technische universität dortmund

Event reconstruction in LHCb

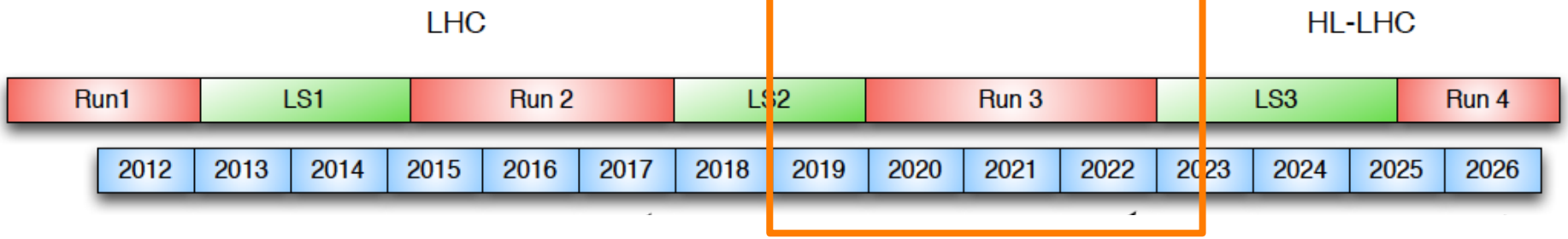
Johannes Albrecht
(TU Dortmund)

HSF workshop

LHC machine evolution

Focus of this talk

Run 1	Run 2	Run 3	Run 4
0.75 x $10^{34}cm^{-2}s^{-1}$	1.5 x $10^{34}cm^{-2}s^{-1}$	2.2 x $10^{34}cm^{-2}s^{-1}$	5 x $10^{34}cm^{-2}s^{-1}$
50ns bunches	25ns bunches	25ns bunches	25ns bunches
pileup ~40	pileup ~40	pileup ~60	pileup ~140 - 200



- Steady increase in machine luminosity both within runs and between runs
- Ultimate goal of 3000fb^{-1} in 10 years of HL-LHC running
  - ☞pp Collision rate of 5.6GHz
- Pileup is the most important metric of event complexity for reconstruction software

	Integrated luminosity		Pileup GPD
	LHCb	GPD	
Run 1	3	25	40
Run 2	10	100	40
Run 3	25	300	60
Run 4	?	+300/a	140-200

Recording of all LHC data?

- Two bottlenecks:
  - **Read-out bandwidth:**
    Fast read-out of O(100M) detector channels
  - **Bandwidth to storage:**
    Offline computing resources for reconstruction, storage

Bandwidth = Event rate x Event size

Event size
ATLAS: 1-1.5 MB
CMS 1-1.5 MB
LHCb ~100kB

Run 2:

Detector readout
ATLAS: ~100kHz
CMS: ~100kHz
LHCb 1MHz

Run 3:

Detector readout
ATLAS: 0.4MHz
CMS: 0.5-0.75KHz
LHCb 40MHz

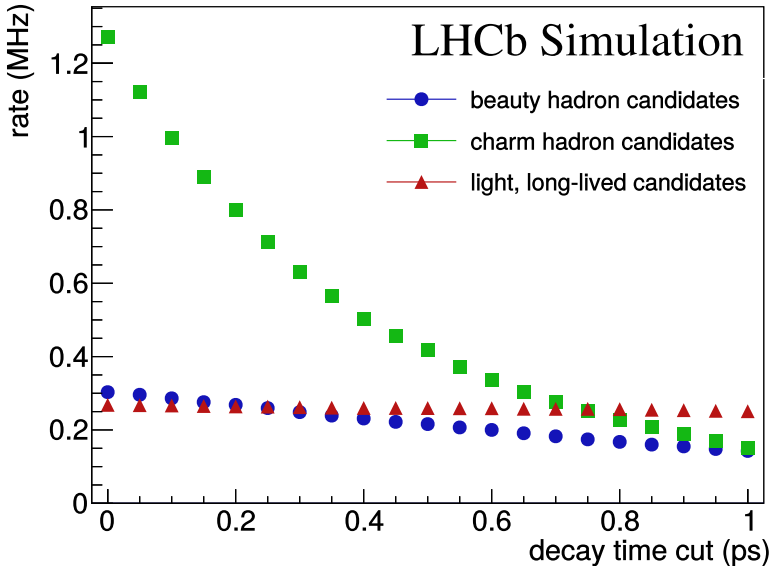
Right now: Data Volume to farm ~ equal in ATLAS, CMS, LHCb
→ Run 3 LHCb will increase it by factor 60
(ATLAS & CMS come for Run 4)

