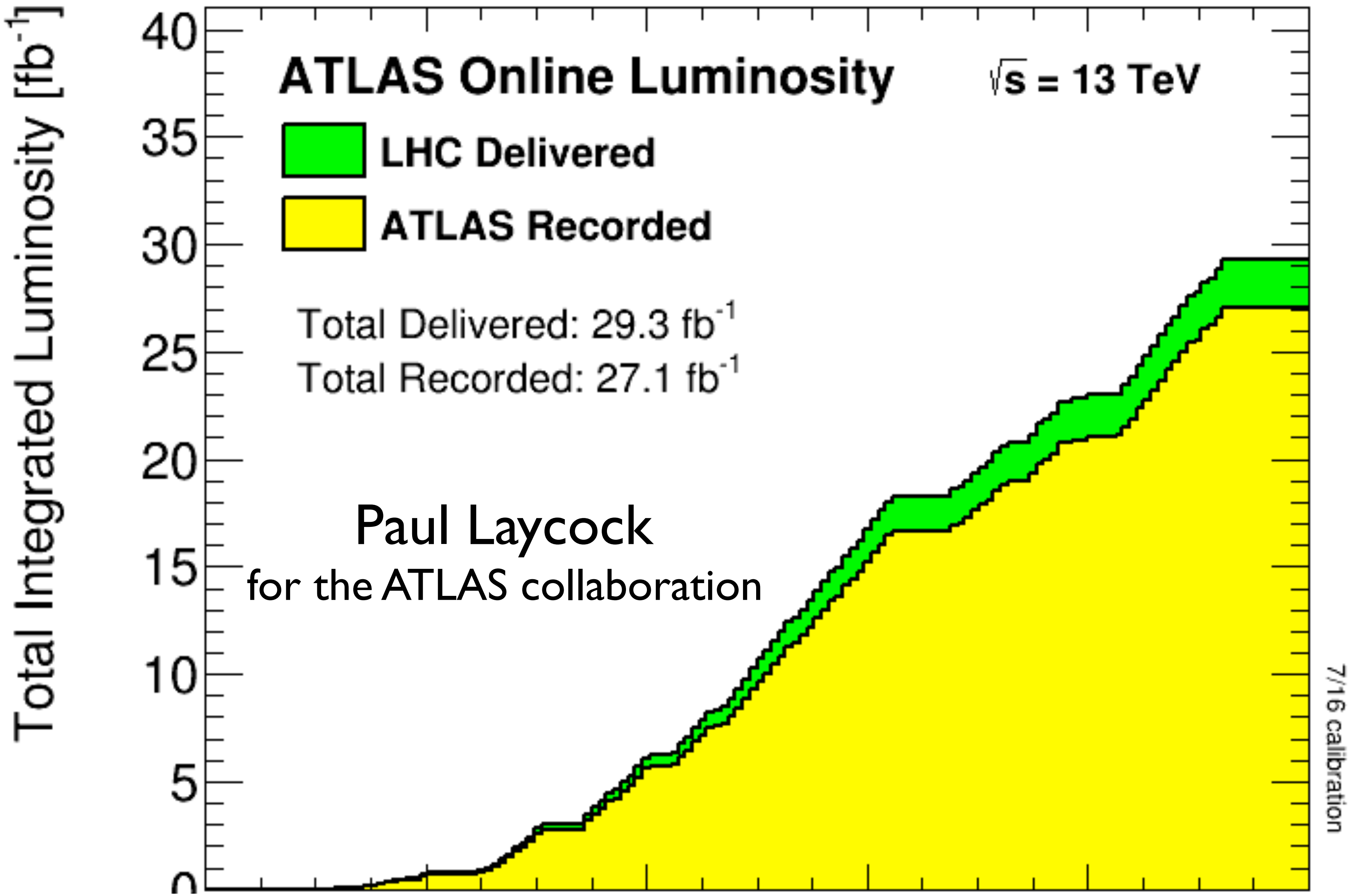
ATLAS Data Preparation in Run 2



Paul Laycock

for the ATLAS collaboration

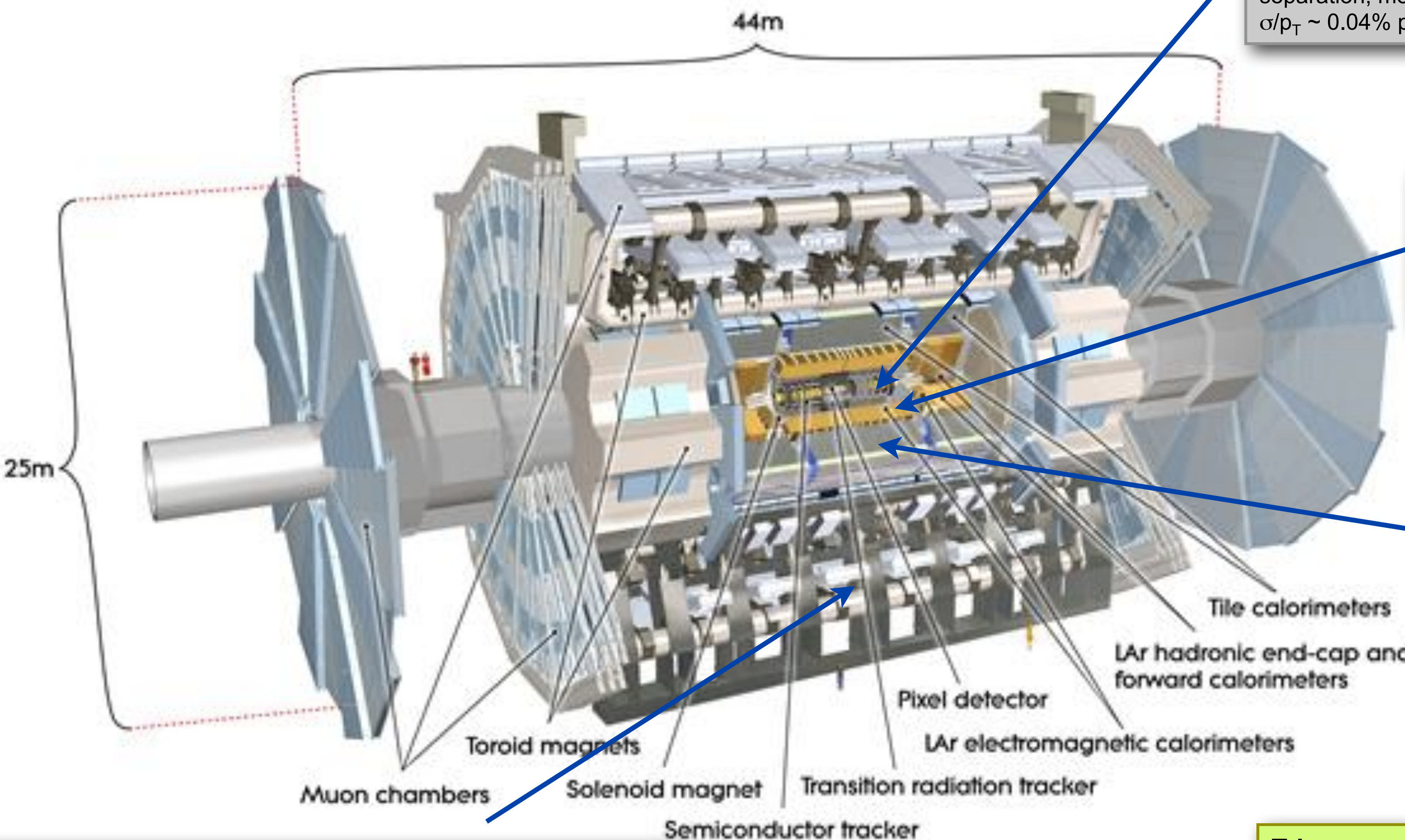
UNIVERSITY OF LIVERPOOL

ATLAS EXPERIMENT

The Atlas Detector

L ~ 46 m, ∅ ~ 22 m, 7000 tons
~10^8 electronic channels

Inner Tracker ($|\eta|<2.5$, B=2T):
Si Pixels, Si strips, Trans. Rad. Det.
Precise tracking and vertexing, e/π separation, momentum resolution:
$\sigma/p_T \sim 0.04\%\ p_T\,(\text{GeV}) \oplus 1.5\%$



