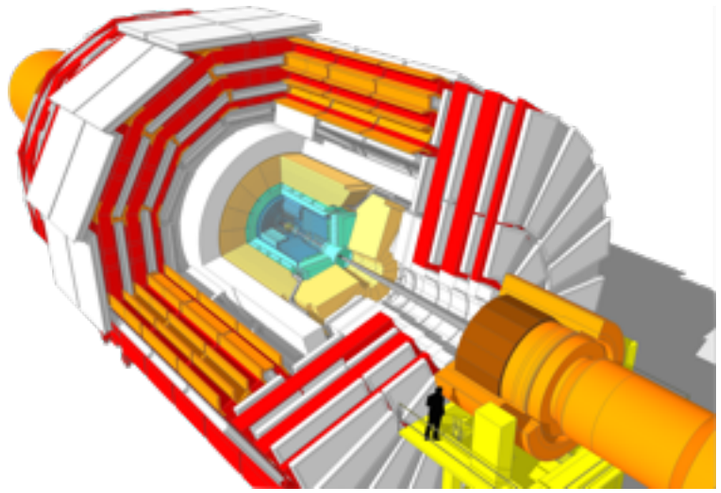
Offline Data Preparation

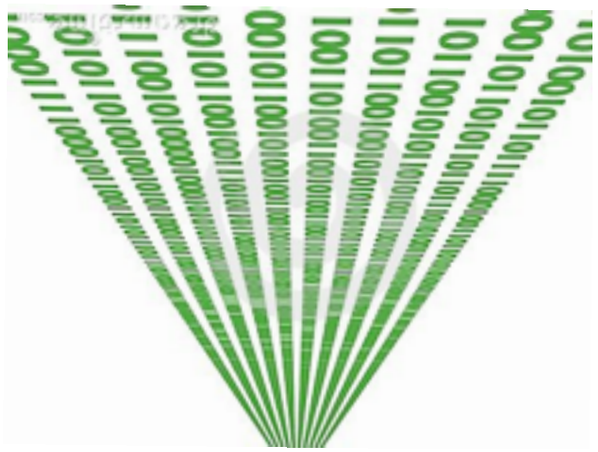
G. Cerminara (CERN)

CMS Induction Session - 10th of July 2015 - CERN

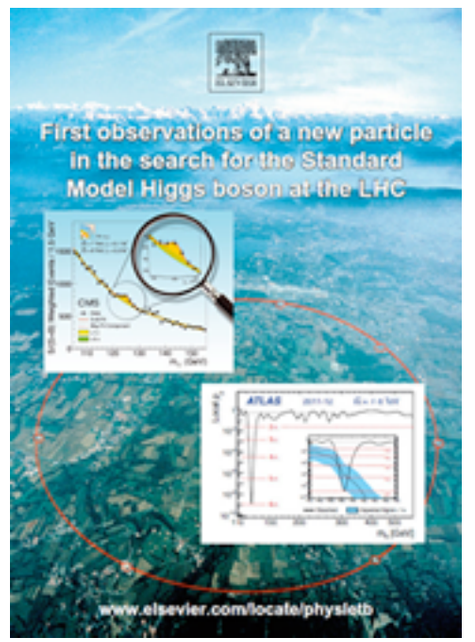
Outline

- The data flow: from P5 to Tier0
  ⟶ prompt-reco &
  Prompt Calibration Loop
- How the data are organized & handled
  - what are the datasets, how to find them, how to move them, how to analyze them
- CMS software ⟶ CMSSW
  - Validation and Data Quality Monitoring
- Data Certification & Luminosity computation
- MC Production
- Alignment & Calibrations





Additional Info on topics
Linked on relevant slides
Contact&Doc



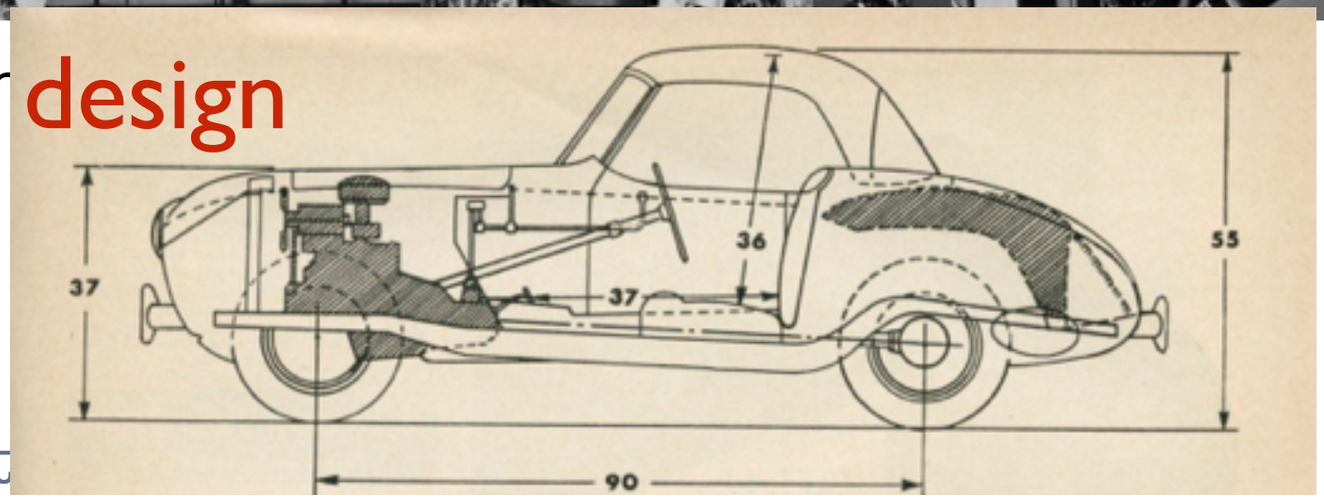
Organization & Coord. Areas

- The topics covered by this talk are the daily bread of 2 coordination areas:
- Offline & Computing
  - software development and integration, event reconstruction and simulation
  - actual data processing and Simulated events generation, events storage and management
- Physics Performance and Datasets (PPD)
  - data certification and quality,
  - Alignment & Calibrations,
  - Software Validation
  - management of Monte Carlo requests
- NOTE: for each topic I will link entry points for documentation and contacts

