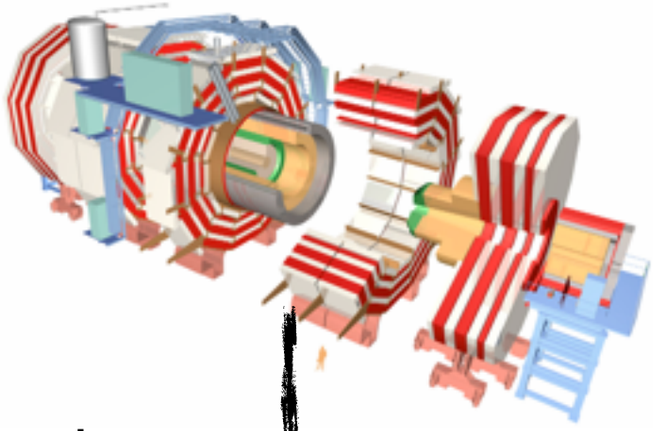
Offline Data Preparation

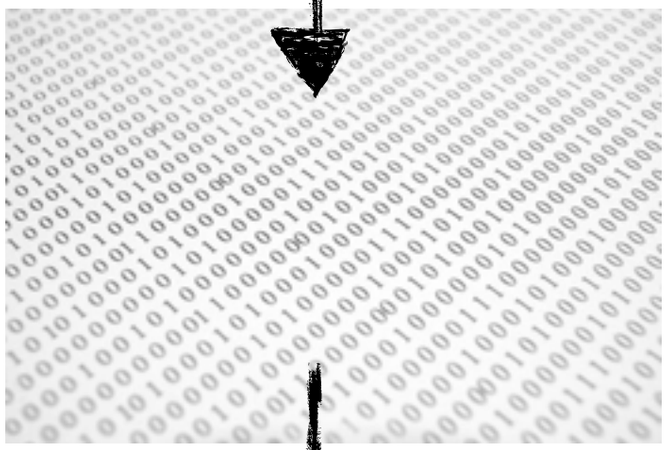
G. Cerminara (CERN)

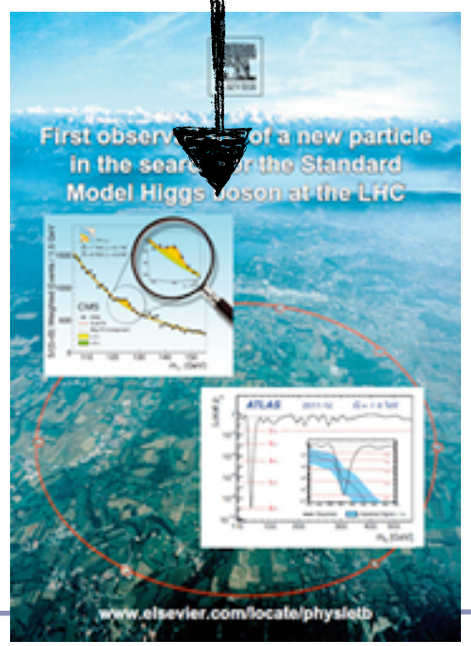
CMS Induction Session - 17th of July 2014 - CERN

Outline

- From P5 to Tier0 (+CAF)
  ⟶ prompt-reco & Prompt Calibration Loop
- Datasets
- Data distribution and Tier structure of our computing
- Finding and moving datasets for your analysis
- CMS software ⟶ CMSSW
- Validation and Data Quality Monitoring
- MC Production
- Alignment & Calibrations
- Data Certification
- Luminosity computation







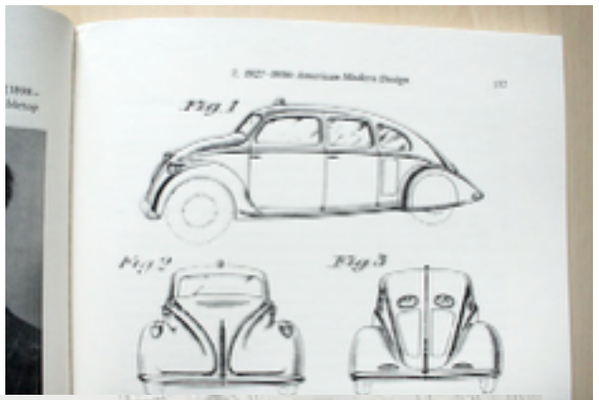
Organization & Coord. Areas

- The topics covered by this talk are the daily bread of 3 coordination areas:
  - Offline:
    - software development and integration, event reconstruction and simulation
  - Physics Performance and Datasets (PPD)
    - data certification and quality, Alignment & Calibrations, Software Validation, Monte Carlo production organization
  - Computing
    - Actual data processing and Simulated events generation, events storage and management
- NOTE: for each topic I will link entry points for documentation and contacts

Organization & Coord. Areas

- The topics covered by this talk are the daily bread of 3 coordination areas:
  - Offline:

→ Design



  - Physics Performance and Datasets (PPD)

→ Product Industrialization

  - Computing

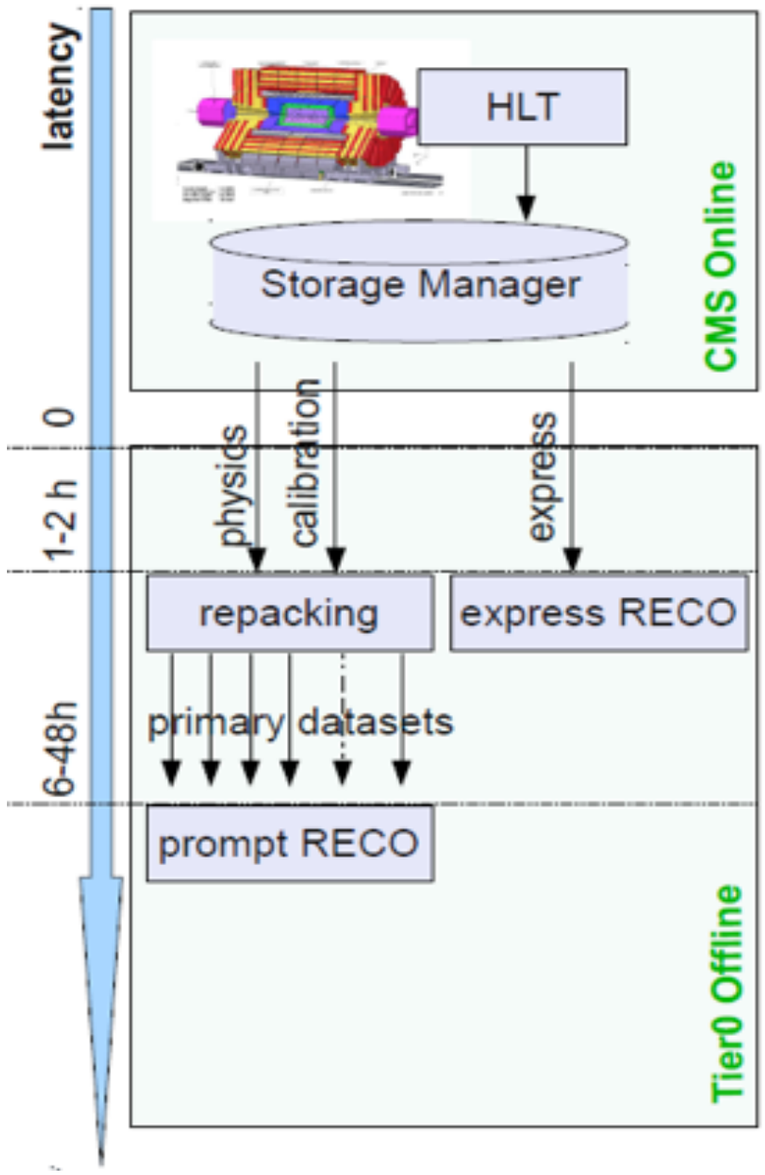
→ Manufacturing

- More complex than this since each area also takes care of the tools for the corresponding step...