EM calorimeter: Pb-LAr Accordion, e/γ trigger, id. and meas., energy res.: $\sigma/E \sim 10\%/\sqrt{E} \oplus 0.7\%$

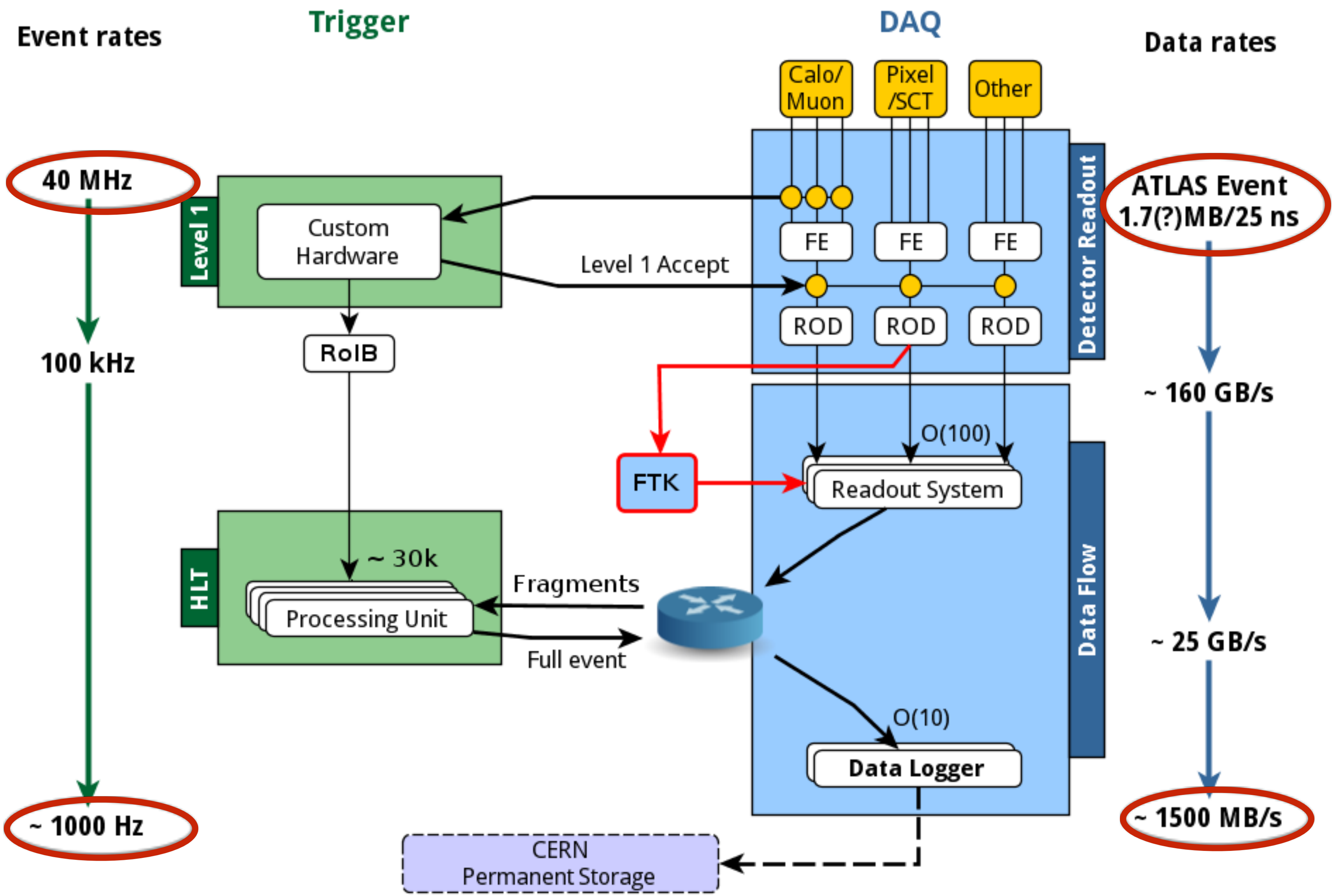
HAD calorimetry ($|\eta|<5$): Fe/scintillator Tiles (cen), Cu/W-LAr (fwd). trigger and meas. of jets and $E_{T,miss}$, energy res.: $\sigma/E \sim 50\%/\sqrt{E} \oplus 3\%$

Muon Spectrometer: air-core toroids with gas-based muon chambers. trigger and meas. with momentum resolution < 10% up to $E_\mu \sim 1$ TeV