Contact&Doc

Organization & Coord. Areas

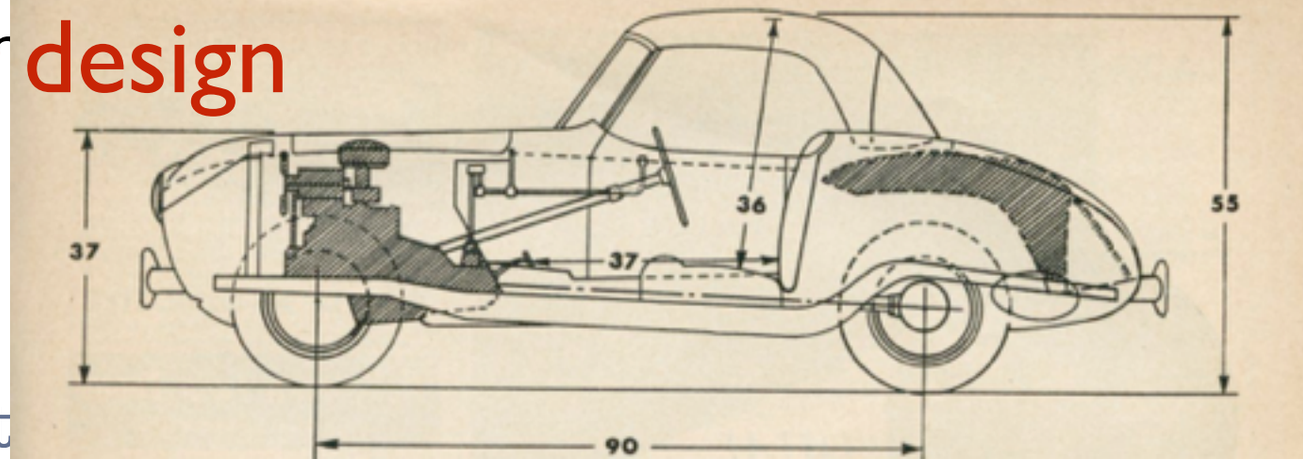
- The topics covered by this talk are design areas:
- Offline & Computing
  - software development and integrat
  - actual data processing and Simulated events generation, events storage and management
- Physics Performance and Datasets (PPD)
  - data certification and quality,
  - Alignment & Calibrations,
  - Software Validation
  - management of Monte Carlo requests
- NOTE: for each topic I will link entry points for documentation and contacts



Organization & Coord. Areas

- The topics covered by this talk are the following areas:
- Offline & Computing
  - software development and integrat
  - actual data processing and Simulated events generation, events storage and management
- Physics Performance and Datasets (PPD)
  - data certification and quality,
  - Alignment & Calibrations,
  - Software Validation
  - management of Monte Carlo requests
- NOTE: for each topic I will link entry points for documentation and contacts

design

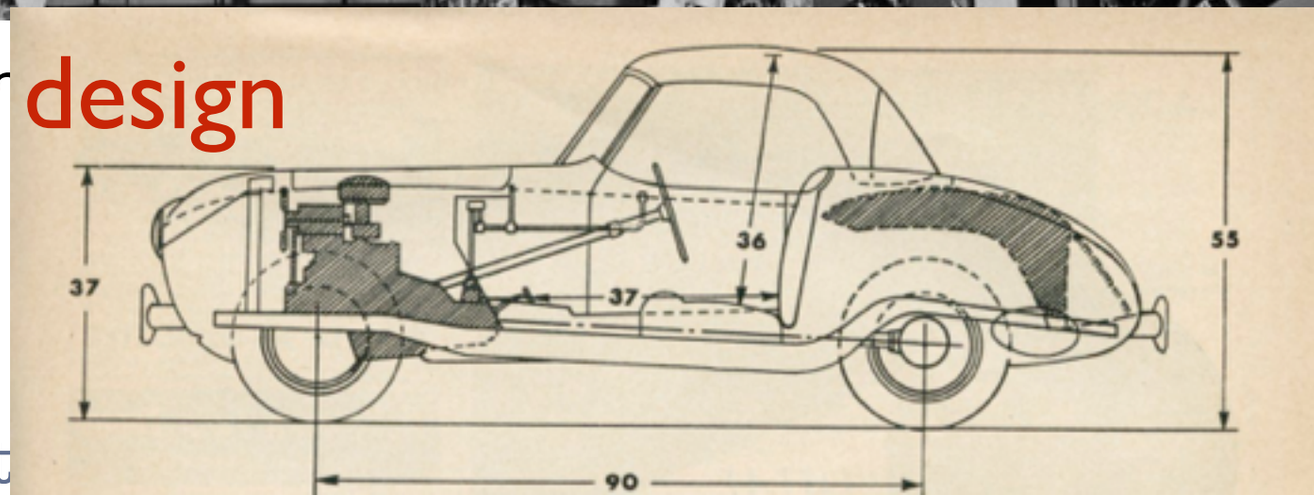


product industrialization

Organization & Coord. Areas

- The topics covered by this talk are

design



- Offline & Computing
  - software development and integrat
  - actual data processing and Simulated events generation, events storage and management
- Physics Performance and Datasets (PPD)
  - data certification and quality,
  - Alignment & Calibrations,
  - Software Validatic
  - management of M

