Triggering and DAQ for a Future Muon Collider Experiment

Isobel Ojalvo
Assistant Professor, Princeton University

Viaduc des Fauvettes

Overview

Review of trigger systems at the LHC

Considerations for a Muon Collider Experiment

Future Technologies and R&D

Going Forward

Triggering at the LHC

At L = 1-2x10^{34}cm^{-2}s^{-1}

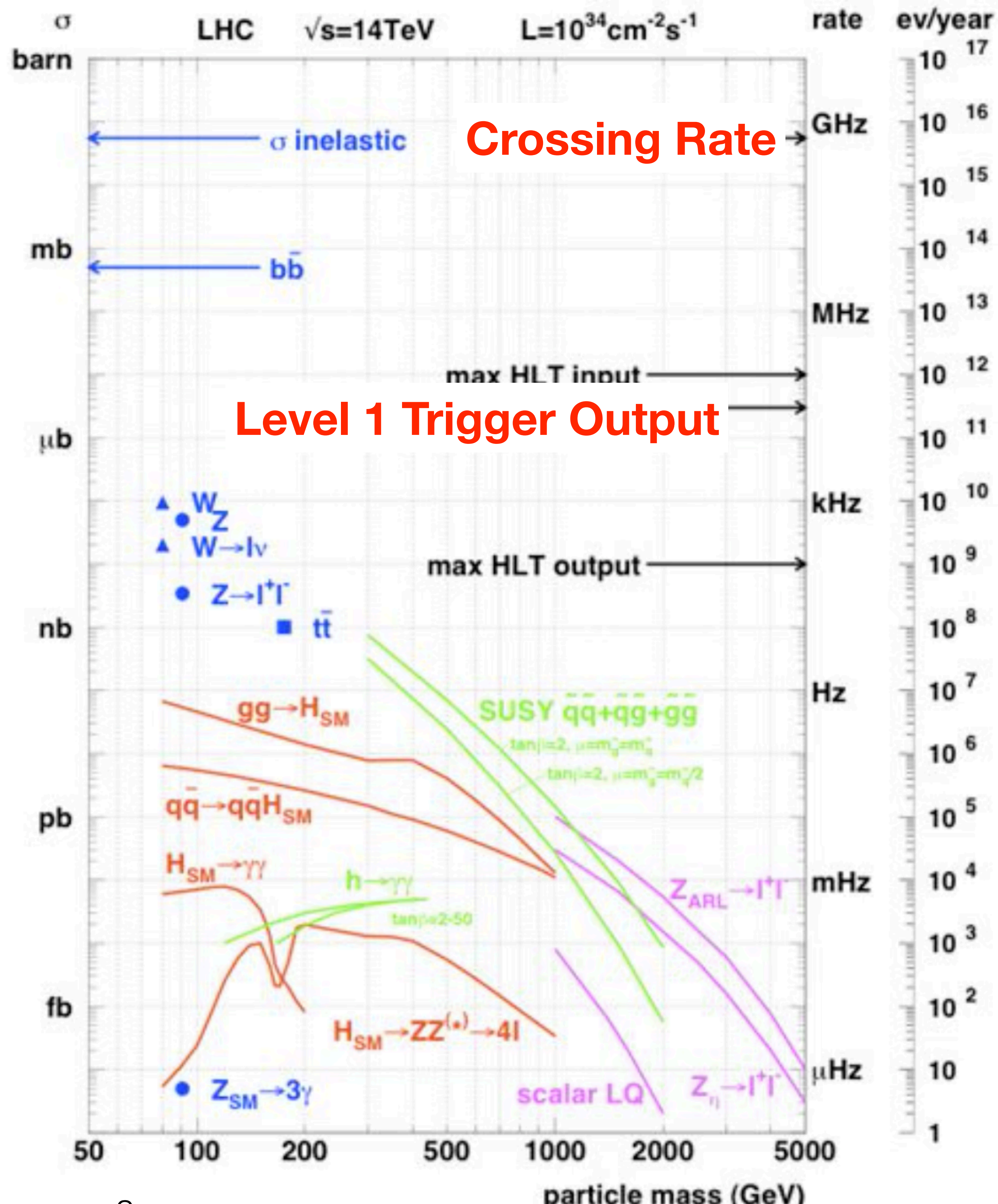
- 40 MHz collision rate
- <40> collisions/bunch crossing
- Higgs: ~0.1 Hz, H4l: 0.1 mHz
- Output data size 5 MB/event

CMS and ATLAS process and **selects events** in stages

- “First Stage” Hardware Triggers with partial detector readout
  - reduce 40 MHz to ~100 kHz
- “Higher Stage” Software Trigger with full detector readout
  - reduce ~100kHz to ~1kHz
  - Tracks fully reconstructed

LHCb and ALICE

- Triggerless readout systems
  - Both have lower luminosity and high purity of ‘interesting’ physics
  - ALICE has a lower crossing rate (600 Pb filled compared to 3k p filled bunches per train



LHCb Experiment

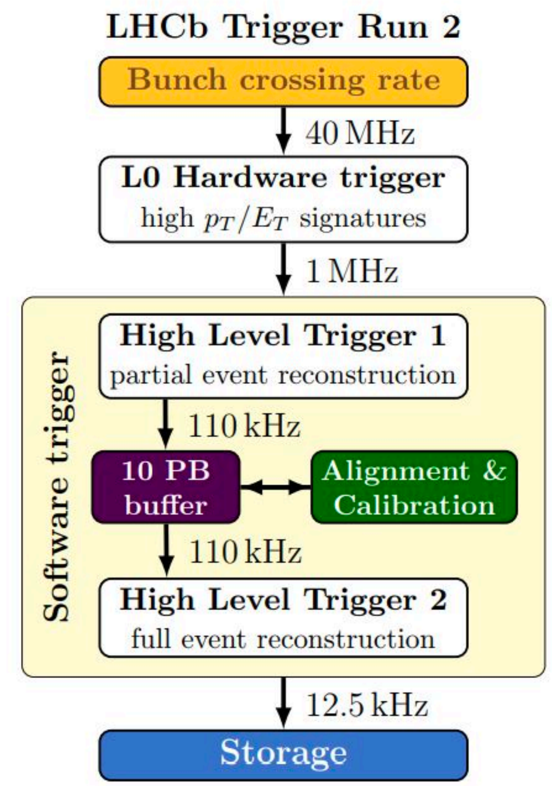
For **Run-3 LHCb removed the hardware trigger** and moved to a streaming approach

- Expecting to take 50 fb-1 of data for Run-3 (CMS/ATLAS expect ~300 fb-1)
- Run at 32 TB/s

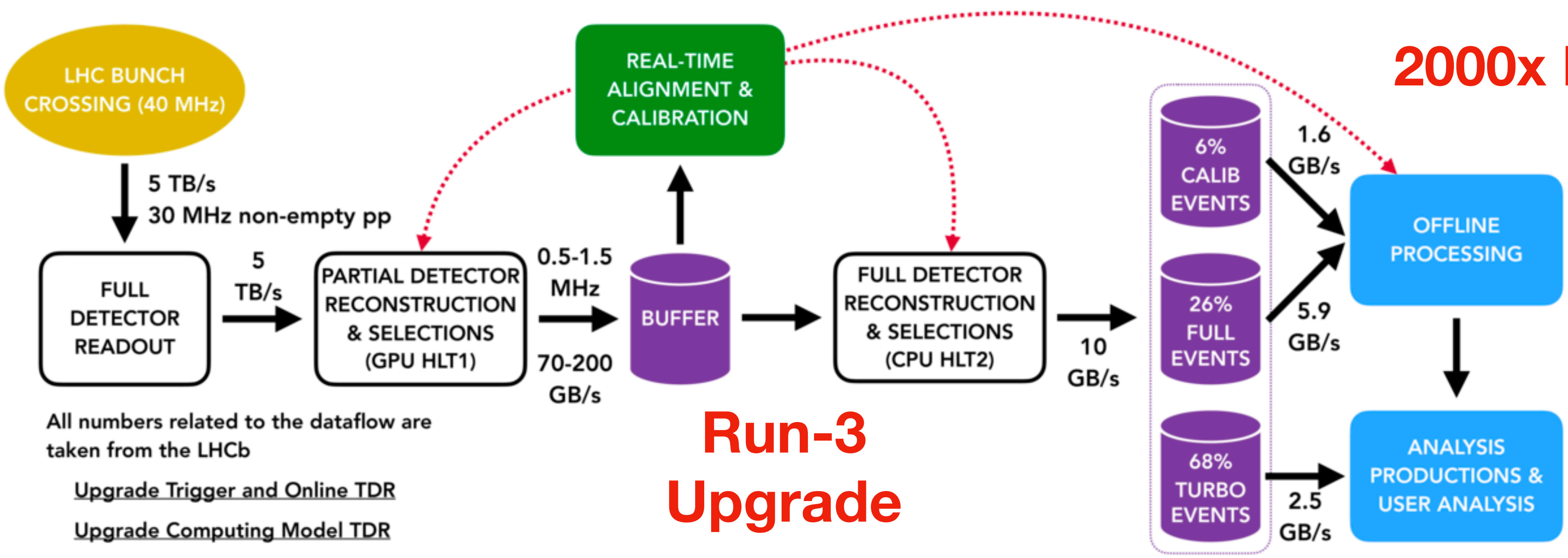
Use “real-time” analysis to reconstruct objects