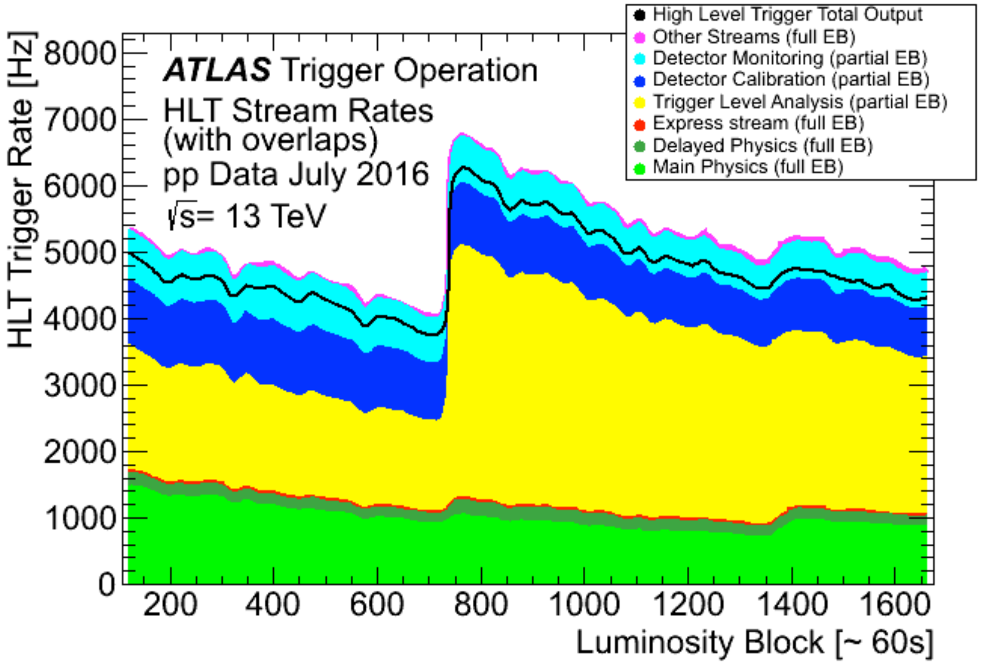
Trigger system: 3-levels reducing the IA rate from 40 MHz to ~200 Hz

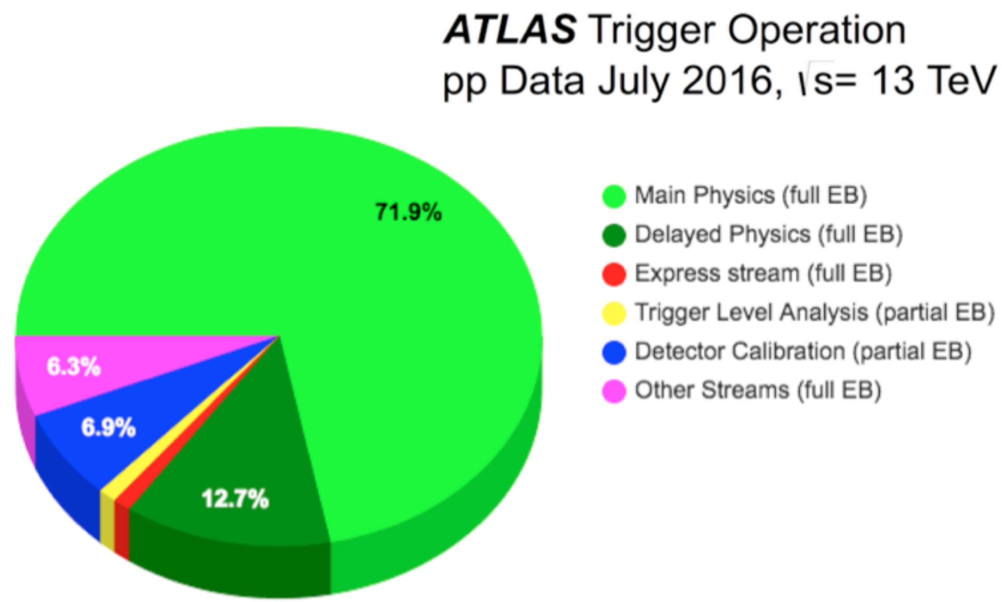
Millions of detector readout channels read out to reconstruct one event

The Atlas Trigger and DAQ



Streams and outputs





Bandwidth dominated by the main physics stream @ 1 kHz

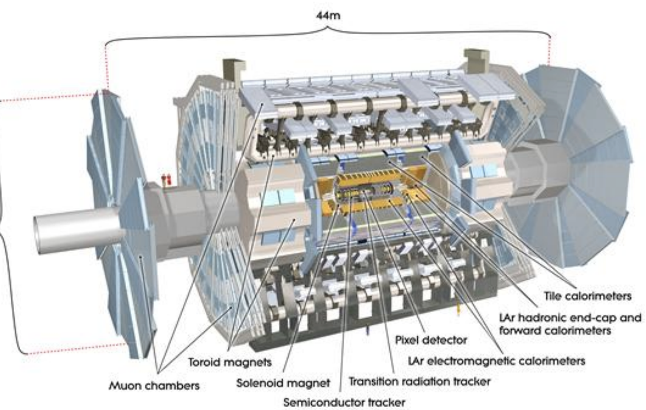
14 full event physics streams

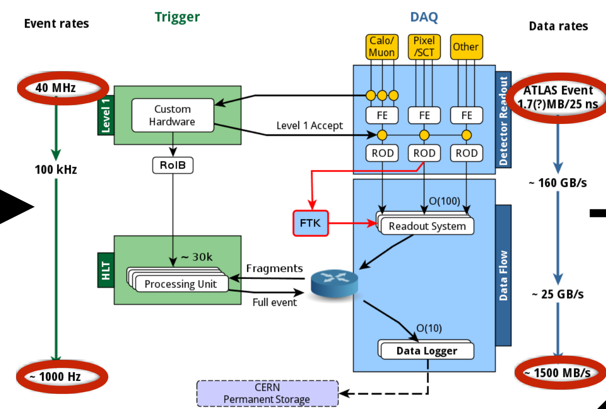
20 partial event calibration and debug streams, rates from < 1Hz to >10 kHz

Mostly processed using centrally managed software configurations

Providing a variety of data formats tailored for each use case

What is Data Preparation?