LHCb's Run 3 signal rates

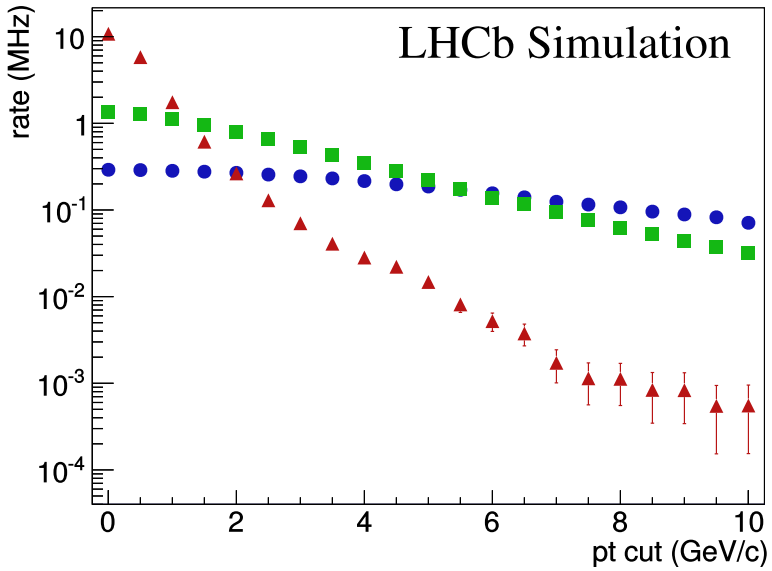
- Challenge: every event contains signal
  - 0.2 b-hadrons, 1.5 c-hadrons, 33 light-long-lived
- Must go beyond rejecting background
  → classify signal and choose wisely

- PT and lifetime alone not sufficient to reduce rate
  → requires all available detector information