- Reduce data from 30 MHz to 1 MHz at **HLT1**
- Utilize GPU co-processors (required adoption of CUDA)
- Real time calibration of detectors
  - Requires updating calibration constants every run for some (but not all) sub-detectors
- Full detector Reconstruction performed at HLT2

Run-2



2000x Data Reduction



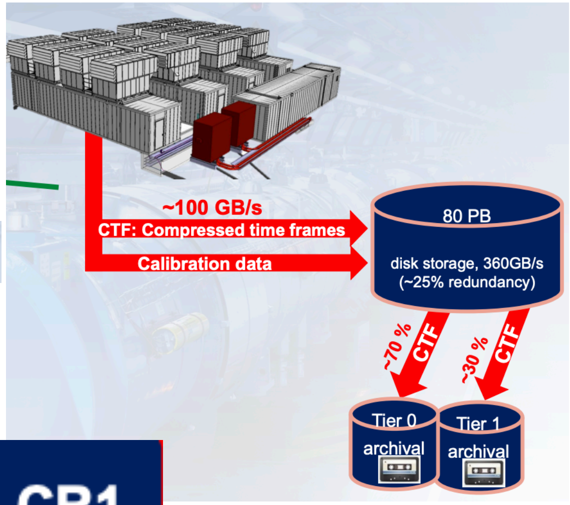
Run-3 Upgrade

talk by C.Agapopoulou

ALICE Experiment

Pb-pb collisions at $6x10^{27}$ $cm^{-2}s^{-1}$, collision rate of 50 kHz, 500 kHz for p–Pb, and 1 MHz for pp collisions