product industrialization

manufacturing

- NOTE: for each to nentation and contacts

From P5 to Offline

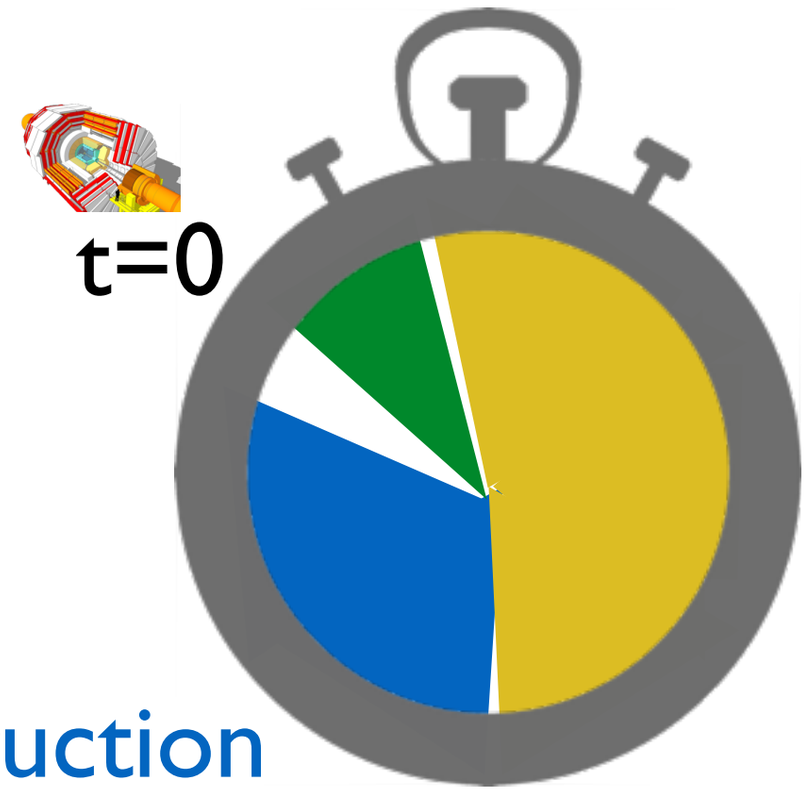
- Data streams & Tier0 (CERN farm) processing
  → specialized for different tasks
  - express (1-2 h) → prompt feedback & calibrations
    ~40Hz bandwidth shared by:
    - calibration (1/2) - detector (1/4) - physics (1/4) monitoring
  - Stream A (bulk) → primary datasets for physics analysis (prompt 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","hotline")
- End of RunI ~300Hz Prompt-Reco + 300-600Hz of "parked" data
  = reconstructed only after the end of the run
- RunII ~1kHz of Prompt-Reco + ~500Hz of data for parking + high rate (but light) scouting streams

Prompt Reco and PCL

Express processing

Data reconstructed for:

- monitoring
- calibration



Prompt Reconstruction

Stream A data are reconstructed consuming calibrations computed by PCL

These are the datasets for analysis

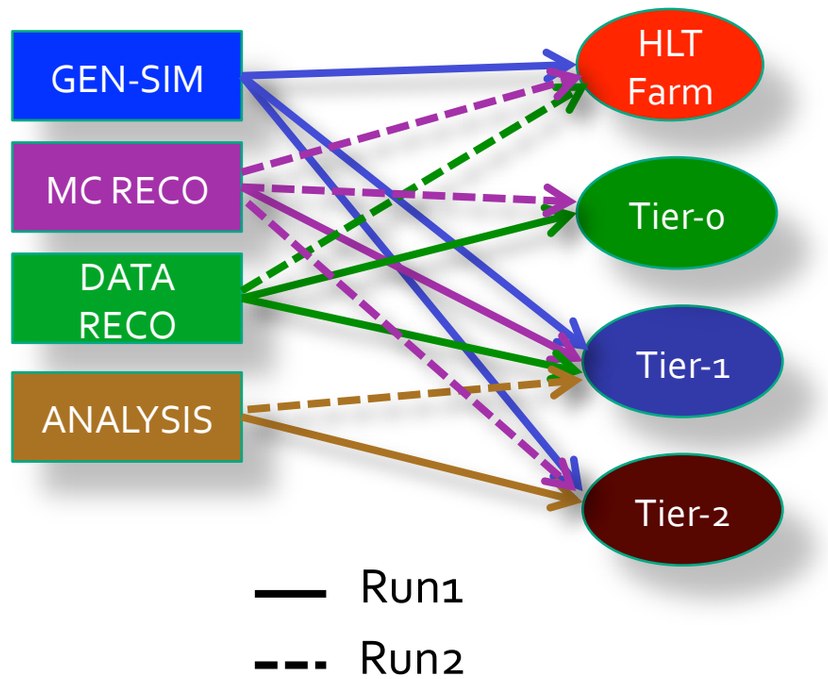
Prompt Calibration Loop (PCL)

Express data are used as input to automated calibration workflows running @ Tier0 (or online)

- beam-spot fix Ls by LS
- ECAL transparency corr.
- SiStrip bad-channels
- SiStrip gains
- SiPixel alignment

Distributed Computing

- "Tiered" computing infrastructure handles distinct use cases
  - **Tier0** (CERN): big farm @ CERN (~15k CPU cores)
    - dedicated to PromptReco (used for other purposes if beam-off)
  - **Tier1**:~10 big centers with large disk/tape cap act (tot. ~25k cores)
    - long term storage and central RECO passes (data and MC)
  - **Tier2**: many smaller centers with disk "buffers" (tot. ~60k 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: fast turn-around analyses or prompt feedback/calibrations
  - Opportunistic resource:
    - big computer centers for limited periods
    - HLT farm when LHC not running (~15k cores)



Contact&Doc

Primary Datasets

- Stream A 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 Tier2 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 (DPG) or Physics Analysis Groups (POG-PAG)
- Dataset & skim definition, is handled by the Dataset Definition Team (DDT) in PPD together with Trigger Study Group (TSG)

Contact&Doc

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 (McM)

Contact&Doc

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`

    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 (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):

Production Campaign + PU scenario + Global Tag

`/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 (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:

Data Aggregation System (DAS): Home | Services | Keys | Bug report | Status | CLI | FAQ | Help

results format: list , 50 results/page, dbs instance prod/global , autocompletion disable Search Reset

dataset dataset=/*/Run2015B*/MINIAOD

Show DAS keys description

mongoDB

Showing 1—26 records out of 26. <first | prev | next | last>

Add filter/aggregator function to the query: grep dataset.created_by + Clear

Dataset: **/Commissioning/Run2015B-PromptReco-v1/MINIAOD**
Creation time: 2015-07-08 19:27:07, Physics group: NoGroup, Status: VALID, Type: data
Release, Blocks, Files, Runs, Configs, Parents, Children, Sites, Physics Groups, py , Subscribe to PhEDEx Sources: dbs3 show