Rates as a function of decay time cut for part. reco. candidates

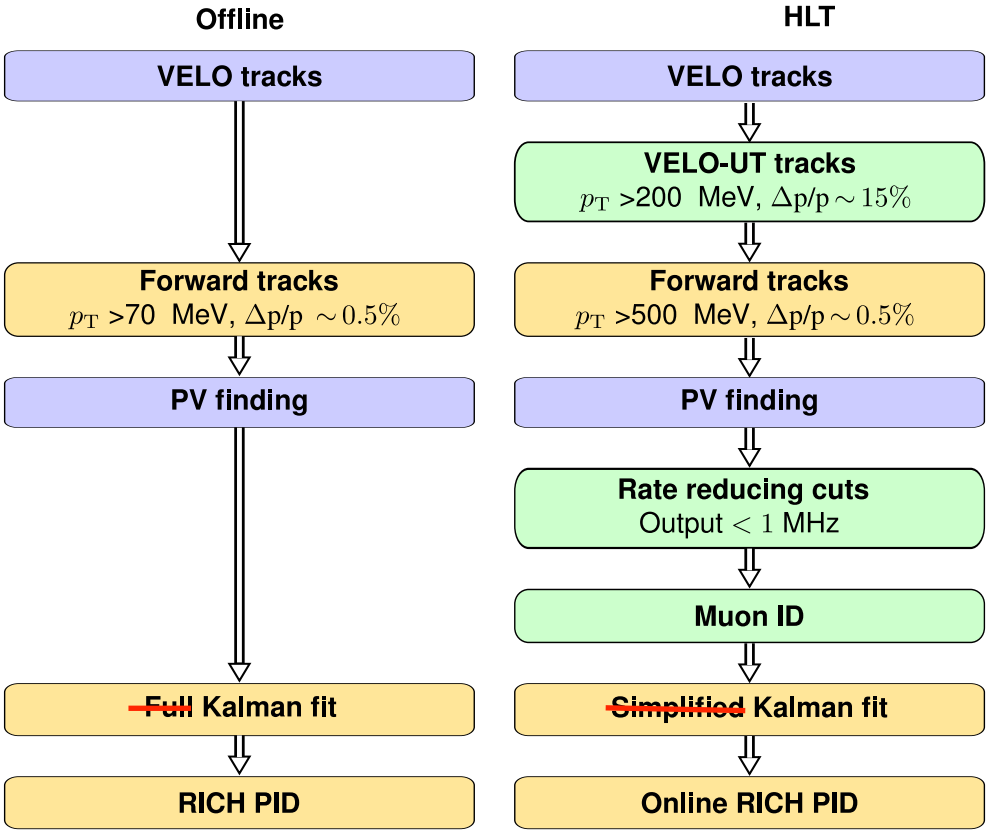


Rates as a function of pT cut for part. reco. candidates

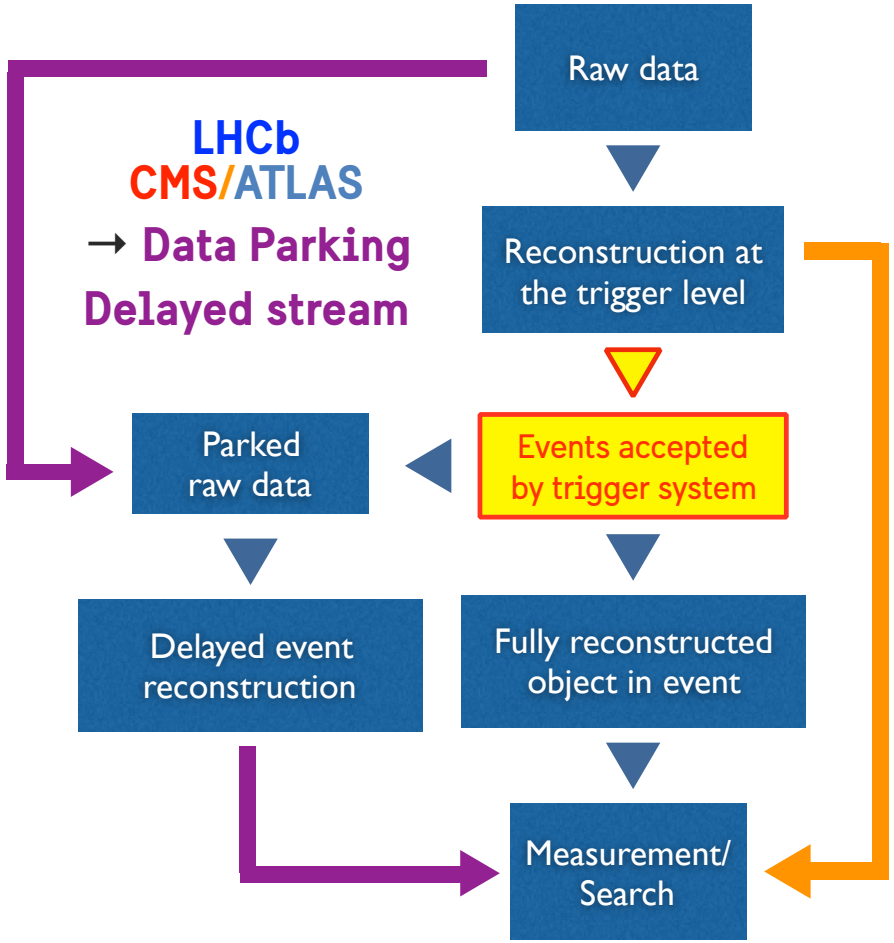


LHCb: 30MHz reconstruction

- Offline-quality tracking in software possible @ 30MHz
  - Uses about 50% of estimated budget (~3MCHF farm)
  - Thanks to upgraded vertex detector & tracker designs
  - Vetex detector outside of dipole magnet
- Converge online and offline reconstructions
  - Already significant steps accomplished for Run 2
- FPGA assisted tracking was tested and found not to be cost effective



Trigger aware analysis



Reduce immediate reconstruction overhead by:
saving data reconstructing it (much) **later**

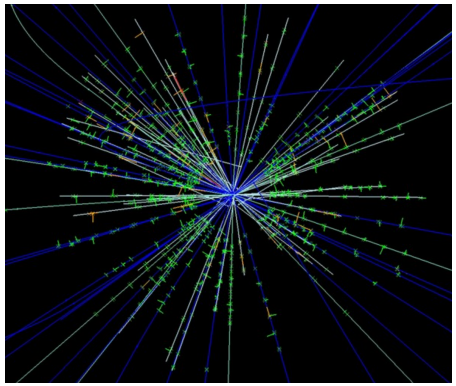
LHCb
CMS/ATLAS
→ Data Parking
Delayed stream

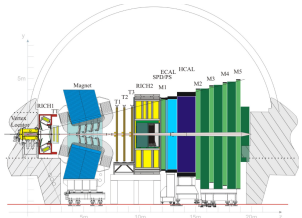
CMS/ATLAS → Data Scouting
LHCb → Turbo Stream

'Bypass' storage limitation (trigger selection) by:
moving **analysis online**
saving **only selected information**
rather than fewer full events

Figure: C. Doglioni

Interlude: calibration and reconstruction





Online event Reconstruction

Events at trigger level good for analysis

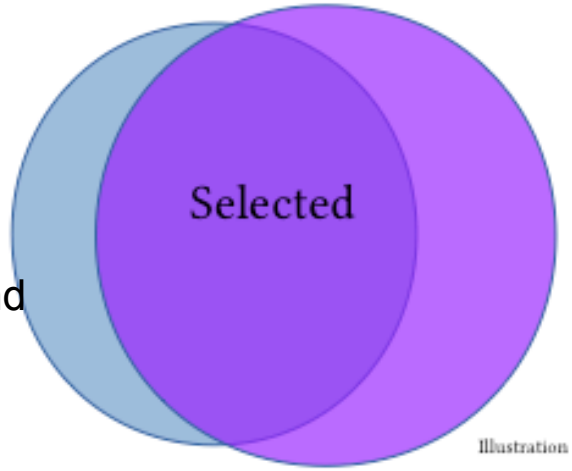
Online calibration of the detector

LHCb: **buffering** data on disk allows for precise detector alignment and calibration