Contact&Doc [4]

From P5 to Offline

- Data streams & Tier0 (CERN farm) processing → specialized for different tasks
  - express (1-2 h) → prompt feedback & calibrations ~40Hz bandwidth shared by:
    - calibration (½) - detector (¼) - physics (¼) monitoring
  - bulk data → primary datasets for physics analysis (prompt reconstruction)
    - delayed of 48h → Prompt Calibration Loop
  - Alignment & Calibration (AlCa) streams
    - dedicated streams & skims for calibration purposes
  - other specialized streams (e.g "data parking", "data scouting")



- End of RunI ~300Hz Prompt-Reco + 300-600Hz of "parked" data
  = reconstructed only @ end of RunI
- Plans for RunII ~1kHz of Prompt-Reco

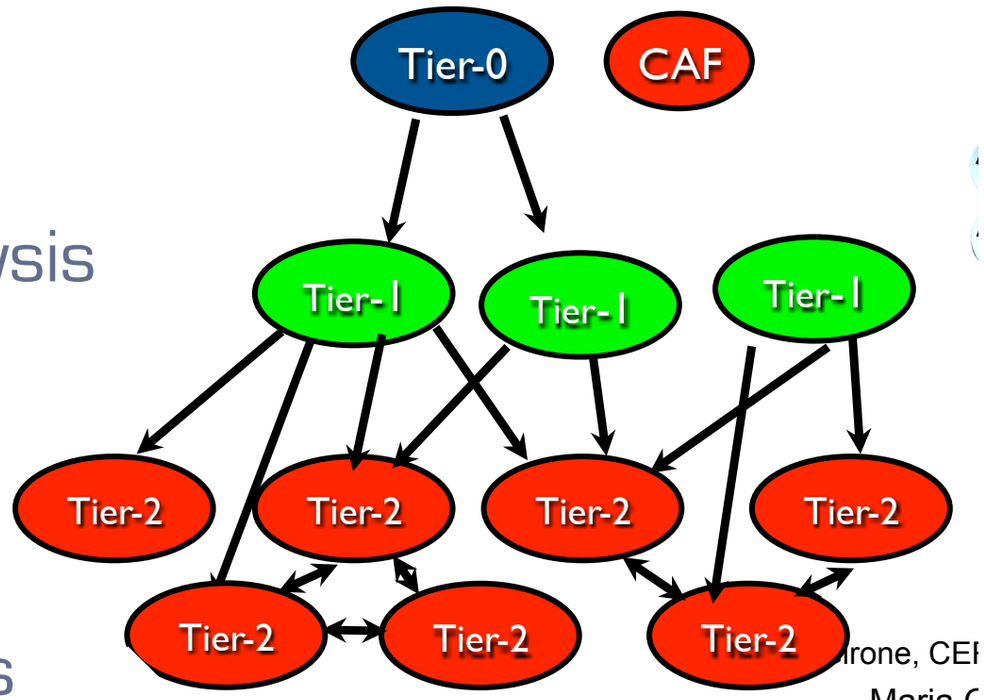
Primary Datasets

- Bulk data split in Primary Datasets (PDs) on the basis of the HLT results ⟶ triggers with similar physics content/use in the same PD
- Constraints from analysis:
  - definition centered on physics objects (e.g. SingleElectron, DoubleMuon, JetHT…)
- Constraints from processing and handling:
  - average event rate approximately uniform across different PDs, to ease distribution at the Tier 2 centers
  - event rate > 10 Hz, to avoid small files & < 200 Hz
- On the top of the primary datasets we can deploy "central skims" ⟶ event reduction through RECO (or HLT based) filters
  - used for Detector Studies or Physics Analysis
- Dataset & skim definition, is handled by the Dataset Definition Team (DDT) in PPD together with Trigger Study Group (TSG)

Contact&Doc [7]

Distributed Computing

- "Tiered" computing infrastructure handles distinct use cases
  - **Tier0** (CERN): big farm @ CERN (~10k CPU cores)
    - dedicated to PromptReco (used for other purposes if beam-off)
  - **Tier1:** ~10 big centers with large disk/tape capacity (tot. ~14k cores)
    - long term storage and central RECO passes (data and MC)
  - **Tier2:** many smaller centers with disk "buffers"(tot. ~32k cores)
    - analysis jobs and MC simulations
      (or central jobs with low IO requirements)
  - **Tier3:** even smaller centers for data analysis
    (no "central" jobs, usually "institute level")
  - **CAF:** Central Analysis Facility:
    dedicated center for fast turn-around
    analyses or prompt feedback/calibrations



Contact&Doc [8]

Finding Datasets

- How do I look for a sample ⟶ Data Aggregation Service (DAS)
  - list datasets and their properties (requestID, sites, run # and LS #....) fetching information from various services
- Anatomy of the dataset name:
  - `dataset = /PrimaryDataset/ProcessingVersion/DataTier`
- Examples:
  - data (prompt reco): `/DoubleMu/Run2012D-PromptReco-v1/AOD`
  - data (re-reco): `/DoubleMu/Run2012D-16Jan2013-v2/AOD`
  - MC (CSA14): `/WprimeToENu_M_3600_Tune4C_13TeV_pythia8/Spring14miniaod-PU20bx25_POSTLS170_V5-v1/MINIAODSIM`
- How do I understand the various re-reco passes and the MC campaigns?
  - Data: https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDataReprocessing
  - MC: the tool for the MC request management https://cms-pdmv.cern.ch/mcm/

Contact&Doc [9]

Finding Datasets

- How do I look for a sample ⟶ Data Aggregation Service (DAS)
  - list datasets and their properties (requestID, sites, run # and LS #....) fetching information from various services
- Anatomy of the dataset name:
  - `dataset = /PrimaryDataset/ProcessingVersion/DataTier`
- Examples:
  - data (prompt reco): `/DoubleMu/Run2012D-PromptReco-v1/AOD`
  - data (re-reco): `/DoubleMu/Run2012D-16Jan2013-v2/AOD`
  - MC (CSA14): `/WprimeToENu_M_3600_Tune4C_13TeV_pythia8/Spring14miniaod-PU20bx25_POSTLS170_V5-v1/MINIAODSIM`

Reference for Generator + Process Information

- How do I understand the various re-reco passes and the MC campaigns?
  - Data: https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDataReprocessing
  - MC: the tool for the MC request management https://cms-pdmv.cern.ch/mcm/

Finding Datasets

- How do I look for a sample ⟶ Data Aggregation Service (DAS)
  - list datasets and their properties (requestID, sites, run # and LS #....) fetching information from various services
- Anatomy of the dataset name:
  - `dataset = /PrimaryDataset/ProcessingVersion/DataTier`

Acquisition Era + Reco Pass

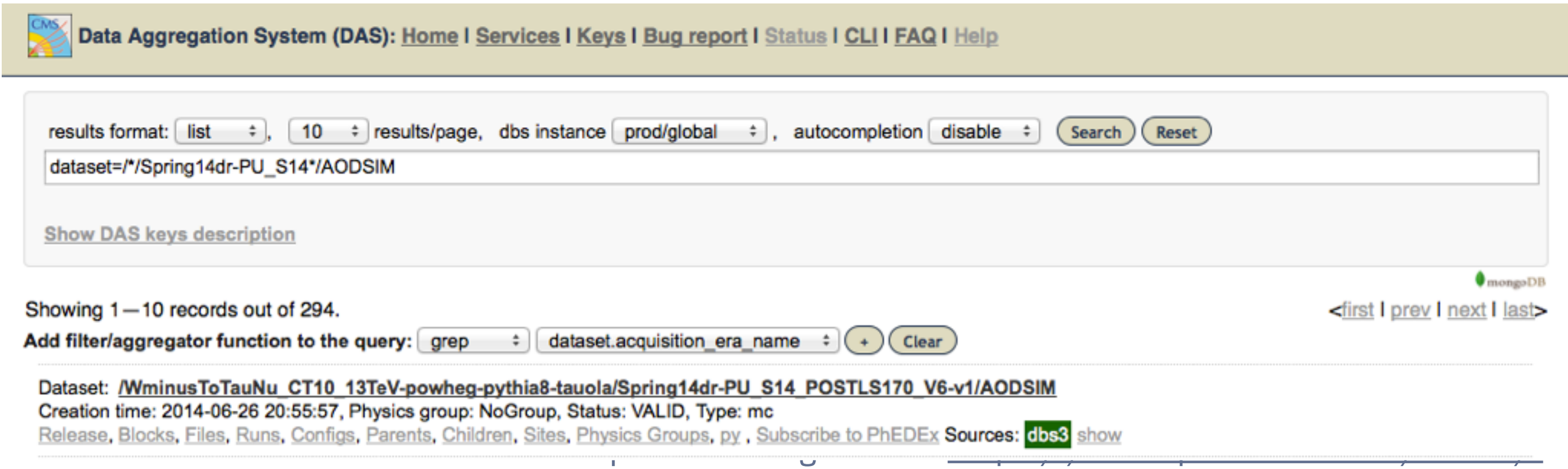
- Examples:
  - data (prompt reco): `/DoubleMu/Run2012D-PromptReco-v1/AOD`
  - data (re-reco): `/DoubleMu/Run2012D-16Jan2013-v2/AOD`
  - MC (CSA14): `/WprimeToENu_M_3600_Tune4C_13TeV_pythia8/Spring14miniaod-PU20bx25_POSTLS170_V5-v1/MINIAODSIM`