- Streaming Readout of all Detectors

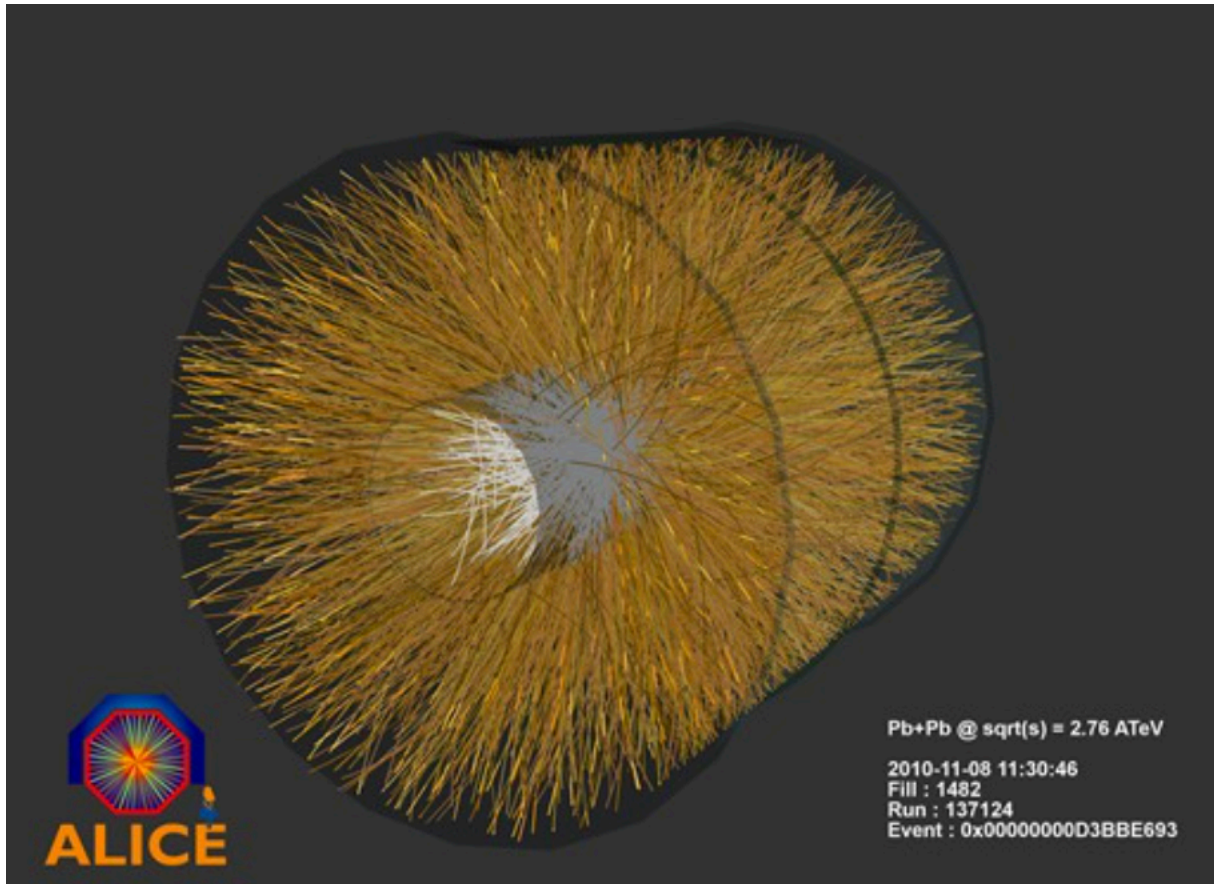


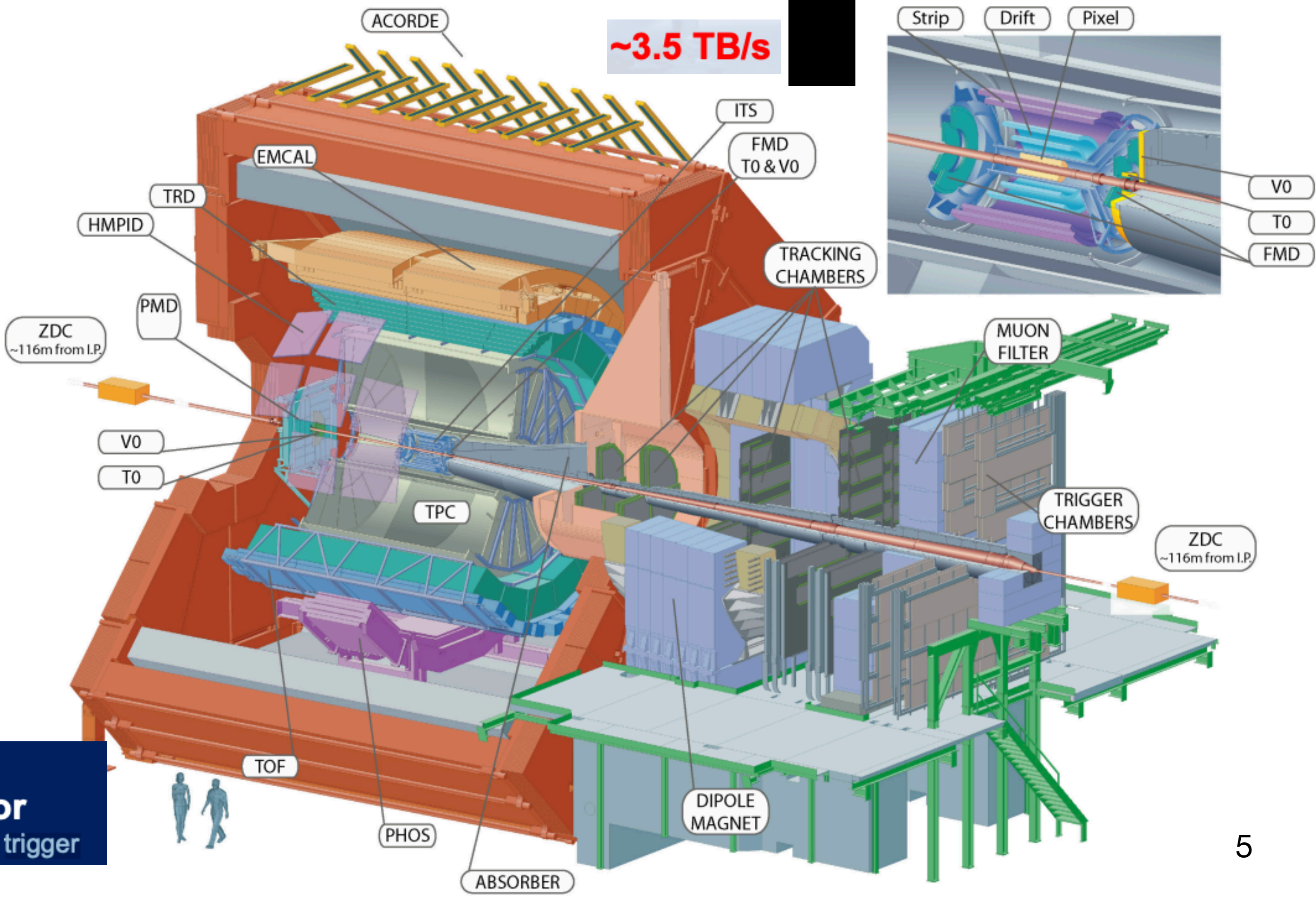
~635 GB/s

CR1

Continuous Raw Data Stream
sent to Processor Farm

~3.5 TB/s



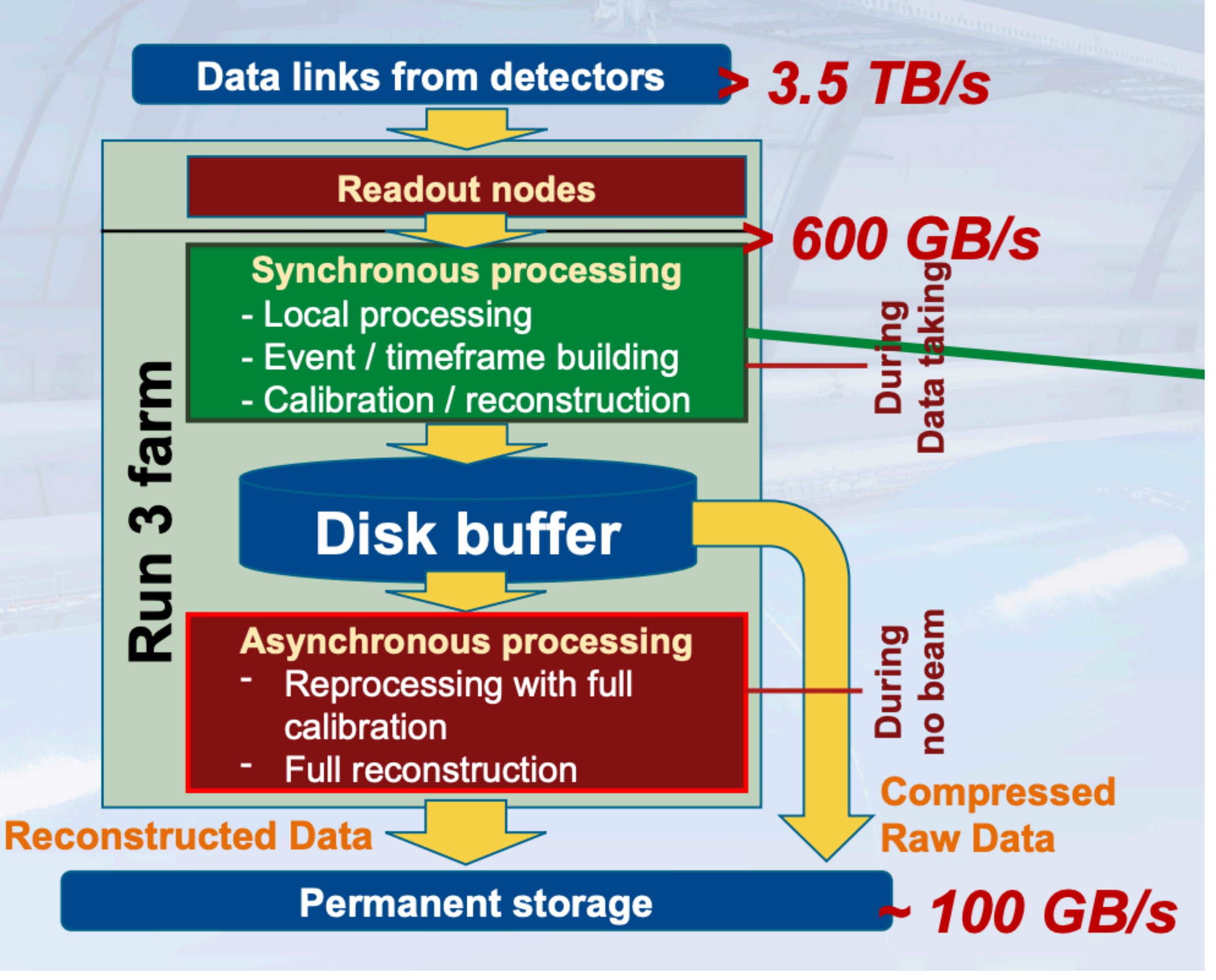


CTP
Central Trigger Processor
Distribution of timing info, heartbeat trigger

ALICE Experiment

Pb-pb collisions at $6x10^{27}$ $cm^{-2}s^{-1}$, collision rate of 50 kHz

- Each collisions is shipped to online systems
- First undergo Synchronous Processing (FPGA based reconstruction)
  - **Data compressed by** storing space point coordinates as residuals to tracks to reduce the entropy and remove hits not attached to physics tracks - 20x compression
    - Possible loss of information if tracks reconstructed incorrectly
  - **Detector Calibration**
  - Full TPC **tracking**
- Next undergo Asynchronous Processing
  - (**GPU based reconstruction**)
  - Full reconstruction, processing, all detectors



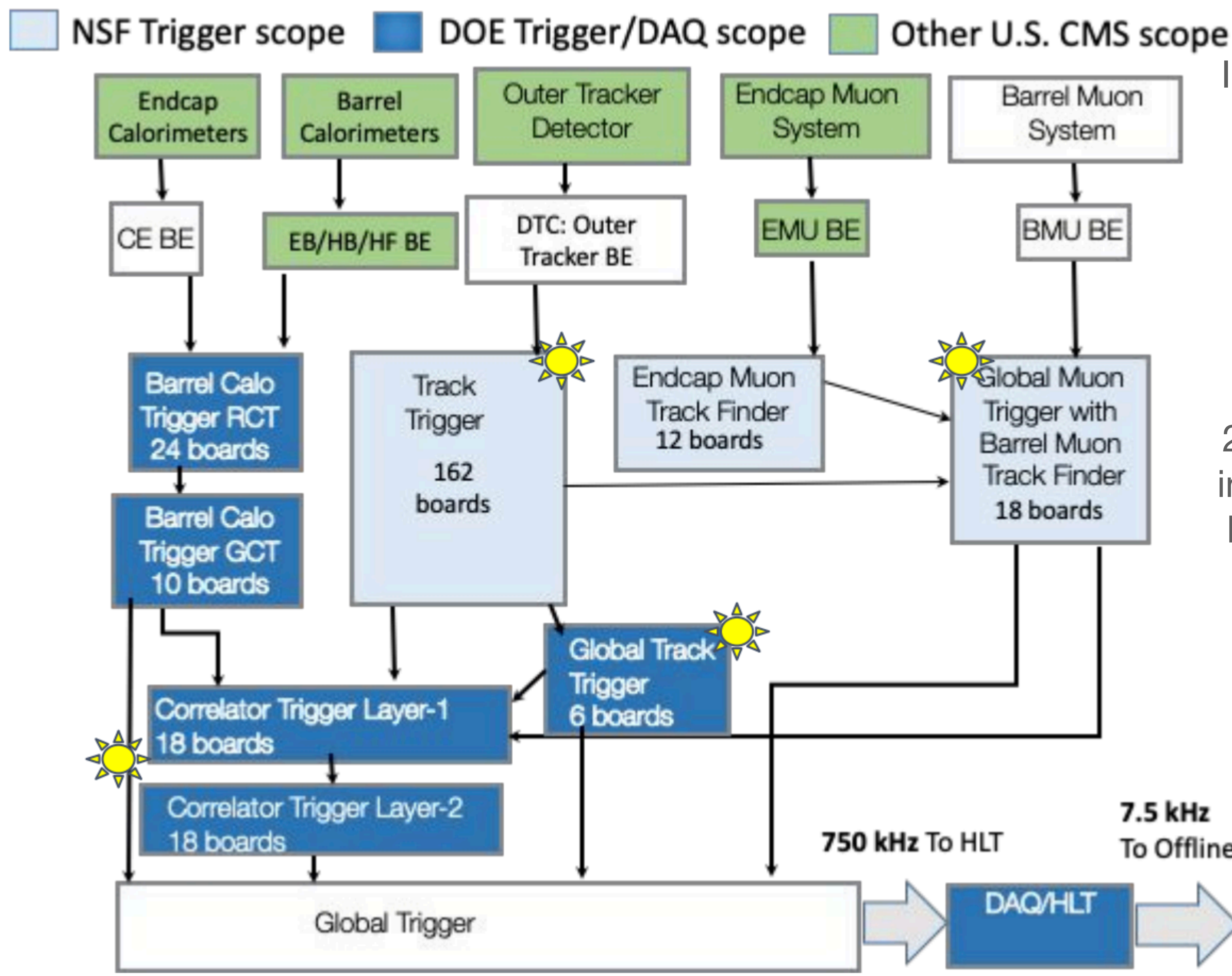
35x Data Reduction

Phase-2 Trigger for CMS

Based primarily on Xilinx Ultrascale+ FPGAs and 25 GB/s optical links interconnected between systems

