Streaming DAQ @ LHCb

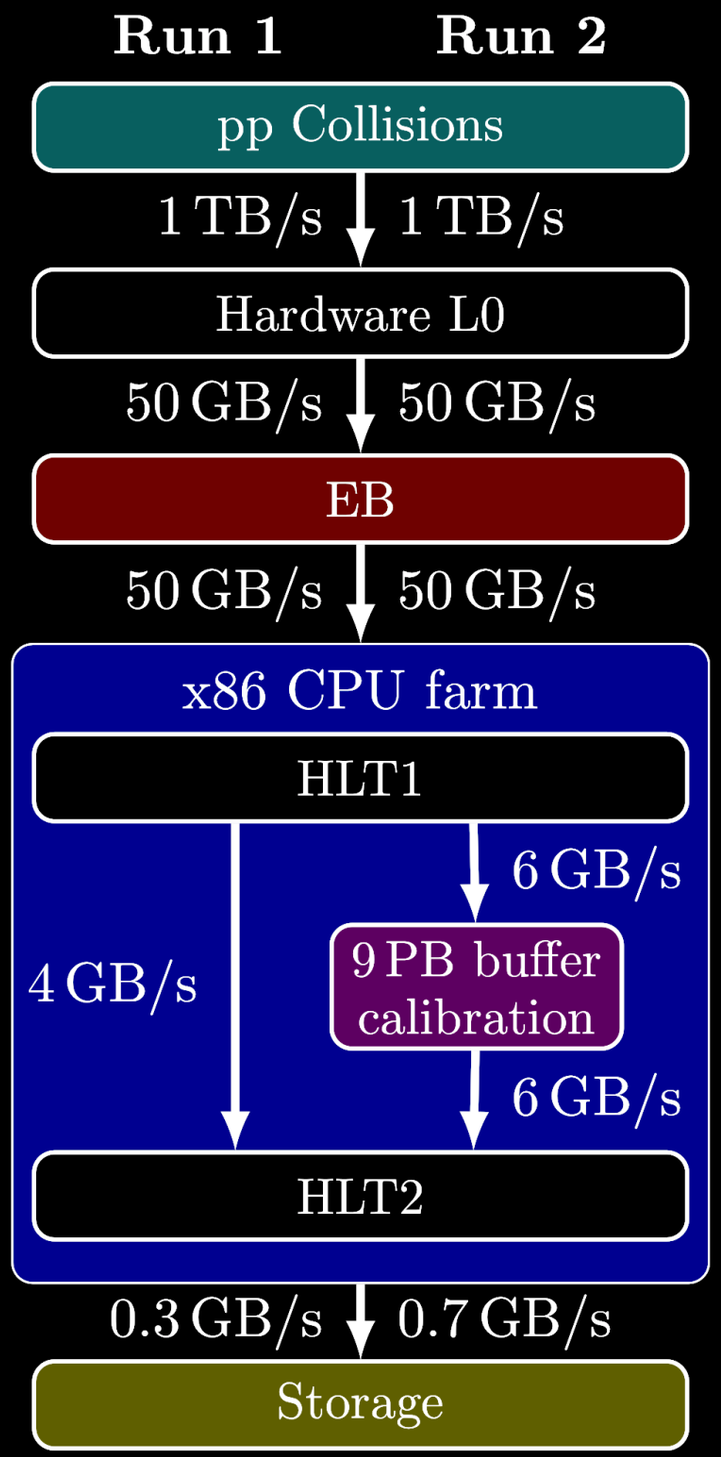
Mike Williams

Department of Physics
Laboratory for Nuclear Science
Statistics & Data Science Center
MIT

December 3, 2019

LHCb if painted by Van Gogh according to a Deep Neural Network.
https://github.com/jcjohnson/neural-style