Prod. Campaign + PU scenario + GlobalTag

- How do I understand the various re-reco passes and the MC campaigns?
  - Data: https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDataReprocessing
  - MC: the tool for the MC request management https://cms-pdmv.cern.ch/mcm/

Finding Datasets

- How do I look for a sample ⟶ Data Aggregation Service (DAS)
  - list datasets and their properties (requestID, sites, run # and LS #....) fetching information from various services
- Anatomy of the dataset name:
  - `dataset = /PrimaryDataset/ProcessingVersion/DataTier`
- Examples:



More on Data Tiers

- Data Tiers: define the event content for the dataset. The most common
  - RAW, RECO, **AOD, AODSIM, MINIAOD, MINIAODSIM**, USER, GEN, FEVT...
- RAW ⟶ is what is repacked in root format @ Tier0
  serves as input for any further processing step
  - new release can always read old RAW files
- RECO vs AOD:
  - RECO needed for dedicated studies and detector commissioning
    ⟶ kept only for 6 months out o prompt-reco
    ⟶ not produced by default in re-reco
  - AOD subset (∼40%) of RECO content meant for analysis
    ⟶ designed to serve 90% of the use cases saving space
- Now we know enough to go one step further: MiniAOD (∼10-15% of AOD)
  - more than a subset ⟶ designed for analysis (more in Luca's talk)

Moving Datasets

- Datasets for analysis ⟶ they need to be on disk @ Tier2/3 (GRID = "where" doesn't matter)
  - **AnalysisOps** usually subscribes the most common samples to "central" space ⟶ no action is needed
  - What if you need something more "exotic"?
- You can ask the transfer @ T2s using the Phedex service
  - **users** can request using the interface (complete datasets or "files blocks")
    ⟶ requests are associated to "groups" and assigned to a given "site"
  - **data manager** approves/rejects the request (usually evaluating the available quota at the destination site)
    ⟶ each PAG/POG group has a person taking care of data-management
  - newly deployed **Dynamic Data Placement** tools
    ⟶ optimal data replication and quota control
- before submitting I want to run on a single file locally
  ⟶ how do I do it
  - most of the files are accessible remotely via XROOTD protocol

Any data,
Any time,
Any where
(AAA)

Contact&Doc [14]

Running on your data

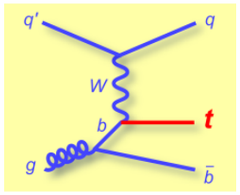
- Once you discovered/moved all the datasets you need to run you analysis ⟶ CRAB

- Luca will give you more details on this

CMS Software

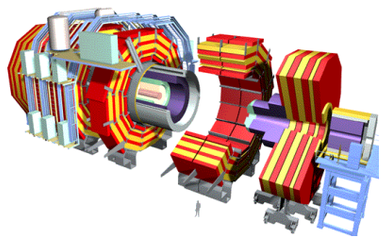
- CMSSW: one release to rule them all
  - GENerator, SIMulation, RECOnstruction ANALYSIS workflows...
- C++ code and configuration handled via Python
  - “git” used for code integration and versioning
- Release schedule follows a “train model”:
  dear developer: catch this train or wait for the next one
  - regular time-table of ~6 months (slightly tuned for major conferences or physics needs)
  - pre-releases are regularly produced while the release is under development
- Feature planning:
  - production releases: main features are decided by the physics/machine constraints goals of the experiment
    - e.g. beginning 2012: speed up of tracking to cope with higher PileUp

Simulation Framework



Physics Generators

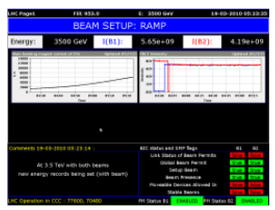
Particle 4-vectors



Geometry/ Material Description

Geant 4

CMS Full Simulation



Simulated Hits from Pileup Interactions

Simulated Hits

Noise Model

Electronics Simulation

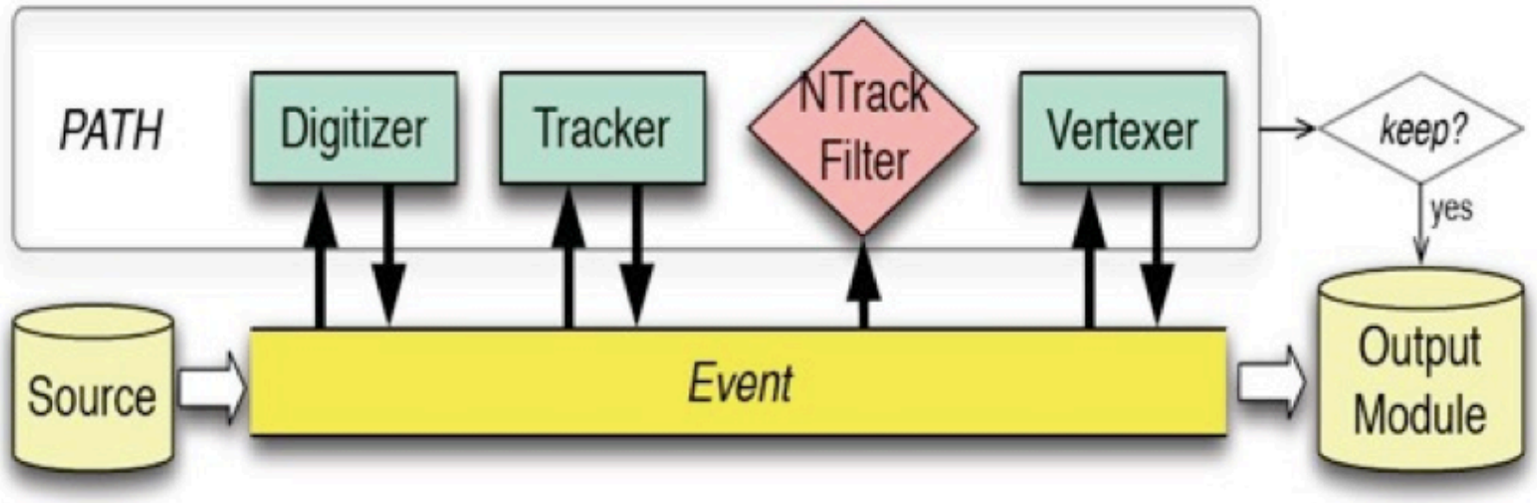
Simulated Raw Data

Alternatively: CMS “Fast Simulation” is a slightly less realistic but much faster simulation of low-level objects (hits, clusters)

Contact&Doc [17]