LHCb/ATLAS/CMS: ~same reconstruction **software and inputs** online and offline

Offline and online reconstruction

Online:
compromise between performance and stringent timing requirements



Offline:
Best available performance, less timing constraints

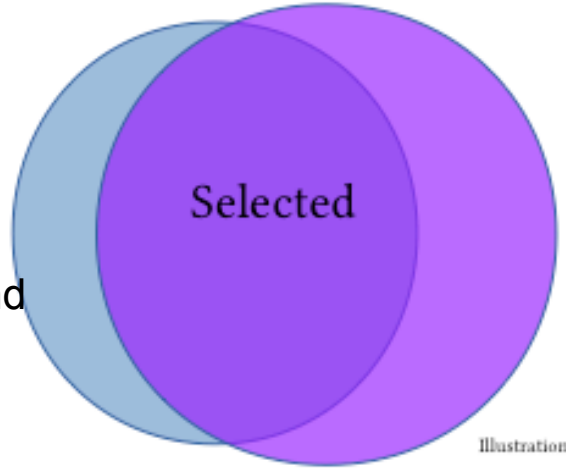
Online-Offline differences in Run 1:
Pattern recognition, alignment & calibration, no hadron PID online, selections

Goal for Run 2:

- Same reconstruction online and offline
- Needs online calibration & alignment

Offline and online reconstruction

Online:
compromise between performance and stringent timing requirements

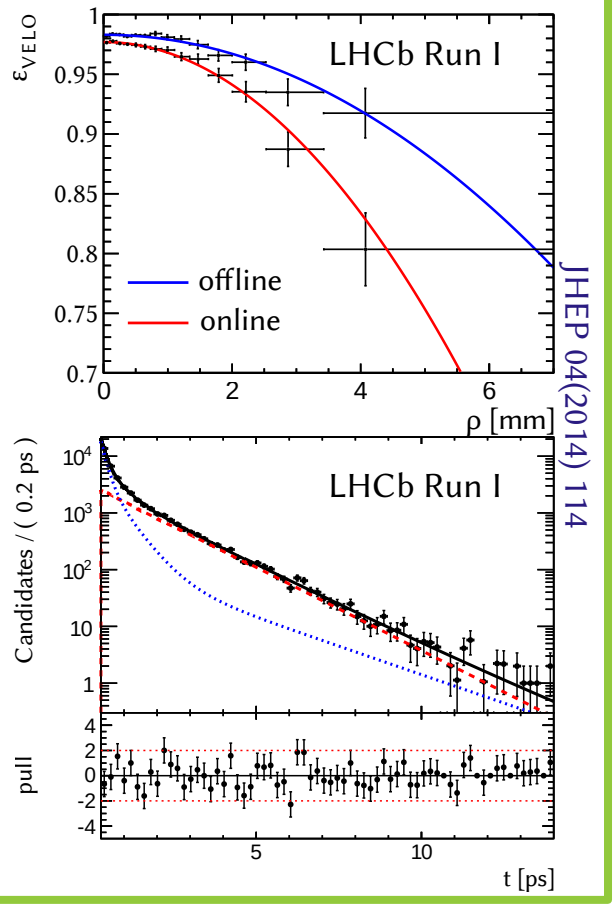


Offline:
Best available performance, less timing constraints

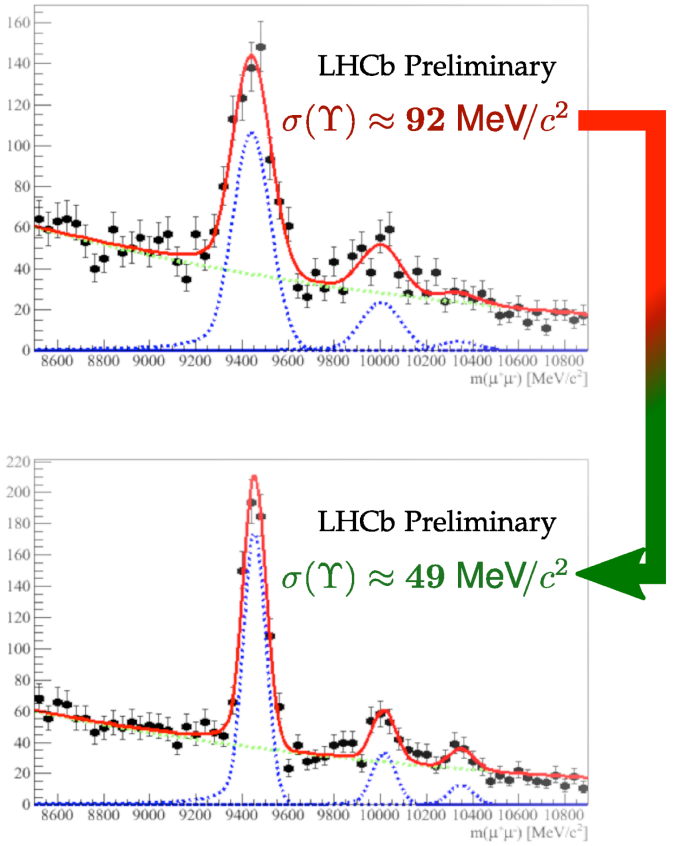
Online-Offline differences in Run 1:
Pattern recognition, alignment & calibration, no hadron PID online, selections

Goal for Run 2:
- Same reconstruction online and offline
- Needs online calibration & alignment

