The CMS Computing, Software & Analysis Challenge

Rainer Mankel (DESY / CERN)
David Futyan (Imperial College)
Christoph Paus (MIT)
for the CMS collaboration

Computing in High Energy Physics Conference 2009
Prague, 24 March 2009

Reasoning & Scope

The Computing, Software & Analysis Challenge 2008

- Full-scale computing, commissioning & physics challenge with large statistics under conditions similar to LHC startup
  - [pre-production of MC samples at various tiers] — Centrally operated
  - prompt reconstruction at T0 — Centrally operated
  - skims for alignment & calibration — Centrally operated
    - ➔ reduced form of reconstructed data, containing precisely the minimal information required as input to a given calibration/alignment algorithm (“AlCaReco format)
  - alignment & calibration “in real time” at the CERN Analysis Facility (CAF) — Alignment & calibration teams
  - re-reconstruction at T1 — Centrally operated
  - physics analysis at T2 and CAF — Physics analysis teams

CSA and CCRC

- CSA08 took place concurrently with the LHC Common Computing Readiness Challenge (CCRC08)
  - additional centrally operated CMS workflows to generate computing load
  - ➔ fixed time scale, no delays accepted
    - all CSA08 production targeted to end on 2-June

See also: Challenges for the CMS Computing Model in the First Year (Ian Fisk)

The Computing, Software & Analysis Challenge 2008 (cont'd)

- This challenge placed strong emphasis on handling alignment & calibration under LHC start-up conditions
- Initial mis-alignments & -calibrations as expected:
  a) before collisions,
  b) after 1 pb^{-1} of data

➔ Situation significantly different from the one at LHC design luminosity (→ Physics TDR)
  - not yet a high rate of "golden" event signatures
    - example: $Z^0 \rightarrow \mu^+\mu^-$ decays for alignment
- Full complexity of many concurrent alignment & calibration end-to-end workflows (with interdependencies)
- Realistic analyses based on the derived constants

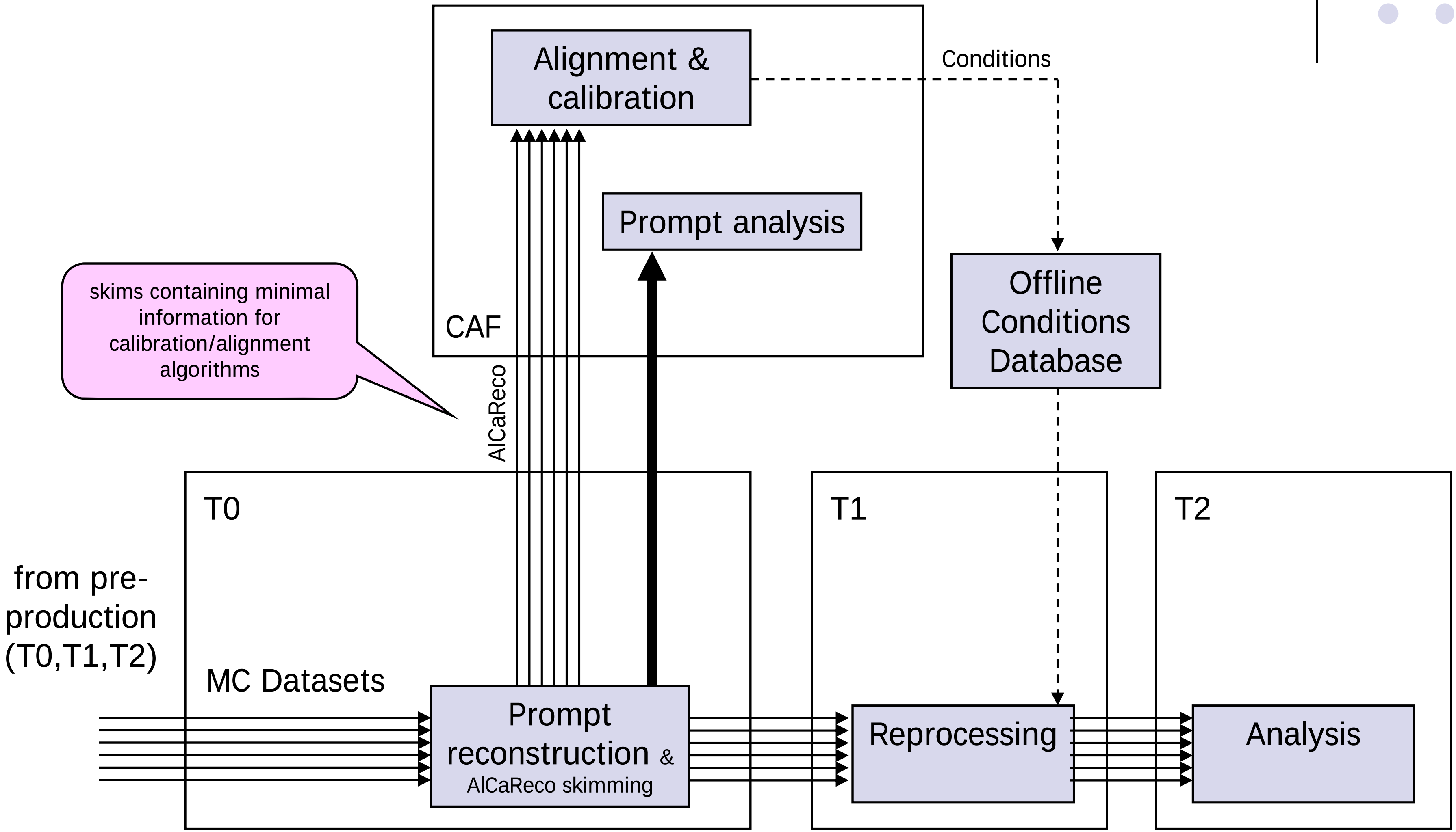
The CSA08 Scenarios

- Assumed two scenarios as they are expected to appear during the beam commissioning of the LHC:

