Real-time analysis workflows: thoughts for IF-04

Caterina Doglioni (Lund University)

Input from: Vladimir Gligorov, Mike Williams, Mike Sokoloff, Conor Fitzpatrick, William Kalderon, Antonio Boveia, Caterina Doglioni, Lukas Heinrich, Baptiste Ravina, Honey Gupta (GSoC), Maurizio Pierini, Javier Duarte, Thong Nguyen - see additional interested groups of people in slide 7

Real-time analysis workflows

Challenge: process and draw inference from large amounts of data (e.g. where the entire dataset cannot be recorded to permanent storage)

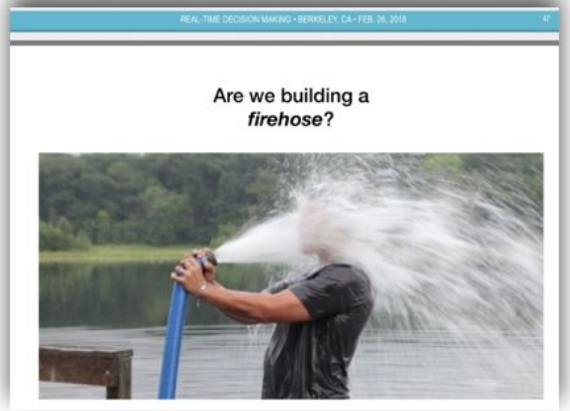
Science cases: cross-collaboration (LHC and beyond).

- For hadron collider experiments: analyse data in real-time to expand the physics program within the same resources e.g. rare new particles buried in large backgrounds
- For others (e.g. astro experiments): move signal / background processing as close to the detector as possible to inform subsequent decisions

C. Fitzpatrick, LHCb

Likely to be the same @ future hadron colliders

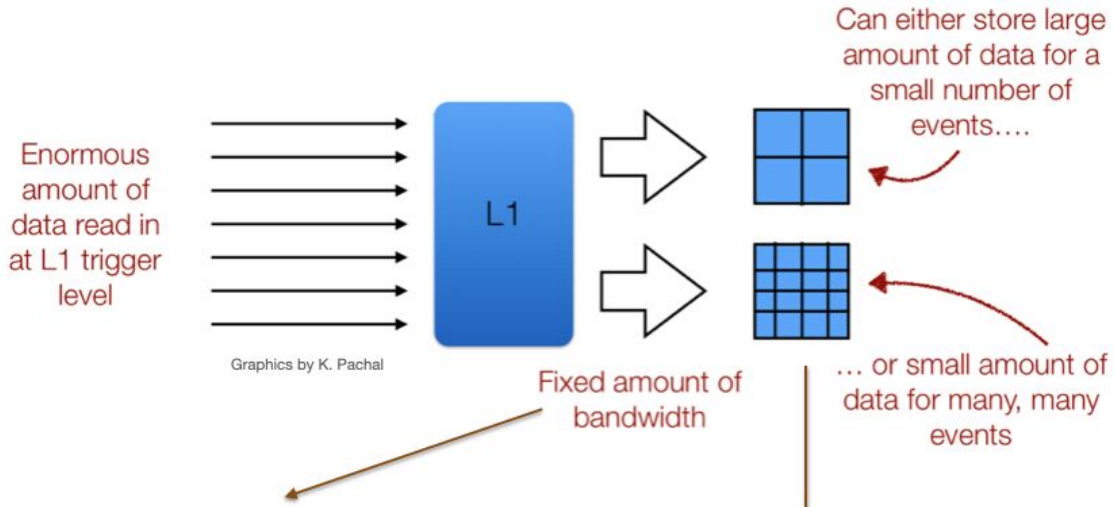
E. Bellm, LSST

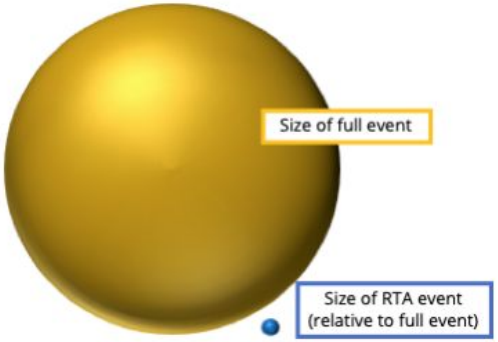


https://zenodo.org/record/3906485#.XxCCxC_Mx3M

Data Scouting (CMS)
Trigger-level Analysis (ATLAS)
Turbo stream (LHCb)

Implementation of a real-time analysis workflow





Perform as much "analysis" as possible @ HLT

- Reconstruction & calibration
- First preselection to skim "backgrounds"

Reduced data formats:

- Only keep final trigger objects (drop raw data)
- Save only "interesting" parts of the detector
- A combination of the two

Present/future real-time analysis focus points

Online (== offline) reconstruction software
- tracking in trigger crucial for pile-up suppression
- use of hybrid computing architectures (accelerators)

ML for real-time analysis: example approaches under evaluation in HEP for HL-LHC
- anomaly detection (e.g. autoencoders) for data quality / detector monitoring
- use of DL for prediction of time-varying calibration sequences
- autoencoders for data compression (for online data analysis and simulation)
- use of accelerators for faster online reconstruction algorithms

For discussion

For backup slides: see https://zenodo.org/record/3906485#.XxCCxC_Mx3M

HEP Software Foundation roadmap for 2020:
https://arxiv.org/abs/1712.06982
Trigger&reconstruction executive summary:
https://arxiv.org/abs/1802.08640
Update for LHCC review, 2020
https://zenodo.org/record/3779250

What to do next?

We would like to continue discussing real-time analysis as one of the principles to build future detectors and their trigger systems

Material from the HEP Software Foundation

- We will submit the LHCC review whitepaper on common software for HL-LHC as a Computational Frontier whitepaper
    - it includes a trigger and reconstruction chapter, should we do something specific to have it included in IF04?
- Collection of ideas/ongoing efforts rather than specific studies

Individual contributions

- Many of us are busy (in our experiments/other Snowmass Frontiers/TGs)...
- How to contribute & collaborate, in practice?
- One possible idea (that some of us are working on): compare “triggerless” lepton colliders and real-time analysis hadron colliders for searches with high-rate background
    - may not be in time to submit a LOI by August 31st, but as work progresses we could present in future IF04 meetings

Interested parties/collaborations (not an exhaustive list)

- People who contributed ideas to these slides (originally from Institut Pascal workshop → biased towards participants of the workshop):
  - Vladimir Gligorov, Mike Williams, Mike Sokoloff, Conor Fitzpatrick
  - William Kalderon, Antonio Boveia, Caterina Doglioni, Lukas Heinrich, Baptiste Ravina, Honey Gupta (GSoC)
  - Maurizio Pierini, Javier Duarte, Thong Nguyen
- HEP Software Foundation
  - Including a Trigger & reconstruction group, conveners: CD, David Lange, Agnieszka Dziurda
- SMARTHEP network
- REALTIME study group (local but interdisciplinary)