Real-Time Analysis in Runs 1 & 2



readout limited to 1 MHz by frontend electronics

EB: event building on 250 dedicated PCs

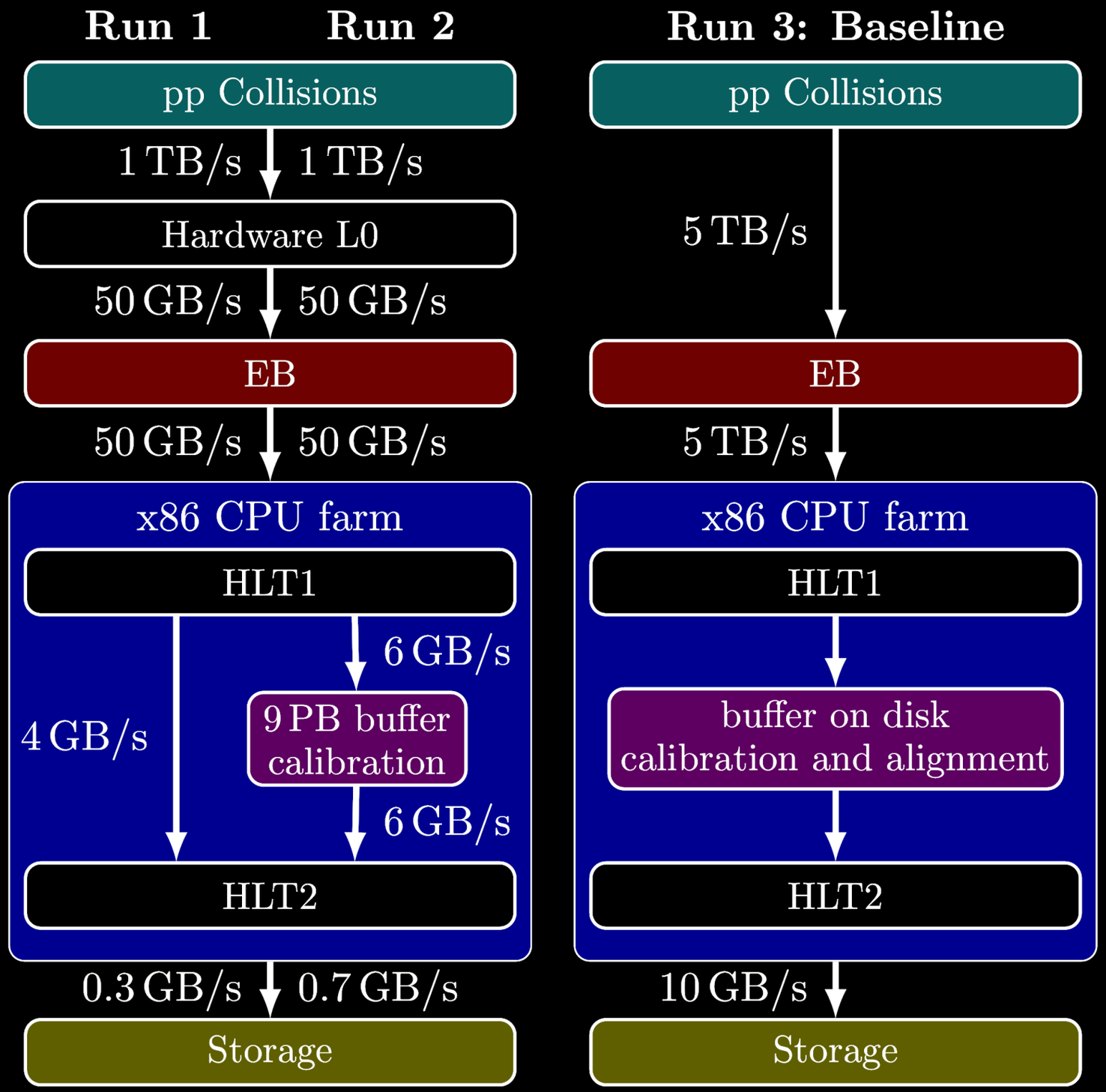
HLT1: tracking for all charged particles with $p_T > 500$ MeV

buffer data online while calibrations run
(can't use RICH detectors without hourly calibration)

HLT2: full "offline" reconstruction (done online)

Performed well beyond design goals in both Run 1 and Run 2.

Real-Time Analysis in Run 3



increase LUMI and upgrade frontend to readout full 5 TB/s

Need 100 GbE here!

Need to do tracking @ 30 MHz (redesigned tracking systems)

Even a 50 PB buffer will fill up in less than a week; HD write speed also a challenge, 100 GB/s input

Need to do full offline-like reconstruction @ O(1 MHz)