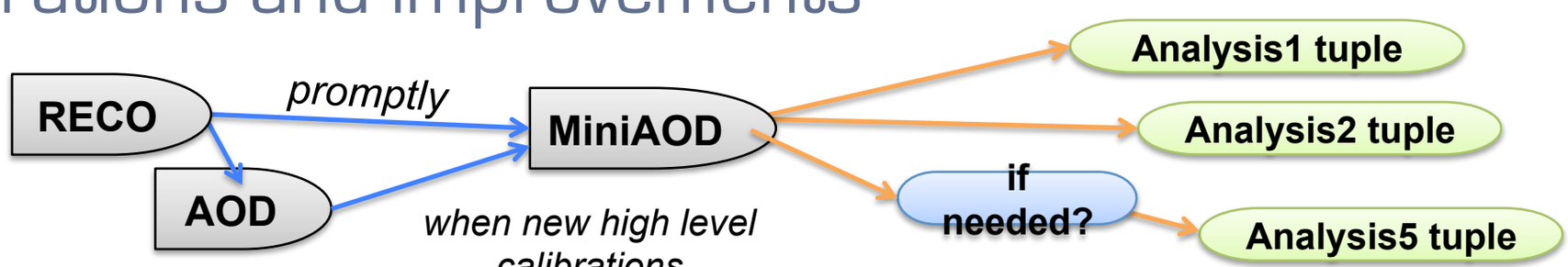
Dataset: **/EGamma/Run2015B-PromptReco-v1/MINIAOD**
Creation time: 2015-07-08 19:20:19, Physics group: NoGroup, Status: VALID, Type: data
Release, Blocks, Files, Runs, Configs, Parents, Children, Sites, Physics Groups, py , Subscribe to PhEDEx Sources: dbs3 show

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 a few months out o prompt-reco
    ⟶ not produced by default in re-reco
  - AOD subset (∼40%) of RECO content meant for analysis
- Now we know enough to go one step further: MiniAOD
  (∼10-15% of AOD)
  - more than a subset ⟶ designed for analysis

MiniAOD

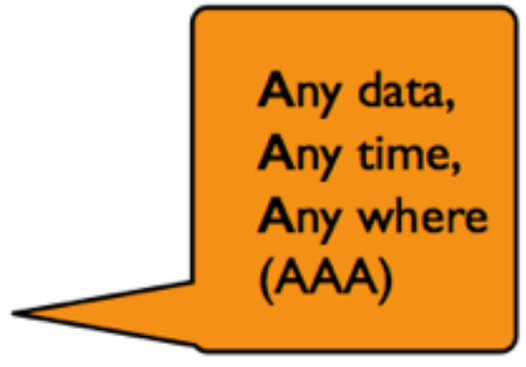
- Lightweight analysis data tier (30-50 kb/ev)
  → serve the needs of ~90% of the CMS analyses
  - High level physics objects (leptons, photons, jets) with detailed information inside
  - Compressed information for all PF candidates, to allow re-clustering jets, re-computing isolation, ...
  - Usable both in full CMSSW and in FWLite/Python
- Centrally produced both for MC and Data
  - can be re-produced from AOD to include newer high-level calibrations and improvements



Contact&Doc

Accessing 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 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 "blocks" of files)
    ⟶ 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)
    ⟶ "standard" analysis use cases addressed by AnalysisOps group, each PAG/POG group has a person taking care of data-management
  - **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 (almost all) datasets accessible remotely via XROOTD protocol
    ⟶ e.g. can run @ CERN reading files @ FNAL



Contact&Doc

Running on Your Data

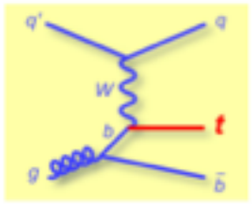
- Once you discovered/moved all the datasets you need to run you analysis ⟶ CRAB

- Jim will give more details on this

CMS Software

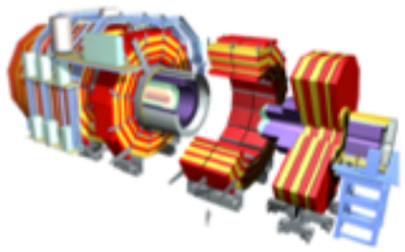
- CMSSW: one release to rule them all
  - **GEN**erator, **SIM**ulation, **RECO**nstruction **ANALYSIS** workflows...
- C++ code and configuration handled via Python
  - "git" used for code versioning and integration
- Release schedule follows a "train model":
  dear developer: catch this train or wait 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: driven by physics/machine constraints & goals
    - e.g. 2014/2015: optimization of Out Of Time PU mitigation for 25ns run

