ATLAS R&D on track reconstruction for Event Filter

S. Veneziano

11/10/24

ATLAS Phase-II T-DAQ

Diagram shows the overall T-DAQ architecture for Phase-II

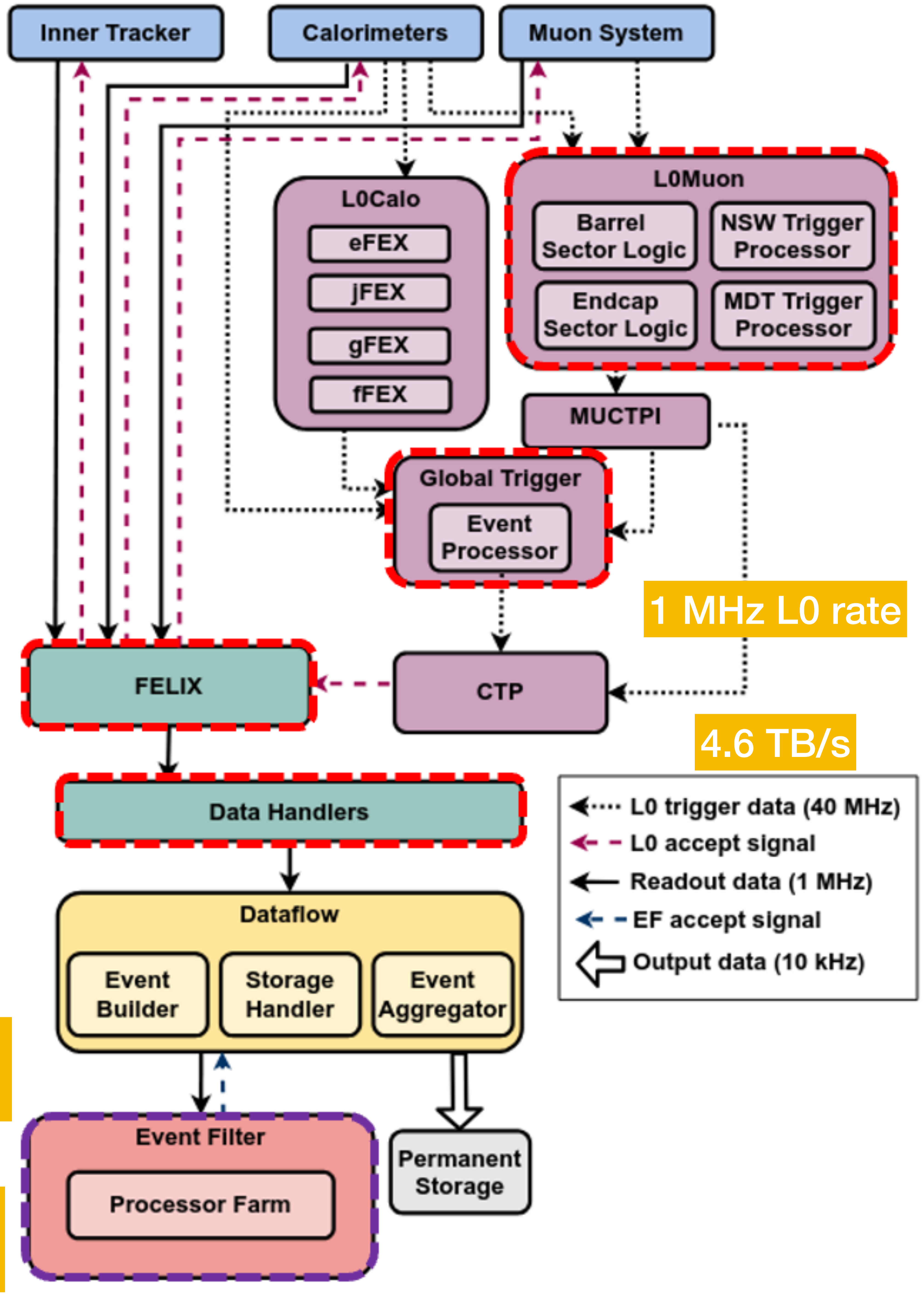
- ATLAS **R&D,** novel **(AI) approaches** and **innovation** for several sub-systems (dotted boxes)

- **Level-0** hardware trigger: **L0Muon** and **Global Trigger**
- **Event Filter** event processing: **Track reconstruction** in the **Inner Tracker** and **Muon detectors,** plus **ACTS** tracking software infrastructure
- **Novel trigger signatures and physics optimisation**