Run 1 example:
Lifetime measurement

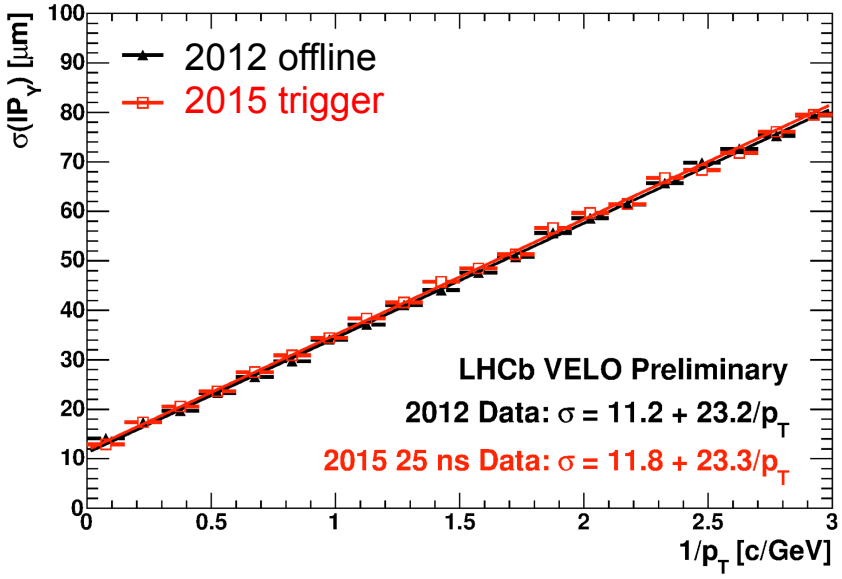


Real-time alignment and calibration

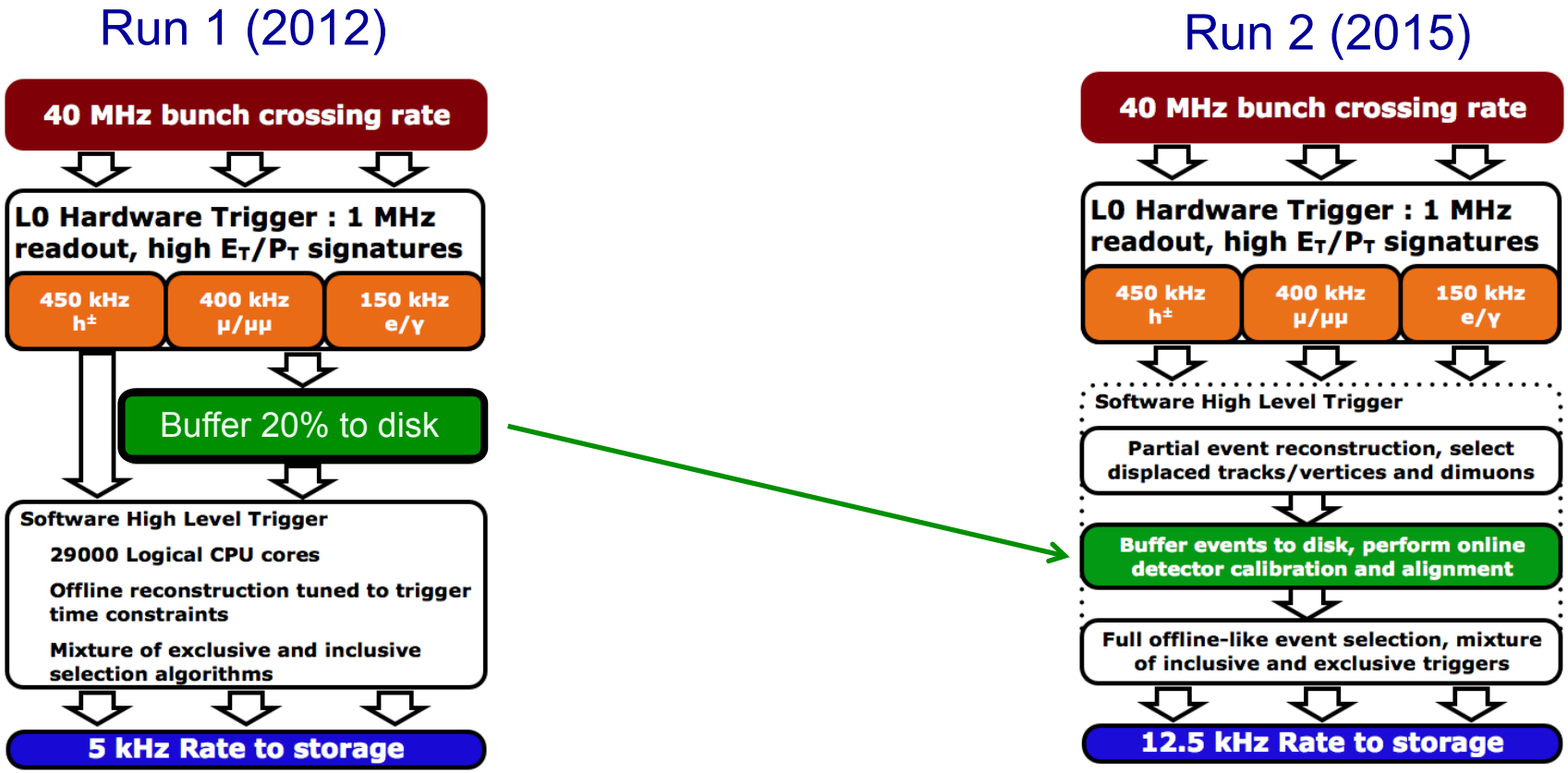


- Benefits of alignment + calibration

- Specific procedures applied for each subdetector
- Monitored in real-time at the control room
- Run 2 Impact parameter resolution in the trigger is now identical to Run 1 offline:



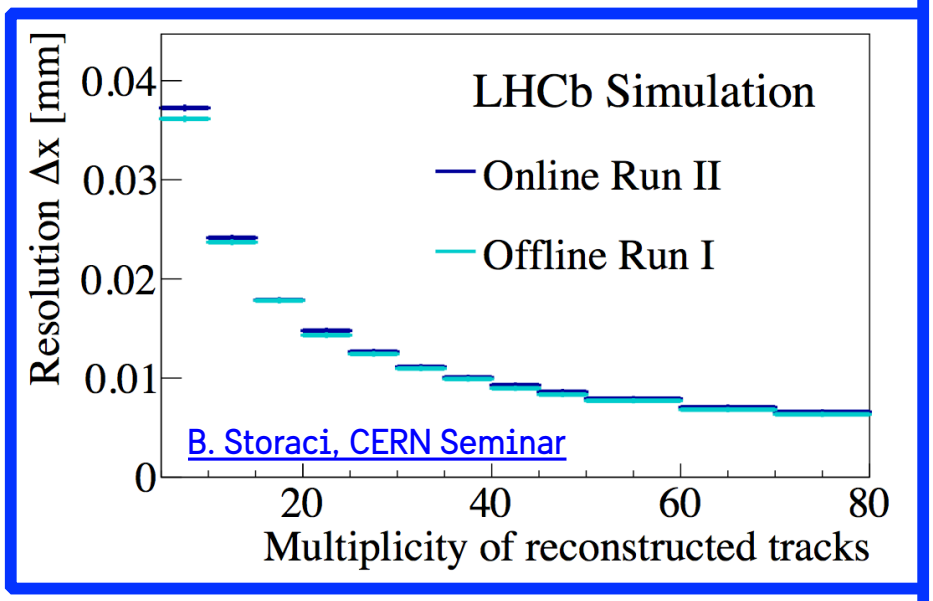
Run 1 → Run 2

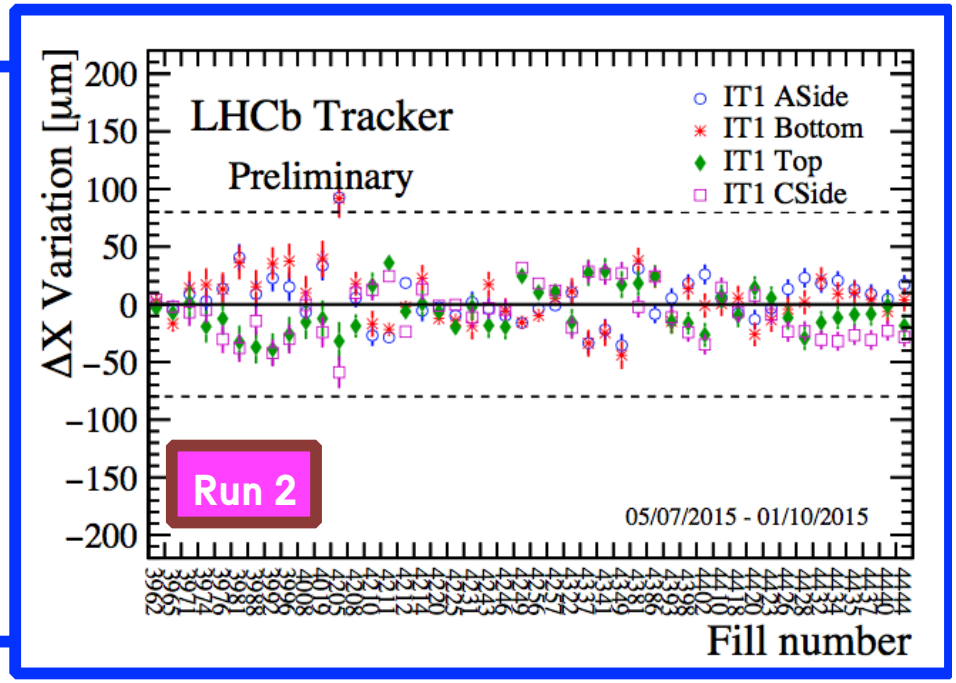


- Extremely large $\sigma_{b\bar{b}}$ and $\sigma_{c\bar{c}}$ in LHC hadron collisions at 13TeV
  - corresponds to **45kHz $b\bar{b}$ and 1MHz $c\bar{c}$** in acceptance
- Trigger system classifies signal to large extend

