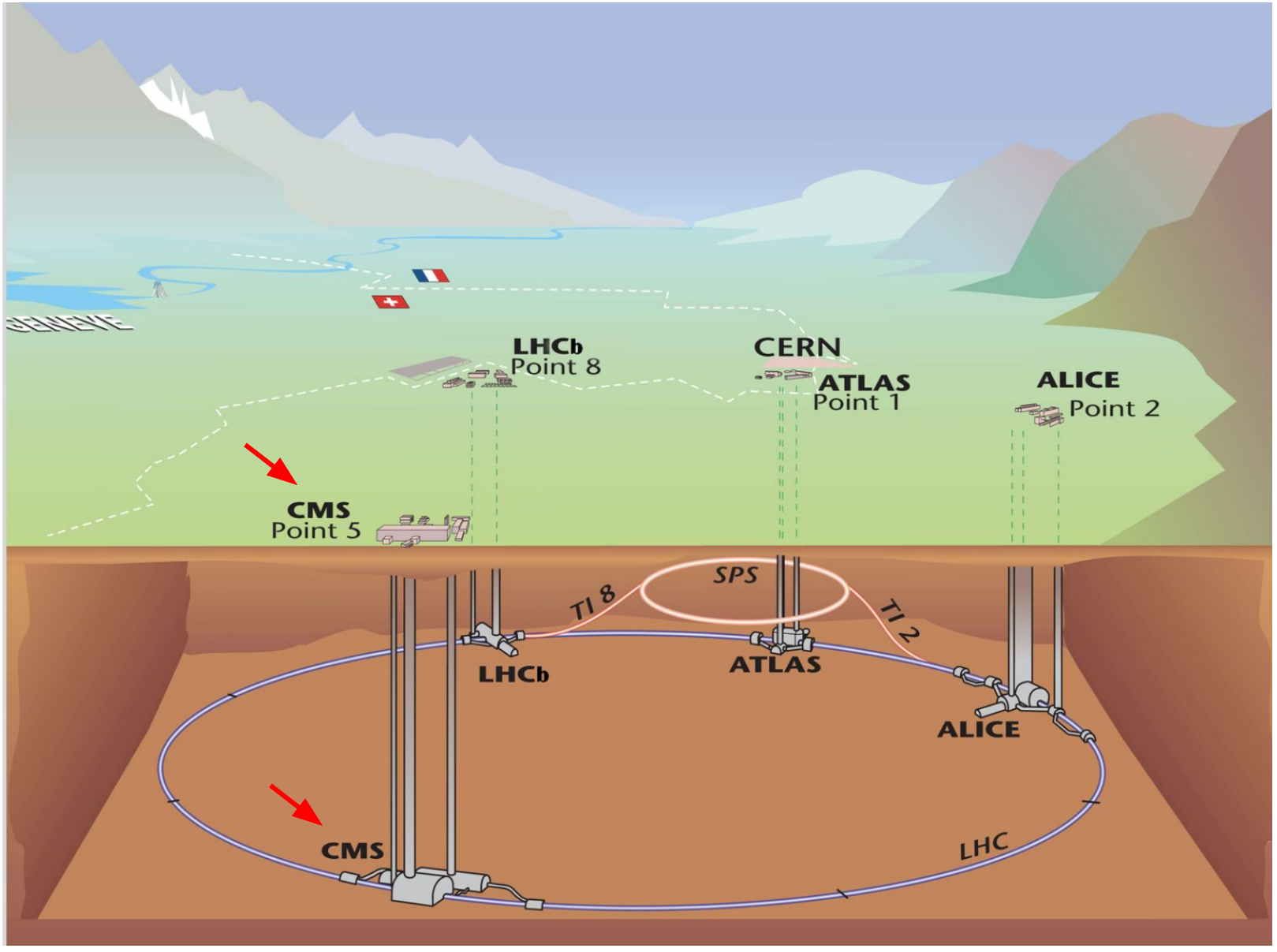
Physics Performance and Datasets(PPD)