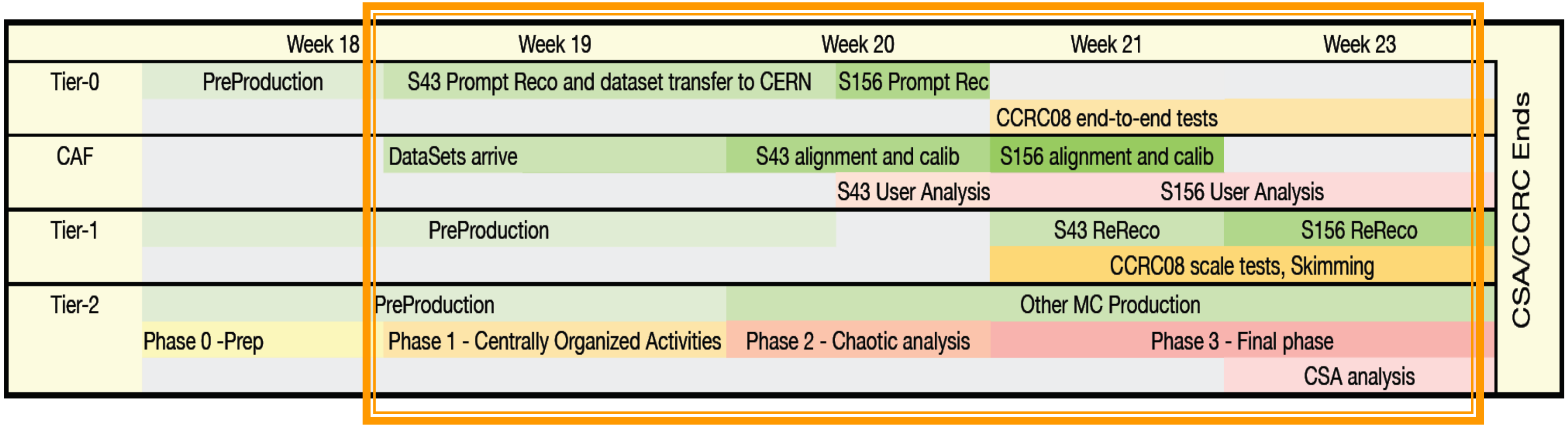
Name	Bunch schema	Luminosity	Duration [effective]	Integrated luminosity	HLT Output Rate	#Events
S43	43x43	$2 \cdot 10^{30}$ $cm^{-2}s^{-1}$	6 days	1 pb^{-1}	300 Hz	150 M
S156	156x156	$2 \cdot 10^{31}$ $cm^{-2}s^{-1}$	6 days	10 pb^{-1}	300 Hz	150 M

- Consequently, the data are governed by low luminosity
  - dominated by minimum bias, jet triggers
  - small sets of high p_T leptons & Z^0 decays
  - non-collision samples:
    - cosmic muons passing tracker
    - HCAL noise

Offline Workflow in CSA08



CSA08/CCRC08 Schedule



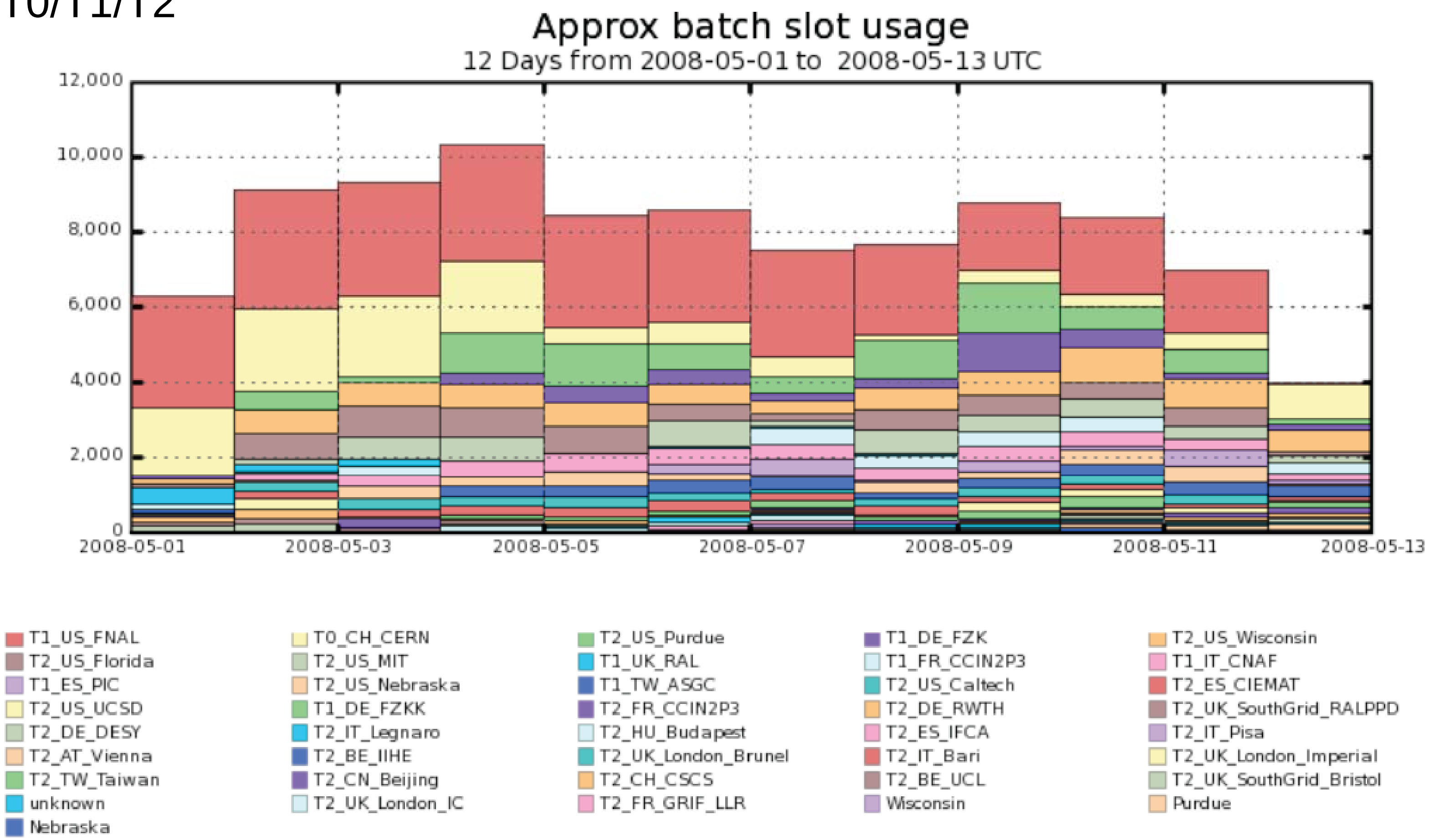
May 08

- Both the 1 pb^{-1} and 10 pb^{-1} data samples are each based on a week of “simulated” data-taking
- Planned O(1 week) for prompt reconstruction
- Target for alignment & calibration: constants ready after 1 week for each sample
- The essential milestones of the CSA08 challenge have been kept