CPU ~8MHS06 (tracking ~6MHS06) in Run4

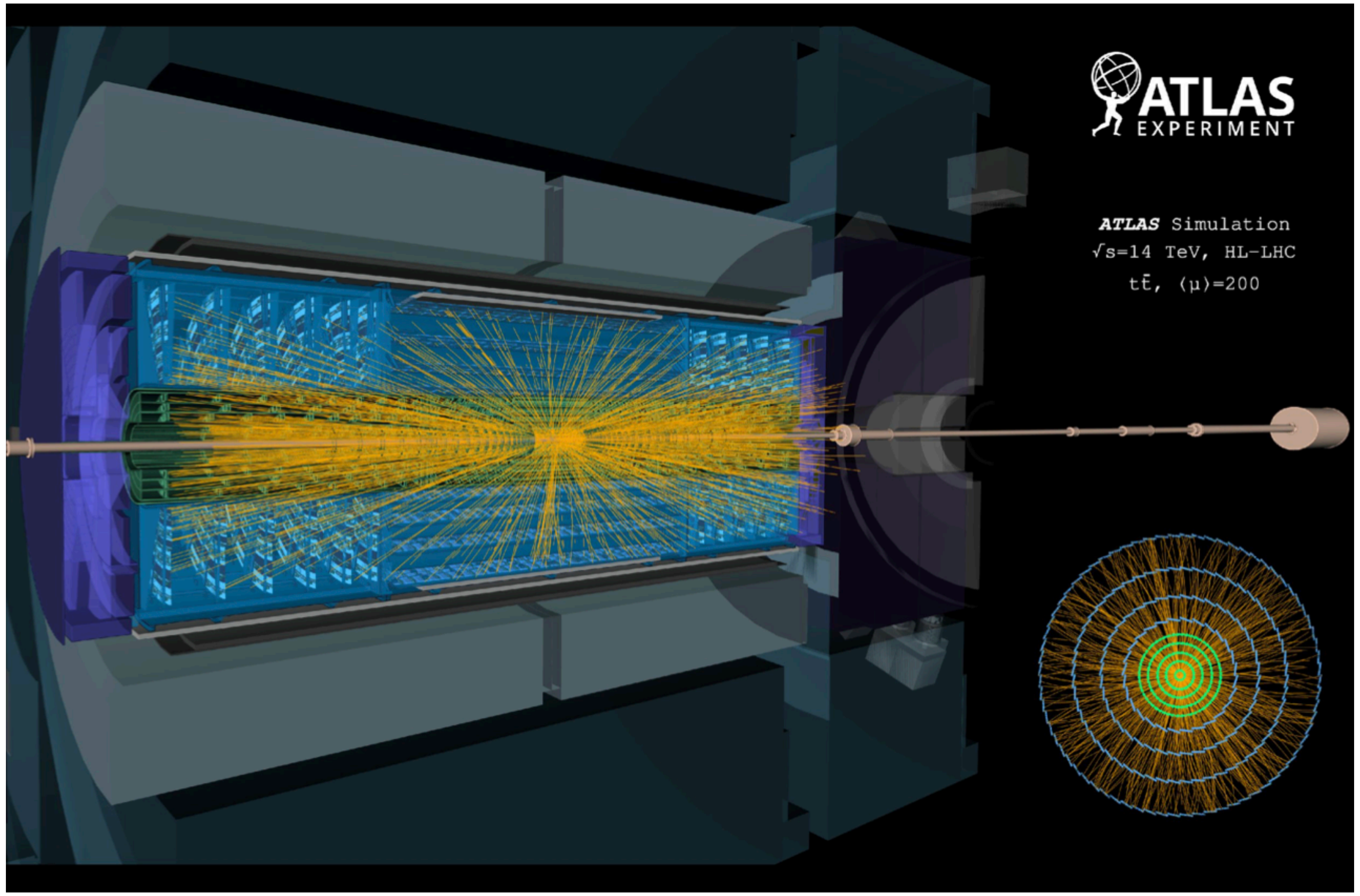
Let us focus only on the Event Filter and the Track reconstruction in the ITk.

Data reduction: 1 MHz ⟶ 10 kHz



Track reconstruction in ATLAS

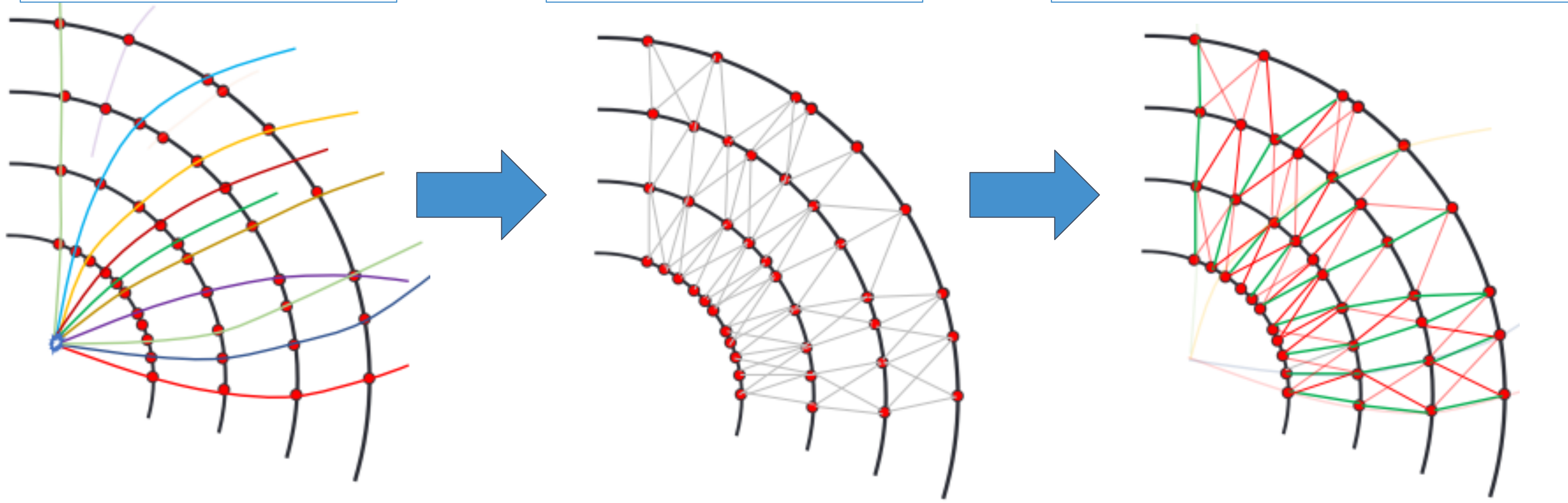
- The ATLAS Phase-II tracker has about 5×10^9 readout channels → 3×10^5 space points / event
  - Data is sparse!
- In a collision event, generated particles leave hits in the detector. Track reconstruction recreates particle trajectories from detector hits.
- An expensive process, especially at high pile-up. HEP community seeks to develop hardware-accelerated, ML-based tracking algorithms.
- We build a machine learning pipeline based on Graph Neural Network (GNN) for track finding under HL-LHC condition (μ = 200).



