Challenges, Requirements and UK's Strengths: LHCb

Conor Fitzpatrick

PPTAP Computing & Software Roadmap Workshop

Beauty2Charm
European Research Council
Established by the European Commission

UK Research and Innovation

Introduction

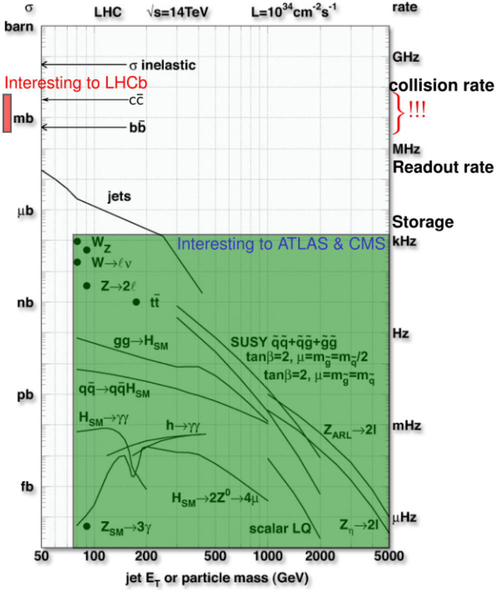
- First, a note on timescales: LHCb is out of phase with ATLAS/CMS:

Original recipe	Runs 1+2	$4 \times 10^{32} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
Upgrade 1	Run 3	$2 \times 10^{33} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
Upgrade 1b	Run 4	$2 \times 10^{33} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
Upgrade 2	Run 5	$1.5 \times 10^{34} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$

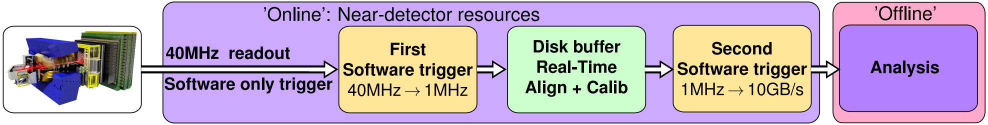
- Upgrade 1b is an intermediate step to make use of the long shutdown for HL-LHC upgrades.
- Major software/computing upgrades are Run 3 (now) and Run 5 with R&D/pilot studies for Run 4.

The challenge for LHCb

- LHCb's operating regime is:
  - Signal dominated, ∼MHz in Upgrade 1, 10's of MHz in U2.
  - hard to trigger efficiently with simple localised signatures
- In Run 3:
  - Triggerless readout to avoid inefficiency of local L0
  - Real-time alignment and calibration of detector (deep buffers)
  - Offline-quality reconstruction online + 'offline' complexity selections as triggers
  - Majority of output will be TLA/Turbo/Scouting trigger objects at ∼10GB/s



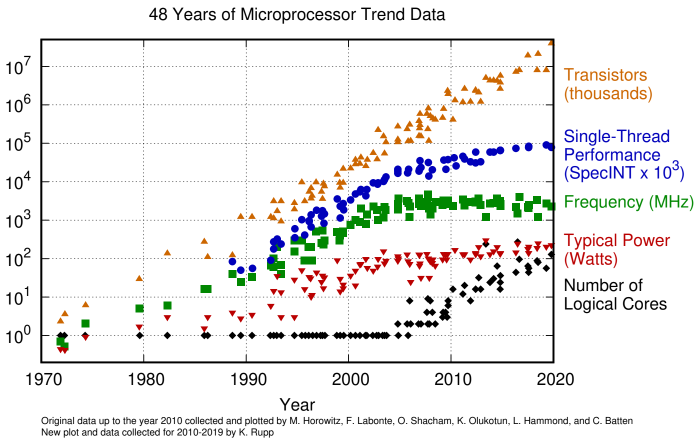
LHCb is compute limited



- If we could trigger and reconstruct all tracks down to $0 p_{\mathrm{T}}$ we would
- Two-stage software trigger performs fast tracking > 500MeV + muons > 80MeV at HLT1, full reco at HLT2 **(UK expertise)**.
- Output is buffered between these stages for alignment and calibration. **(UK expertise)**
- Huge signal rates are also an offline computing challenge:
  - Need to simulate datasets proportional to data collected **(UK expertise)**
  - Need an offline analysis infrastructure to efficiently process and manage data **(UK expertise)**
  - LHCb has used its x86 trigger farm for simulation out of fill throughout Run 2. Equivalent to 50% extra grid capacity.