Computing Performance

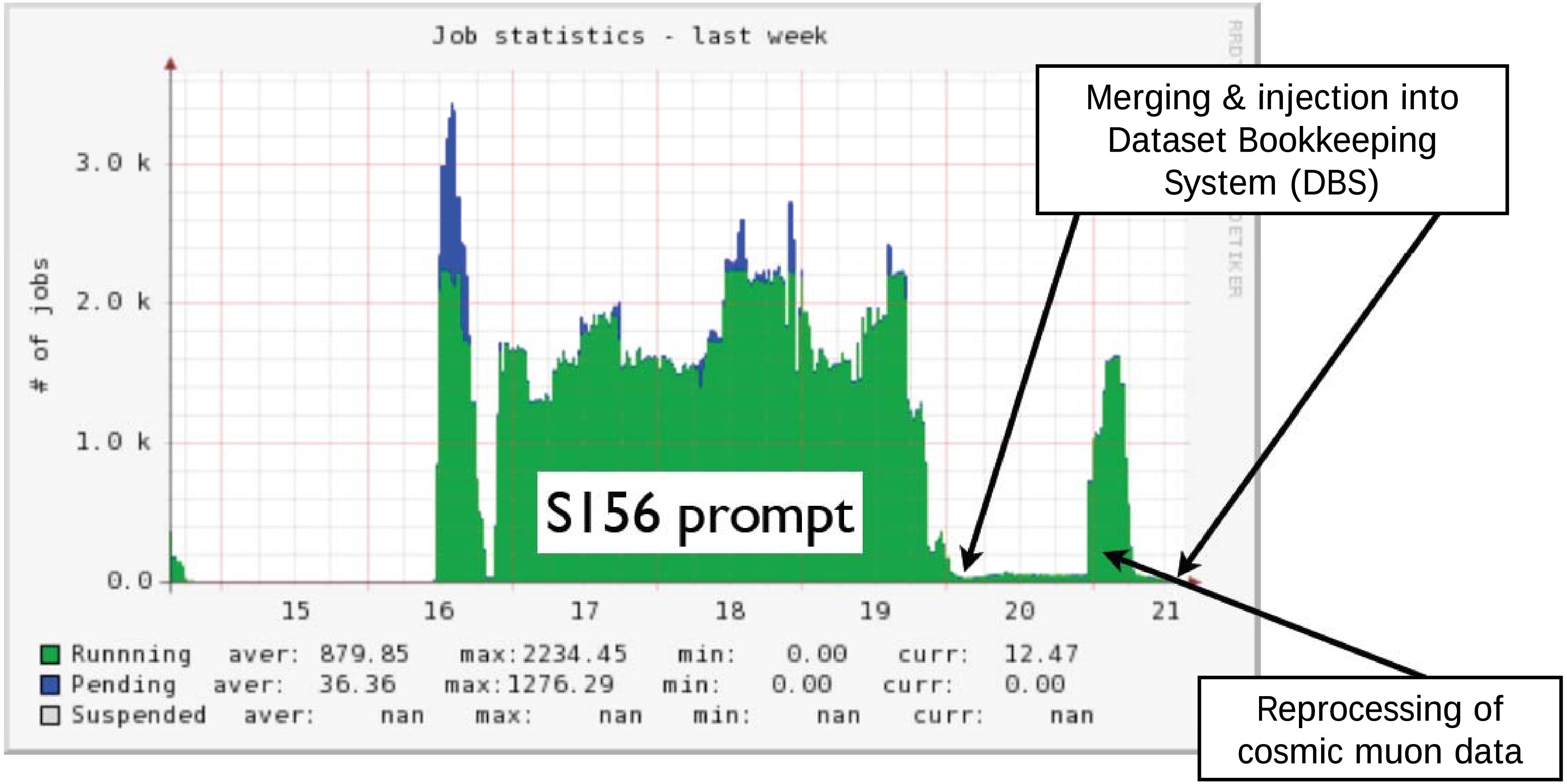
Computing Performance: Pre-Production (Event Simulation)

- On average ~8000 concurrent jobs, at all WLCG tier levels: T0/T1/T2



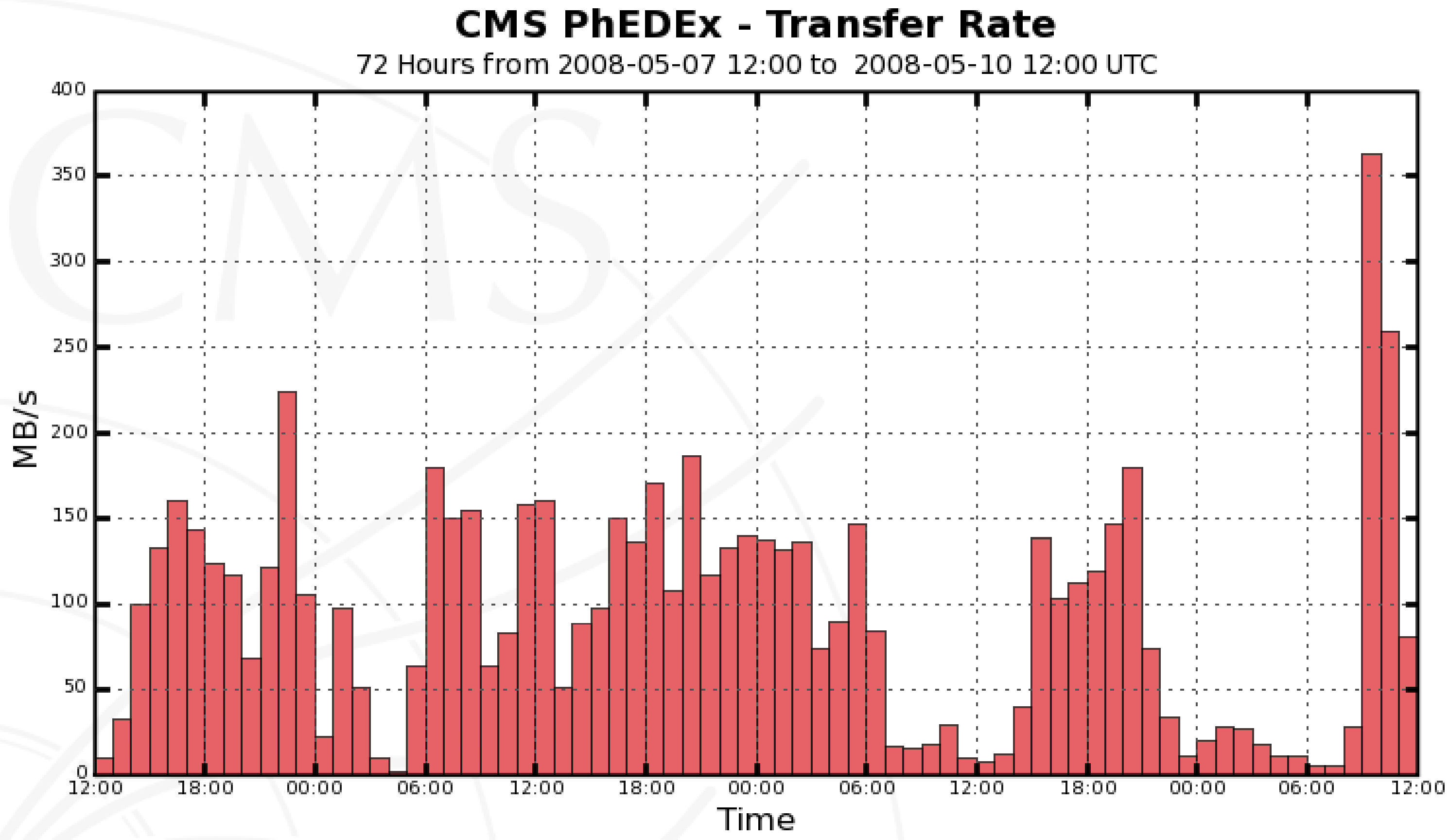
Computing Performance: Prompt Reconstruction (at T0)