Charged particles leave hits in the detector

Represent the data using a graph

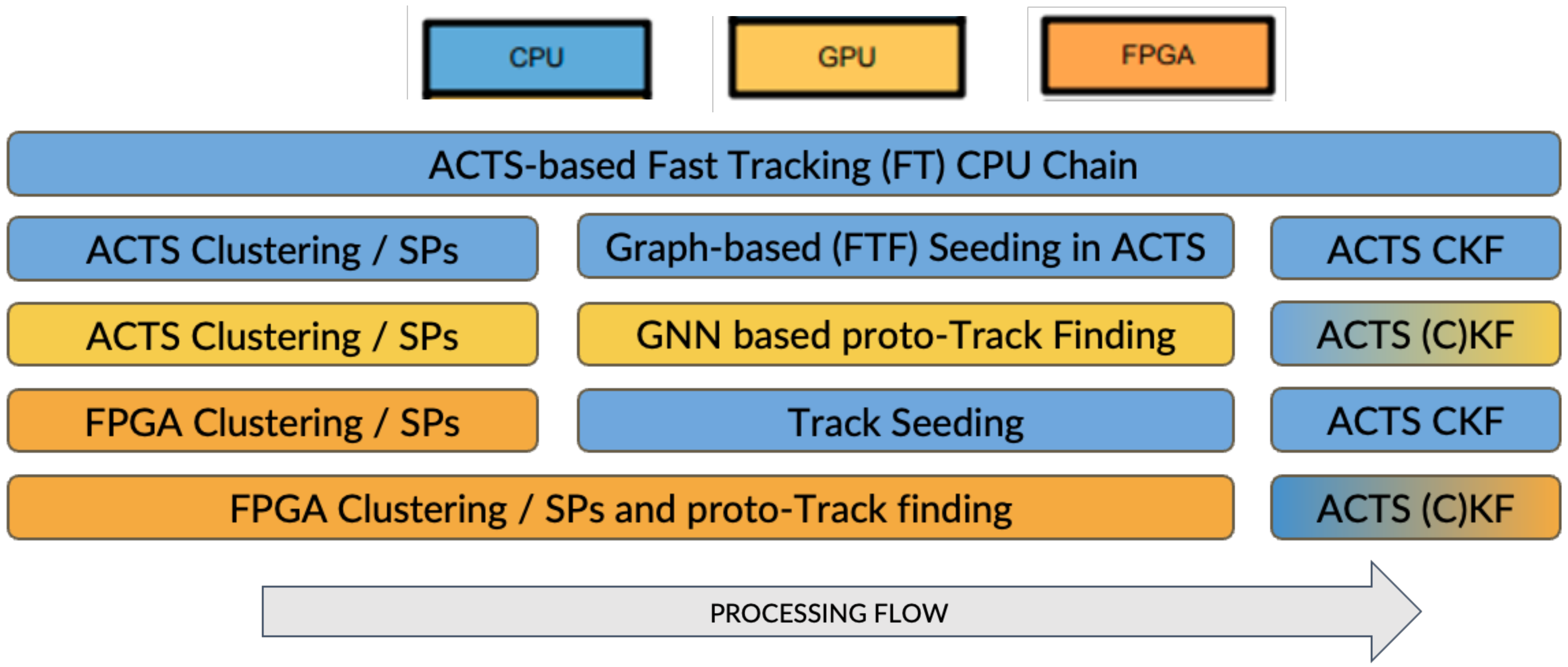
Classify node edges in high or low probability to belong to real tracks



Event Filter: ITk Reconstruction

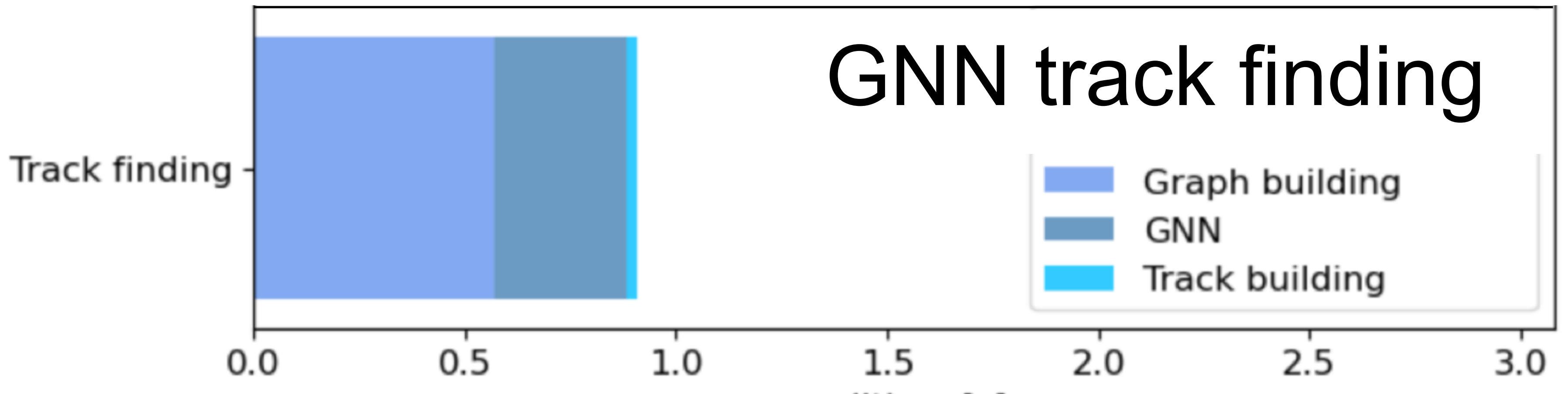
ATLAS to decide Event Filter technology in 2025