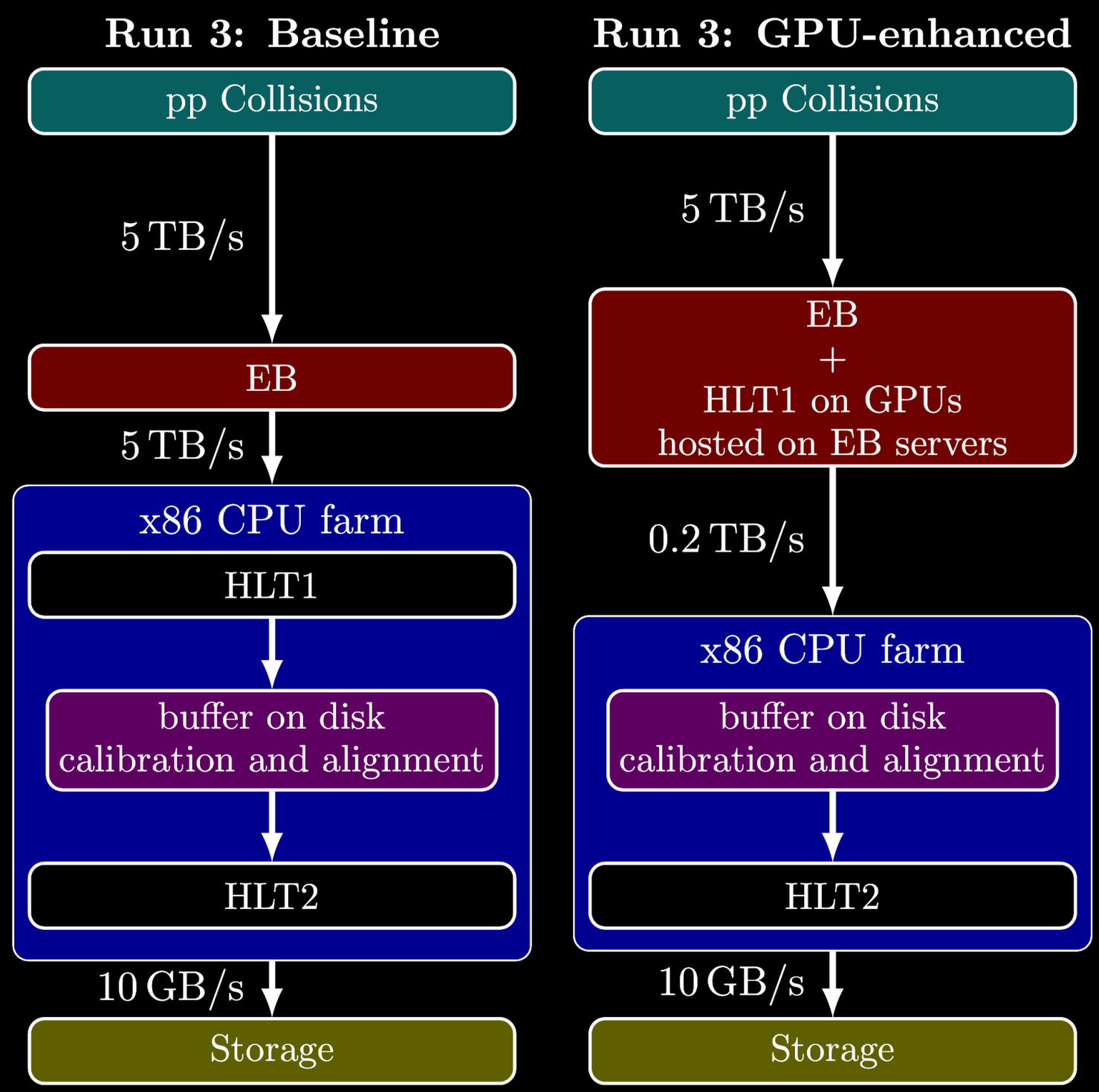
Removing the hardware trigger completely will greatly improve our physics performance!

Real-Time Analysis in Run 3

Alternative design: Use GPUs installed on EB machines to run the full HLT1 (full working proto-type application can do this with no p_T threshold @ 30 MHz on 500 GPUs).

No need for 100 GbE on EB PCs, can instead use on-board network cards; freed up PCIe slots host the GPUs.

Network savings ~ total cost of GPUs, plus GPU hardware is improving much faster than CPUs (plus all CPUs free to run HLT2). Only downside is FTE to develop & maintain CUDA reconstruction.



GPU-enhanced option currently being considered.

Summary

- LHCb is upgrading its tracking systems and frontend electronics — and increasing its luminosity — to enable reading out every event (no hardware trigger), a total of 5 TB/s.

- Transferring such a large data volume around puts a lot of strain on the networking. Must keep the system as compact as possible! (Distances of transmission one of main cost drivers.) In addition, networking no longer underutilized (not optimal); need to be careful about matching I/O to FLOPs in the system.

- Using higher-FLOP hardware like GPUs can reduce network costs (many times fewer GPUs needed to run tracking than CPUs; more compact). Our model avoids GPU-CPU I/O bottleneck by doing everything on the GPU (i.e. not using them as co-processors).

- Important to design tracking systems to facilitate faster tracking algorithms (e.g. nothing that has ambiguities that must be resolved later).

- HDD not improving as fast as everything else; buffering data is becoming more difficult, write speeds are a challenge.

- Bottom line: Hardware-less triggering is a challenge, but we are on track to meet this — and many interesting physics channels will see close to 10x improvement in signal efficiency (e.g. dark photon decays).