Non-standard workflows

Marco Montella, Tobias Fitschen, Caterina Doglioni

Talk structured around **EPPSU document questions:** long-shot questions that may help us make cross-experiment statements for future colliders.
They require a minimal amount of research (e.g. asking around about experimental/readout strategies), output could be a couple of paragraphs for the final document (with lots of citations)
→ volunteers/collaborators welcome, we can work within the HSF Trigger&Reco group!

PR slide for non-standard workflows: real-time analysis

Not covering here: non-standard workflows for LLPs, only non-standard workflows in trigger streams

Traditional data analysis is **asynchronous**:

First record and store data,
then reconstruct/analyze it

Real-time data analysis

Analyse data as soon as it is collected

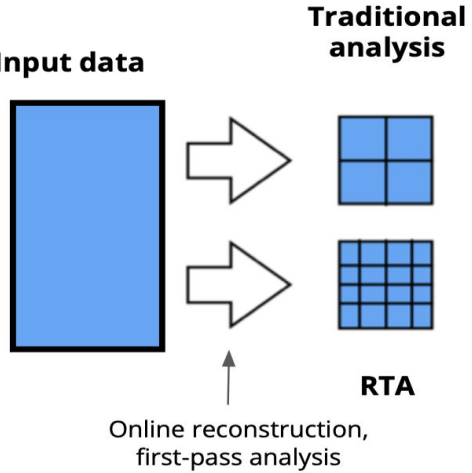
→ only **store smaller amounts of information**

→ reduce time-to-insight

→ accelerate decision making

(achieved using ML and hybrid architectures)

Also: connections with industry as this is not only "our" problem



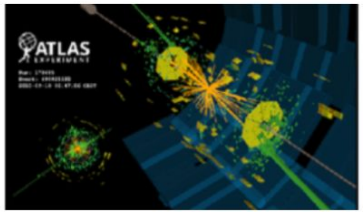
Example: store final-state information

→ save more data

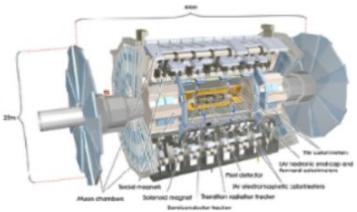
This concept will continue to be relevant for even more data-intensive future colliders

Where are the limitations that RTA can overcome?

(example of a LHC General Purpose Detector)



Detector readout to hardware trigger



L1 hardware trigger

Event selection (trigger)

Object reconstruction and calibration

Data analysis

CPU for processing events

large but limited processing power

Disk/tape to store events

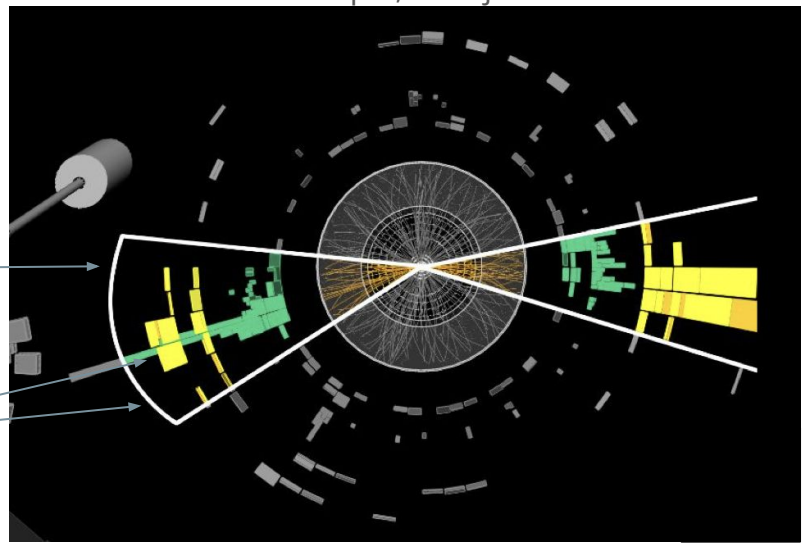
EPPSU document question: What will be the limitations of future collider / nuclear physics / HI experiments?