- 150 M events reconstructed in less than 4 days



Computing Performance: Data Transfers into CERN

- Pre-production: transfers from various T1+T2 into CERN
- Driven by production. (Not saturating capacity)



Alignment & calibration workflows in CSA08

Note: workflows were performed "in real time"
→ no additional optimization possible

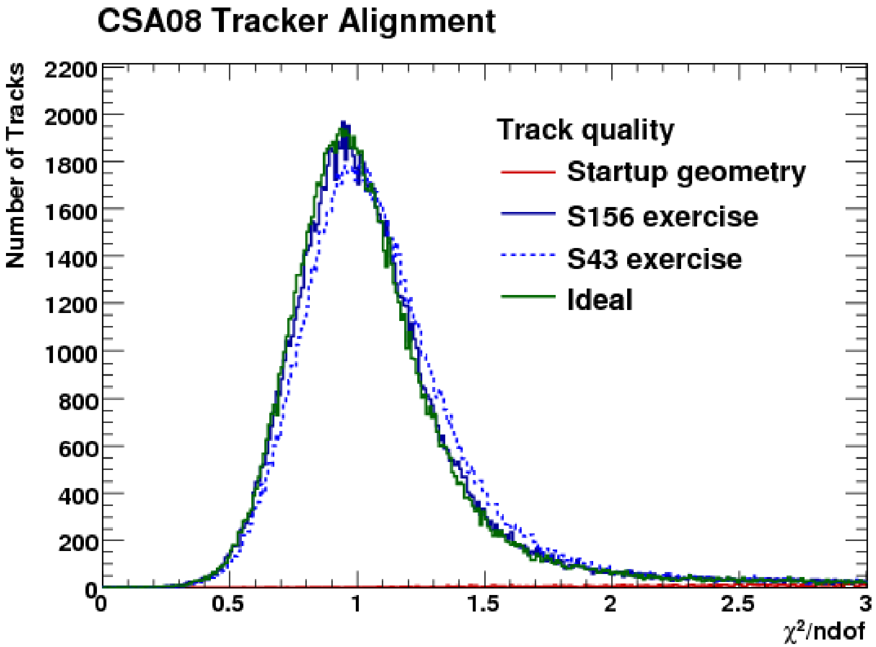
Alignment & Calibration in CSA08

- The following alignment & calibration workflows were performed:
  - Tracker alignment with MillePede-II, HIP & Kalman filter algorithms
  - Muon system alignment with MillePede-like and HIP algorithms
  - ECAL calibration exploiting ϕ-symmetry, & using response from $\pi^0 \rightarrow \gamma\gamma$ and $Z \rightarrow ee$ decays
  - HCAL calibration exploiting ϕ-symmetry, single-pion response & balancing with di-jet signatures
  - Muon drift tube calibration: time pedestals & drift velocity
  - Pixel tracker calibration: Lorentz angle
  - Strip tracker calibration: Lorentz angle & cluster charge
  - Determination of beam spot (before & after alignment)

See also: Commissioning the CMS Alignment and Calibration Framework (David Futyan) [Poster]

Tracker Alignment

- Several algorithms used:
  - HIP (Hit and Impact Point)
  - Kalman filter
  - MillePede-II (shown)
- Results:
  - 1 pb^{-1} (S43): only minimum bias (6.6M) and muon (p_T>5 GeV) samples used
  - 10 pb^{-1} (S156): cosmics, muon (p_T >11 GeV) and di-muon samples added



➔ Significant improvement of track quality

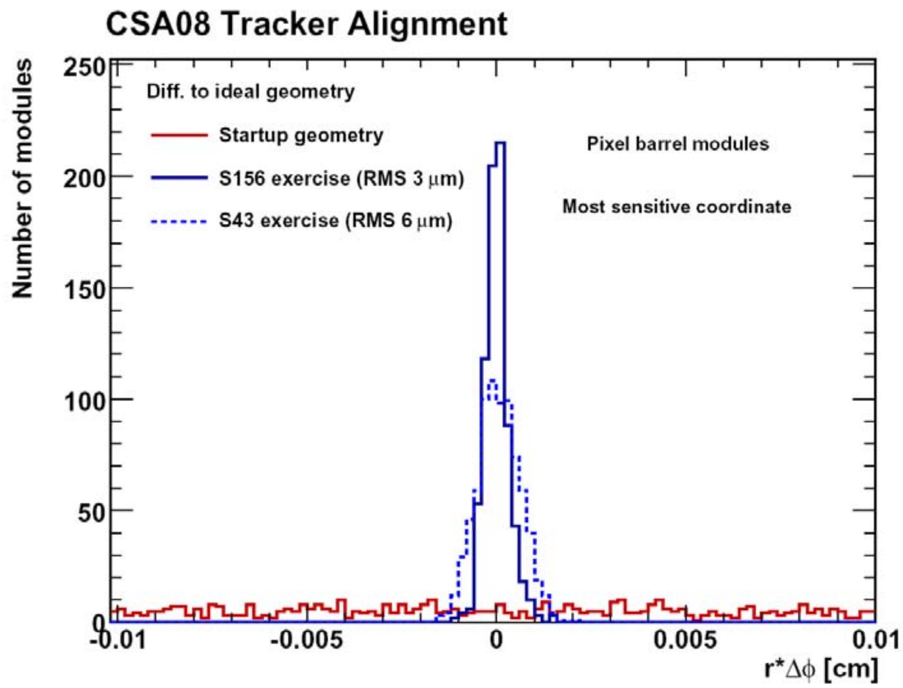
➔ distribution of track χ^2 / n_{DF} already close to ideal