First time tracking available at CMS

Possible due to careful detector design



Input at ~50 TB/s
40 MHz Rate

25 GB/S optical links interconnect between FPGA-based boards

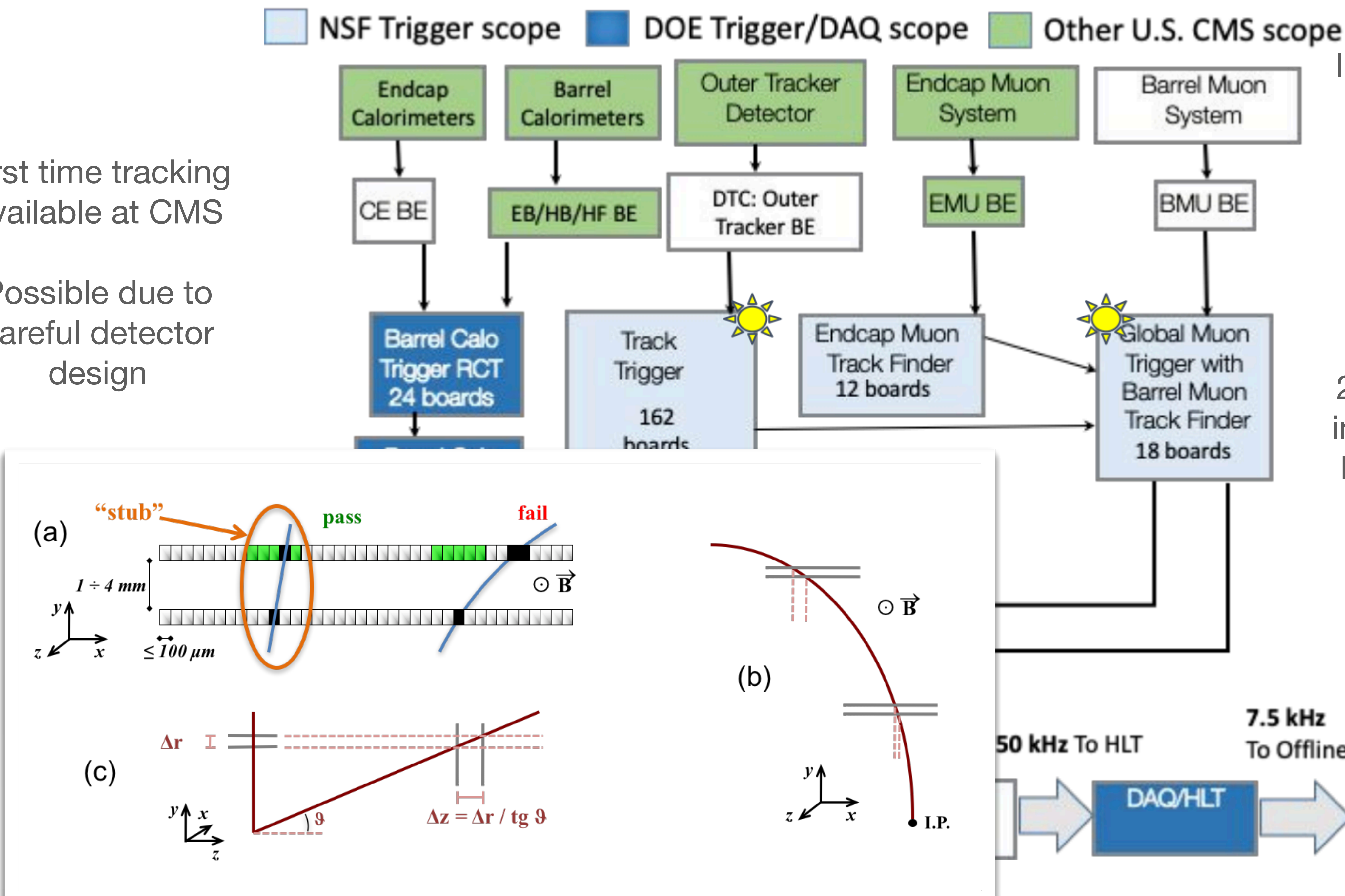
5-10MB/Event

Phase-2 Trigger for CMS

Based primarily on Xilinx Ultrascale+ FPGAs and 25 GB/s optical links interconnected between systems

First time tracking available at CMS

Possible due to careful detector design



Input at ~50 TB/s
40 MHz Rate

25 GB/S optical links interconnect between FPGA-based boards

5-10MB/Event

Physics Requirements for a Muon Collider

Consider the Physics we hope to access at a 10 TeV Muon Collider:

Assume a 10kHz crossing rate

- Higgs+X production at 10 TeV is expected to have a rate of <0.1 Hz
- WW production via vector boson fusion will be at 1 Hz level
  - **Signatures:** High Energy Electrons, Muons, Taus and Jets
    - Jets should be relatively high energy when originating from W or H
  - 'Easy' to differentiate these signatures from BIB
- **Important to define:** What could be lost from down selection of events?
  - Photons, MET, low p_T objects
  - Exotic signatures - disappearing tracks, LLPs

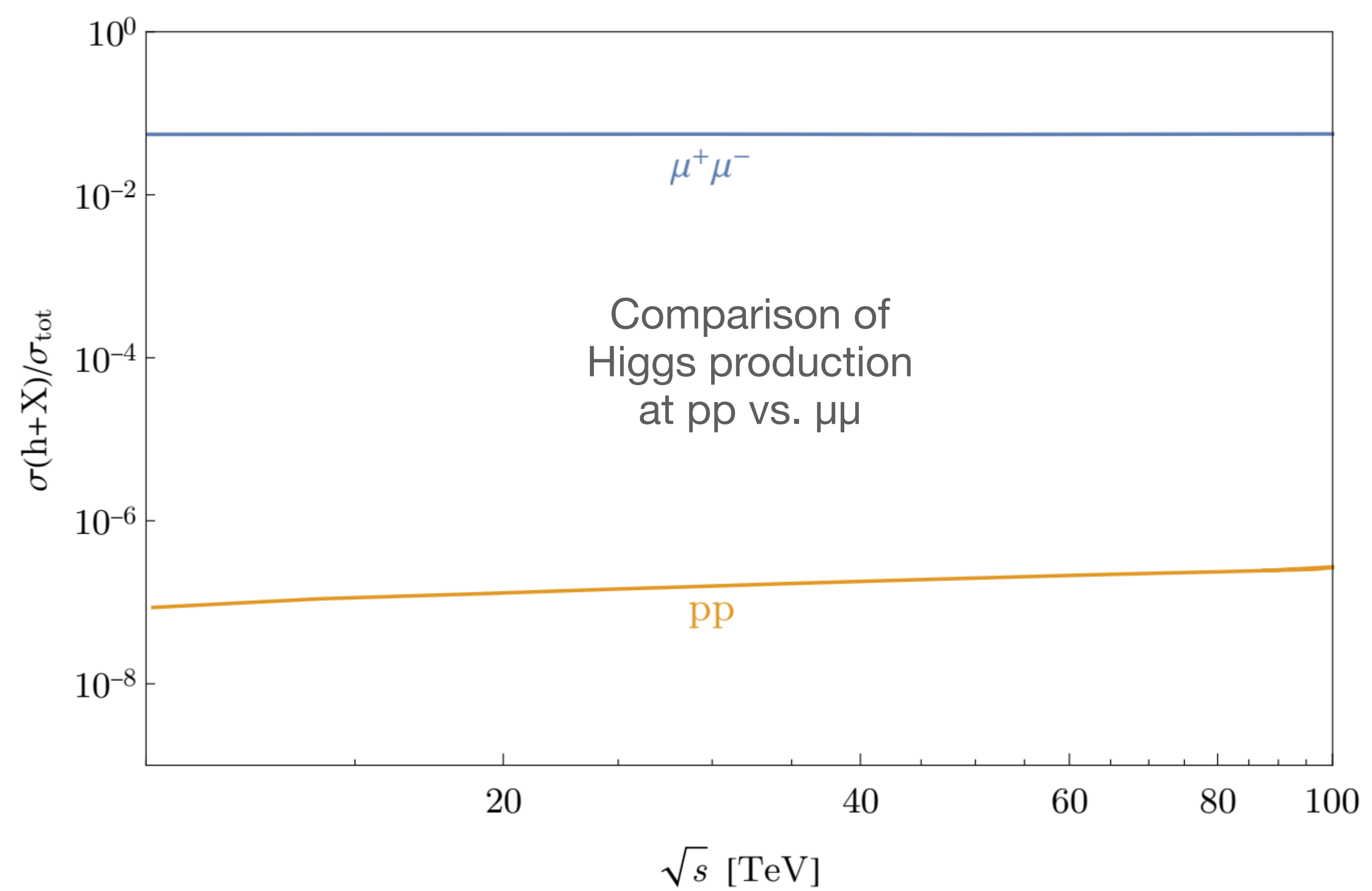


Figure 4: Higgs production cross section $\sigma(h + X)$ as a fraction of a representative "total" cross section σ_{tot} for $\mu^+\mu^-$ and pp colliders. For $\mu^+\mu^-$ colliders, we compute Higgs production using the LO cross section for $\mu^+\mu^- \to h + \nu\bar{\nu}$, while the "total" cross section σ_{tot} is taken to be the rate for single electroweak boson production, which is dominated by VBF production of W, Z, h, γ at these energies. For pp colliders we take the Higgs production cross section to be the N3LO cross section for $gg \to h$ [50] presented in [51], while the "total" cross section σ_{tot} is taken to be the $pp \to b\bar{b}$ cross section computed by MCFM [52].