- Ingenuity, but realism: Moore's law is already not the answer to all our computing needs

Non-standard workflow options: a shopping list (with some LHC bias)

- Continuous Readout (ALICE/LHCb/nuclear physics)
  - Read out at 40 MHz (e.g. w/o a L1 trigger)
- Data Scouting (CMS) / Turbo Stream
- (LHCb) / Trigger Level Analysis (ATLAS)
  - Save HLT-recorded jets only
- Selective Persistency (LHCb) / Partial Event Building (ATLAS)
  - Save 4-momenta of jets + raw data behind it
- Data parking (CMS) / Delayed Stream (ATLAS)
  - Save everything, reconstruct later

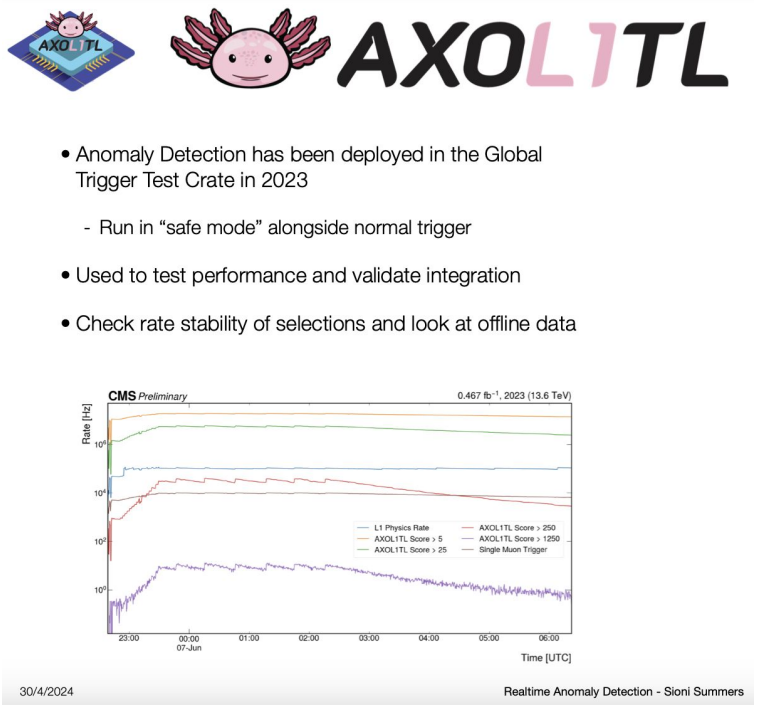
Example, with jets



EPPSU document question: What can we learn in the cross-talk between LHC / nuclear physics / HI experiments? What could we collaborate on?

Where do we go from here?

- All experiments going towards some form of **data analysis at 40 MHz**
  - Can make use of ML on edge architectures
  - Goal can be outlier detection, but also beyond…see also Dylan Rankin's talk today



S. Summers, EuCAIFCon 2024
Also: https://arxiv.org/pdf/2312.10009

EPPSU document question: What will we learn from HL-LHC in terms of RTA and its applications, and what will be left to innovate at future colliders?

A note on sustainability

- Interest in non-standard workflow from **energy sustainability** perspective
  - [from talks + discussions at this workshop] Even though we may be able to have a “green” energy grid, energy will still be scarce and expensive
- Examples of links between non-standard workflows and sustainability:
  - Find the right computing architecture for the problem
  - Reduce consumption of algorithms for constrained / edge architectures
- Change in perspective: learn to use fewer resources from online, *then* apply to offline
  - This won’t work everywhere, but we could list examples of where it works?
  - Also: how not to fall into https://en.wikipedia.org/wiki/Jevons_paradox (doing more with less means doing a lot more with the same resources, instead of doing the same with much less)

EPPSU document question: Does it make sense for our field to link lessons learned from real-time analysis to energetic sustainability?

Where sustainability is discussed for EPPSU

Cross-topic, if you want to focus on software/computing/ML you can join:

- The general Sustainable HEP effort mailing list and mattermost and meetings
  - Overall perspective, will include recommendations for computing
- The JENA (=**J**oint **E**CFA **N**uPECC **A**PPEC **A**ctivities) software working group meeting
  - A whitepaper has been drafted on common software and computing issues with a mention of sustainability, and will probably be used as some form of input to EPPSU
- The JENA EuCAIF (=**Eu**ropean **C**onsortium for **AI** in the **F**undamental sciences) for FAIR / sustainable machine learning
  - Activities are just starting, when2meet here (fill by tomorrow), for the first December meeting