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

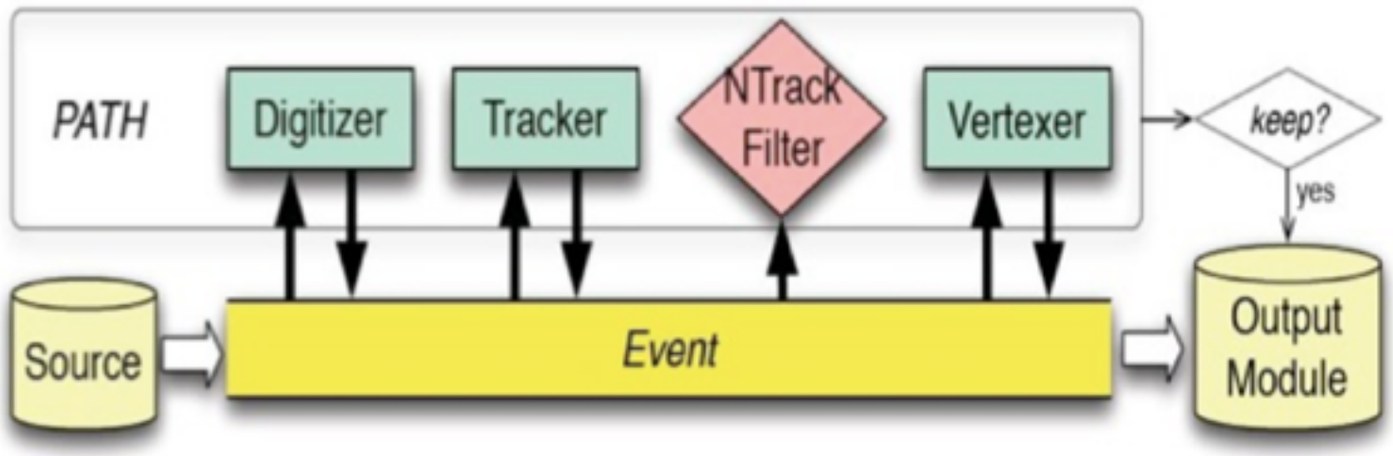
Simulated Raw Data

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

Contact&Doc

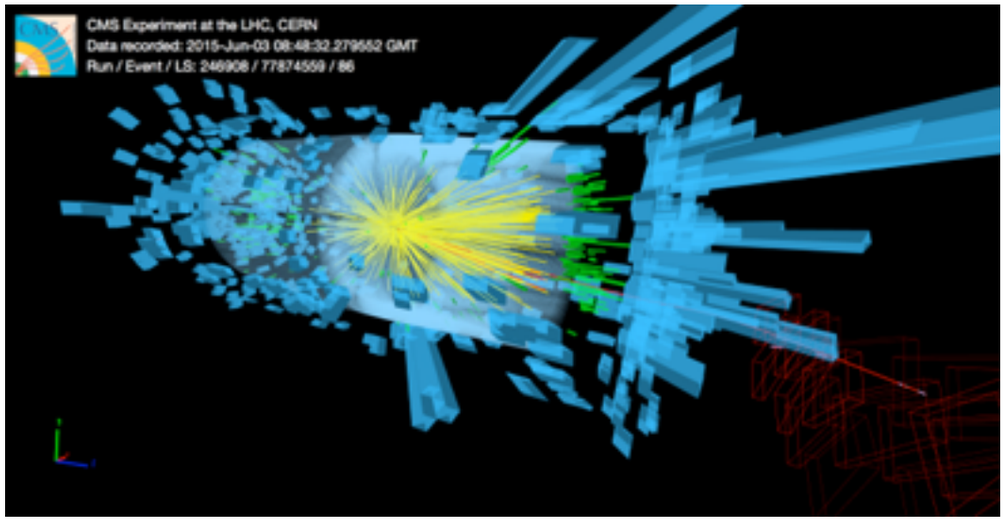
Reconstruction

- CMSSW framework:
  - event reconstruction algorithms are implemented as “modules” communicating via the “Event”