Data Preparation

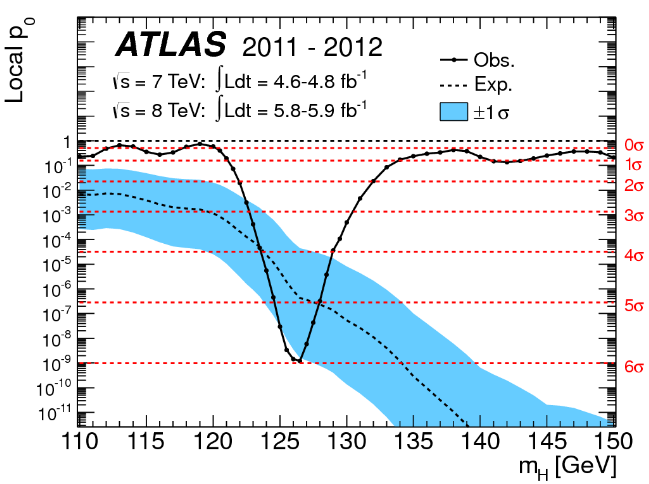
Data Quality Assessment

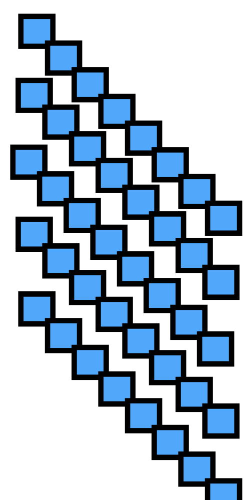
Calibration

Primary Reconstruction

Good Run List

AOD

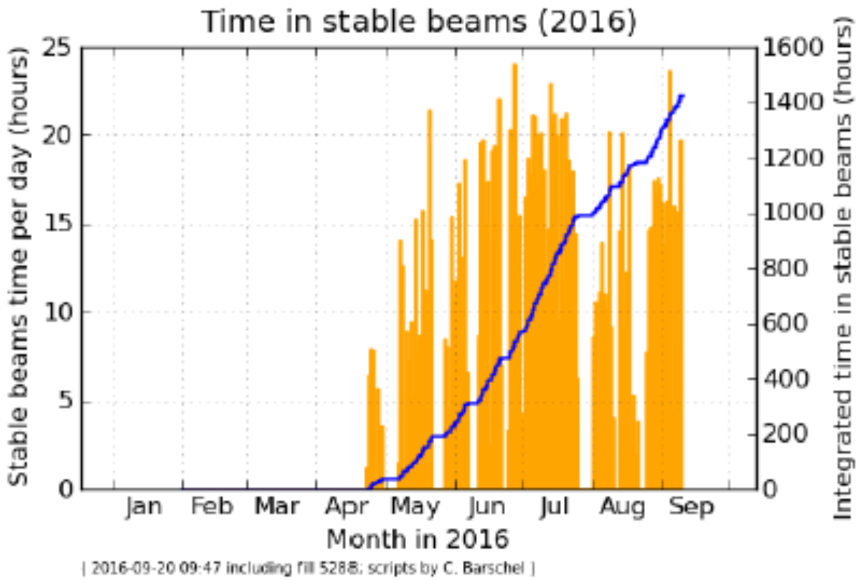


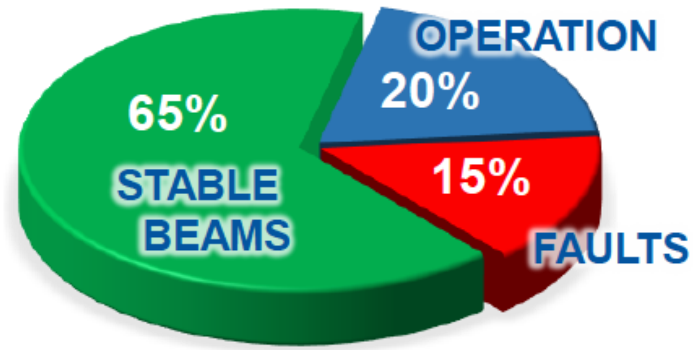


Dr Paul Laycock

UNIVERSITY OF LIVERPOOL

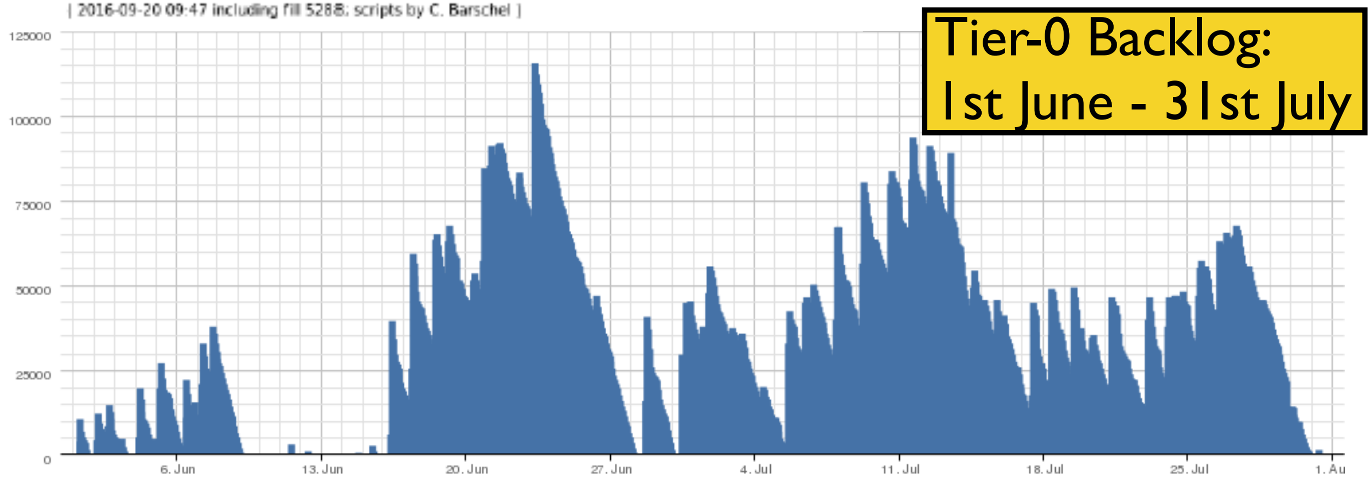
Run 2 Challenges





mid-year availability,
physics production only
(no commissioning, MD, ...)

LHC uptime reached ~80%



Despite improvements in hardware, software and workflows, we had backlogs at the Tier-0

- Pileup: Averaged over all fills, pileup is similar to 2012 (broad distribution with mean ~20)
- Peak pileup increasing towards end of year and expected to be able to reach 50 next year