Coordination area

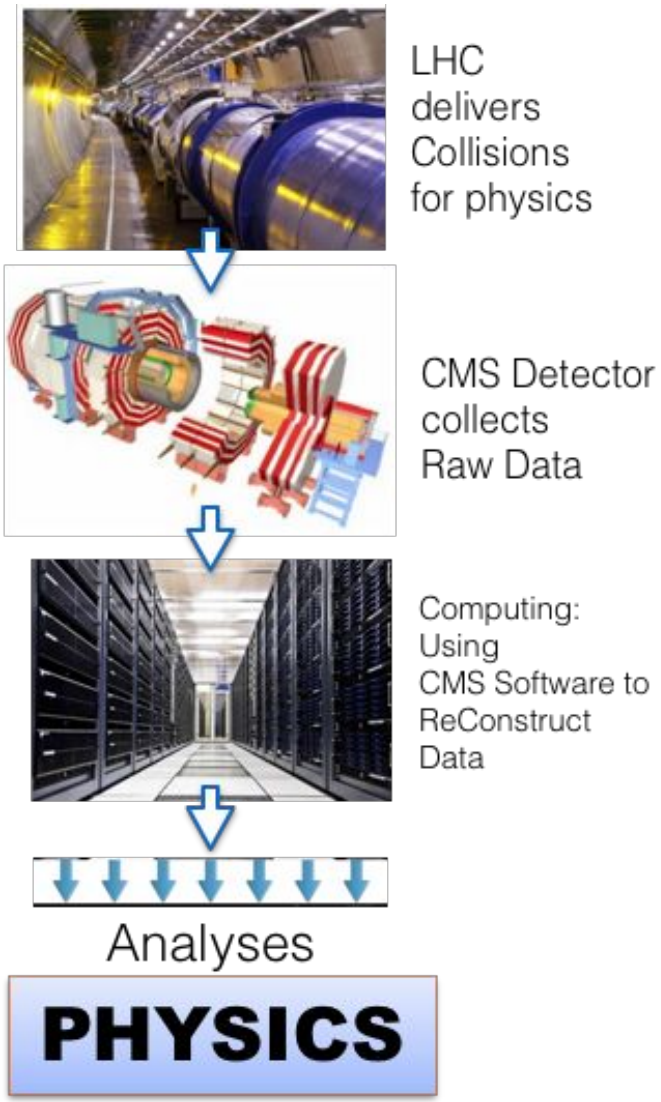
and

DQM-DC Group

CERN, LHC and Experiments

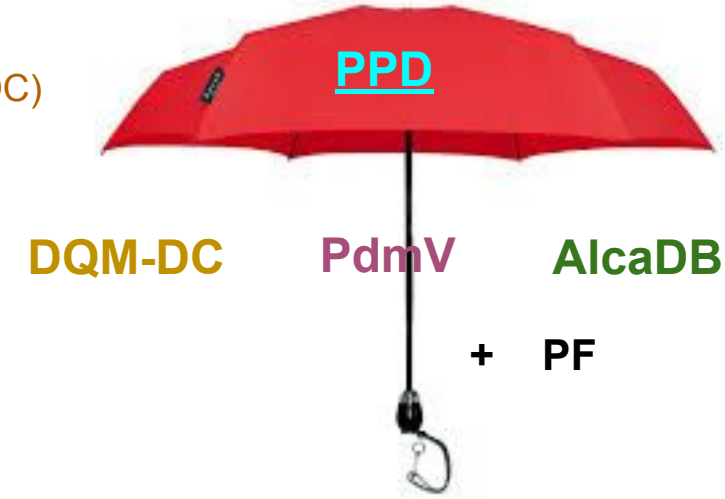


LHC and CMS Experiment
From P5 (online) → offline → Physics



What is PPD?

- The go between for what happens at the “P5 Data Taking” and what happens in “Computing/Physics”.
- **Lives in the both online/offline world**
  - Online is dealing with RAW data and dealing with information coming off the detector
    - Collect high quality data as efficient as possible
  - Offline is dealing with:
    - (re)processing of data and producing simulated events
    - Certification of the processed data
  - For both online/offine
    - **Provide alignment/calibration constants**
    - Validate software used by HLT/DQM/T0 and more
- To accomplish above there are 4 groups:
  - Data Quality Monitoring & Data Certification (DQM & DC)
  - Physics Data & Monte Carlo Validation (PdmV)
  - **Alignment Calibration & Database (AlCaDB)**
  - Particle Flow (PF) shared with PC



(DQM-DC)

Data Quality Monitoring & Data Certification Group

https://twiki.cern.ch/twiki/bin/view/CMS/DQM

DQM-DC group is responsible for

- the online data quality monitoring operation, including the 24/7 DQM shift operation
- Develp/maintain/support the DQM tools' (CMSSW(DQM part), DQM GUI, RunRegistry, DQM^2, HDQM, ….)
- Coordinating/Delivering the central Data Certification

DQM Tools: DQM Gui, Run Registry (RR)

