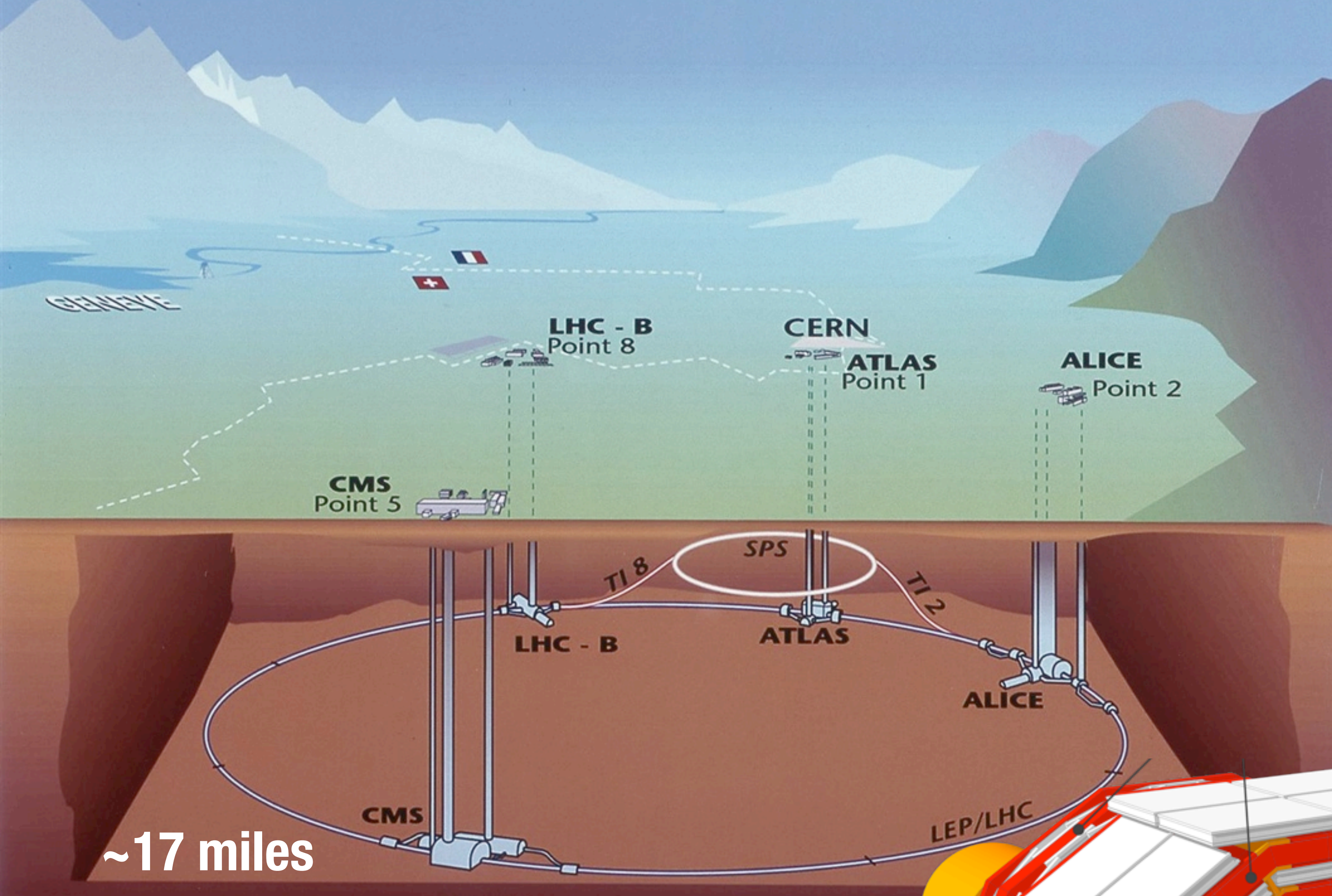
Track-triggering at CMS for the High-Luminosity LHC