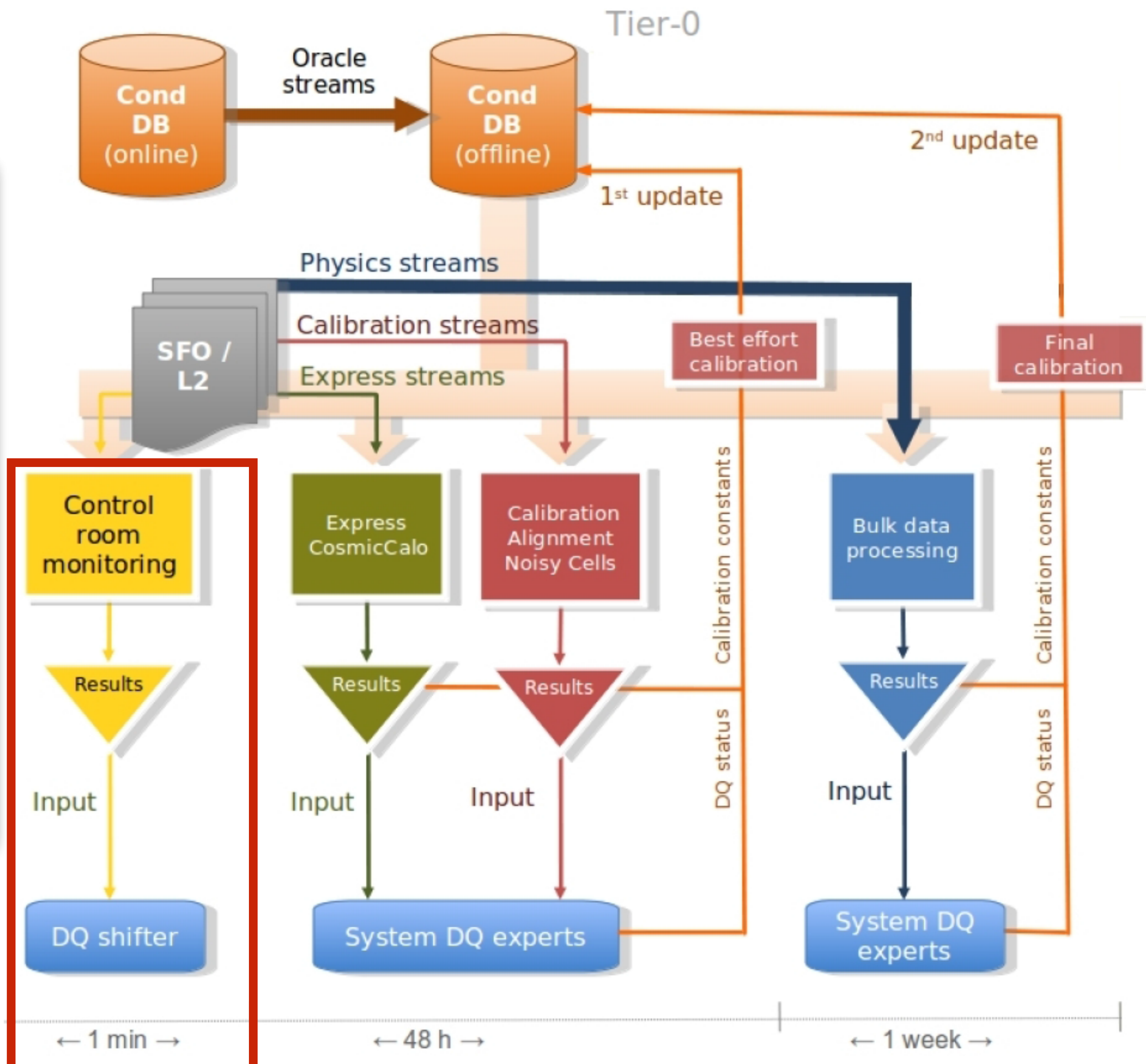
Control room monitoring

Dedicated machines in the ATLAS Control Room

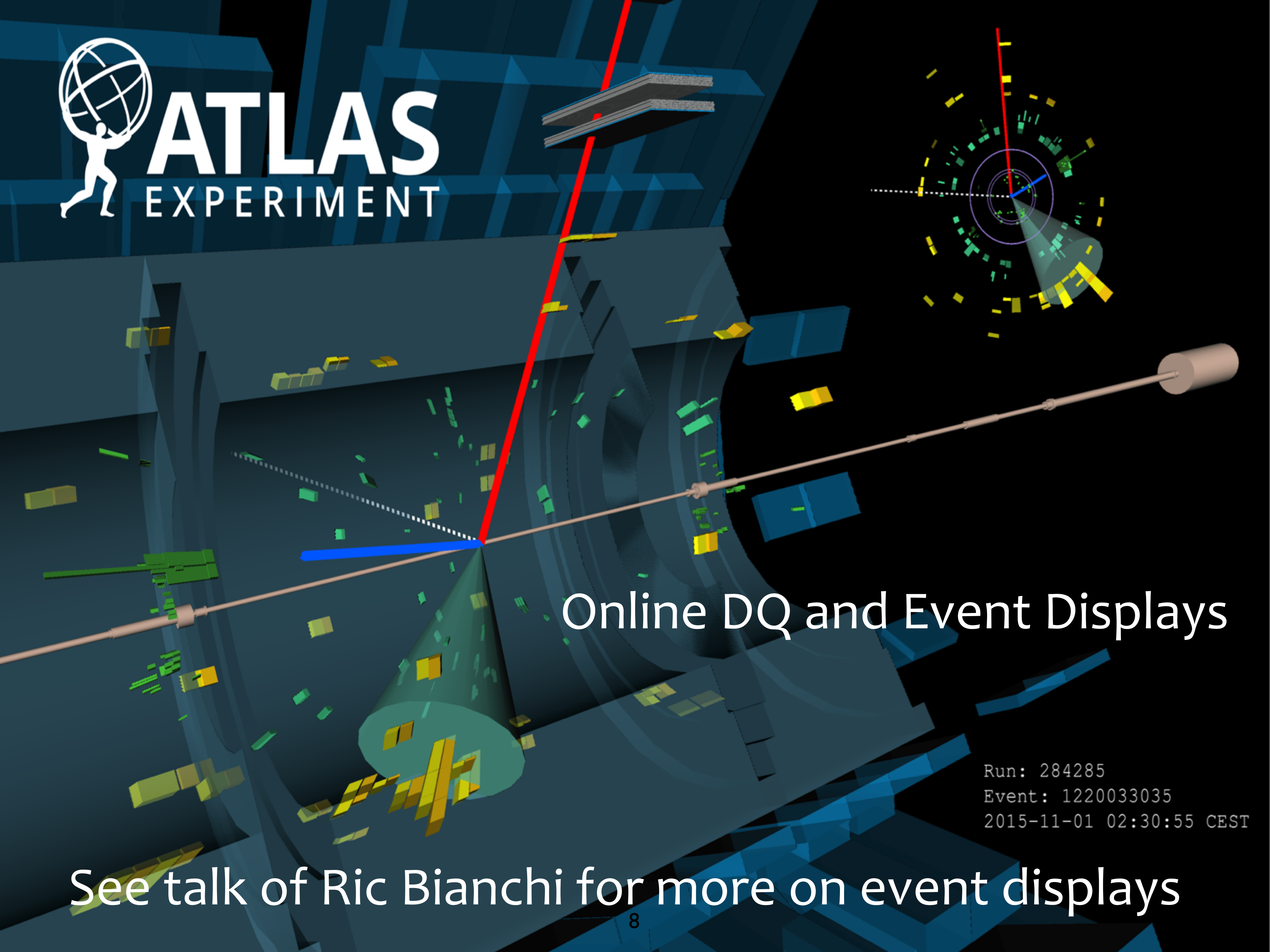
Running reconstruction producing data quality histograms