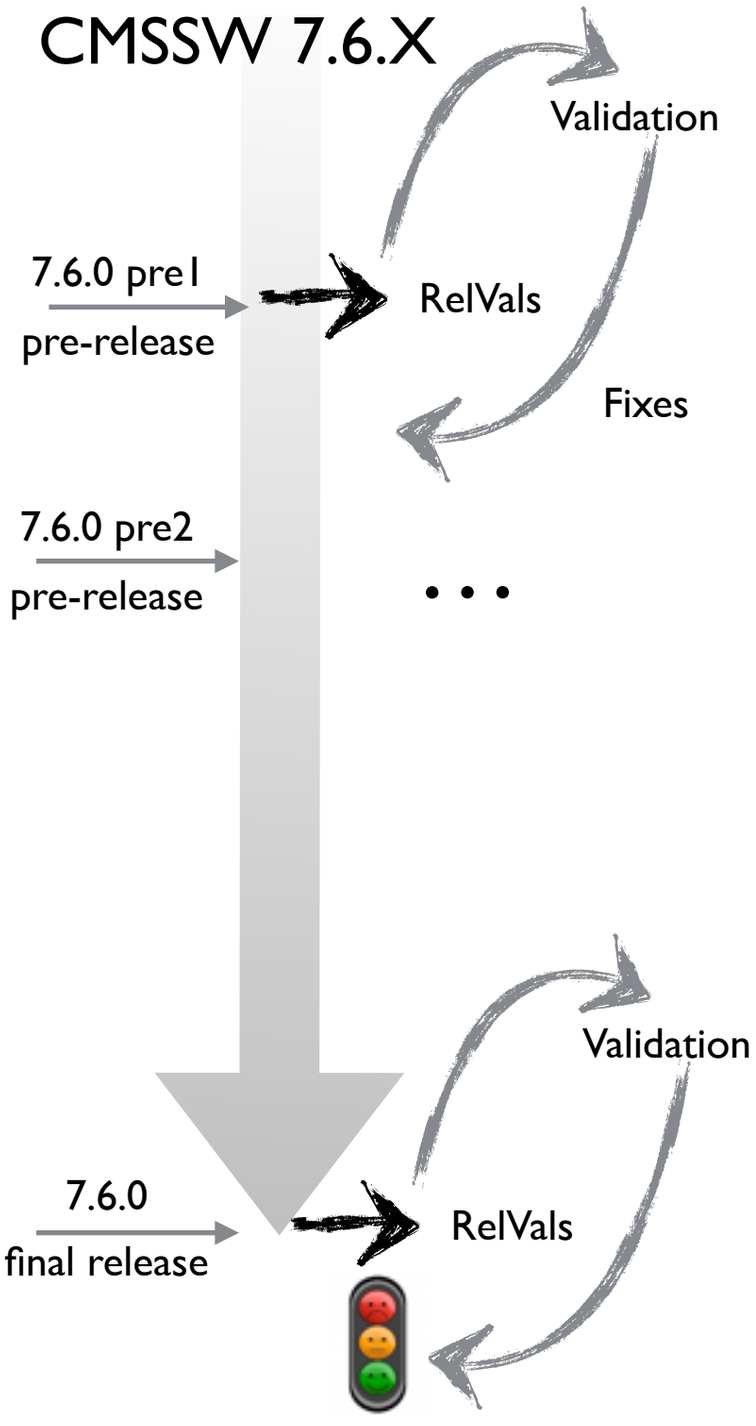
AOD/MiniAOD

- Output of reconstruction (AOD/MiniAOD)
  ⟶ “physics objects” in ROOT format



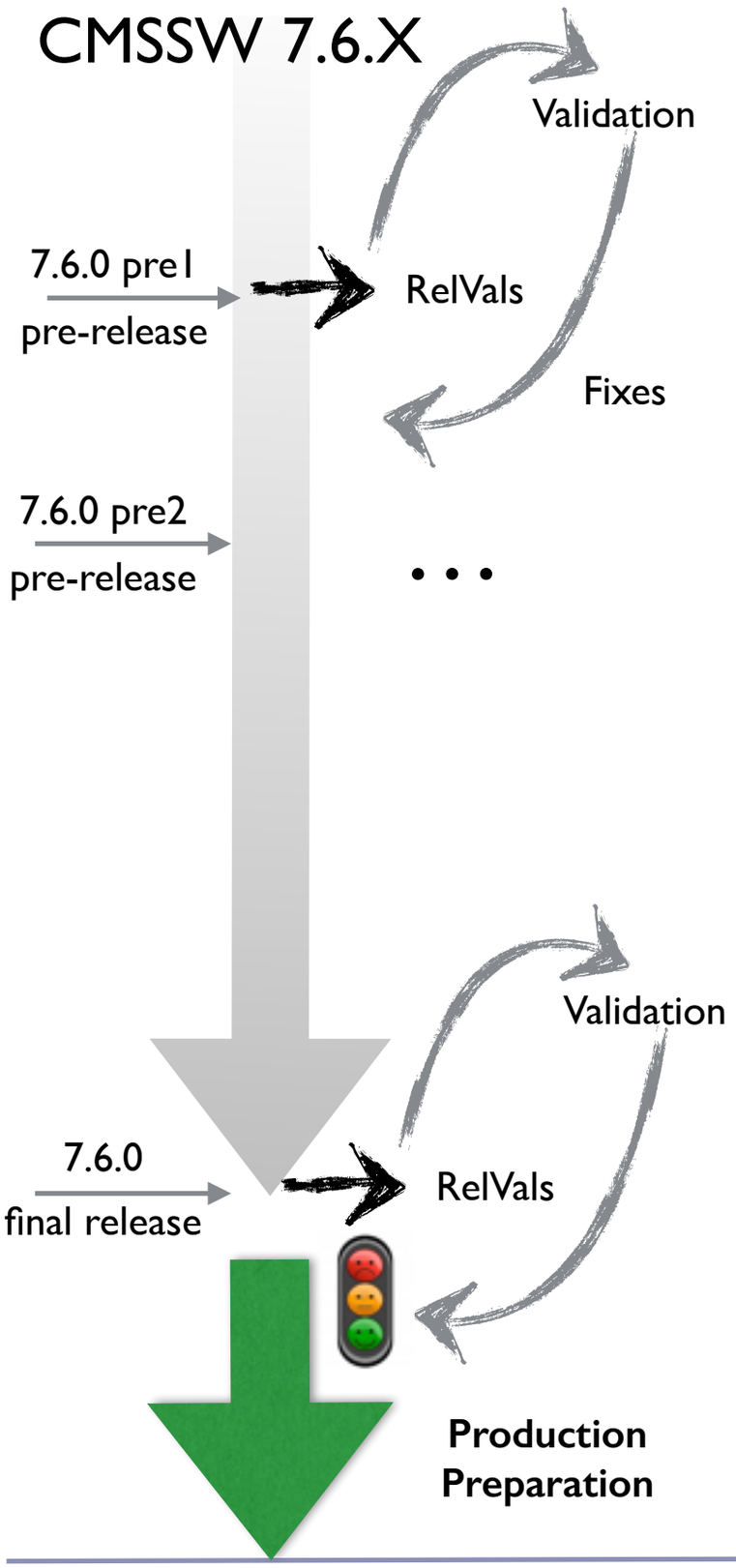
  - Tracks, muons, electrons, photons, jets, etc...

READY FOR PHYSICS?



- Release integration 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:** iterative process performed all along release cycle (from pre-releases to final version)
  1. DPG, POG and PAG validation experts check the plots
  2. PPD/PdmV group coordinates 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

Ready for Physics?



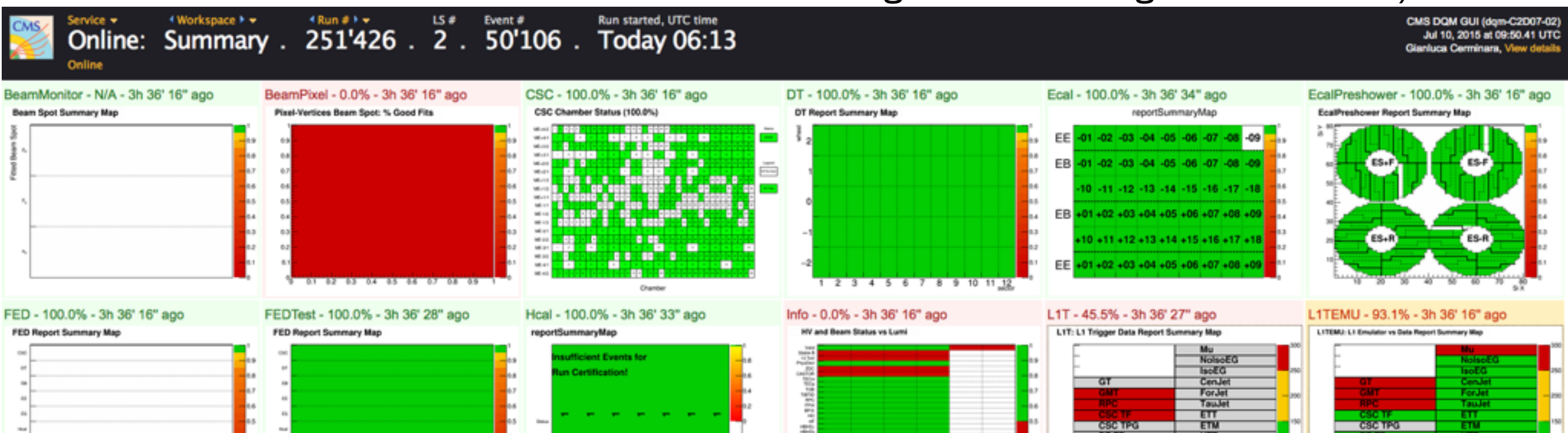
- Release integration bound to Quality Assurance Tests ⟶ Data Quality Monitoring (DQM)
  - unit tests & regression tests
  - small scale production tests: **Rel**ease **Val**idation Test (**RelVal**) producing DQM plots
- Once a major release (X.Y.Z) is green-lighted ⟶ start **preparation of the campaign** (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