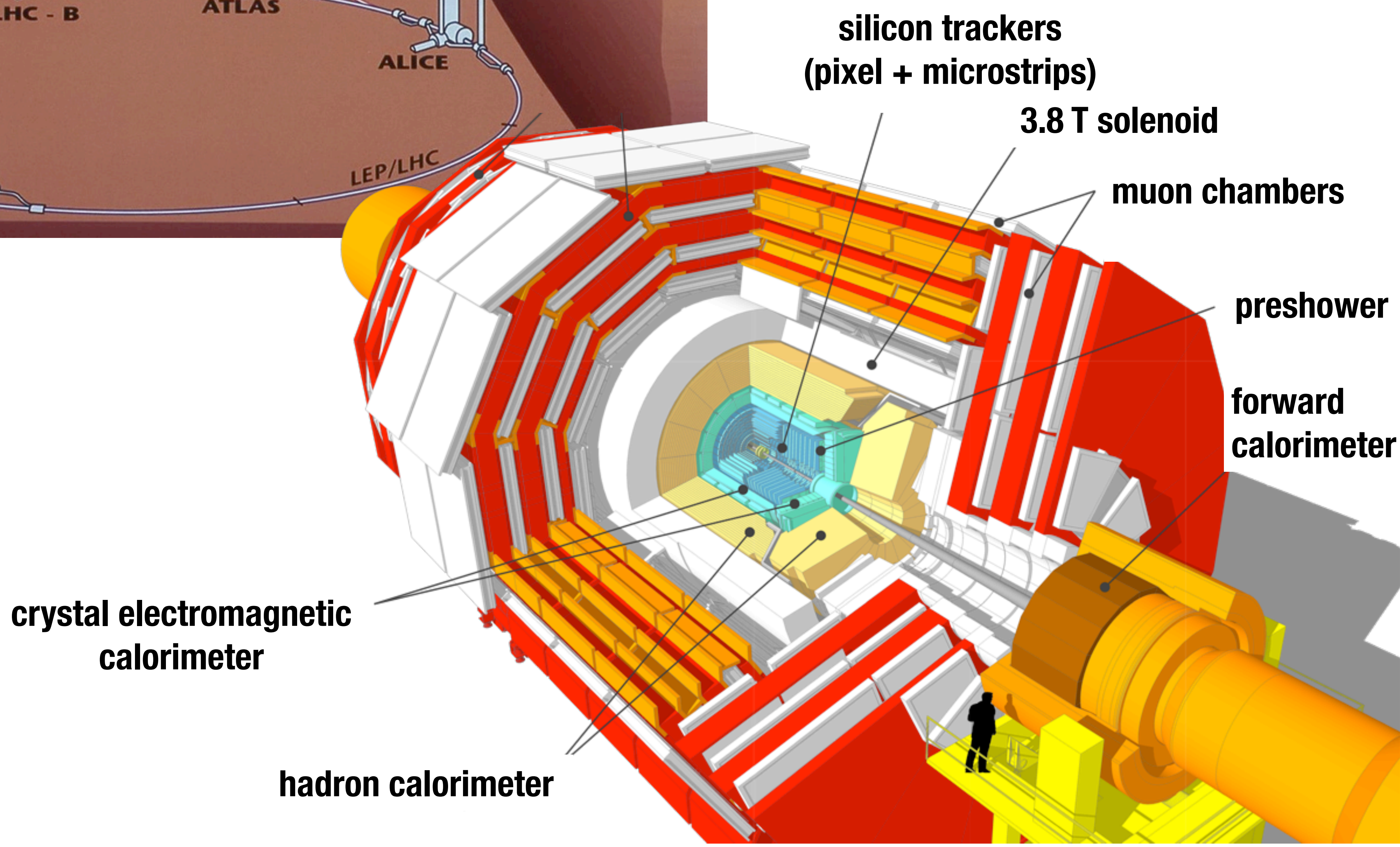
Louise Skinnari (Cornell University)