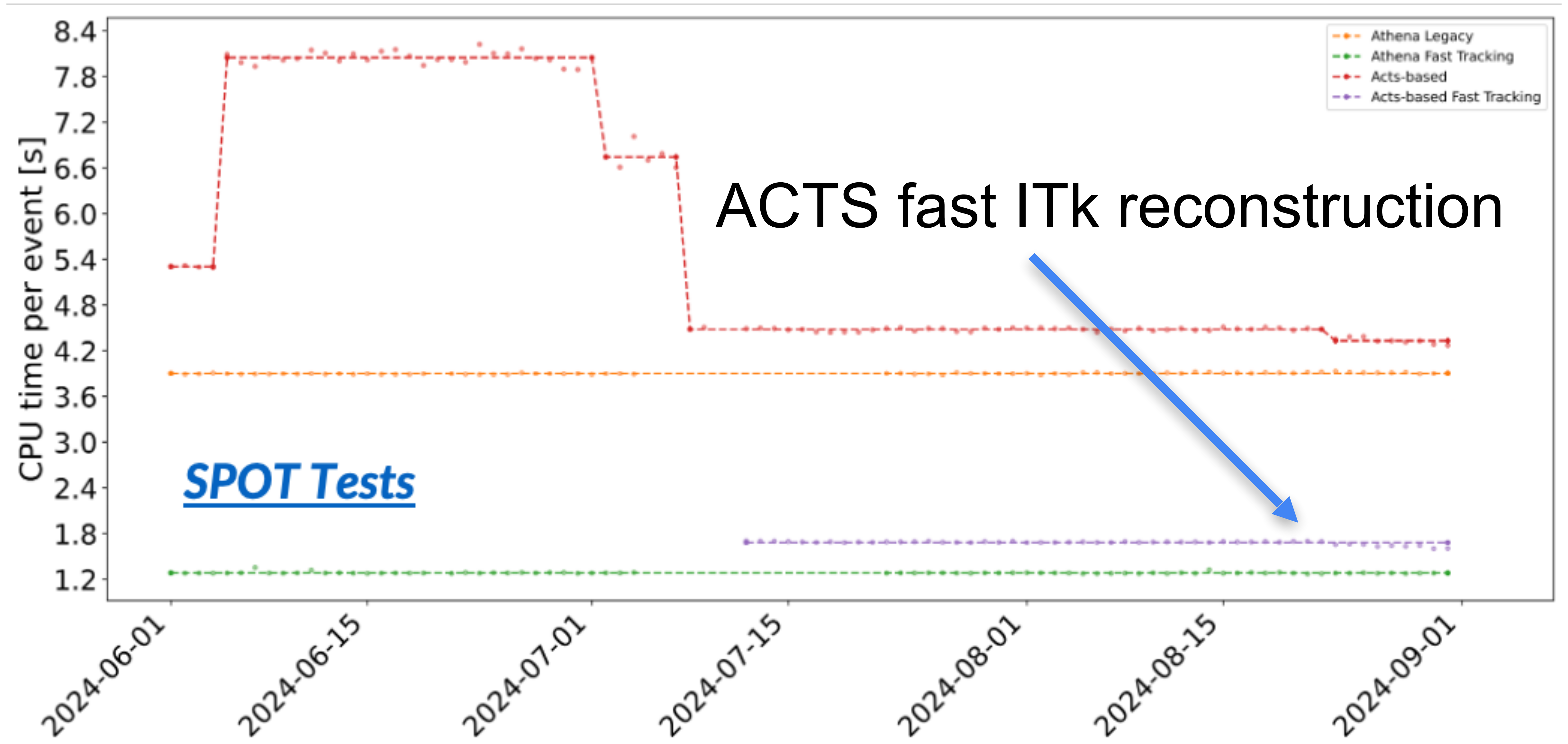
- R&D for number of chains
  - Investigate use of **GPU** and **FPGA** processors and studying **GNNs** as an alternative track finding approach
    - **Novel approaches to the reconstruction of charged particles trajectories using Graph Neural Networks (GNN) are being actively developed**
- Bringing **ACTS fast tracking** (general software framework) to production level



Enabling integration and R&D work

- Development **support** for CPU aspects of FPGA chains
- ACTS fast ITk reconstruction now integrated in automated ATLAS **SPOT CPU monitoring**
- Interfacing GNN based track finding with CPU based reconstruction steps
- First results of **GNN track finding on A100**