See also: Application of the Kalman Alignment Algorithm to the CMS Tracker (Edmund Widl) [Poster]

Tracker Alignment: Accuracy

- Precision relative to true geometry, after undoing global shifts & rotations
  - quality of internal alignment of these structures

	rφ precision [μm] from MillePede-II		
Tracker Subsystem	**Startup***	**S43 (1 pb^{-1})**	**S156 (10 pb^{-1})**
Barrel Pixel	105	6	3
Tracker Inner Barrel	482	24	10
Tracker Outer Barrel	106	30	23
Forward Pixel	120	48	48
Tracker Inner Disks	445	48	38
Tracker End Cap	92	29	26



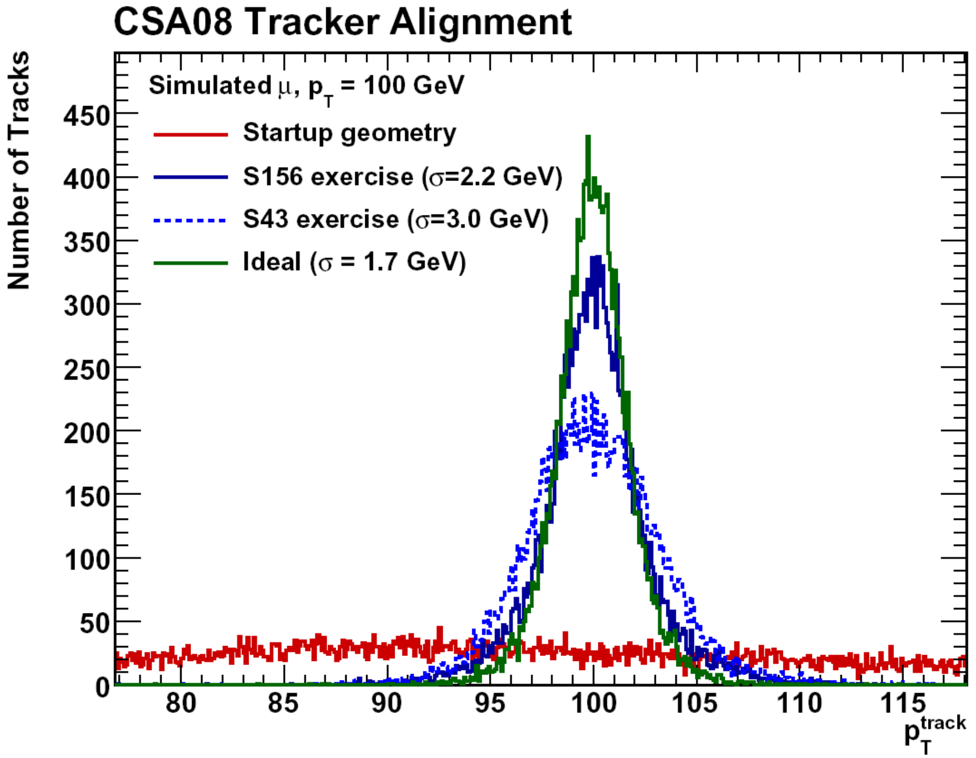
*The expected "startup" alignment will be revised according to the results of extensive data-taking with cosmic muons

Tracker Alignment (cont'd)

- p_T resolution at high momentum very sensitive to coordinate resolution & thus to alignment
  - also systematic effects (e.g. due to weak modes) can show here

➔ Visible improvement (Gaussian fits):

MillePede S43	3.0%
MillePede S156	2.2%
Ideal	1.7 %



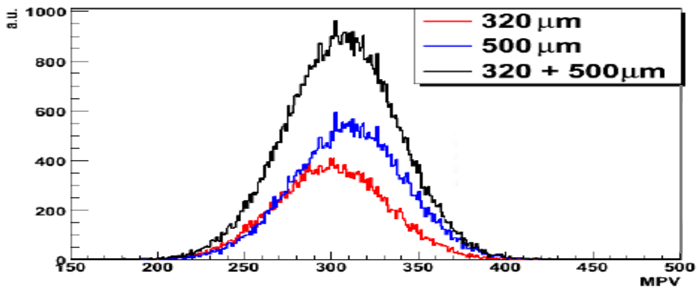
➔ Underlines crucial rôle of cosmics

Tracker Calibration

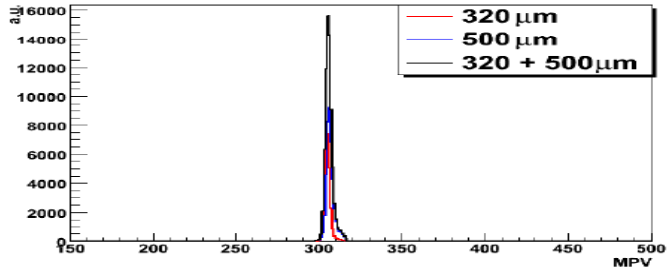
- Cluster charge calibration of the strip tracker
  - artificial mis-calibration: 5% in S156 (10% in S43)
  - 23 M minimum bias events
  - fit most probable value (MPV) of cluster charge spectrum for each sensor (Landau) → calibration factor
  - sharp peaks after calibration, calibration accuracy <1%
- Lorentz angle calibration of pixel tracker
  - using “grazing angle” technique
  - applied on global muon tracks
  - error of global fit 0.1%

See also: The CMS Tracker calibration workflow: experience with cosmic ray data (Simone Frosali) [Poster]