UCL HEP Seminar, April 7, 2017



LHC

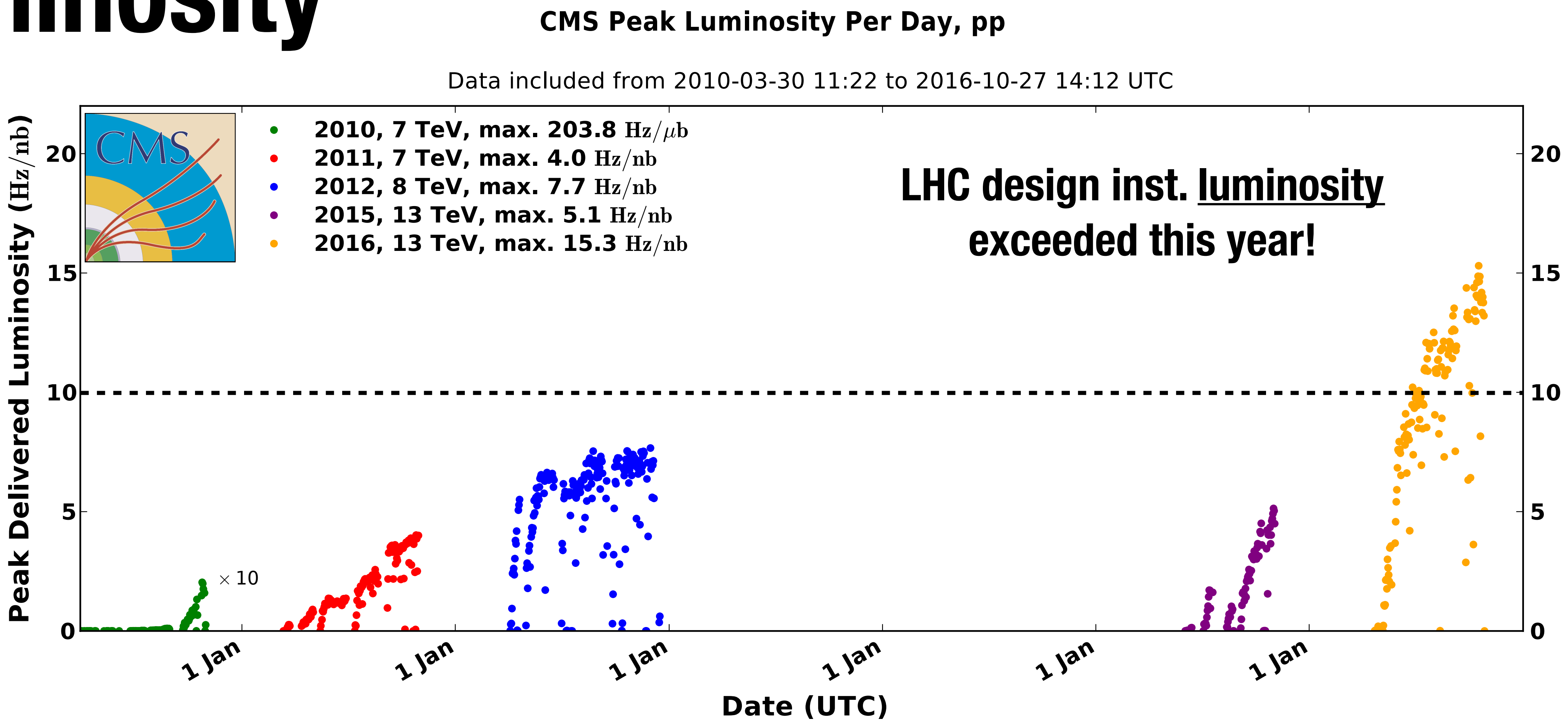
2808 bunches of ~10^{11} protons
crossing frequency: 40 MHz
operating @ 13 TeV