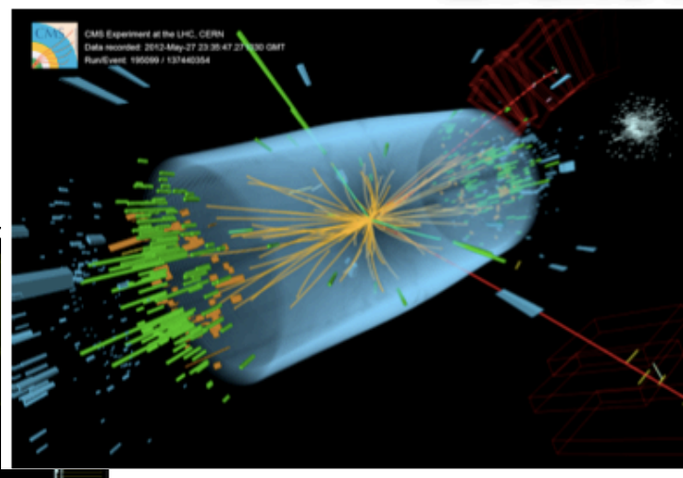
Reconstruction

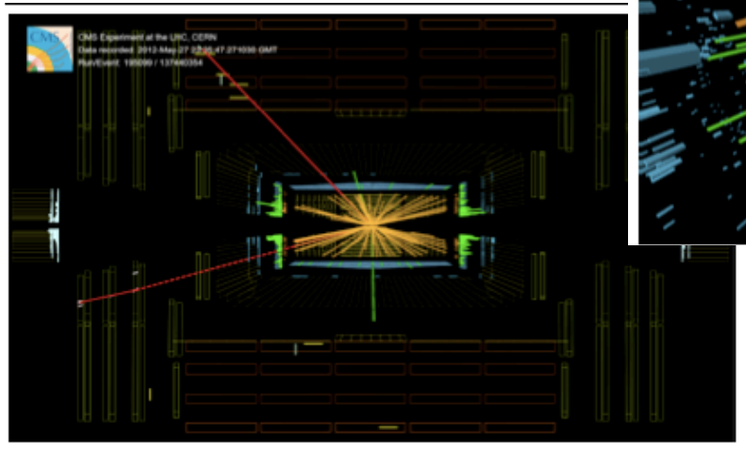
- In CMSSW framework, event reconstruction algorithms are implemented as “modules”, which communicate via the “Event”



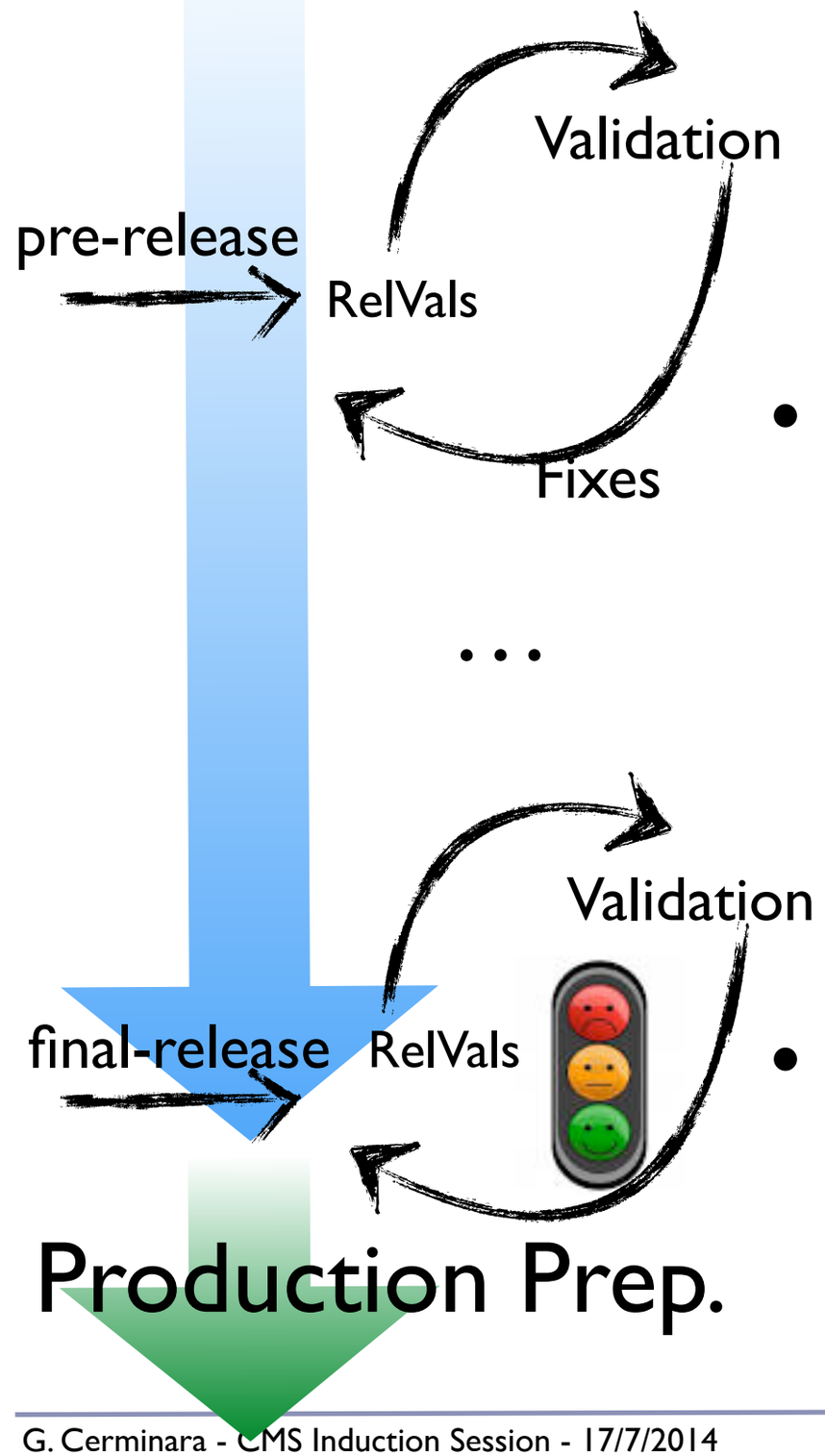
Analysis Object Data (AOD)

- AOD data format ⟶ particle-flow driven “physics objects” in ROOT format
  - Tracks, muons, electrons, photons, jets, etc...





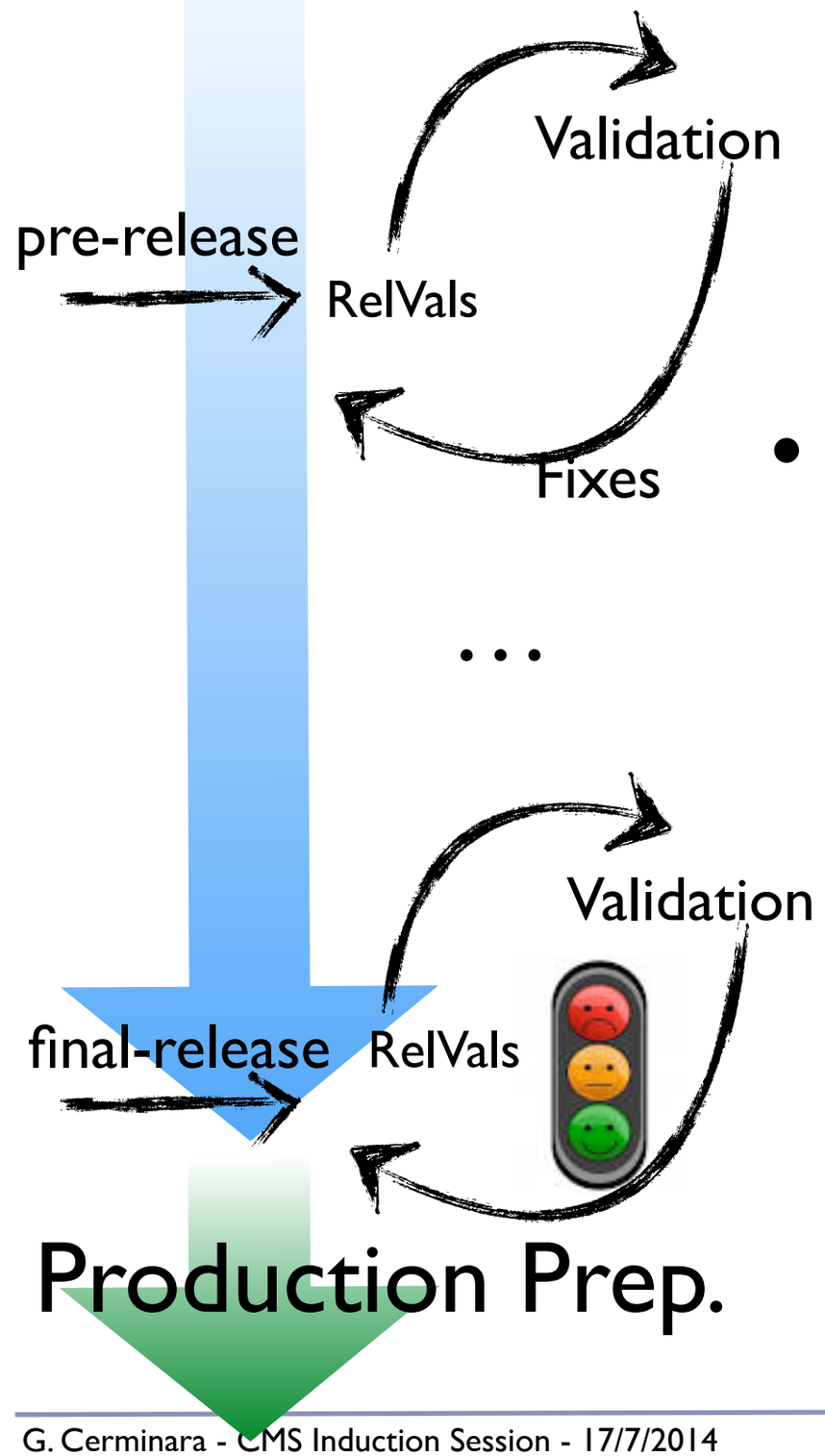
READY FOR PHYSICS?