Considerations for a Muon Collider Experiment

For a Muon Collider Experiment:

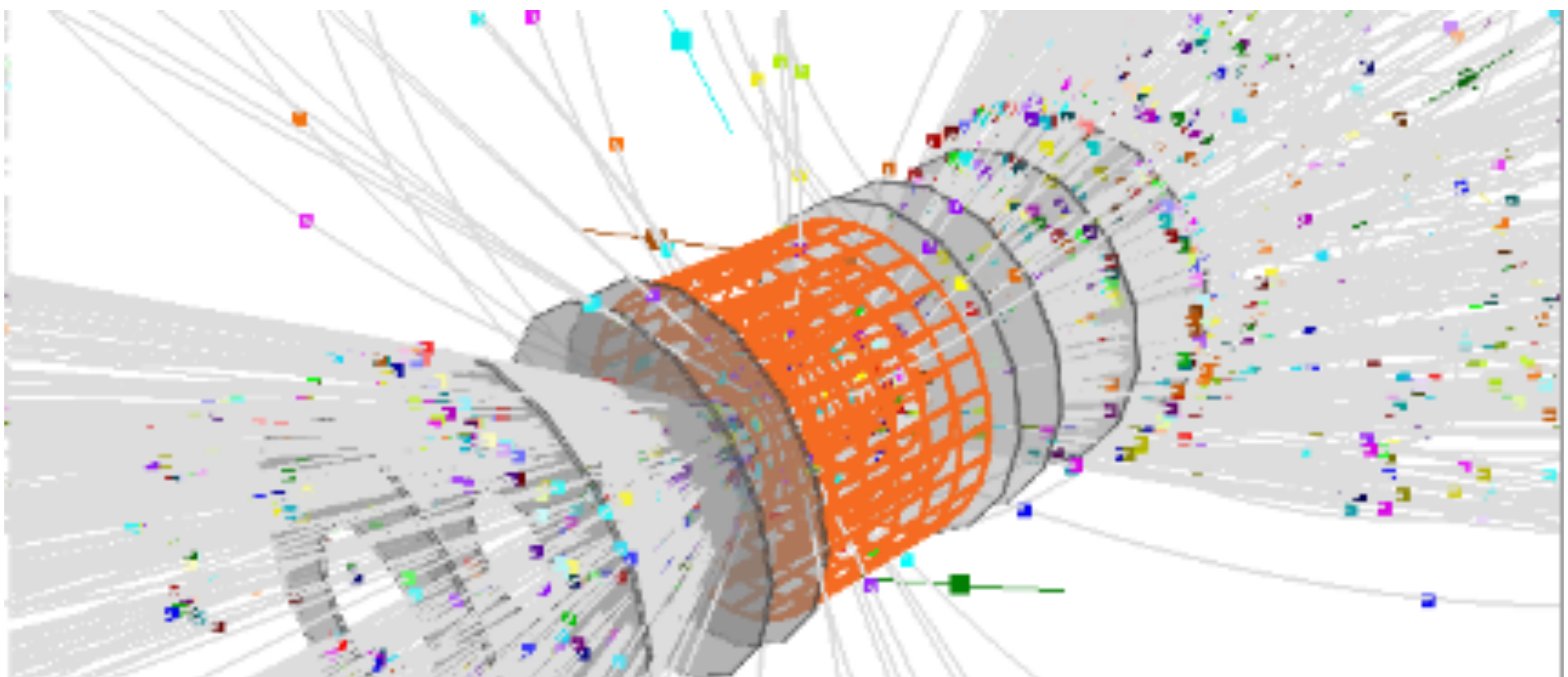
- Targeted Luminosity $10^{34} – 10^{35}$ cm^{-2} s^{-1}
- 100 kHz bunch crossing rate -> new event every 10 microseconds

Major Consideration: Large Beam Induced Background

- Real time reconstruction that exploits techniques for BIB reduction in real-time
- Could benefit from "smart" detectors that have position, time, and amplitude measurement

Data rates are dominated by the tracking system and calorimeters

- **Data size** = **Number of Hits x 32 bits per hit** - Hits counted in a 1 nanosecond readout window
- **Data rate** = **Data size x Bunch Crossing Rate**

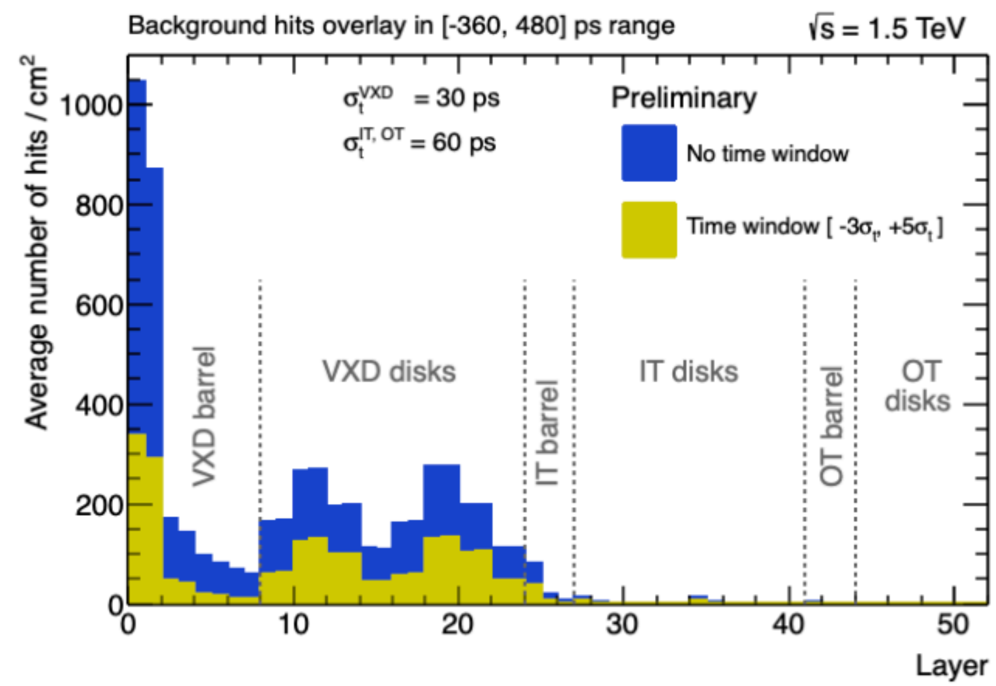


Requirements for a Muon Collider Experiment

Raw calculations of data rate from S. Jindariani

Assume module size of 20 cm²

- With 50x50 microns pixel size, ~800k pixels per module with 1 ns window
- 1% occupancy, up to 8k hits per module in the inner vertex tracker
- 32 bits to encode x/y/amp/time