Frequent updates from offline data quality monitoring to provide best feedback

Dedicated DQ shifter to raise the alarm



Dr Paul Laycock



Online DQ and Event Displays

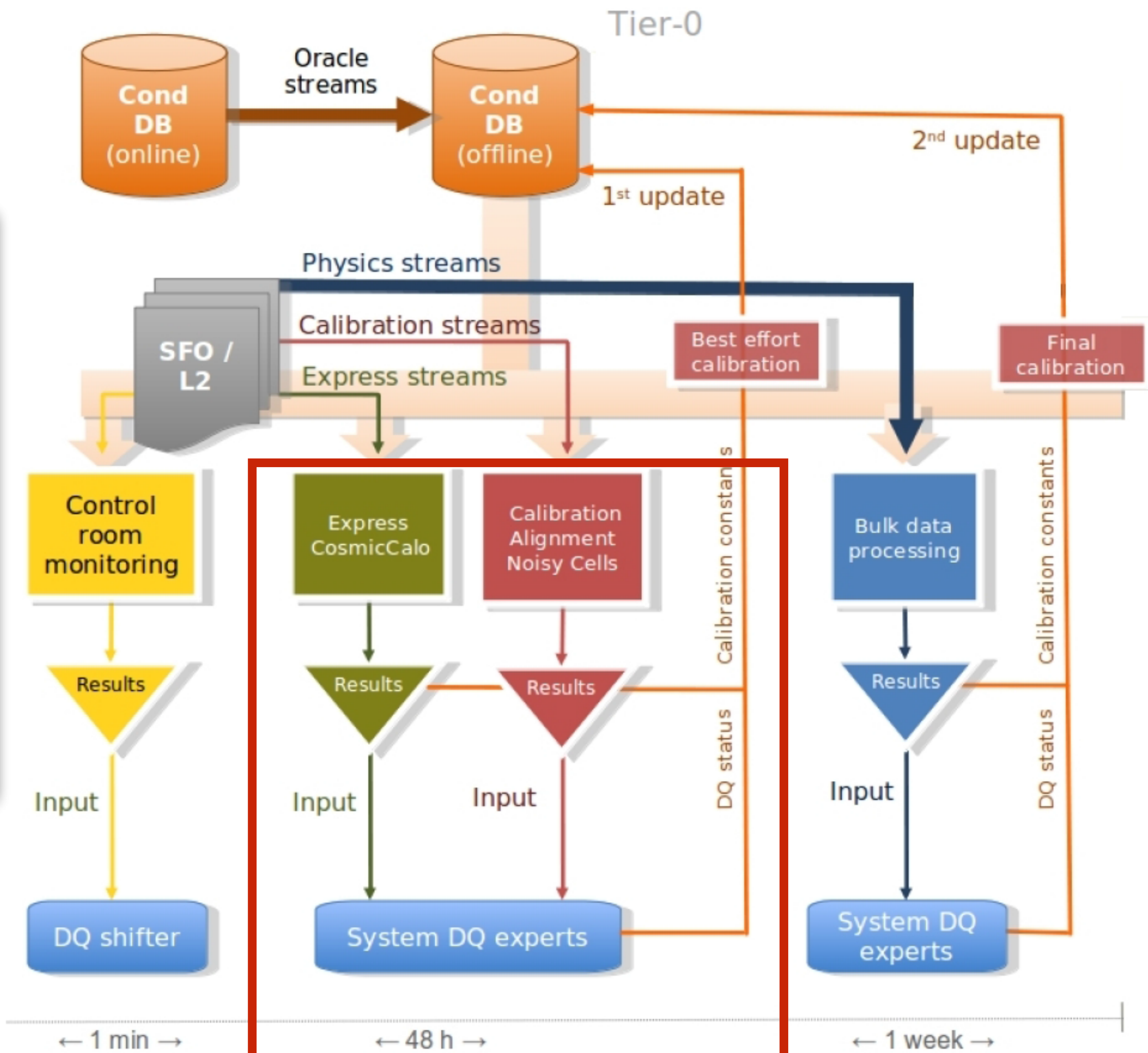
See talk of Ric Bianchi for more on event displays

Prompt calibration and data quality

Tier-0 ~10k dedicated slots

First pass processing takes ~12 hours for fast data quality checks on ~2% data

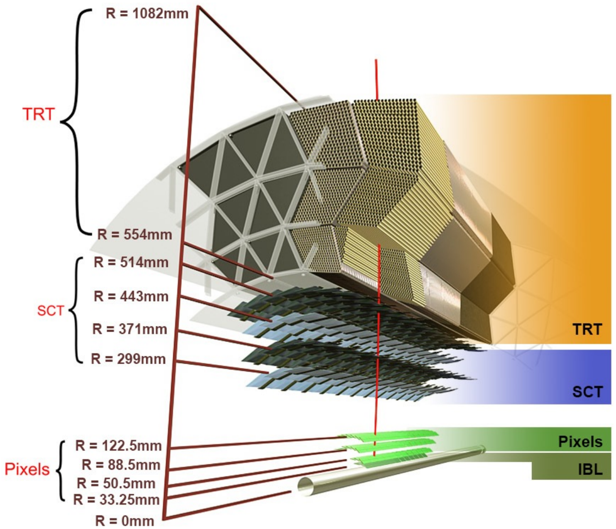
Calibration streams processed to provide updated conditions within 48 hour calibration loop