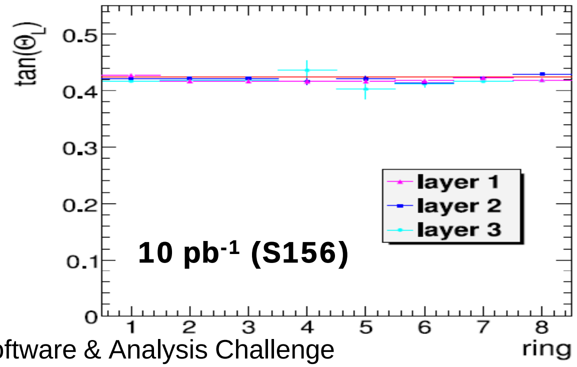
MPV before calibration



MPV after calibration

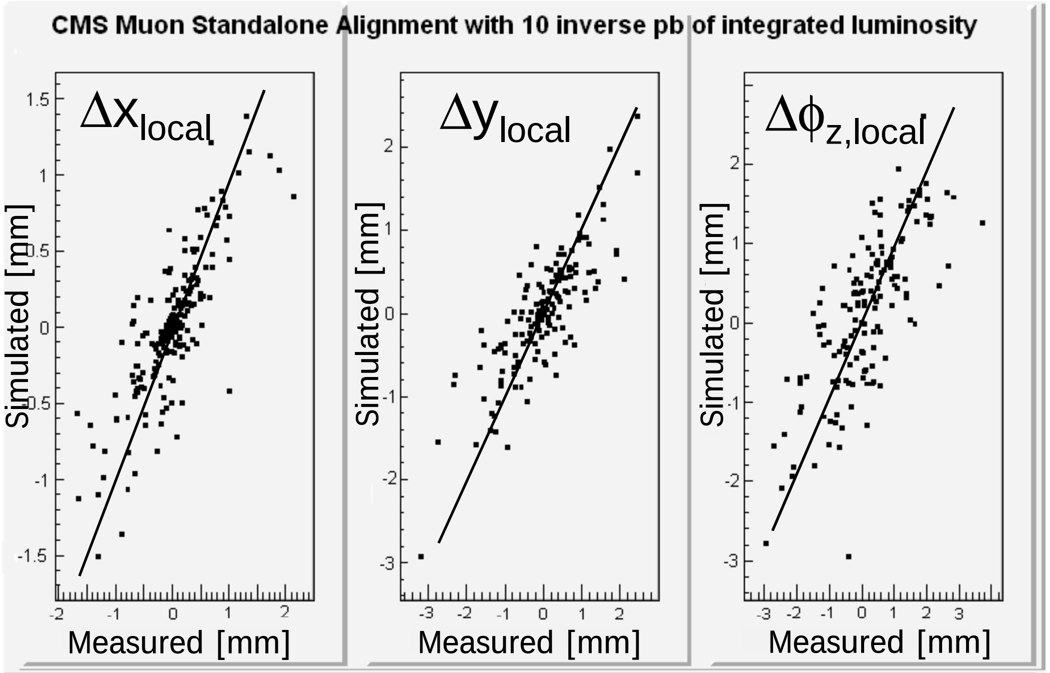


Lorentz angle after calibration



Muon System Alignment I

- Caveat: normally we expect to need 50-100 pb^{-1} to align the muon system
- Try internal alignment of barrel muon system using Millepede-like algorithm
- With 10 pb^{-1} sample, see first correlation between measured and simulated misalignments



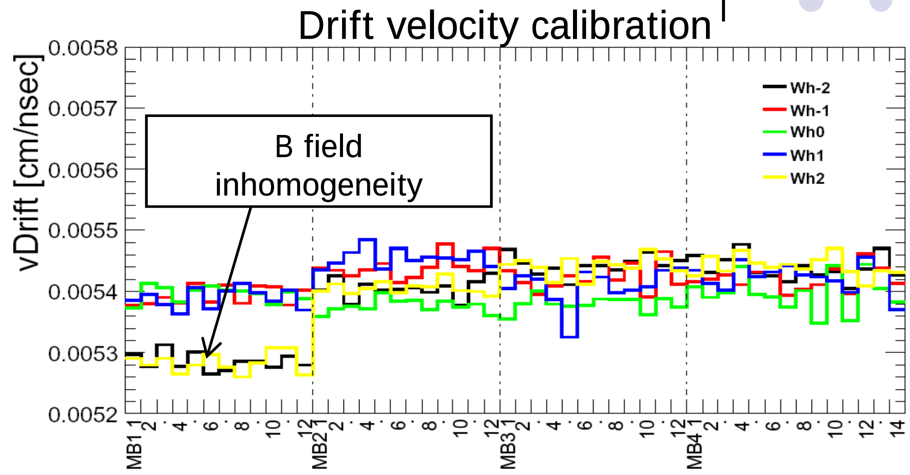
- Typical accuracy 700-800 μm in measurement direction
  - ➔ as expected, limited by number of high-p_T muons
  - ➔ need more integrated luminosity for accurate alignment
- Also alignment of muon chambers with tracker as reference (HIP algorithm) successfully operated

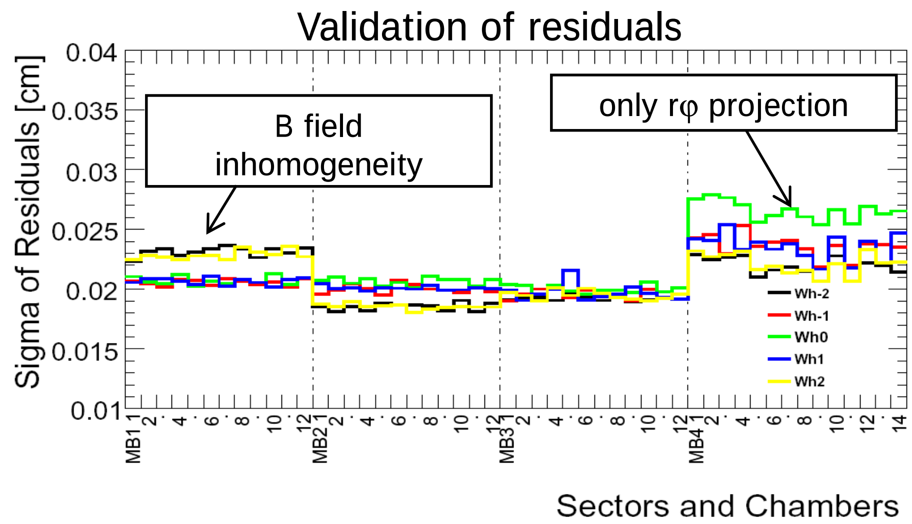
See also: The CMS Muon System Alignment (Pablo Martinez)