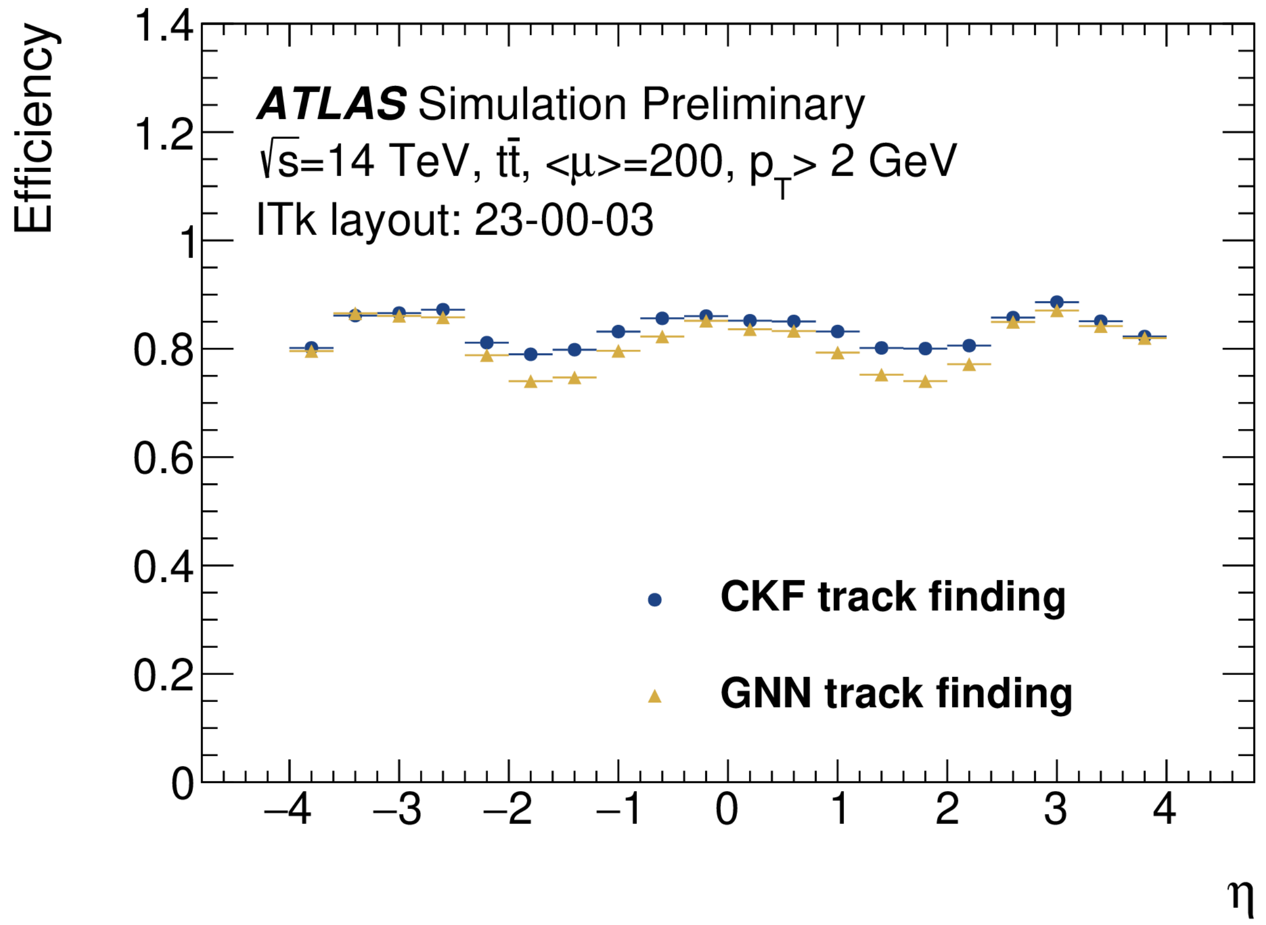


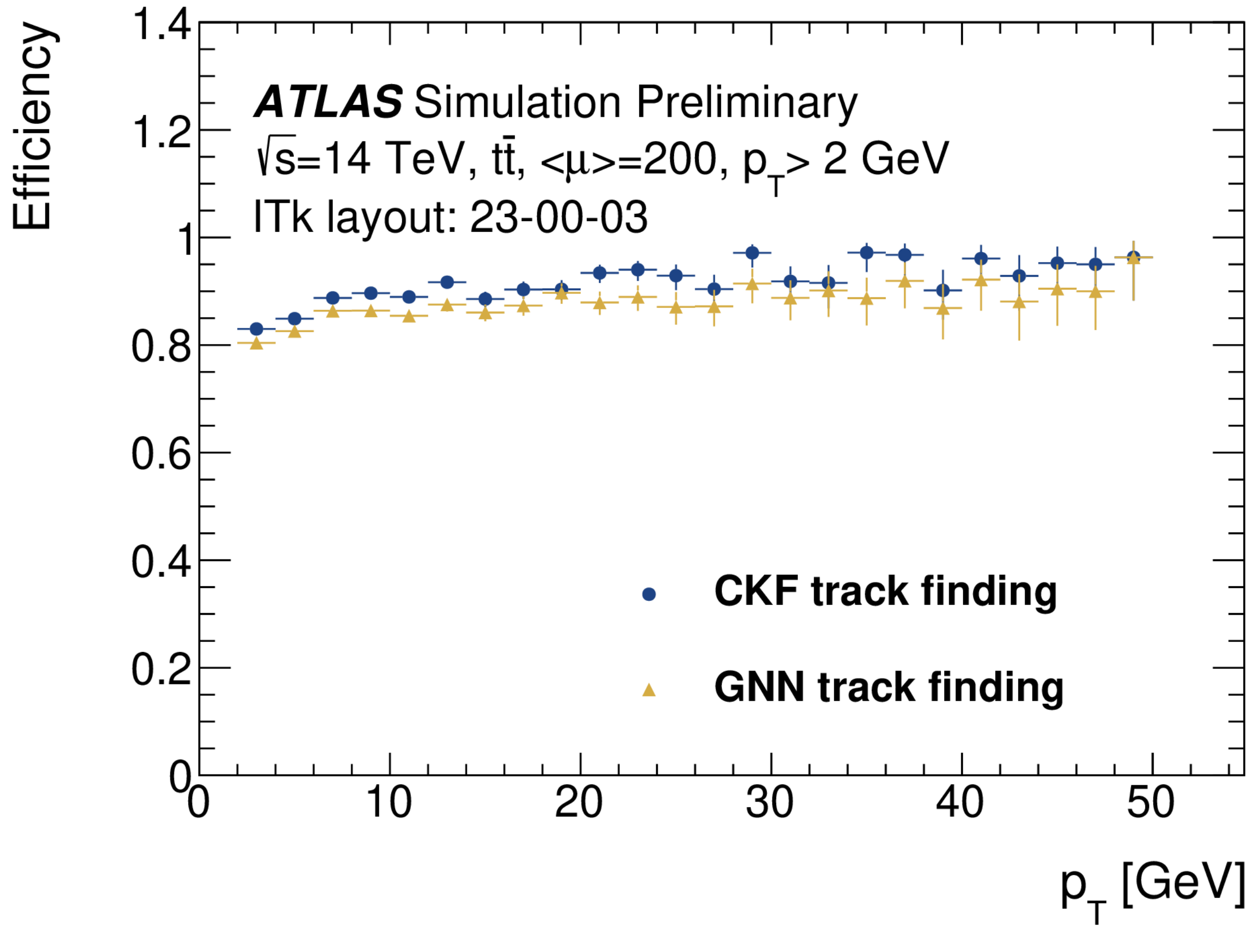


ATLAS Collaboration, IDTR-2023-06, October 2023 (link)