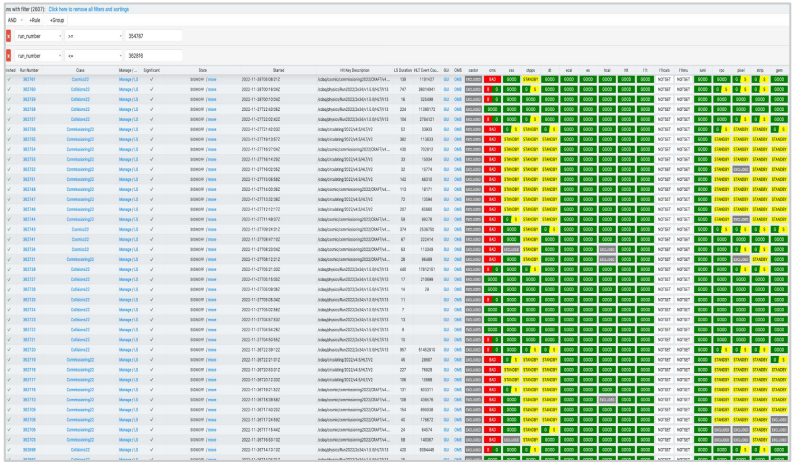
DQM GUI is a tool to display DQM plots produced by RECO (or any CMSSW workflow). Run Registry is a tool to book-keep Data Quality information. They are used both in online and offline environment

- **Online**: process events selected by HLT to display variables in the control room w/ very low latency
  - Live monitoring of detector performance during data taking (DQM GUI, RR)

- **Offline**: process all events while they are simulated or reconstructed and fill diagnostic plots for detailed monitoring of the performance
  - Data certification (DQM GUI, RR)
  - Validation+verification (DQM GUI)

DQM GUI

Run Registry (RR)



Data Certification (DC)

Certification of data carries out by teams of experts in detector & physics objects. The central DC team coordinate the certification experts and produce JSON files (indicating "good data" for physics analyses)

- **Golden**: require all sub-detectors/POGs to be "GOOD"
- **Muon-only**: no requirements on calorimeters
- **DCS-only**: require only tracker to be read-out & powered

The format of file is in *JSON format*

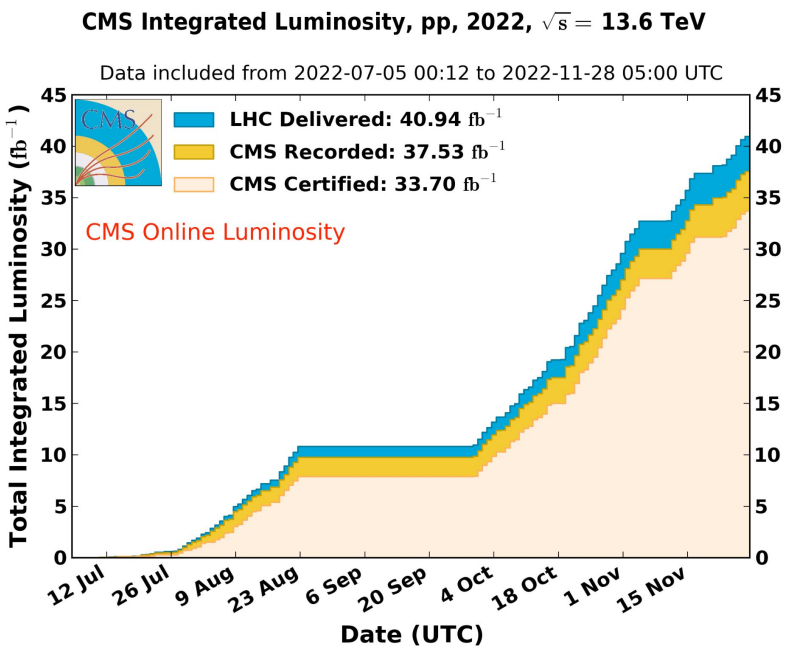
- Weekly for PromptReco
- After each major reprocessing

Ongoing development work:

- NanoDQMiO (ls by ls certification)
- ML based monitoring/certification
- Completing New GUI

Recent DQM-DC Tutorial
About JSON files, & NanoDQMIO

Golden JSON from the complete 2022 PP Data




```
{"273158": [[1, 1279]], "273302": [[1, 45
"273408": [[1, 6]], "273409": [[1, 309]],
113], [115, 412]], "273448": [[1, 391]],
"273493": [[1, 233]], "273494": [[1, 192]
[[1, 173]], "273725": [[83, 252], [254, 2
"274159": [[1, 43]], "274160": [[1, 207]]
"274240": [[1, 40], [42, 82]], "274241":
```