Calibration of Muon Drift-Tube Chambers

- Time pedestal calibration
- Drift velocity calibration
  - using “mean timer” method
- Homogeneous results for drift velocity of ~54.2 μm/ns
  - as expected, lower values for inner chambers of wheels near end cap regions (non-linearities due to inhomogeneous stray field)
- Analysis of residuals from 3D segments gives measure of resolution after calibration
  - as expected, higher values for inner chambers of wheels near end caps regions (non-linearity), and for MB4 chambers (only one projection available)

See also: Calibration of the Barrel Muon DT System of CMS (Silvia Maselli) [Poster]

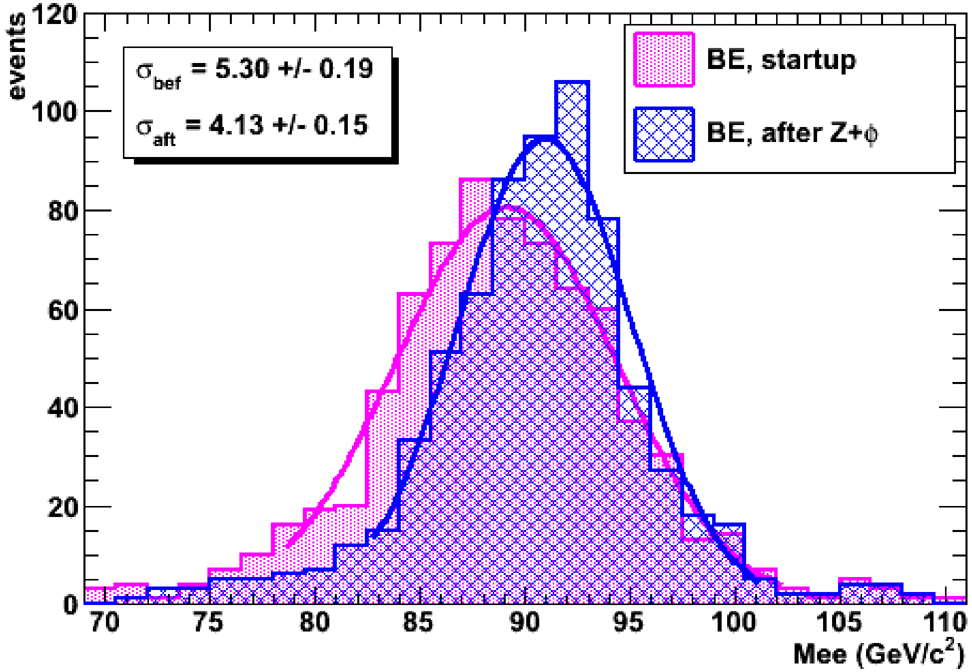




Calibration of Electromagnetic Calorimeter

- At startup, ECAL will already be pre-calibrated at a level of ~1.5 % (barrel) and ~10% (end caps)
- Exploiting the ϕ-symmetry of minimum bias events, the residual mis-calibration in the ECAL end caps is reduced to a few percent soon after startup
  - 20 M minimum bias events used (10 pb^{-1} sample)
- Z decays with one electron in barrel and one in end caps validate inter-calibration of barrel and set absolute energy scale

$Z \rightarrow e_{\mathbf{B}arrel}\ e_{\mathbf{E}ndcap}$



Physics Analyses Based on CSA08 Data Samples

- Physics analyses were carried out in four main areas:
  - measurement of charged particle spectra & analysis of the underlying event
  - early observation of muons, measurement of the di-muon mass spectrum, observation of J/Ψ, Υ and Z resonances
  - early observation of electrons, observation of the Z resonance
  - early observation of jets, their corrections and the extraction of early jet physics
- These analyses were carried out:
  - during CSA08 using prompt S43 + S156 reconstruction, and re-reconstructed S43 data
  - during the 2 weeks following CSA08 using re-reconstructed S156 data

➔ Important validation of alignment & calibration constants

Lessons

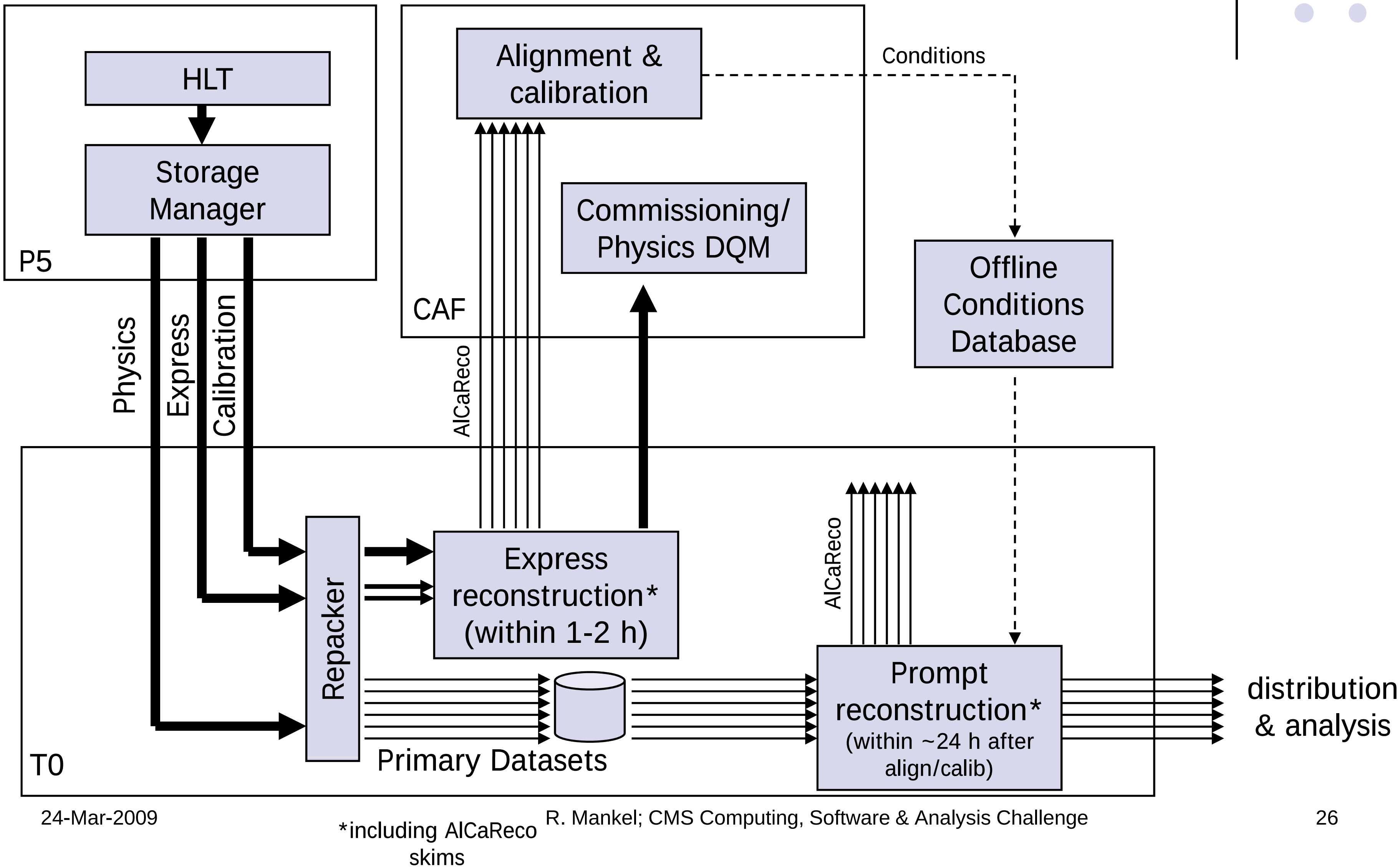
- Computing
  - though pre-production & prompt reconstruction were partly concurrent, overall traffic was still manageable
  - overhead in merging & registration procedures observed → corrected
- Alignment & calibration
  - interdependencies turned out to be very important
    - tracker alignment & muon system alignment
    - tracker alignment & Lorentz angle calibration
    - beam spot & alignment
  - ➔ these were properly addressed in the 10 pb^{-1} workflows
  - all alignment & calibration workflows technically fit into a 24h window
    - important for prompt calibration workflow
- Note: due to e.g. the extended runs with cosmic muons, in several aspects CMS initial alignment & calibration in reality will be better than assumed for CSA08