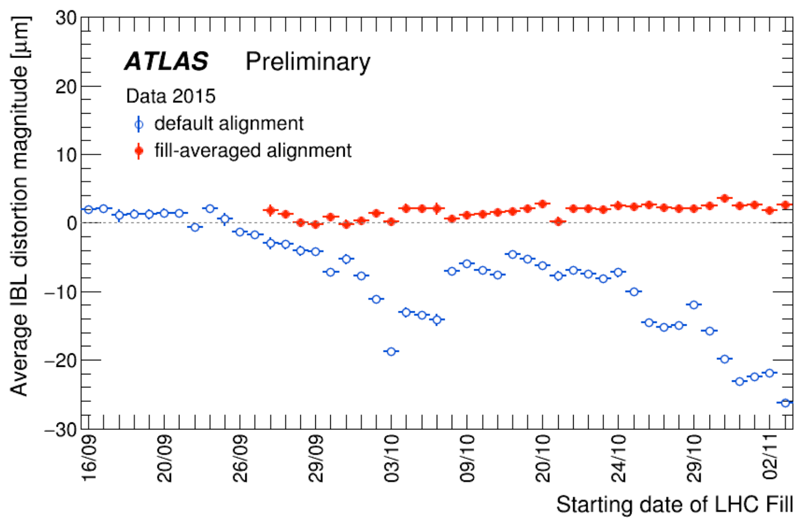


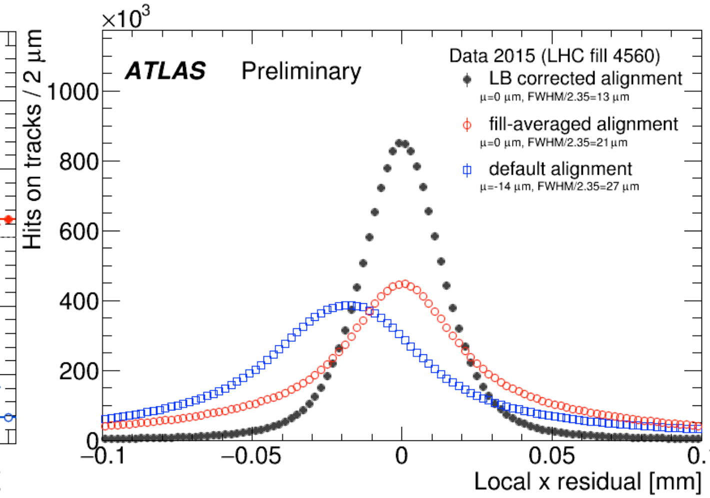
Dr Paul Laycock

ATLAS EXPERIMENT

Example: Prompt ID Alignment







An example: The Insertable B-Layer (IBL) exhibits temperature-driven distortion

Problem: Significant misalignment which biases the beamspot determination
- Worsening problem would render the data unusable for physics

Solution: Implemented a procedure to promptly correct alignment and beamspot
- Alignment procedure requires four iterations - resource heavy - centrally managed
- Subsequently improved to correct all inner detector package movements within a run

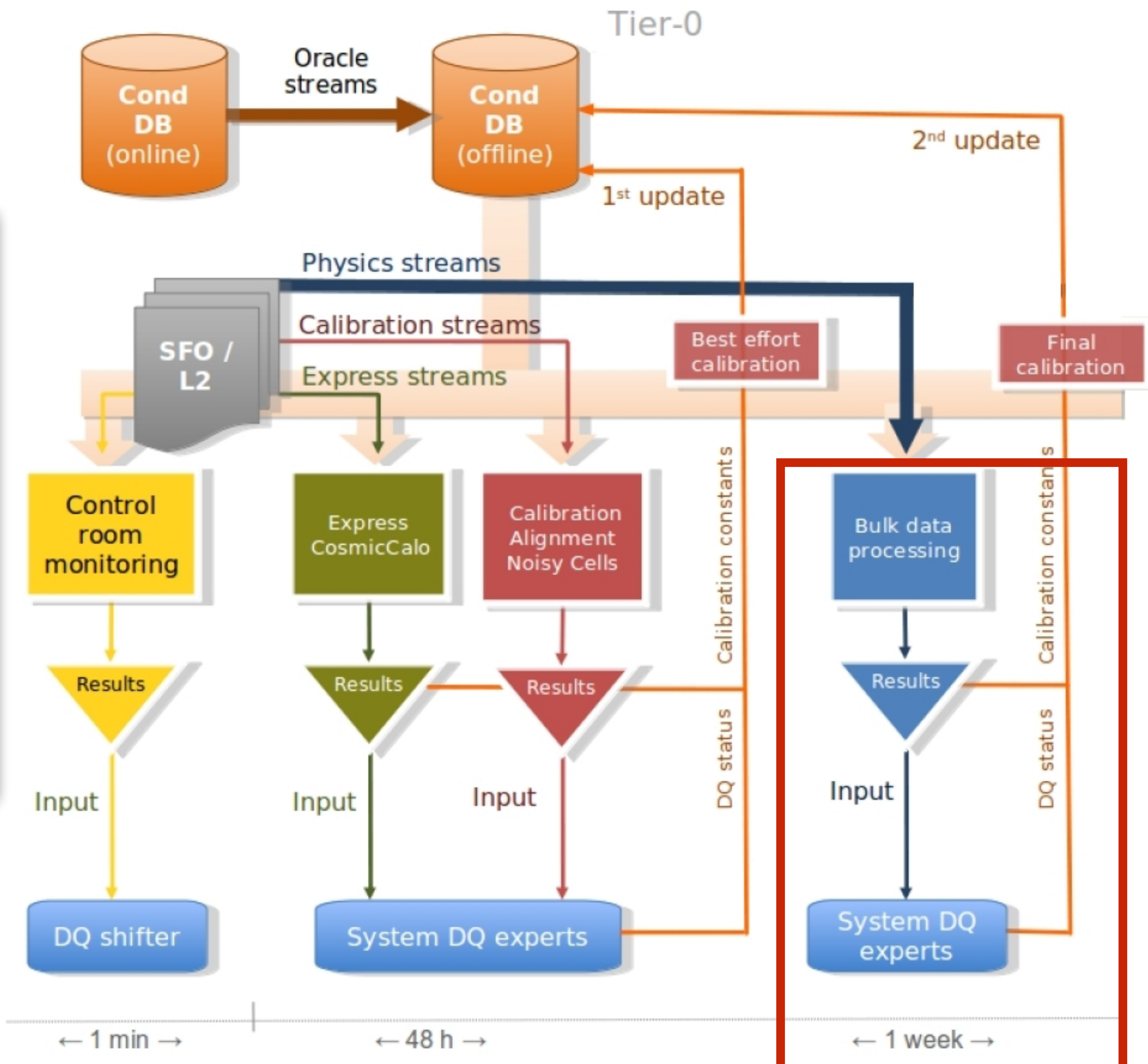
Centrally manage as many calibration procedures as possible to optimise resource usage