Contact&Doc

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

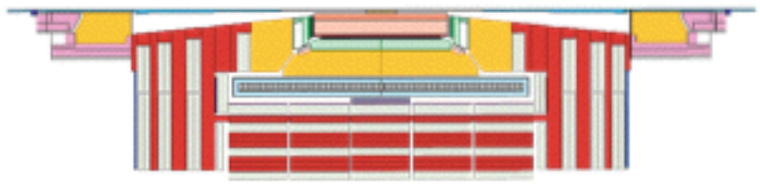


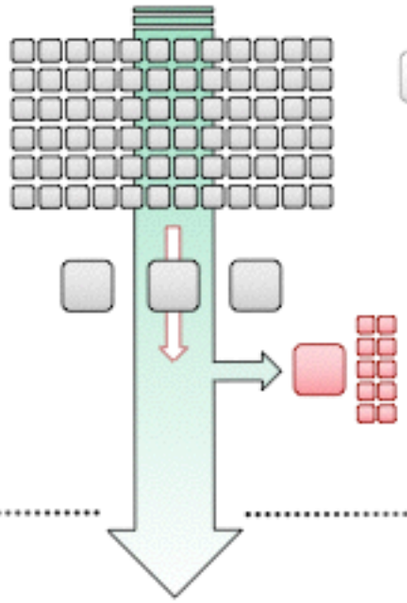
Data Certification

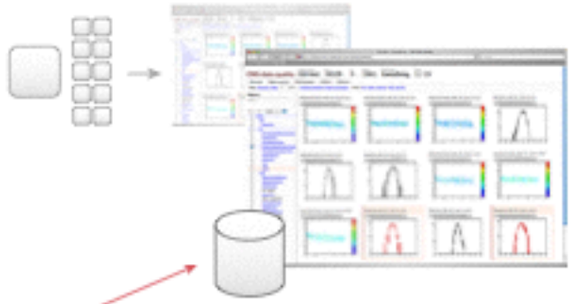
- Complexity of detector and offline processing requires continuous monitor of the quality of the data
- Team of Detector and Physics Object experts check DQM plots for each run and select those usable for analysis
  - look for unexpected effects that could affect analysis level quantities (e.g noise spikes, dead areas of detector, 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

Contact&Doc

Data Certification







Detector & Operation Efficiency

Online

Offline

Conditions

Conditions

Run Registry

certification JSON

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

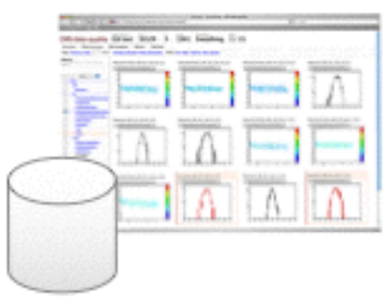
Tier-0

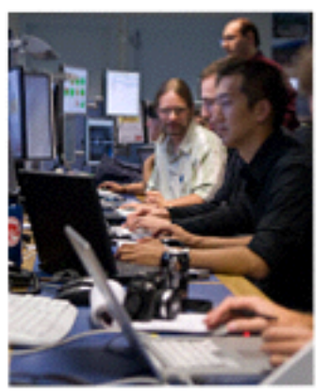
Tier-1s

CAF

Release Validation

Simulation Validation





daytime

2/day

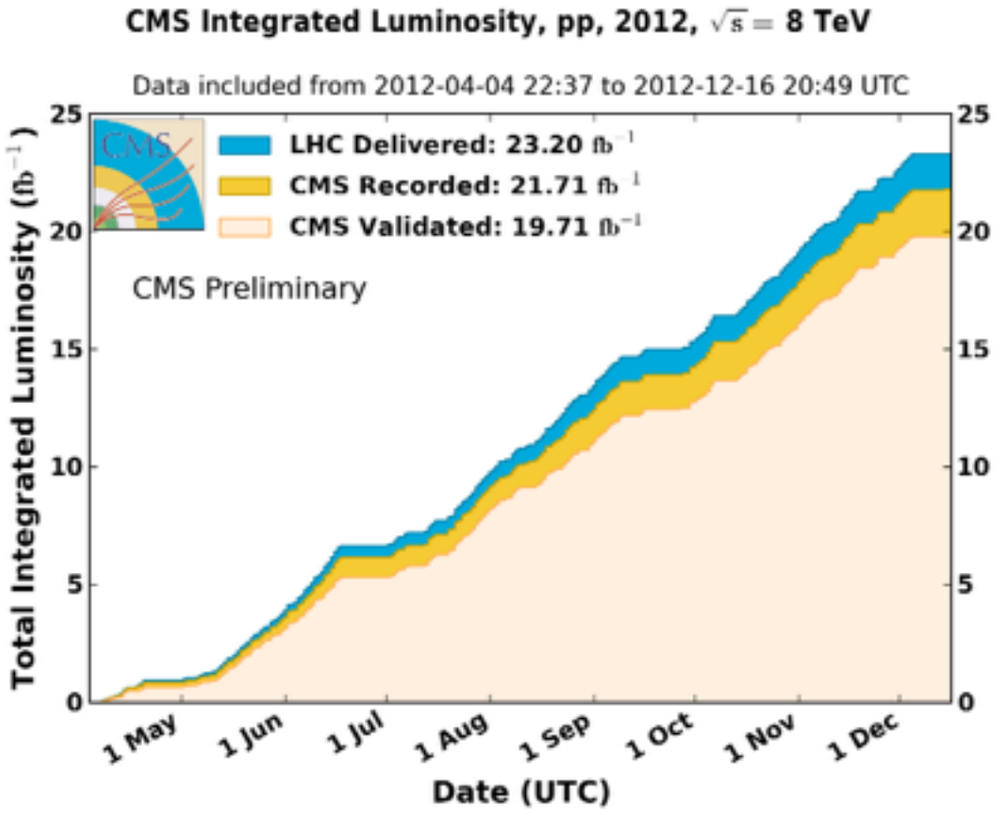
1/day

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 calorimeters
  - **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