H. Torres of behalf of the ATLAS Collaboration, Proceeding of Connecting the Dots 2023 (link)

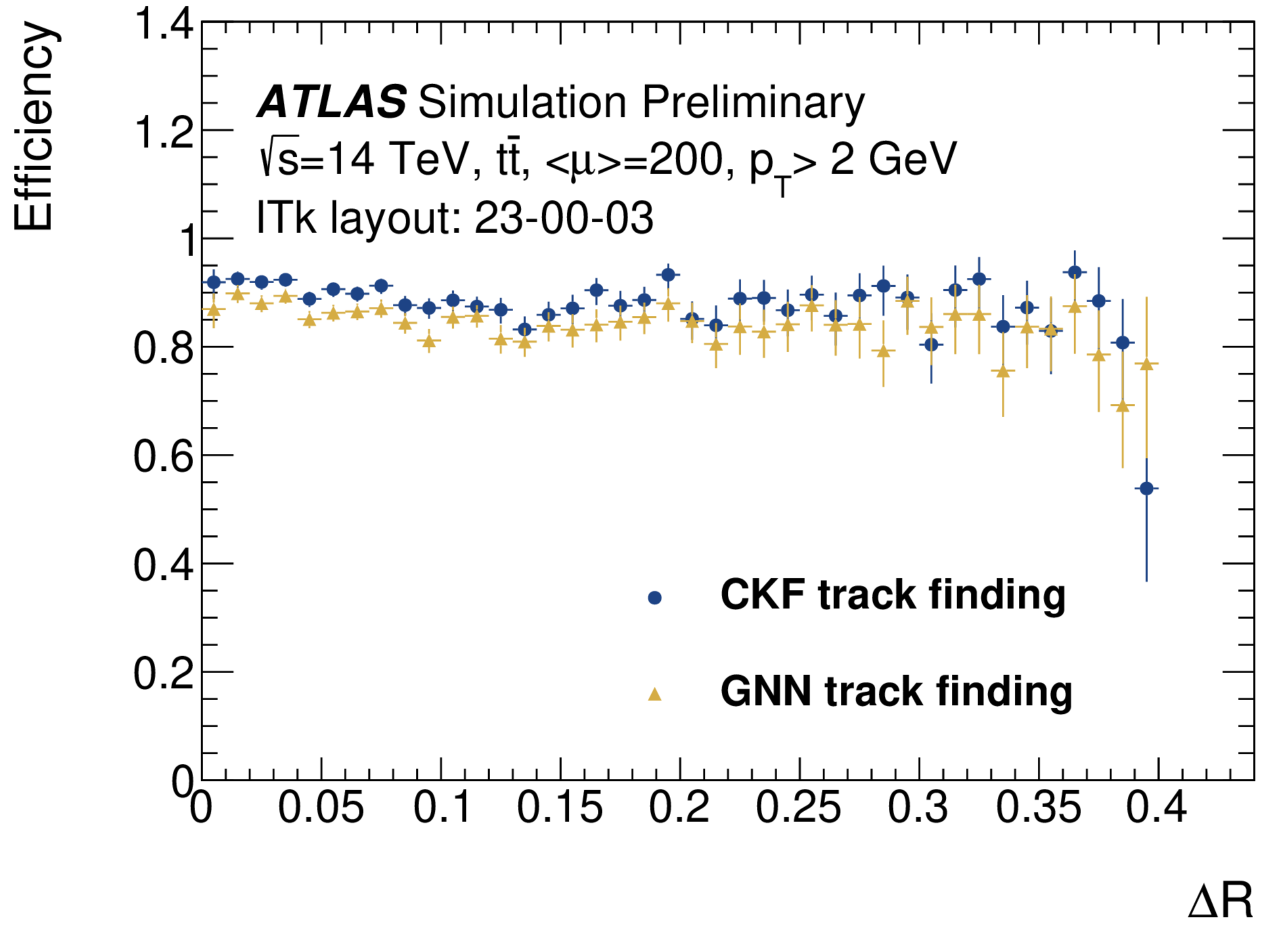
Current physics performance

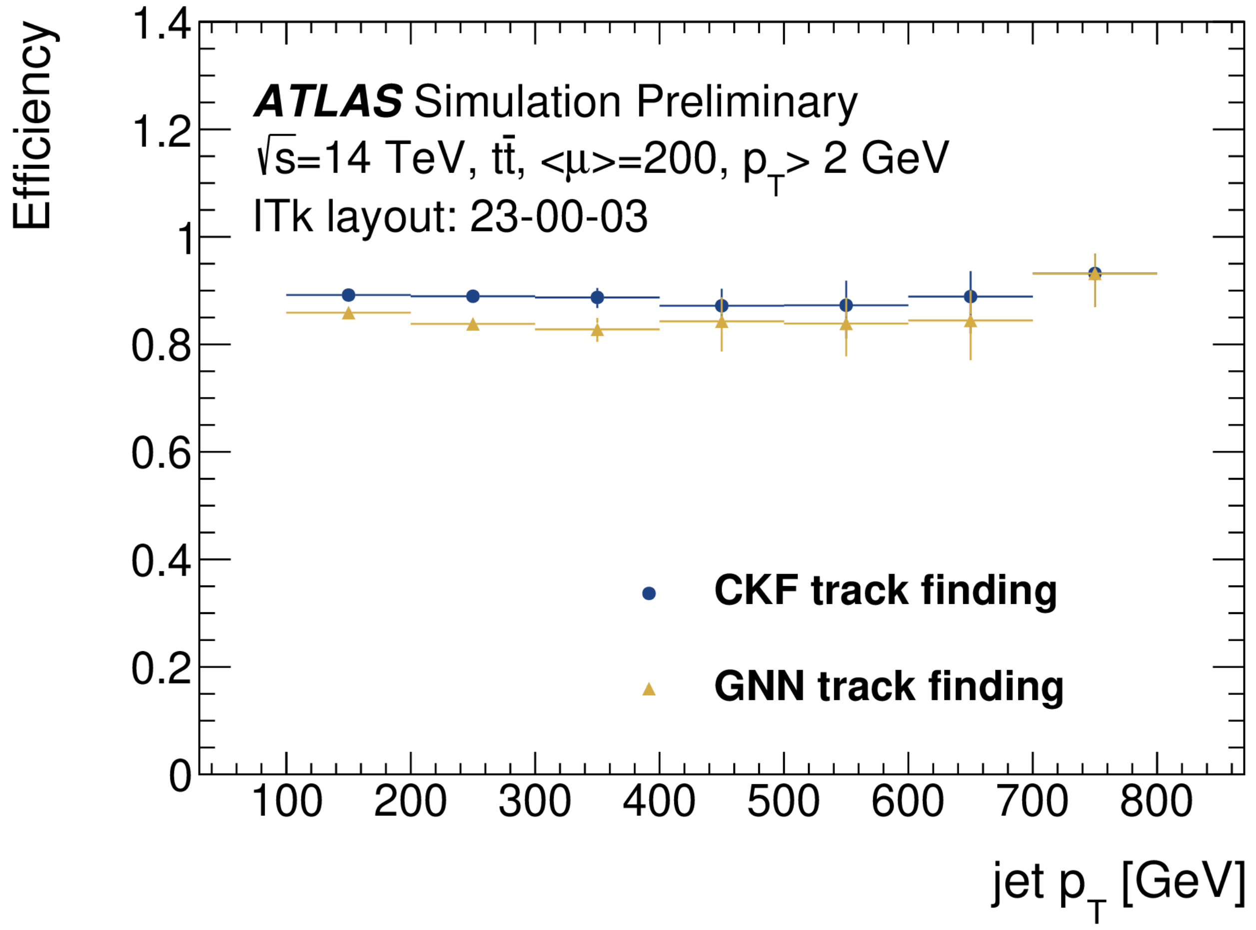




Performance is already very close to standard techniques, much space for exceeding state-of-the-art performances and to finally gain on event filter overall compute power and:

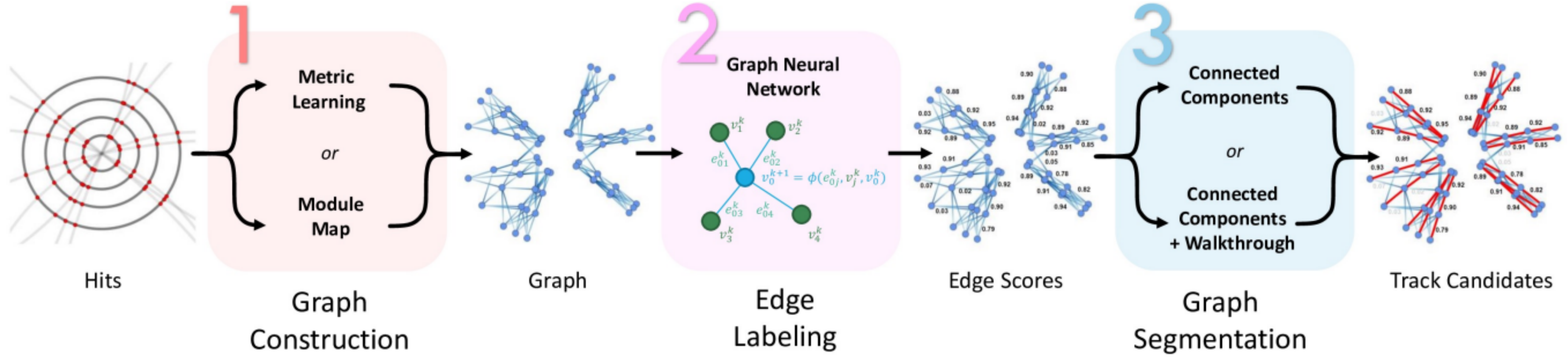
- **farm physical size**
- **cost**
- **power consumption**



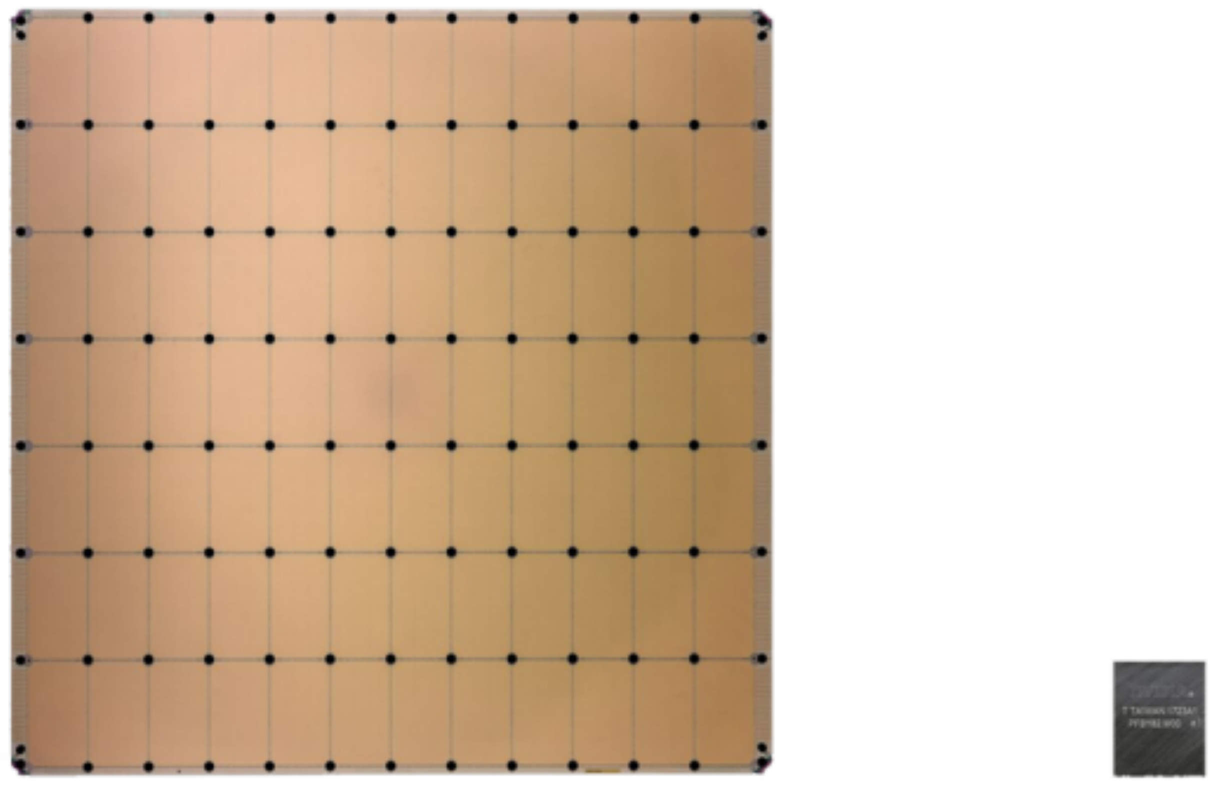


Tracks inside jets

GNN 4 ITK with Cerebras WSE-3



- Current employed hardware limits this line of R&D:
  - O(weeks of training) on Nvidia A100
    - → **Slow Model training turn-around**
  - Hw memory constraints limit number of parameters and possible graph size → **Difficult scaling up of our models**
  - Very large combinatorics to deal with → **Limit Investigation of new model architectures**
- The characteristics of Cerebras WSE-3 has the potential to unlock these R&D avenues



Cerebras WSE-3
4 Trillion Transistors
46,225 mm² Silicon

Largest GPU
80 BillionTransistors
814 mm² Silicon

cerebras [WSE-3 Datasheet]

Backup

LHC Timeline