Bulk data processing and data quality

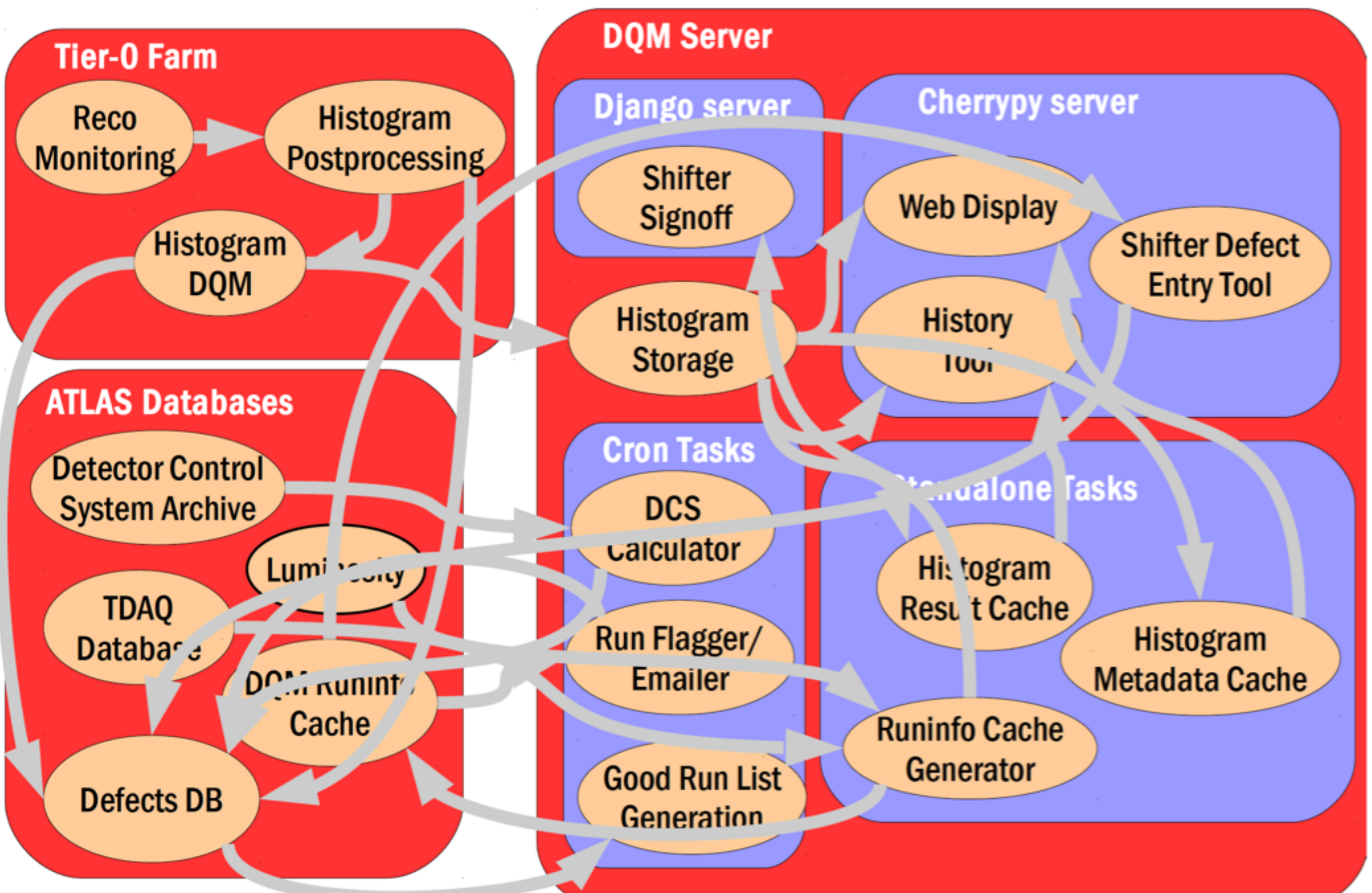
Tier-0 ~10k dedicated slots

Launch bulk processing of all physics data after 48 hour calibration loop

Typically
~60M events per day
~60 TB per day



Data Quality Monitoring Framework



Histograms production running at Tier0 as part of core reconstruction

DQ assessment using web display for histogram presentation

Data rejection down to one minute

Defect database records problems

Final DQ logic combines defects to flag bad data

Output a Good Run List

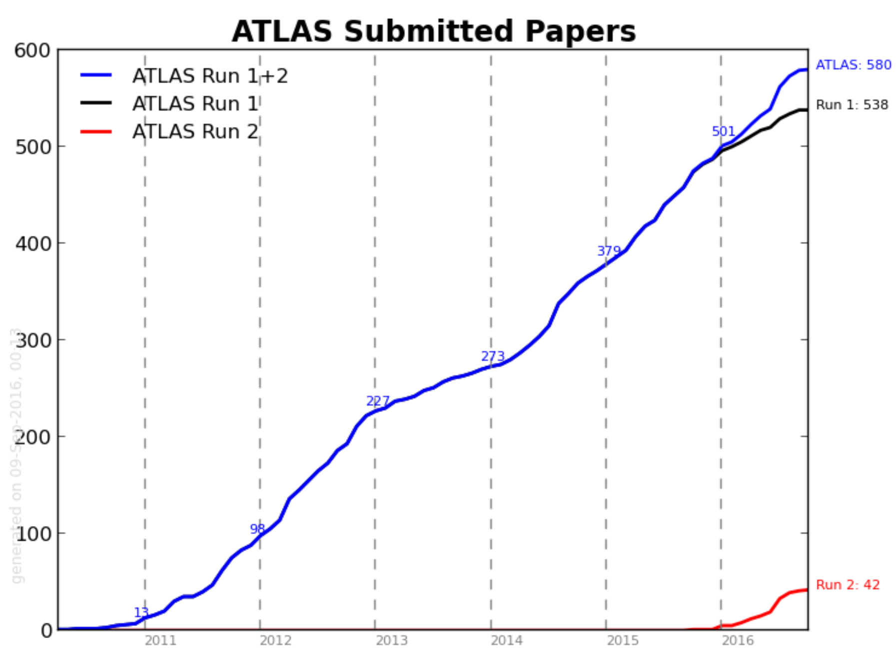
Data Quality and Publications

ATLAS pp 25ns run: April-July 2016										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
98.9	99.9	100	99.8	100	99.6	99.8	99.8	99.8	99.7	93.5
Good for physics: 91-98% (10.1-10.7 fb^{-1})										

Luminosity weighted relative detector uptime and good data quality efficiencies (in %) during stable beam in pp collisions with 25ns bunch spacing at $\sqrt{s}$=13 TeV between 28th April and 10th July 2016, corresponding to an integrated luminosity of 11.0 fb^{-1}. The toroid magnet was off for some runs, leading to a loss of 0.7 fb^{-1}. Analyses that don't require the toroid magnet can use that data.

Data Quality efficiency for the ICHEP dataset

High efficiency in general (Toroid magnet problem made a significant dent in this dataset but now this is around 2%)



Publication-ready data available after around one week

Also thanks to success of the new ATLAS analysis model

Summary

- The impressive LHC performance this year raised serious challenges
  - The ATLAS Data Preparation met the challenges head on !
- A similar overall workflow to that used successfully in Run 1
  - Prompt processing of a fraction of data for fast DQ and calibrations
  - Bulk processing launched after 48 hours
- More workflows, especially critical and resource-heavy, have been centralised
  - New ones added to improve detector performance
- Data Quality assessment critical to ensure good DQ for physics
  - Improved DQ in the control room to catch problems as early as possible
  - Improved offline DQ workflows and tools for final assessment for physics
- High DQ efficiency and fast turnaround for physics