- Release integration is bound to Quality Assurance Tests ⟶ Data Quality Monitoring (DQM)
  - unit tests & regression tests
  - small scale production tests: **Rel**ease **Val**idation Test (**RelVal**) producing DQM plots
- **Validation** is an iterative process and in performed all along the life cycle of the release (from pre-releases to final one)
  1. DPG, POG and PAG validation experts check the plots
  2. PPD/PdmV group coordinate validation campaigns
  3. sign-off on quality of release and calibrations
- differential validation compares plots of each release w.r.t last one
  - each cycle takes 4-5 days to have the samples + 1 week for feedback from the validators

Contact&Doc [19]

Ready for Physics?

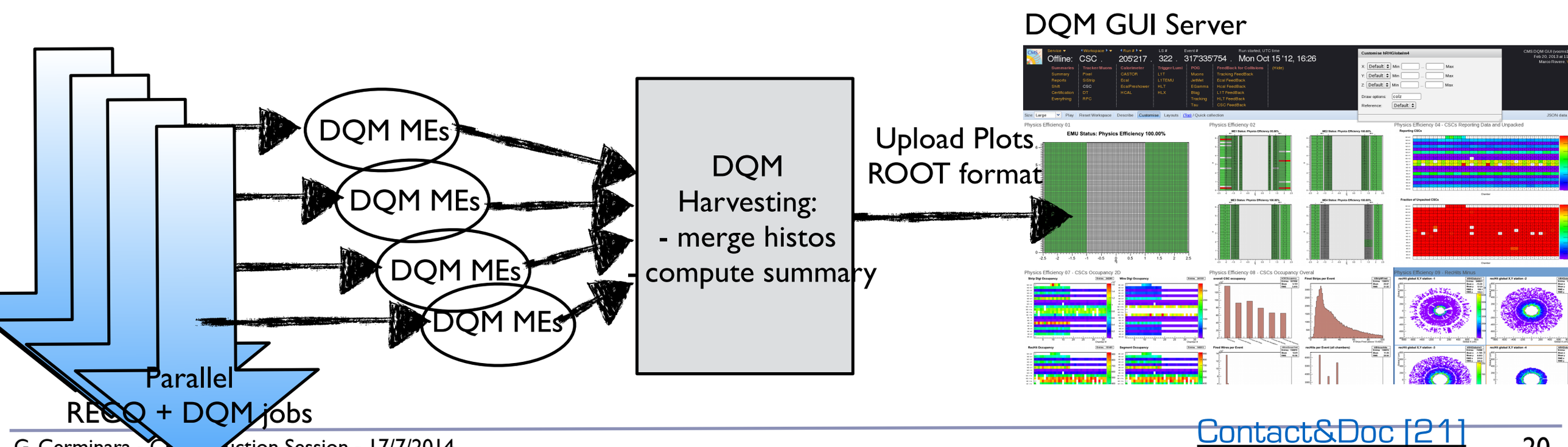


- Release integration is bound to Quality Assurance Tests ⟶ Data Quality Monitoring (DQM)
  - unit tests & regression tests
  - small scale production tests: **Rel**ease **Val**idation Test (**RelVal**) producing DQM plots
- Once a major release meant for production (X.Y.0) is cut out, PPD starts the **preparation of the campaign** (= re-reco or MC production):
  - finalization of the alignment and calibration conditions (and their validation)
  - finalization of the parameters for the Pile-Up overlay (PU scenario)
  - preparation of the injection machinery for the central processing by computing

Data Quality Monitoring

- DQM is the tool to produce plots while running RECO (or any CMSSW workflow)
- 2 main applications:
  - **online** ⟶ samples events after HLT and plots quantities with very low latency
    ⟶ live monitoring of detector performance during data taking
  - **offline** ⟶ reads all events while they are reconstructed
    ⟶ data certification
    ⟶ release validation

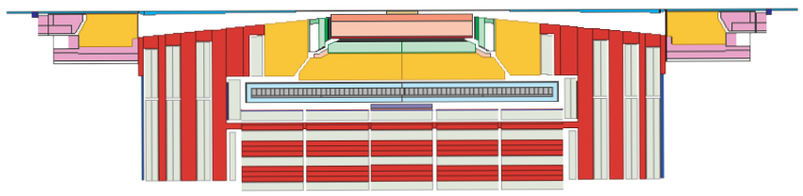
- DQM GUI ⟶ front-end to browse histograms for a given dataset/run