Summary

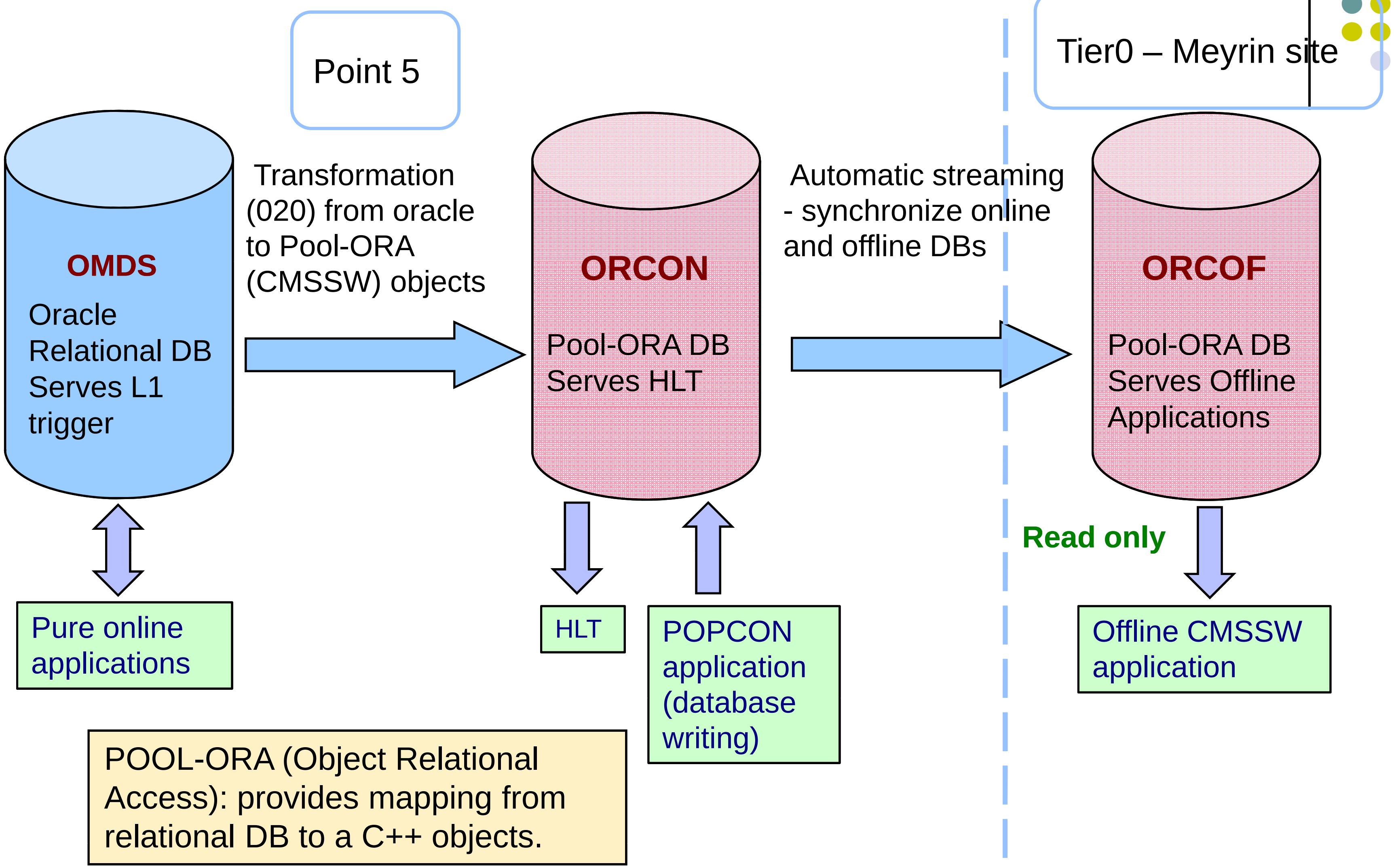
- CSA08 has successfully demonstrated significant components of the CMS computing workflow
- In particular the alignment & calibration framework has been successfully proven
  - 1 pb^{-1} & 10 pb^{-1} exercises completed on time by all sub-detectors
  - all required constants uploaded to the production database
  - re-reconstruction could proceed on schedule
- Organizational challenges were mastered
  - complexity of a large number of workflows
  - inter-dependencies between workflows
  - management of database conditions
- Realistic physics analysis performed with low latency
  - preparation for early observations with LHC

Additional Material

CMS Design Offline Workflow (with Prompt Calibration)



The Conditions Database



Frontier

- For reading, ORCON and ORCOF are accessed via an intermediate caching layer called Frontier
  - Each database query is cached on the Frontier squids (http based proxy servers) to avoid the database itself being overloaded with repeated requests to access the same tables
  - T0 has 4 squids, FNAL has 2, all other T1, T2 sites have a single squid