LHCb: offline vs online reconstruction

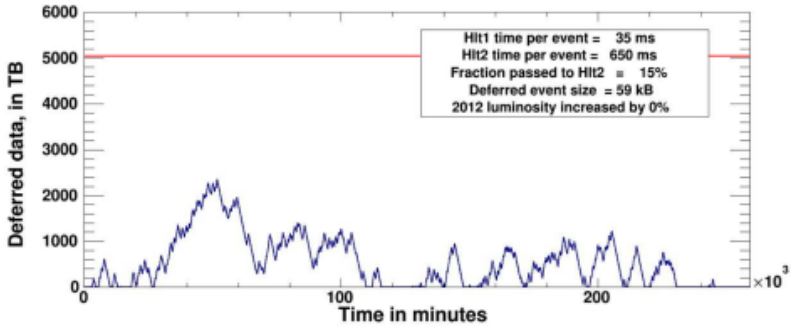
Run-2: online and offline reconstructions **effectively the same**





Real-time calibration in **Turbo Stream**:
in minutes, **realign** with fresh data and **update** constants if needed
Possible thanks to **HLT/calibration farm computing power**

Disk space sufficient to store 160 hours of HLT1 events (60kB,150kHz)



Summary ?

- LHCb performs its big upgrade after Run2
  - Data rate in farm now: ~ same as ATLAS /CMS
    → Run 3 factor 30 increase!

- Signal dominated:
  We need to select the **right signal** to reduce data rate

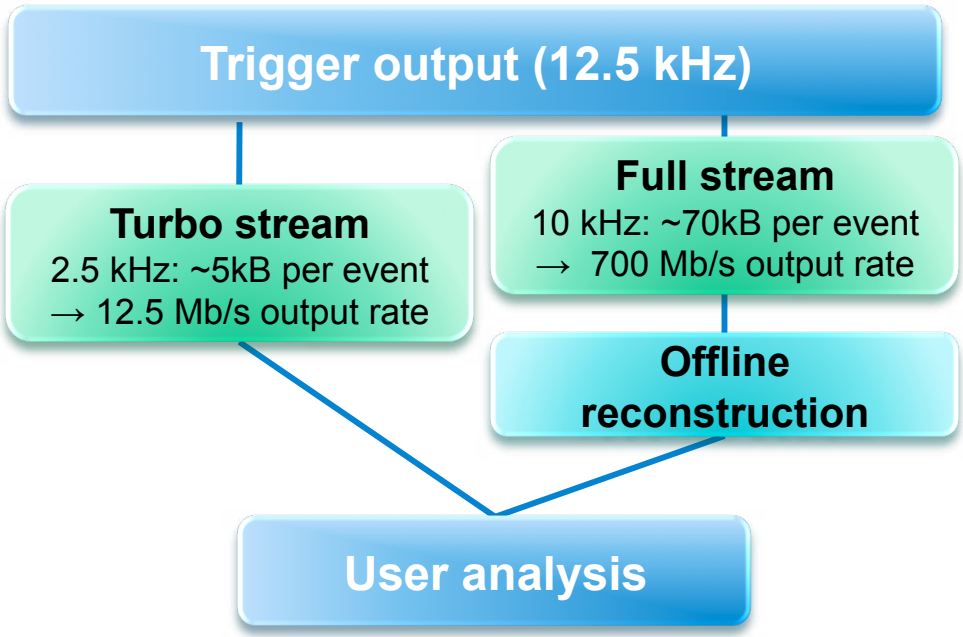
- Hence, we need to reconstruct the full event in real time
  - Real time calibration and fast reconstruction
  - CPU growth???

LHCb: Turbo stream

- **Turbo** stream:
  - more rate, less detail
  - no offline recalibration
  - keep only subset of objects
    (event size:)
  - 24h turnaround in analysis

- **Full** stream:
  - fewer but more precise events
  - can reprocess offline
  - keep full event
    (event size:)

Run 2 numbers:



Turbo Stream: 20% of the trigger at 2% of the cost

Physics cases: signal dominated channels
(charm, strange but also hard to select channel like $\tau \rightarrow \mu\mu\mu$)

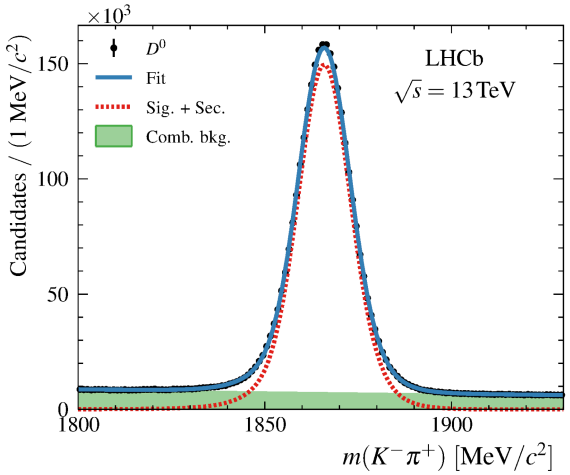
LHCb data in Run 3

- Proof of principle done in 2015: two paper published on turbo stream
  - Charm cross section at 13 TeV
  - J/ψ and b cross section at 13 TeV