Run

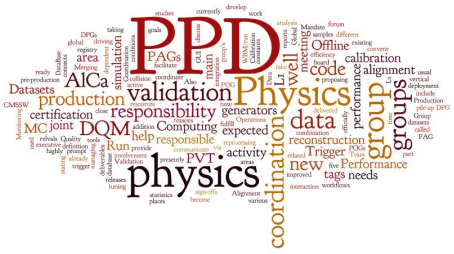
Range of lumisection

DC Results: [Run2] [Run3]

Additional slides

PPD Organization

14-Dec-2022



PPD Coordination
Kaori Maeshima
Jordan Martins
L1

Resource Manager
Kirsti Aspola

Database Officer
Giacomo Govi

PC

Physics Data & MC Validation (PdmV)
Bugra Bilin
Kirill Skovpen
Sunil Bansal
L2

Data Quality Monitoring & Certification (DQM-DC)
Rosamaria Venditti
Emanuele Usai
Shin-Shan "Eiko" Yu
L2

Alignment Calibration & Database (AlCaDB)
Francesco Brivio
Helena Malbouisson
Tamas Vami
Saumya Saumya
L2

Particle Flow (PF)
Kenneth Long
Juska Pekkanen
L2

L3

DPG, POG and PAG Validation
Nadeesha Wickramage, JinFeng Liu
Anup Kumar Sikdar,
Meena Meena
Ansuree Vijay

Monte Carlo Request Management
Ozgun Kara, Andrey Pozdnyakov,
Manuel Sommerhalder,
Samadhan Kamble,
Sadhana Verma

Development
Geovanny A. González-Rodríguez

Dataset Definition Team
You-Ying Li

L3

Software Development & Support
Michel Succar Medina

Data Certification
Geetanjali Chaudhary
Aliya Nigamova

Online shift manager
Adrienne (Dee) Hahn

Online operation manager:
Petr Mandrik

ML infrastructure
Luka Lambrecht
Petr Mandrik

L3

Software Coordinator
Yuan Chao
Chris Misan

AlCa-TSG Responsible
Purbita Prova
Mohamed Darwish

AlCa prompt condition validation
Pritam Kalbhor
Matheus Macedo

Prompt Calibration Loop Manager
Tongguang Cheng
Tapio Lampen

CondDB webmaster
Andres Cardenas

Contacts

Joint PF-JME-BTV ALCA contact
OPEN

TSG contact
OPEN

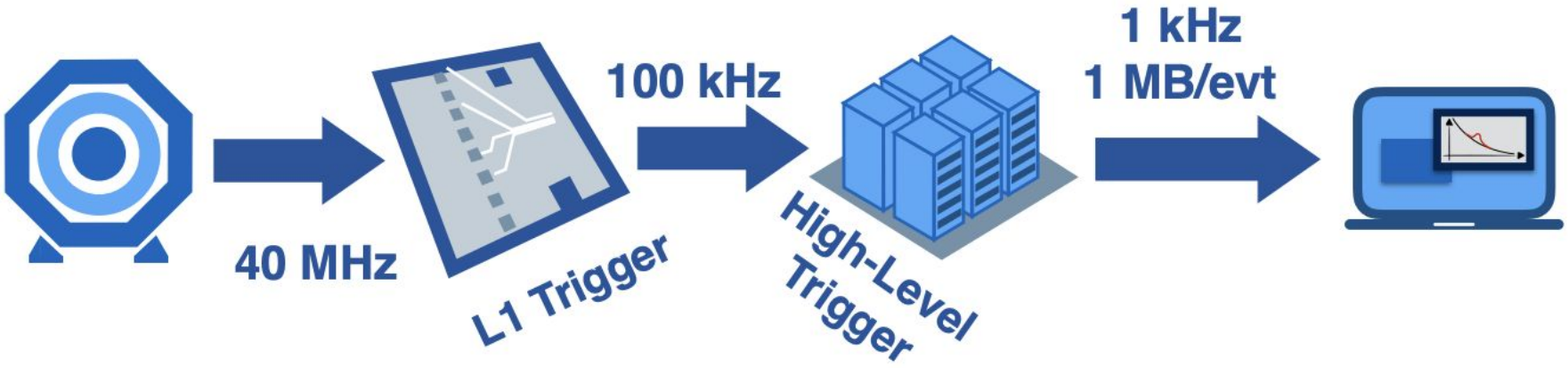
ML contact
Yu Zhang

Validation contact
Peter Young

RECO contact
Lauren Hay
OPEN

DPG, POG & PAG contacts and Experts

Data from P5 to Offline



Trigger rate and overlap between PDs