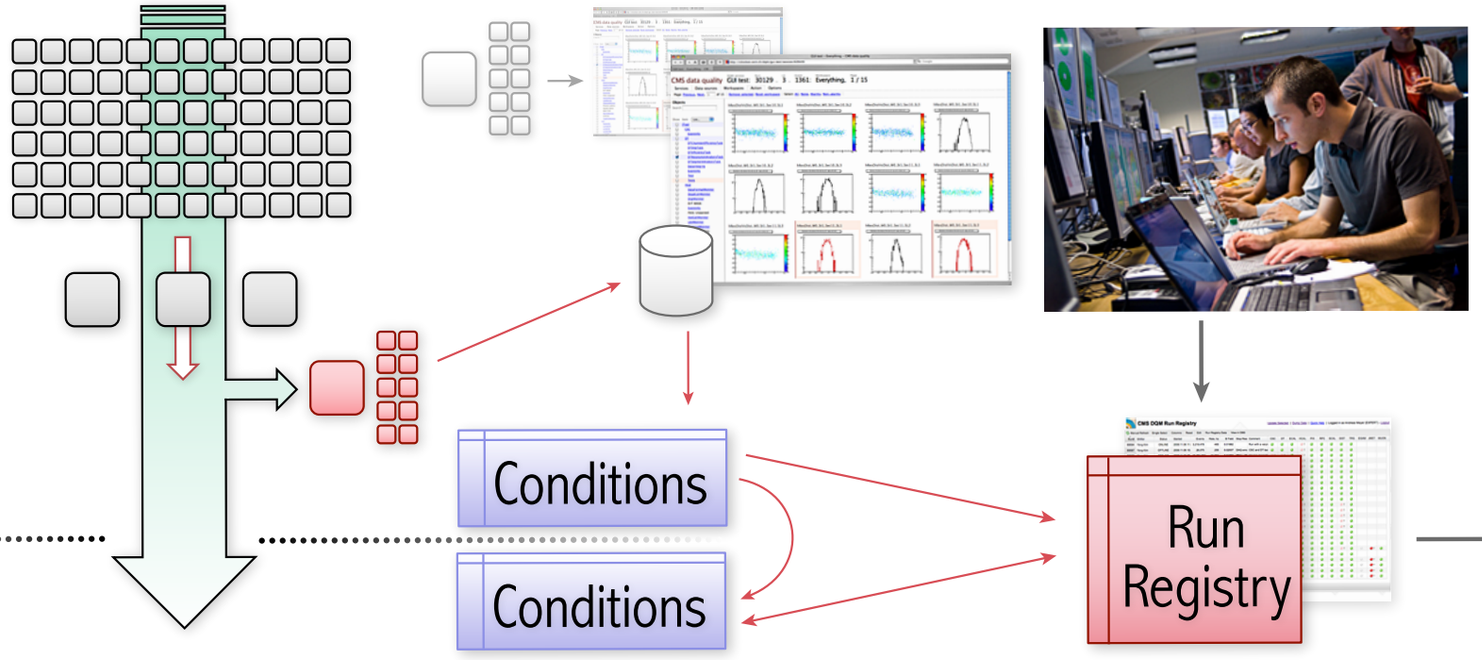


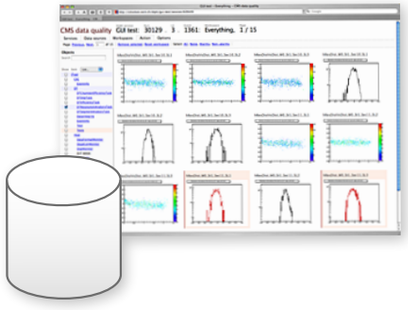
Data Certification

- Complexity of detector and offline processing requires continuous monitor of the quality of the data
- Team of Detector and Physics Object experts check DQM plots for each run and select those usable for analysis
  - look for unexpected effects that could affect analysis level quantities (e.g noise spikes, dead areas of detectr, problematic calibrations...)
- During data taking ⟶ continuous certification of prompt reco datasets
- PPD/DQM-DC team → coordinates the validation activity
  - Reports @ PPD General Meeting on Wednesday 14:00 -16:00

Data Certification







Online

Offline

Tier-0

Tier-1s

CAF

Release Validation

Simulation Validation

Conditions

Conditions

Run Registry

CERN 24/7

Detector & Operation Efficiency

certification JSON

```
{"132440": [[157, 378]],
 "132596": [[382, 382], [447, 447]],
 "132598": [[174, 176]],
 "132599": [[1, 379], [381, 437]]
}
```

CERN daytime

2/day

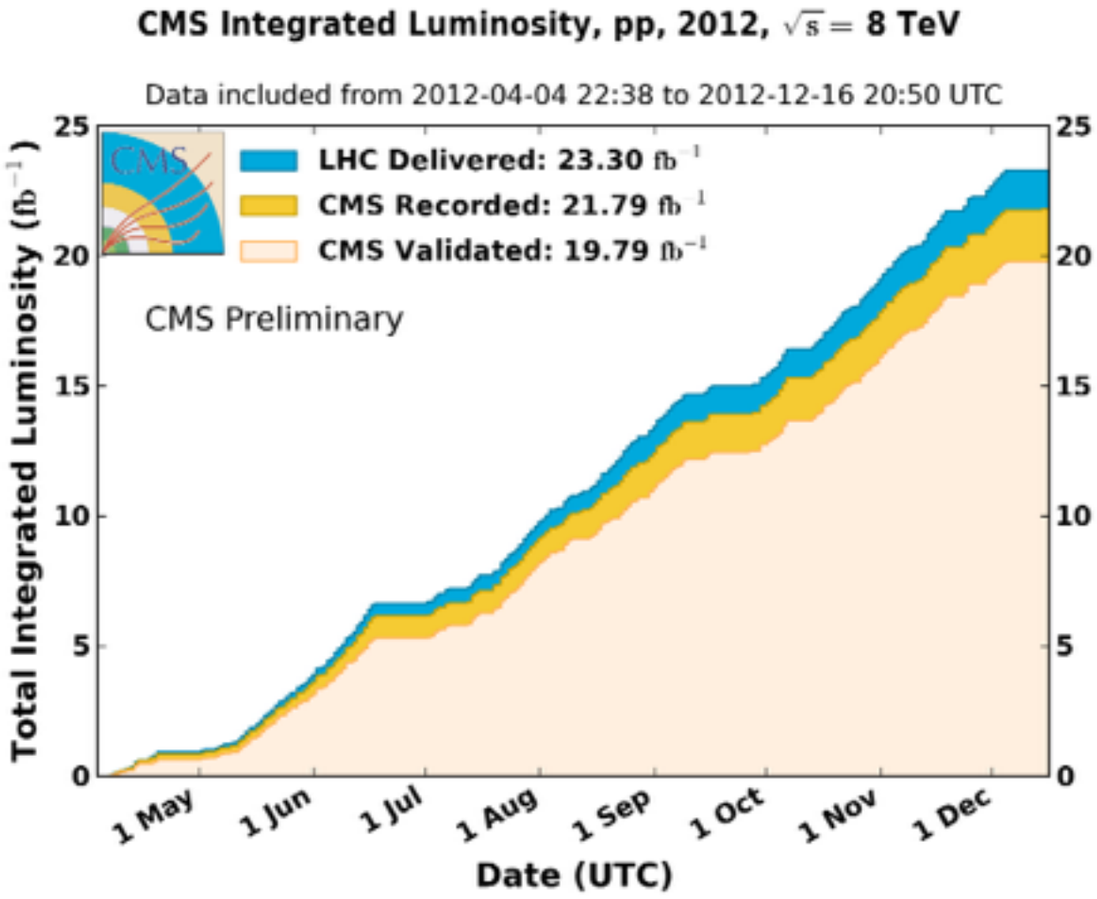
DESY 1/day

Physics Data Certification

Data Processing | Visualisation | Certification | Sign-off

Certification & Analysis

- Selection of LumiSections (LS) (= 23s of run) considered GOOD for physics
  - distributed in JSON format
    ⟶ weekly for PromptReco
    ⟶ after each major re-reco pass
- Several "flavors" of the JSON file:
  - **golden** ⟶ requires all sub-detectors/POGs to be "GOOD"
  - **muon-only** ⟶ no requirements on calo
  - **DCS-only** ⟶ only requires Tracker to be powered
- How do I use the JSON file:
  - to be used in CRAB to run only on CERTIFIED LSs of your dataset



CMS Preliminary Results: Apr-Dec 2012 proton-proton collision runs									
Tracker		Calorimeters			Muon Spectrometer			Magnet	Operational
Pixel	SST	ECAL	ES	HCAL	CSC	DT	RPC		
98.9	99.6	98.6	99.5	97.2	99.3	99.8	99.4	98.6	99.2
All good for physics: 91%									

Luminosity weighted fractions (in %) of data certified as good for physics analysis relative to 21.79fb^{-1} of data recorded by the CMS experiment during 2012 proton-proton collisions at √s=8TeV between April 4th and December 17th.

Contact&Doc [24]

Luminosity

- How to compute luminosity for your analysis
- You need to know the luminosity of ALL the LSs you run on and ONLY those
  - CRAB reports the LS successfully processed by your jobs in the same JSON format used for certification

certification JSON


```
{"132440": [[157, 378]],
 "132596": [[382, 382], [447, 447]],
 "132598": [[174, 176]],
 "132599": [[1, 379], [381, 437]]
}
```