- Extremely fast turnaround
  - Analysis presented 1 week after data taking

Charm production at 13TeV

Submitted to JHEP

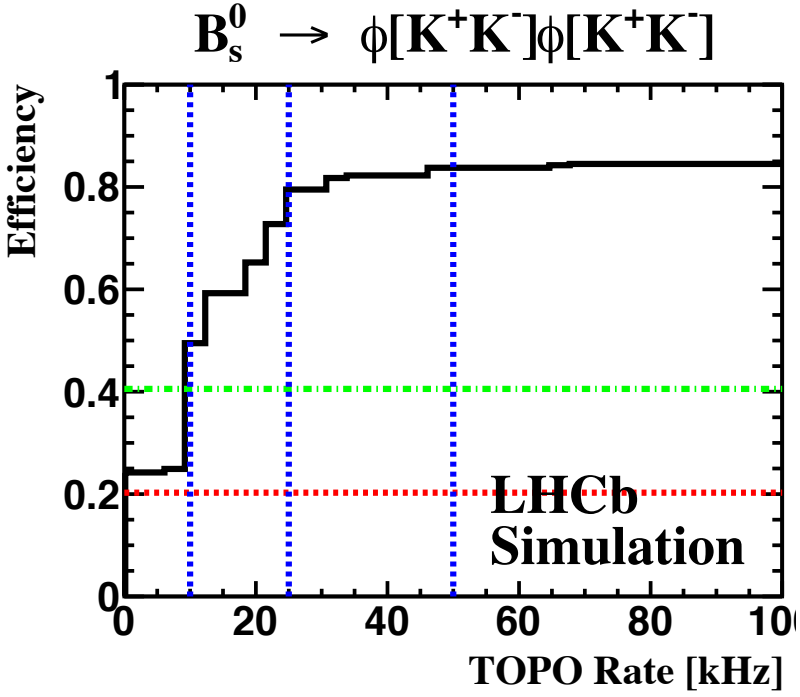


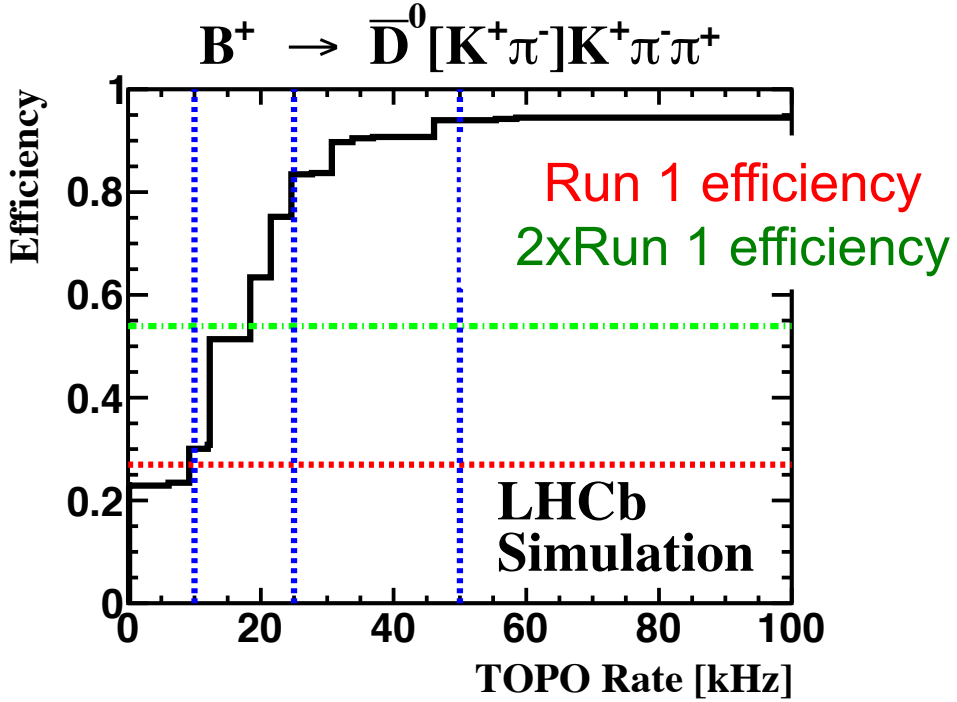
- Extremely pure signals in real time!

- Interesting to compare with the ALICE strategy of keeping lossy compressed RAW
  - LHCb plan to keep full RAW for the most interesting events; compressed RAW for others and no RAW at all for others
  - No longer limited to binary decisions about RAW data

One slide on physics

- LHCb: proof-of-principle inclusive trigger for Run 3 (smart trigger utilizing BDT)





→ **Full software trigger allows a factor 4 in efficiency for hadronic modes**

→ @high output bandwidth or partial event saving

LHCb-PUB-2014-031