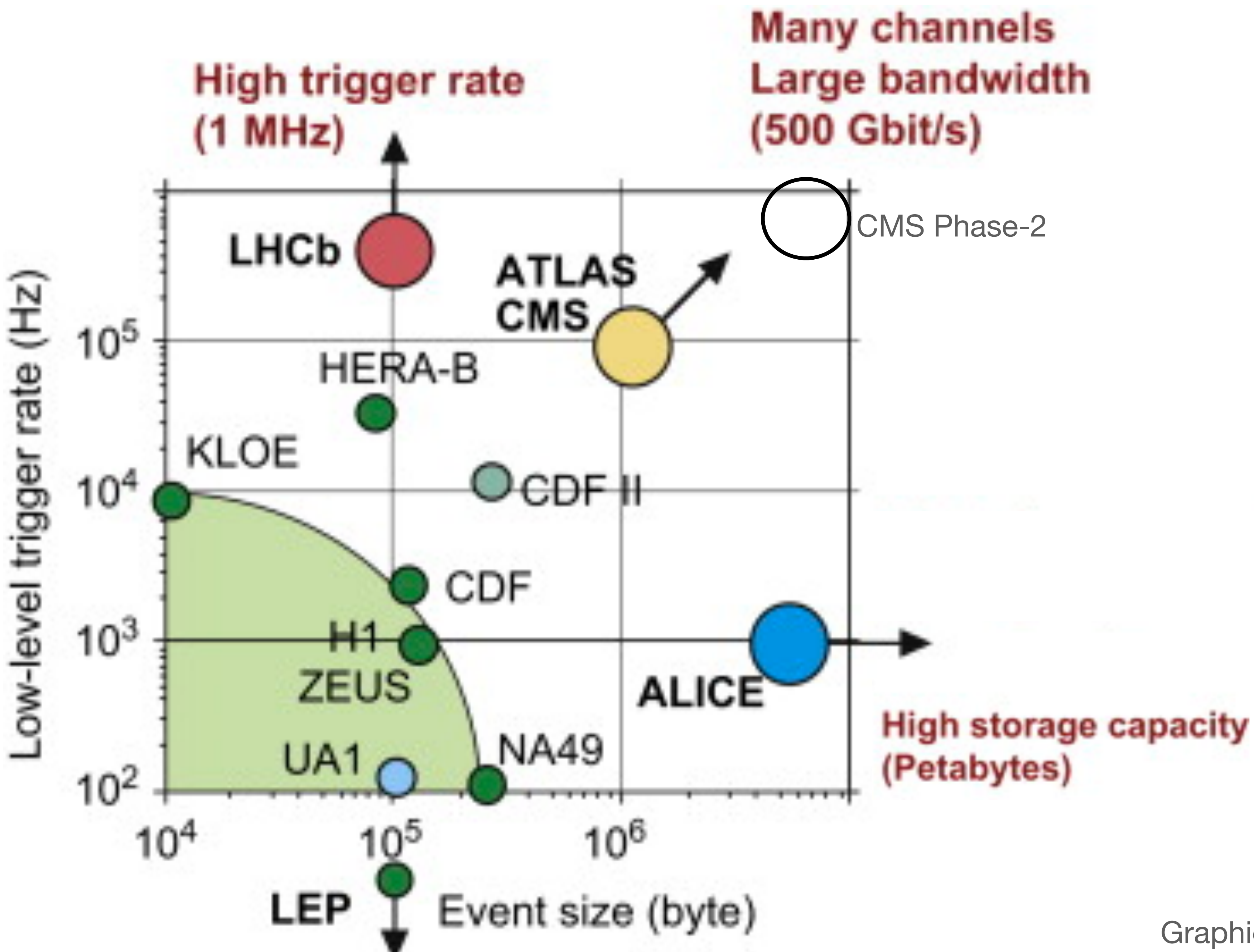
Data rates: 8k hits * 32 bit * 100 kHz * 2(safety factor) ~ 50 Gbps per module (20 cm2) ~10 Gbps per FE Chip (4 cm2)

- **Double the rate compared to HL-LHC FE chip. Requires R&D.**
    - But should be achievable in ~10-20 years

More online handles should be explored: **Data compression**, some **front-end clustering**, **p_T module based suppression** (preliminary estimates indicate x5 rate reduction), real-time calibration

Where does that put the Muon Collider?

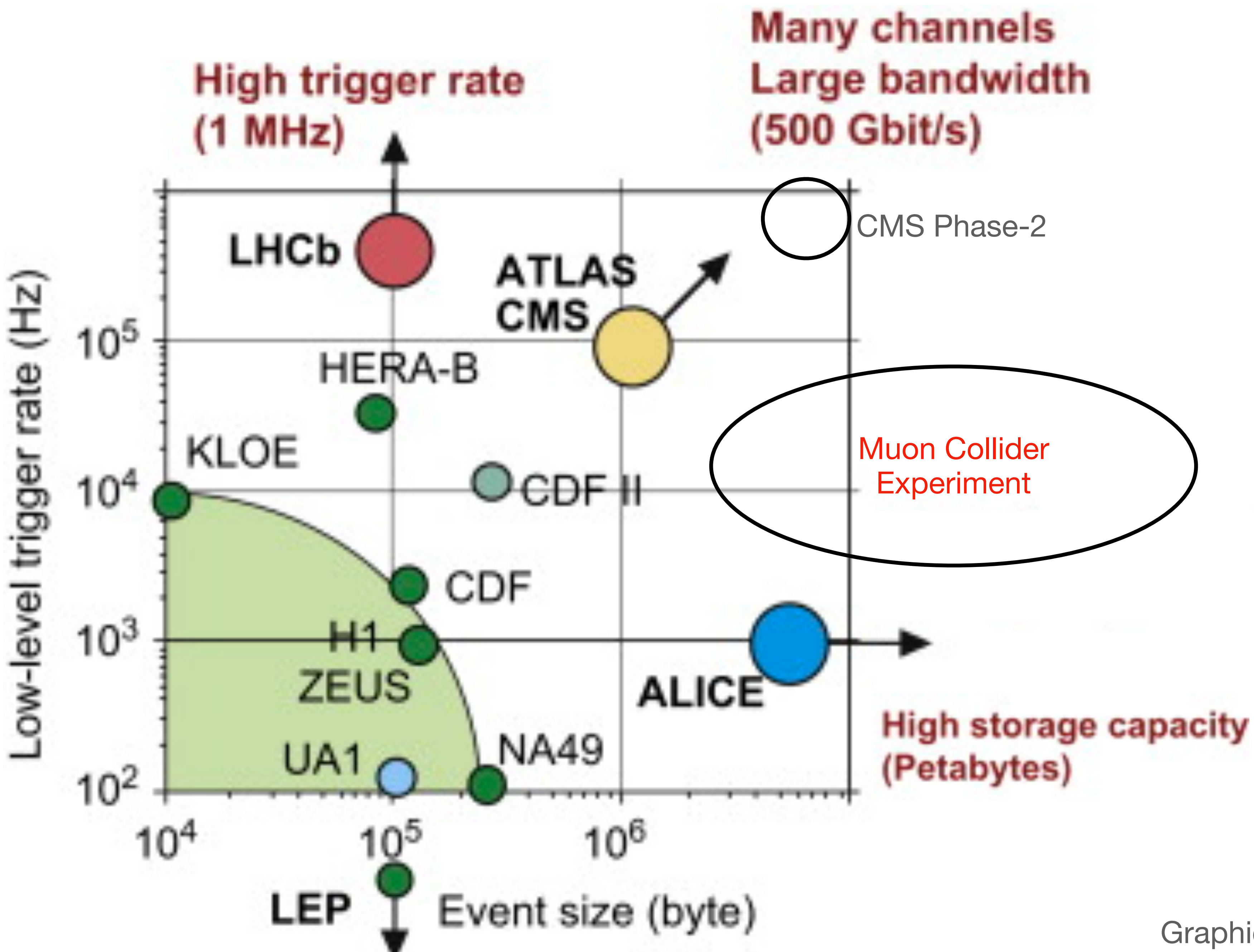
Comparing an experiment at a Muon Collider to other collider experiments



Graphic from
paper by M. Jeitler

Where does that put the Muon Collider?

Comparing an experiment at a Muon Collider to other collider experiments



Graphic from paper by M. Jeitler

Schemes for Muon Collider

Different schemes explored for event size reduction

- lpGBT at the LHC provide bandwidth at up to 10GB/s for detector readout
  - Can this increase to 20GB/s in 20 years? - increasing the bandwidth is a common goal across future colliders
- Estimate 10k links with 20 GB/s bandwidth
  - Less if aggressive schemes for on detector reduction utilized (pointing towards interaction point)