CRAB

CRAB JSON

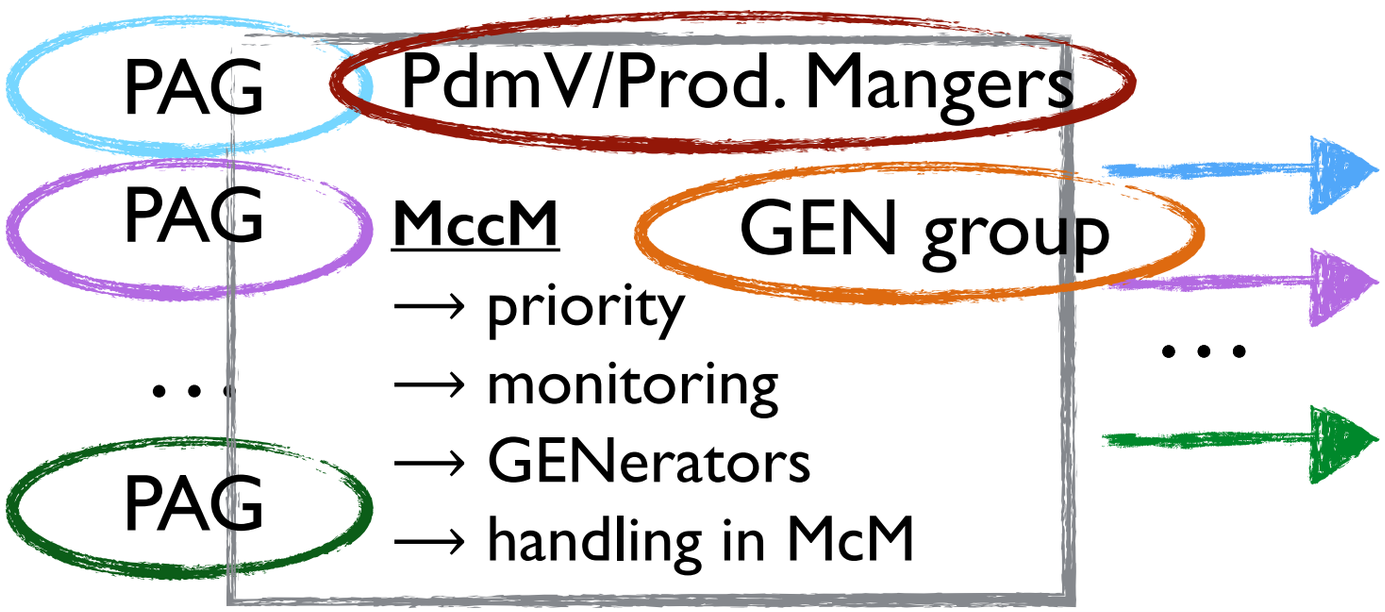

```
{"132440": [[157, 378]],
 "132596": [[382, 382], [447, 447]],
 "132598": [[174, 176]],
 "132599": [[1, 379], [381, 437]]
}
```


- The Lumi POG provides tools to compute the luminosity starting from the JSON file:
  - `lumiCalc2.py`: reports HF measurements of LHC delivered,CMS recorded luminosity and recorded luminosity in a given hlt path
  - `pixelLumiCalc.py`: pixel reconstructed CMS recorded luminosity and recorded luminosity in a given hlt path
- NOTE: using directly the certification JSON can be problematic for several reasons:
  - LS that failed prompt reco or re-reco will appear in the JSON
  - LS that failed in your jobs will appear in the JSON

Contact&Doc [25]

MC Production

- MC production matches data processing and targets specific conferences/events
  - usual scale for 1 year of data to be analyzed is ~4 Gevents
- Production organization is a complex task:
  - O(1000) requests from analyzers with different physics goals
  - prioritization to cope with resource budget & conference timeline
- Activity coordinated by PPD/PdmV with Production Managers & Computing Operations
  - each PAG has MC contacts taking care of requesting sample needed by the group (after consultation with conveners)
  - requests are discussed in weekly Monte Carlo Coordination Meetings (MccM) on Thursday @ 15:00
  - tool for bookkeeping & monitoring of campaigns and single requests ⟶ McM



Computing "CompOps team"

Calibration Workflows

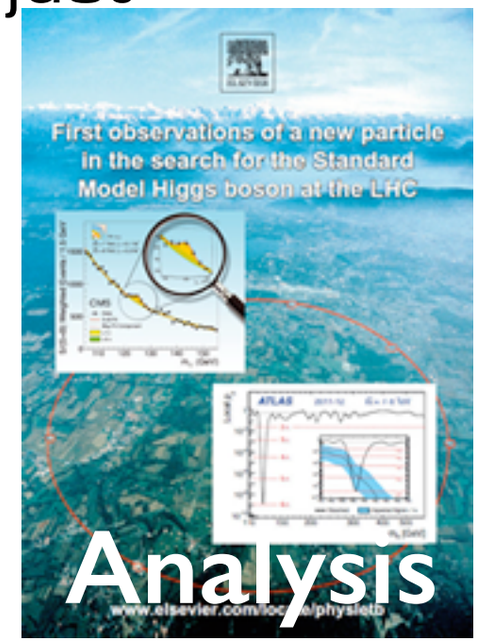
- Provide most up-to-date conditions @ all stages of the data processing Different workflows depending on the time scale of updates:
  - **quasi-online calibrations** for HLT and express (e.g beam-spot used by HLT)
  - **prompt calibrations**: monitor/update conditions expected to vary run-by-run (or even more frequently) ⟶ exploit 48h delay of prompt-reco
  - **offline workflows** for re-reco passes and analysis level conditions (e.g Jet Energy Corrections, Muon Momentum Scale)
- A consistent set of conditions for a given task (i.e. HLT, prompt-reco, analysis... etc) is provided by a **GlobalTag** (GT)
  - non trivial interdependencies between calibrations
  - ~300 sets of parameters are needed in any point in time to calibrate and align CMS
- Dedicated GTs for analysis are made available by AlCaDB team ⟶ to be used with consistent dataset

Contact&Doc [27]

Outlook

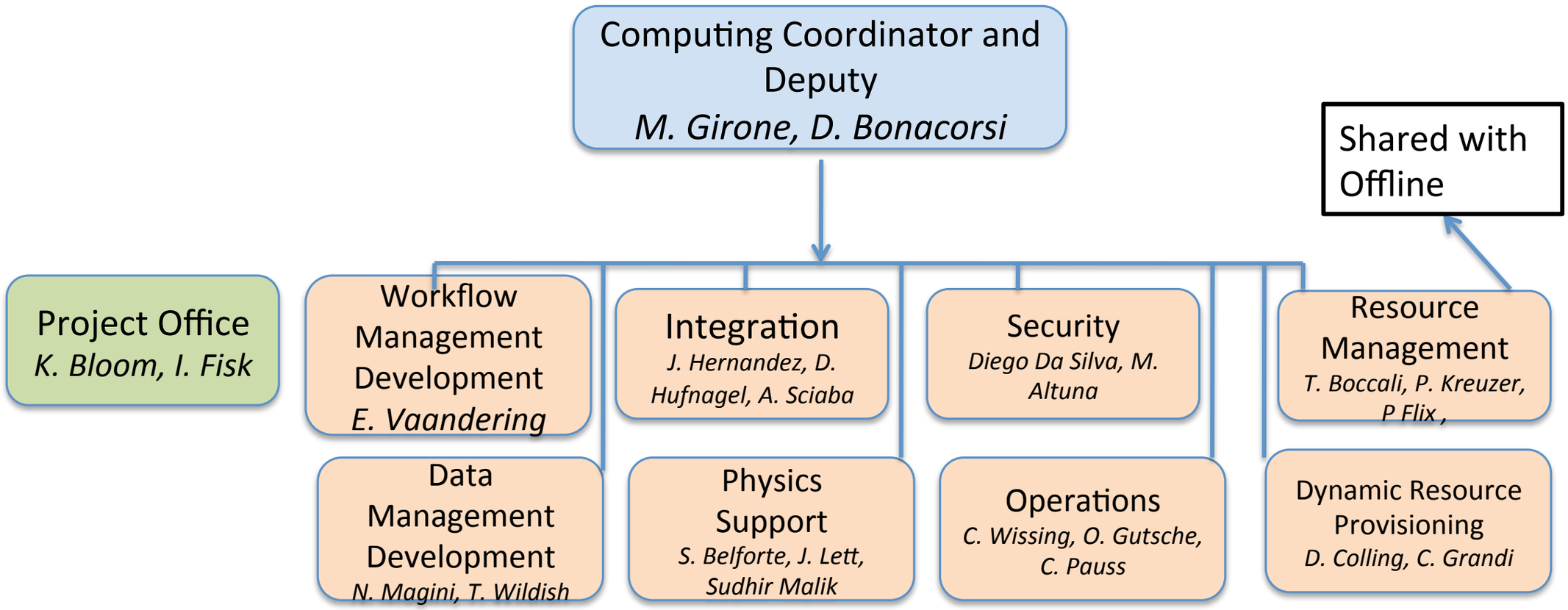
- (some) Physicists tend to think that the universe is just

