Data Preparation for the CMS Detector at 8TeV at the LHC

Jean-Roch Vlimant (CERN)
on behalf of the CMS Collaboration

ICHEP 2012, Melbourne, VIC, Australia
4-11 July 2012
Outline

- Overview of CMS Offline Operations
  - Data And Monte-Carlo Validation organisation
  - Software Development for 2012 LHC collisions
  - High Level Trigger preparation
  - Alignment and Calibration for data taking and simulation
  - Computing infrastructure upgrade for increases Pile-Up
  - Data Certification of highly important collision data
- Conclusions and outlooks
Certification: masking of luminosity block (~1min worth of data) based on detector conditions and performance of event reconstruction

Validation: verification that changes introduced are preserving or improving the Physics through put of data taking, from detector, object reconstruction and analysis point of view.

♫ are major parts of a new organization group emerged in 2012 from experience gained with Offline operation dedicated to Physics Performance and Dataset.
Validation On Data and Monte-Carlo

- Automation of workflows and procedure towards simplifications of validation tasks
  - Central production, automated comparison, centralized communications
- A strong core of people involved in validation on Monte-Carlo and Data have been joined by collaborators from Physics groups to scrutinize the area of the phase space after analysis selection.
- Book keeping tools have been developed to simplify the work of collaborators, and increase the visibility of their work.
- Consolidating the workflows on under-covered area of production
- Working towards the validation of the software for the upgrade of the CMS detector, to adapt to upgrade of the LHC in the coming decade.
Software Development for 2012

- Trying to cope with ever-improving LHC operation, more luminosity, more added pile-up events, more complex events
- Improvement in computing performances of the CMS software
- Physics performance of event reconstruction unaffected under technical modifications; Improved under algorithm development.
  - Phase 1 in 2011 to cope with increased luminosity
  - Phase 2 early 2012 to prepare for increasing luminosity and favor increased trigger rate.
- Main gain was achieve in tracking algorithm optimization
- Algorithm optimization and redesign, compiler architecture, memory management improvements, root version all played a constructive interference role

Event processing time remained constant with increased pile-up (<30s/evt)
High Level Trigger Upgrade for 2012

- Computing upgrade for 50% more CPU available (150ms/evt) to cope with inevitable increase of processing time of events, and allow for more bandwidth to tape.
- Useful Trigger rate is limited to how much can be reconstructed in “real time” (~300Hz average).
- Effective trigger rate can be used for dataset, which will be reconstructed during 2013 shutdown (~300Hz average)
- Improvement in computing performances of the CMS software decreased limitation on CPU extensive algorithm
- Reduced trigger rate obtained with improved resolution at all levels
  - Improved reconstruction algorithms, calorimeter calibration, noise rejection, cross triggers, pile-up subtraction, particle flow jets, noise cleaning
  - See J.Brook and S.Beauceron presentations for more details
Calibration for 2012 Data Taking

- Optimization of alignment and calibration producers played a role in the improved computing performance of CMS software in 2012.
- Technical improvements in database and monitoring simplified detector operation.
- Conditions were scheduled and delivered
  - Prior: Delivery of improved conditions when rolling in the software release for 2012 data taking and Monte-Carlo productions.
  - During: Experience gained in 2011 for Prompt Calibration Loop (PCL). New calibration introduced, with monitoring to protect the events reconstruction in “real time” from being corrupted.
  - After: Electromagnetic Calorimeter calibration has been derived for a timely reprocessing of critical dataset.

![Distribution of the median of the residuals in TPE](image1)

#modules

![Relative E/p scale](image2)

![CMS 2012 Preliminary](image3)
Computing Infrastructure for 2012

- Improved computing management system was finalized and commissioned 2012 data.
- Improved traceability and monitoring of production.
- Computing operation was re-organized based on previous experience from previous years of operation.
- Simplified operation and improved coordination showed great success
- Resource management permitted to deliver more than the prediction
- Lesson learned from 2012 already turned into future projects for improved productivity
Monte-Carlo Production

- Events Simulation Production is staggered
  - Event generation @ 8 TeV: everything had to be redone (4 Billion events)
  - 2012 specific pile-up profile to maintain statistical power of simulation
  - Initial batch of samples for early data comparison
  - Prioritized samples for ICHEP
  - Pre-approval cross-checks samples in last weeks
- Improved coordination from analysis to production allowed for deliveries for all high priority analysis.

![Graph showing rate of production in M/month for Summer12 in GEN-SIM and related samples](image-url)
Data Certification For ICHEP

- Optimization of data certification producers played a role in the improved computing performance of 2012
- The certification workflow has been unchanged for 2012, since the latency resides mostly in the time it takes for shifters, experts, and conveners to coordinate the best understanding of the detector and physics content
  - Prompt reconstruction is certified weekly
  - Reprocessings are certified within 2 weeks
- All data until June 19\textsuperscript{th} (before technical stop of LHC) has been certified and delivered to analyst for this conference.
Summary and Outlook

- Validation of new software, new conditions, and new trigger menus on Data and Monte-Carlo has been consolidated to stream-lined deployment of ever-better CMS data taking envirronement.

- A large number of small improvements on CMS software with regards to computing time and memory consumption have been put together to sustain production in 2012 LHC operation mode; without compromising the Physics through put.

- Improved detector calibration and alignment were properly scheduled and timely delivered to ensure the best data quality for CMS.

- Improved computing management mechanism has been commissioned for better managed CMS resource and increased productivity towards delivery of the crucial dataset of 2012 data.

- Planning for consolidation and looking ahead into the deployment of software and conditions for the CMS upgrade.