CMS

Compact
Muon
Solenoid

Luminosity

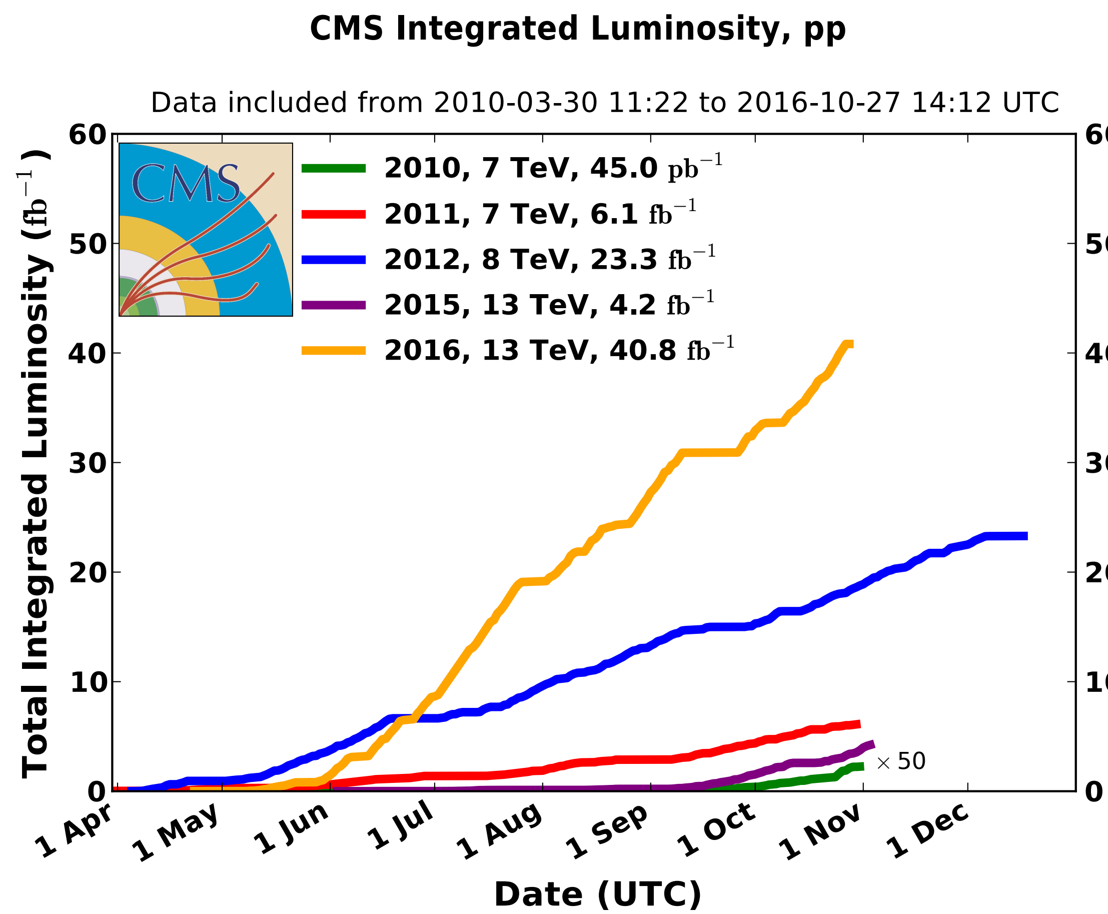


instantaneous luminosity

$$N = \sigma \times \underbrace{\int L \, dt}_{\text{integrated luminosity}}$$