&



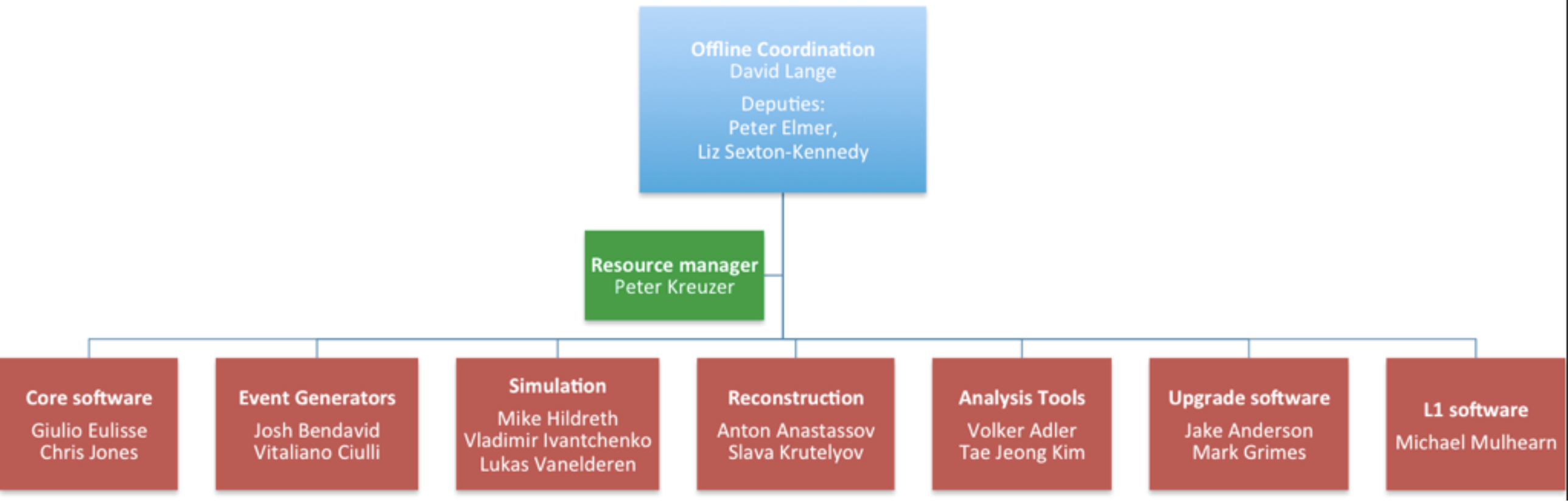
- This is almost true…however to exploit at best the hardware and publish top quality analysis you need more
  - software, data preparation and computing reached in CMS an unprecedented level of complexity
    ⟶ if you don't realize it it just means that they work fine!
- Plenty of opportunities for exciting work in these fields:
  - a single person can actually have a large impact (and have fun!)

Computing Organization



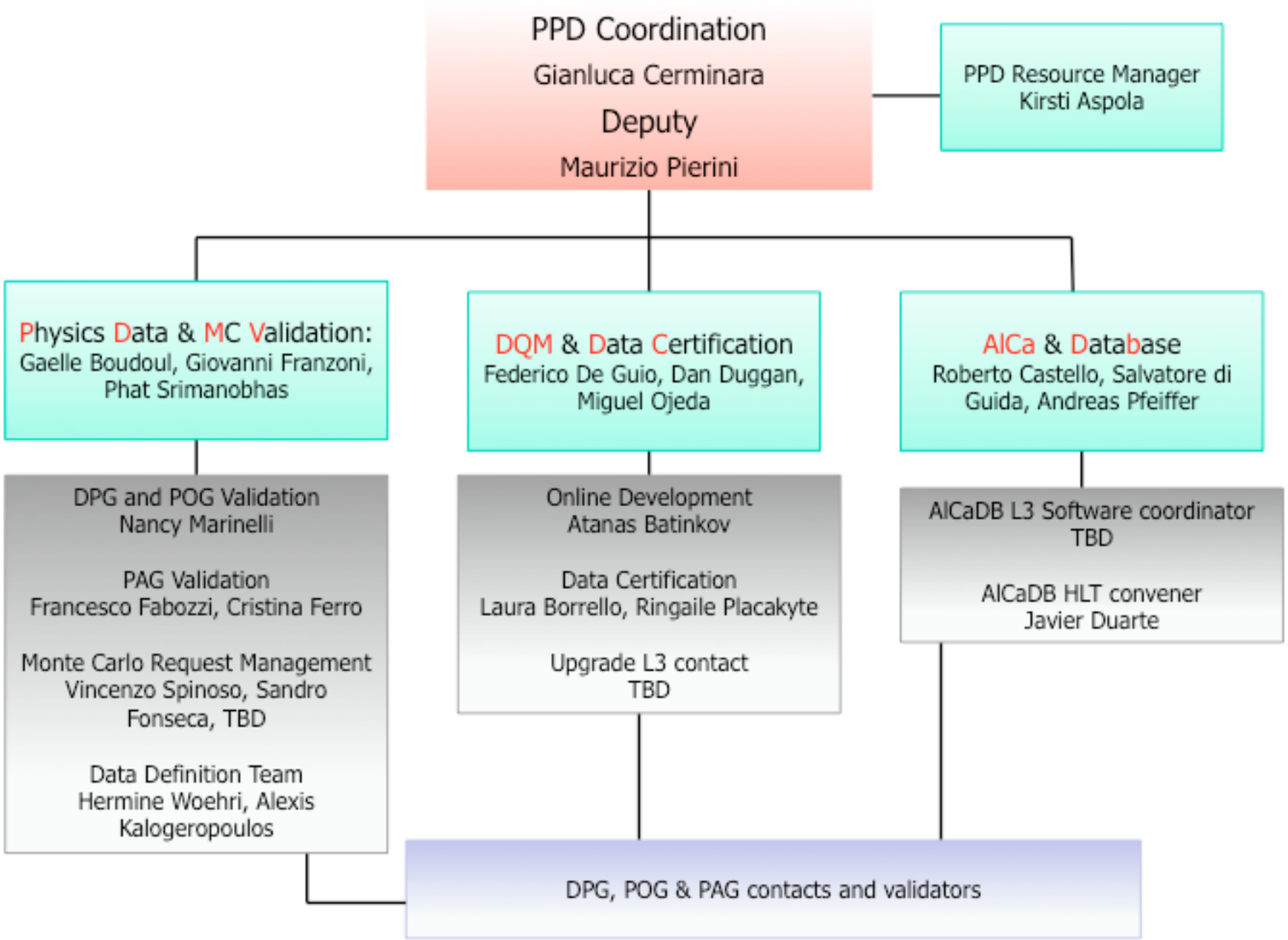
Offline Organization

CMS Offline Management for 2014



PPD Organization

Physics Performance & dataset (PPD) management



References

	Contacts	Documentation
[4]	cms-ppd-coordinator@cern.ch cms-offline-coordinator@cern.ch cms-computing-coordinator@cern.ch	PPW Main Twiki Offline Main Twiki Computing Main Twiki
[7]	hn-cms-dataset-definition@cern.ch	DDT Twiki
[8]		Computing Model Workbook
[9]	hn-cms-computing-tools@cern.ch	DAS
[14]	hn-cms-phedex@cern.ch	XROOTD doc Phedex - Phedex Workbook
[17]	hn-cms-offlineAnnounce@cern.ch hn-cms-relAnnounce@cern.ch	Offline Workbook SW Guide
[19]	hn-cms-relval@cern.ch hn-cms-physics-validation@cern.ch	PdmV Twiki
[21]	hn-cms-evfdqmannounce@cern.ch	DQM Twiki
[22]	hn-cms-data-certification@cern.ch	DQM-DC Twiki RunRegistry
[24]		Twiki JSON File
[25]	hn-cms-luminosity@cern.ch	Twiki LumiCalc
[26]	hn-cms-prep-ops@cern.ch	PdmV Twiki McM
[27]	hn-cms-alca@cern.ch	AlCaDB Twiki GlobalTag Twiki