Doing more with less

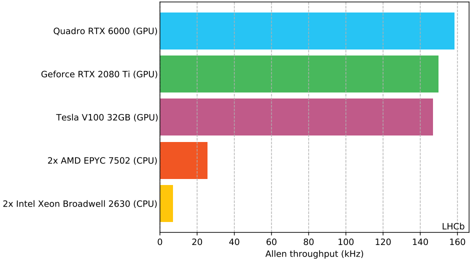
- For Run 3, 30 $\times$ the HLT1 input rate without 30 $\times$ the cash is a challenge

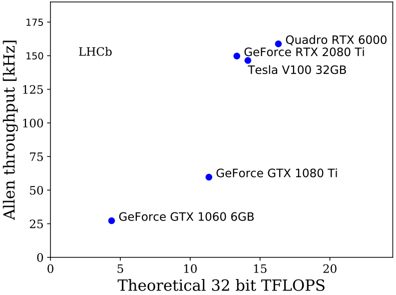


- Processing technologies have transitioned from higher CPU frequencies to incresed parallelism.
- Requires a dramatic change in how we design and run our software
- LHCb took two routes for its Run 3 trigger:

An HLT1 exclusively on GPUs

- R&D efforts for Run 3 followed two technology options:
  - CPU: LHCb-TDR-016
    - Transition to a fully multithreaded HLT1 & HLT2
    - Exploit vectorisation where possible, restructure data formats
    - Make use of a lightweight scheduler to maximise CPU utilisation
  - GPU: LHCb-TDR-021
    - Implement entire HLT1 reconstruction and triger on GPU
    - Rewrite all HLT1 algorithms in Allen, a new CUDA framework
    - For now, keep HLT2 on CPU



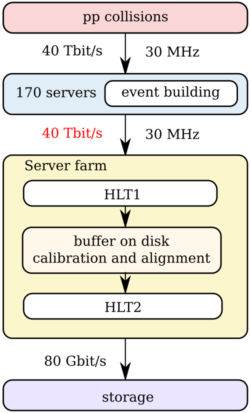


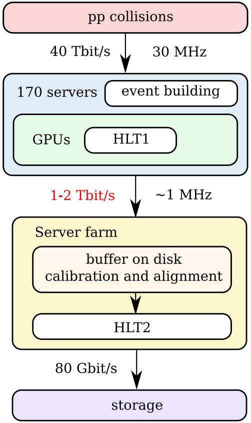
- After delivery of both options, we performed a global cost optimisation to determine the Run 3 baseline

C. Fitzpatrick

Cost considerations

- Events are built on dedicated nodes in both scenarios
- These are then processed by HLT1 on the filter farm (CPU) or GPU cards (Allen)





- Significant cost saving comes from reduced network infrastructure → GPU HLT1 adopted as baseline `arXiv:2105.04031`
- Performance scaling (previous slide) shows promise for expansion with future GPU generations

C. Fitzpatrick

July 19, 2021

Requirements going forward

- Experience with GPUs provides some ideas for future R&D:
- Baseline for Run 5 is to port the entire HLT to GPU but R&D ongoing on alternative technologies.
- Disadvantage is that simulation can't (yet) make use of GPUs but:
- No application is off-limits for GPUs anymore. The limitations are in software engineering expertise.
  - Need significant RSE personpower to develop faster, more efficient generators + simulation frameworks. Currently the largest consumer of processing for LHCb.
  - Also need local support and training for physicists to develop their analyses/selections for GPU
  - Have to ensure HLT can be run on the grid for simulation. HLT1 AMD/HIP port is promising but full **Heterogeneity** is worth further study

Conclusions

- Future computing requirements are known and will continue to be a bottleneck for LHCb
- Future computing *evolution* is harder to predict: The GPU trigger has shown that making sure all our software can be adapted to take maximum advantage of future trends is key
  - **Knowledge retention** is a concern: Run 1+2 software was $\sim$10 years old by end of Run 2. PDRA timescales and lack of career path for RSE-type activities is detrimental.
  - **Knowledge exchange** is also important: All aspects of software from trigger to offline analysis needs dedicated RSE effort to develop, maintain, improve and train.
- LHCb's UK expertise covers trigger, simulation and offline analysis but this is driven by physicists and precarious ECRs
- More efficient use of computing in the same budget means more physics for LHCb, and it's better for the environment: arXiv:2106.07701
- This points to group based, long-term RSE effort.