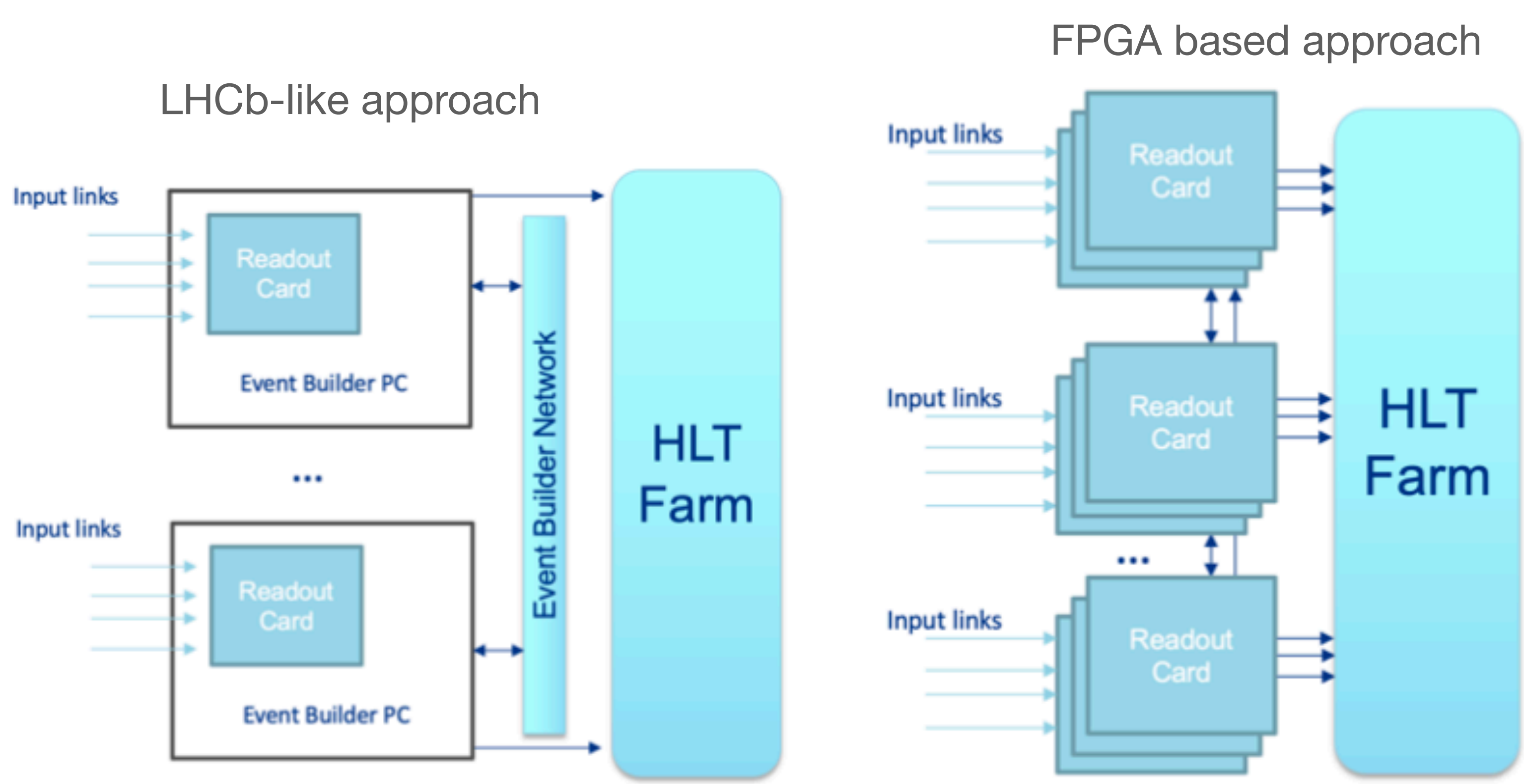


Fig. 15: Schematic illustration of possible ways to structure the TDAQ system. The one on the left shows an LHCb-like approach with a software Event Builder. The one on the right uses hardware boards to structure event data and pass them to the HLT farm.

Promising Technologies and R&D Directions MuC

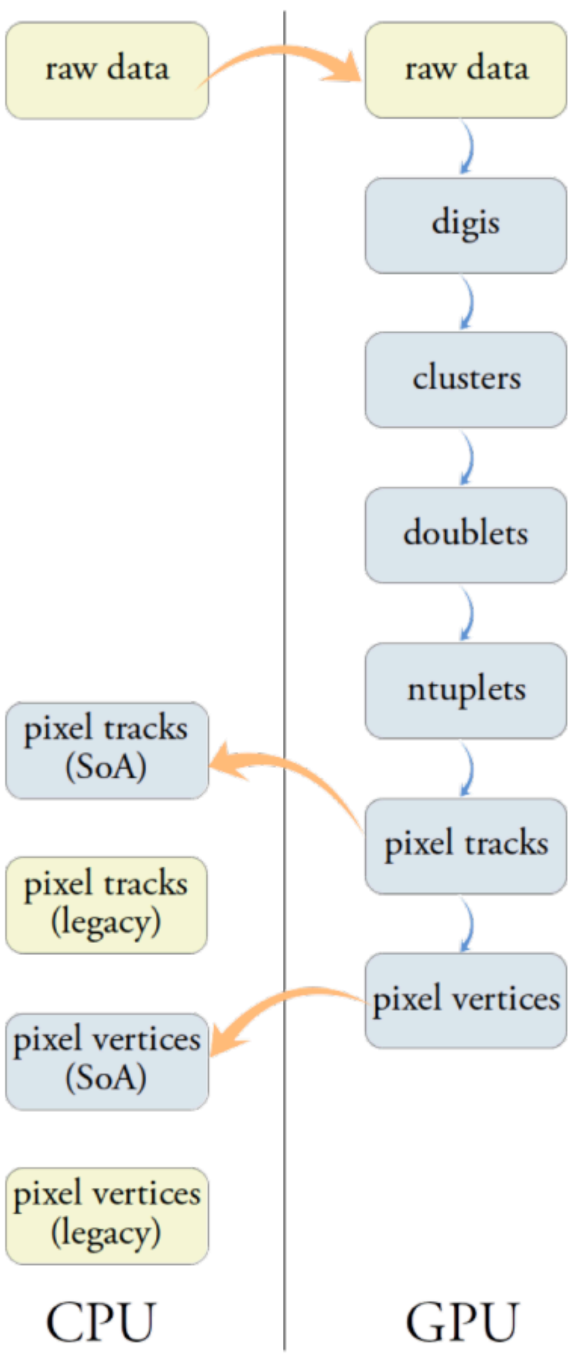
Reconstruction Co-Processors

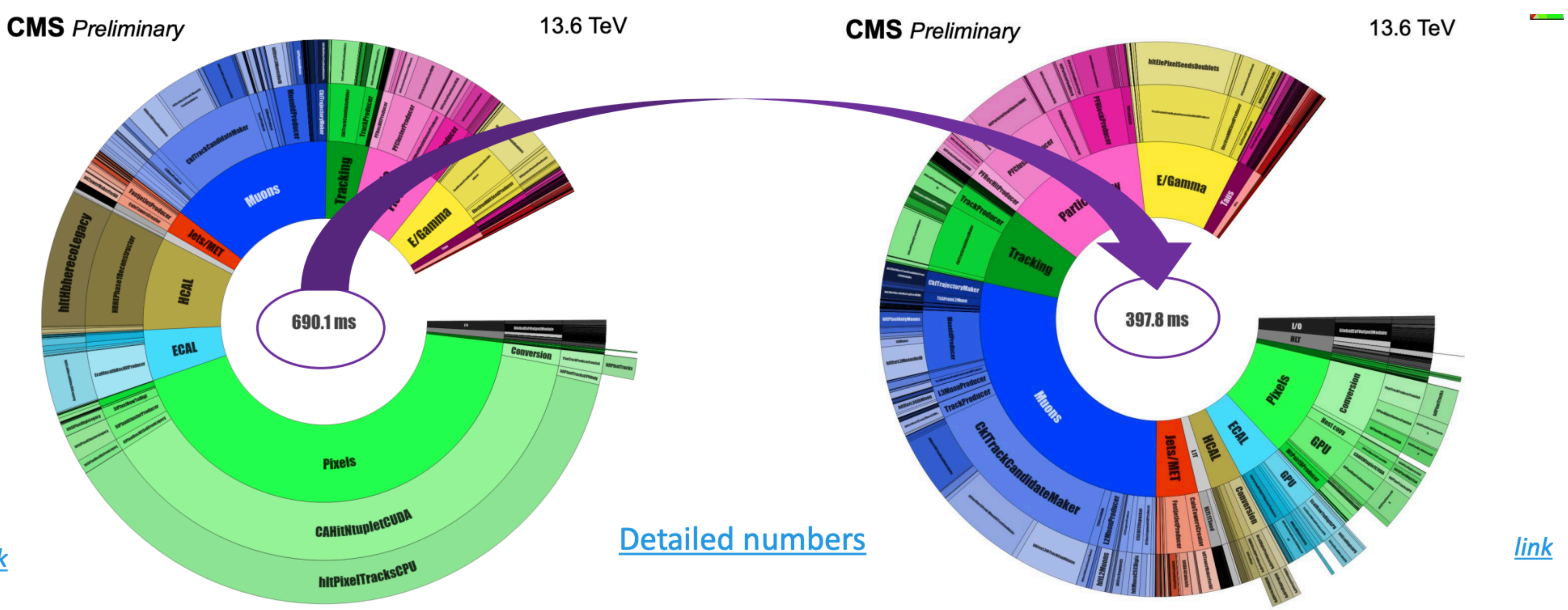
LHCb, ALICE and CMS are using GPU co-processors at HLT for LHC Run-3 (2021-2025)

- CMS shows an average HLT processing time per event 690 ms (CPU) 397 ms (CPU+GPU)

FPGA coprocessors are being explored by a few groups for specific applications

- Great for ML algorithms, bandwidth constraints





link

Detailed numbers

link

Average time per event for CPU Only Configuration

Average time per event for CPU + GPU Configuration

Snowmass Summary of TDAQ for Future Colliders

Application of Machine Learning to DAQ systems, particularly considering **co-design of hardware** and **software** to apply ML algorithms to real-time hardware will make future experiments operationally efficient

- **Seek out funding opportunities from a wide variety of sources**

Design TDAQ system architectures to enable more intelligent aggregation, reduction, and streaming from detectors to HLT and offline data processing

- Continue to leverage advances in industry (FPGAs, GB/s optical links, GPUs, core-processors)
- Continuoius