LUMINOSITY: HOWTO

- 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}


```
{ "251027": [[1, 60], [61, 61]],
  "251028": [[1, 18], [19, 21], [141, 141]],
  "251143": [[1, 16], [17, 19]],
  "251147": [[1, 24], [29, 29], [30, 187]]
}
```


→ CRAB →

CRAB output {JSON}


```
{ "251027": [[1, 60], [61, 61]],
  "251028": [[1, 18], [19, 21], [141, 141]],
  "251143": [[1, 16], [17, 19]],
  "251147": [[1, 24], [29, 29], [30, 187]]
}
```


- Lumi POG provides tools to compute luminosity starting from JSON file:
  - `lumiCalc2.py`: reports HF measurements of LHC delivered, CMS recorded luminosity for LSs in the JSON ⟶ only for Run1
  - `lcr2.py`: temporary solution for Run2, will soon be replaced by
  - `brilCalc.py`: should provide same functionality of `lumiCalc2.py`
- NOTE: using directly the certification JSON can be problematic for several reasons:
  - LS that failed prompt reco or re-reco will appear in the cert. JSON
  - LS that failed in your jobs will appear in the cert. JSON

Contact&Doc

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 Request Managers, GEN conveners & 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

McM WebService

MccM Meeting
⟶ priority
⟶ monitoring
⟶ GENerators
⟶ McM

Computing "CompOps" team

Contact&Doc

Calibration Workflows

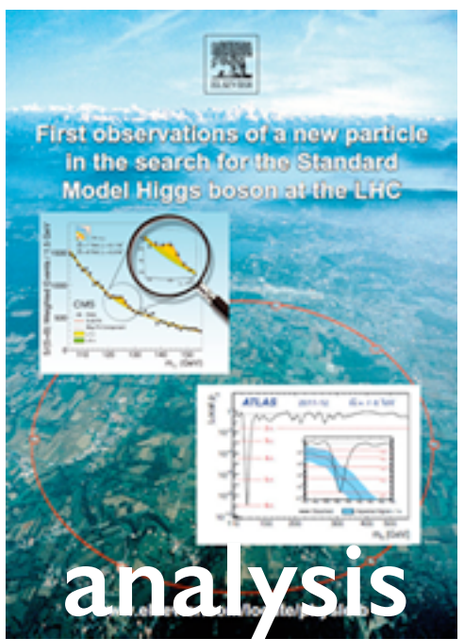
- Provide most up-to-date conditions @ all stages of the data processing
  Different workflows depending on the time scale of updates:
  - **quasi-online calibrations** for HLT and express (e.g. beam-spot used by HLT)
  - **prompt calibrations**: monitor/update conditions expected to vary run-by-run (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 to calibrate and align CMS
- Dedicated GTs for analysis are made available by AlCaDB team
  ⟶ to be used with consistent dataset

Contact&Doc

Outlook

- (some) Physicists tend to think that CMS 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!)

Off&Computing Organization

Offline and Computing Co-Coordinators
M. Girone, D. Lange

Offline & Computing Management Board
D. Bonacorsi, P. Elmer, O. Gutsche, L. Sexton-Kennedy

Project Office
T. Boccali, I. Fisk

R&D
D. Bonacorsi, P. Elmer

Physics Support
S. Belforte, J. Lett, S. Malik, J. Hernandez

Data Management Development
N. Magini, T. Wildish

Workflow Management Development
E. Vaandering

Dynamic Resource Provisioning
D. Colling, C. Grandi, D. Hufnagel

Operations
C. Wissing, C. Pauss

Security
D. Da Silva, M. Altuna

Core Software
C. Jones, S. Muzaffar

Event Generators
J. Bendavid, R. Covarelli

Simulation
M. Hildreth, V. Ivantchenko, S. Sekmen, N. Vanelderen

Reconstruction
S. Krutelyov, C. Vuosalo

Analysis Tools
T. Jeong Kim

Upgrade Software
D. Elvira, M. Grimm

L1 Software
M. Mulhearn

PPD Organization

PPD Coordinators

PPD Resource

Physics Data & MC Validation (PdmV)
Gaelle Boudoul:

Data Quality Monitoring & Data Certification (DQM-DC)

Alignment Calibration & Database (AlCaDB)
Salvatore Di Guida:

DPG and POG Validation
Hengne Li;
TBD

PAG Validation
TBD

Monte Carlo Request Management
Helena Malbouisson;
Carley De Oliveira Martins;
Yong Yang;

Online Development
Broen Van Besien;

Offline Development
Dmitrijus Bugelskis;

Software Coordinator
TBD

AlCa-HLT Contact
Javier Duarte;

DPG, POG & PAG contacts and

References

Page	Contacts	Documentation
3	cms-ppd-coordinator@cern.ch cms-offcomp-coordinator@cern.ch	PPD Main Twiki Offline Main Twiki Computing Main Twiki
9		Computing Model Workbook
10	hn-cms-dataset-definition@cern.ch	DDT Twiki
11	hn-cms-computing-tools@cern.ch	DAS
16	hn-cms-physTools@cern.ch	MiniAOD Workbook
17	hn-cms-phedex@cern.ch	XROOTD doc Phedex - Phedex Workbook
20	hn-cms-offlineAnnounce@cern.ch hn-cms-relAnnounce@cern.ch	Offline Workbook SW Guide
22	hn-cms-relval@cern.ch hn-cms-physics-validation@cern.ch	PdmV Twiki
24	hn-cms-evfdqmannounce@cern.ch	DQM Twiki
25	hn-cms-data-certification@cern.ch	DQM-DC Twiki RunRegistry
27		JSON File Twiki
28	hn-cms-luminosity@cern.ch	LumiCalc Twiki Lcr2 Twiki
29	hn-cms-prep-ops@cern.ch	PdmV Twiki McM
30	hn-cms-alca@cern.ch	AlCaDB Twiki GlobalTag Twiki