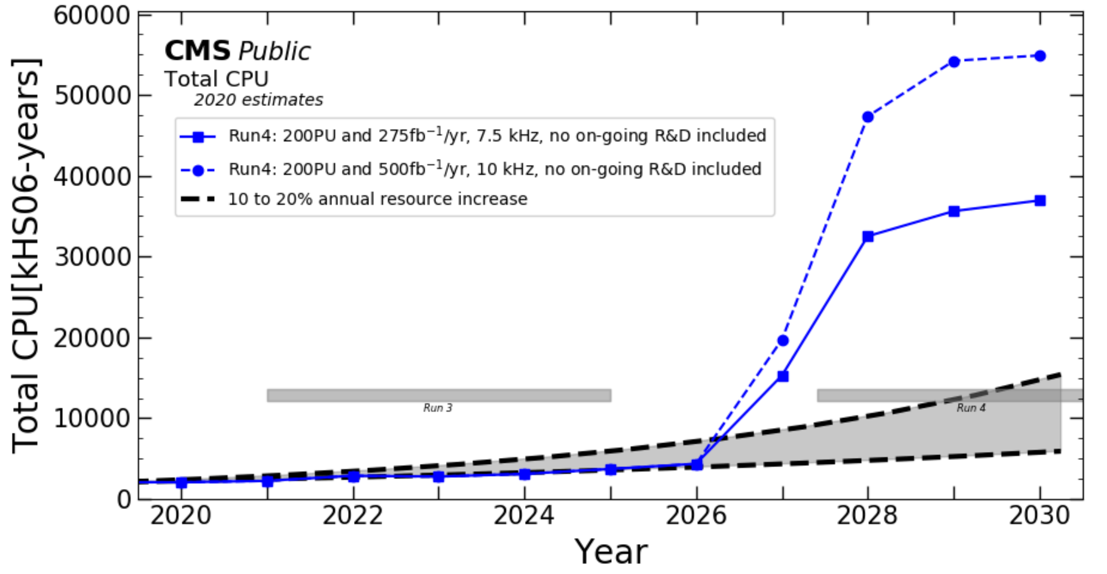
Crucial to **develop improved readout technologies to increase data bandwidth** and operate in extreme environments while fitting into the power and material budgets

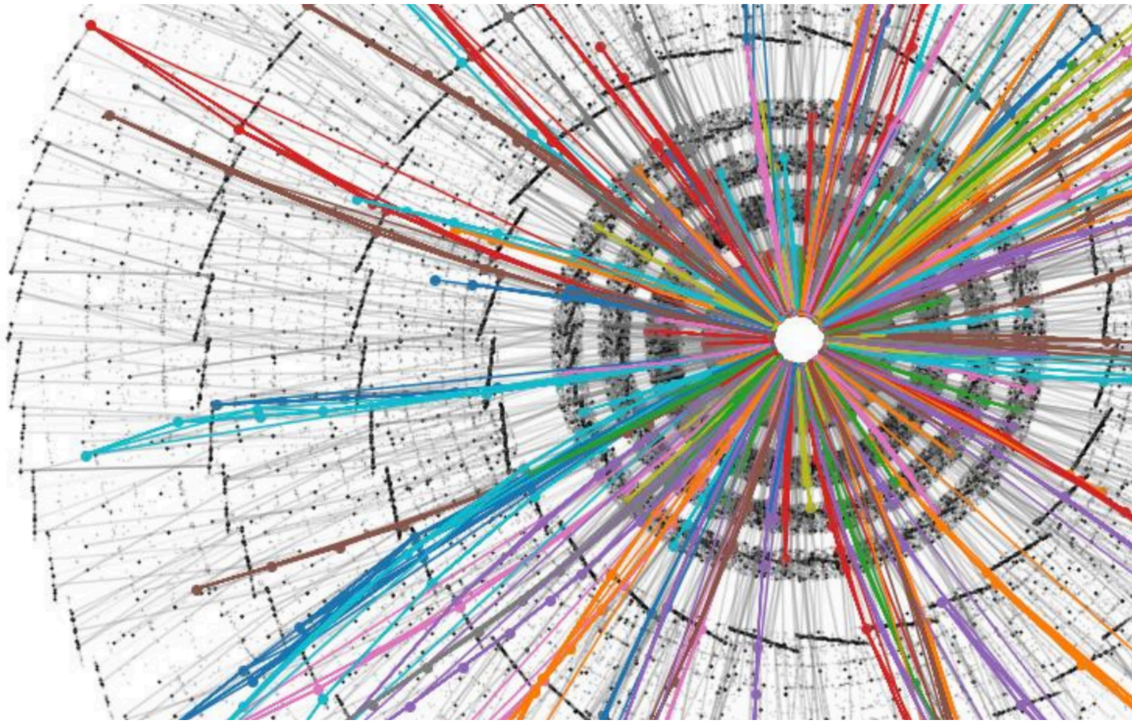
- Radiation tolerance is not a typical requirement for high throughput readout in industry

ML Techniques for Reconstruction

Useful to consider updating reconstruction techniques

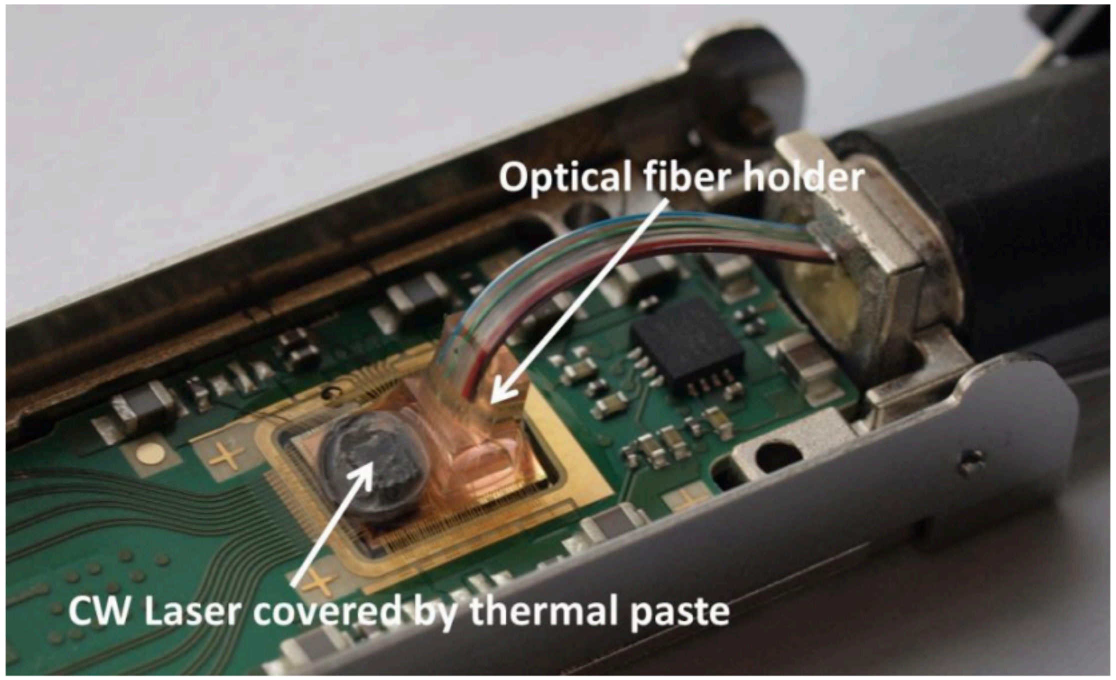
- Traditional tracking (Kalman Filter) algorithms scale worse than quadratically with the number of hits
- **Graph-based approaches are well suited** since tracking data can be naturally encoded as a graph structure
- **Graph Neural Networks (GNNs) consider local information between pairs of hits to learn relationships** between them in order to "connect the dots" and infer tracks
    - ML techniques typically **scale linearly** with the number of hits as they do not contain many interactive loops
- Note, **none of this is 'magic'**, even proof-of-principle, implementation is a lot of work
    - However, many public and private foundations are interested in fast ML applications, these should be leveraged for R&D





R&D: Readout Technologies

VCSEL (Vertical cavity surface emitting laser)**-based links (50GB/s)** and **Si-photonic links in development**



Wireless data transmission also an interesting option

- broadcast for control and configuration of detectors without bulky copper cables

FPGAs here to stay for the foreseeable future due to ease of design, synchronous data throughput and high customizability

- **Connectivity between FPGAs must go beyond PCIe gen5 bandwidth**
  - Further explore CPU-FPGA hybrids
    - Will links provide the bandwidth needed for real-time detector read out?

Significant design time needed for FPGA based firmware

- Continue use and development using High Level Synthesis based firmware design and GPU co-processing

Conclusions

It is **far too early to make any final decisions about trigger and DAQ** at a future experiment on a muon collider

- BUT as detectors and full-reco simulations are being developed it is important to think about **real time data reduction schemes**
  - **The CMS track-trigger is an excellent example of this**
  - **Asking ~ 1-2 FTEs/year to embed trigger/DAQ experts with detector design and full simulation groups (this can be split up to 0.25 FTE/system)**
- Also important to take note of reconstruction bottle necks and consider mitigation strategies for fast reconstruction

Be prepared to assimilate the latest computing technologies as they become available

- FPGAs and GPUs changed the way we trigger at the LHC

Identify trigger/DAQ R&D needs that are unique to collider physics

- Radiation tolerant high speed data transfer methods, what else?