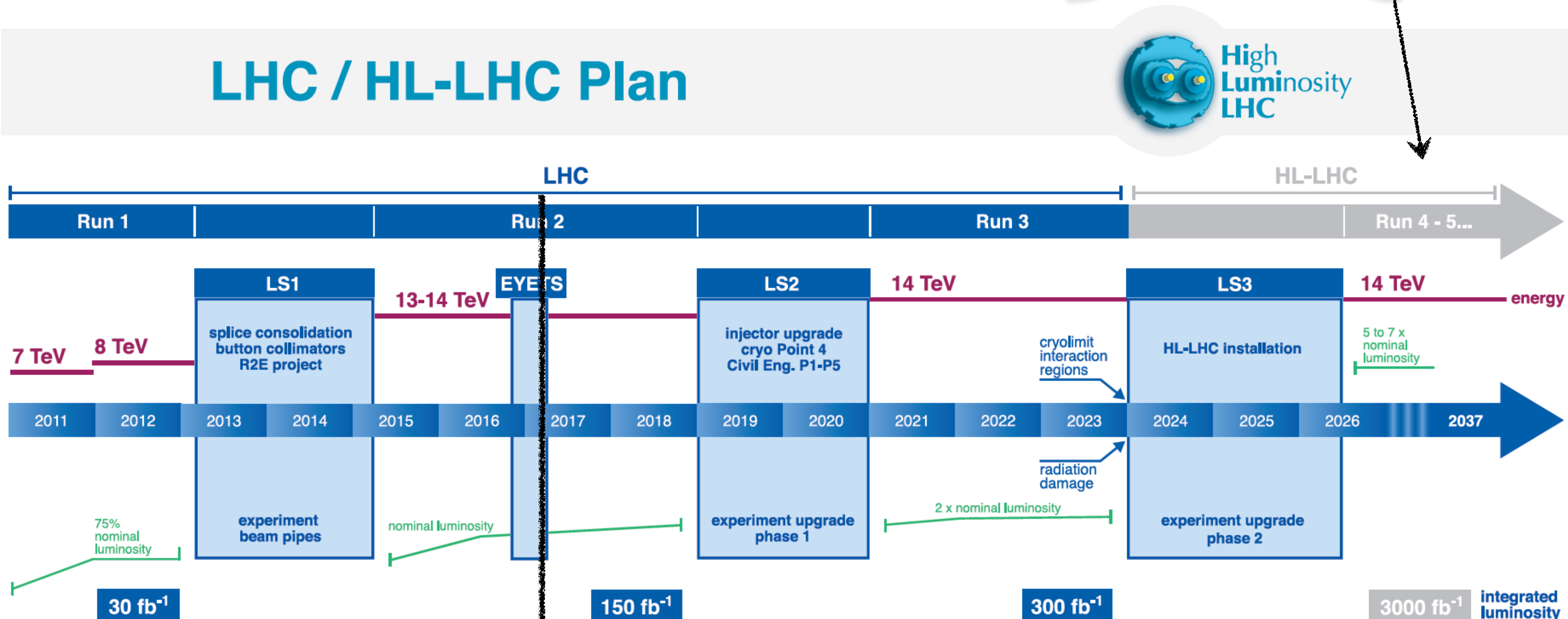
number of events

cross section



LHC → High Luminosity LHC

Peak luminosity:
~7.5x10^{34} cm^{-2}s^{-1}



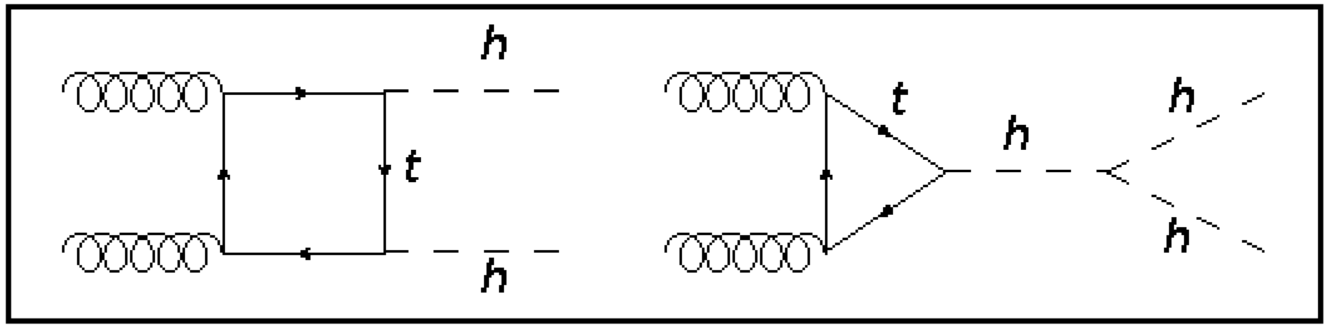
TODAY

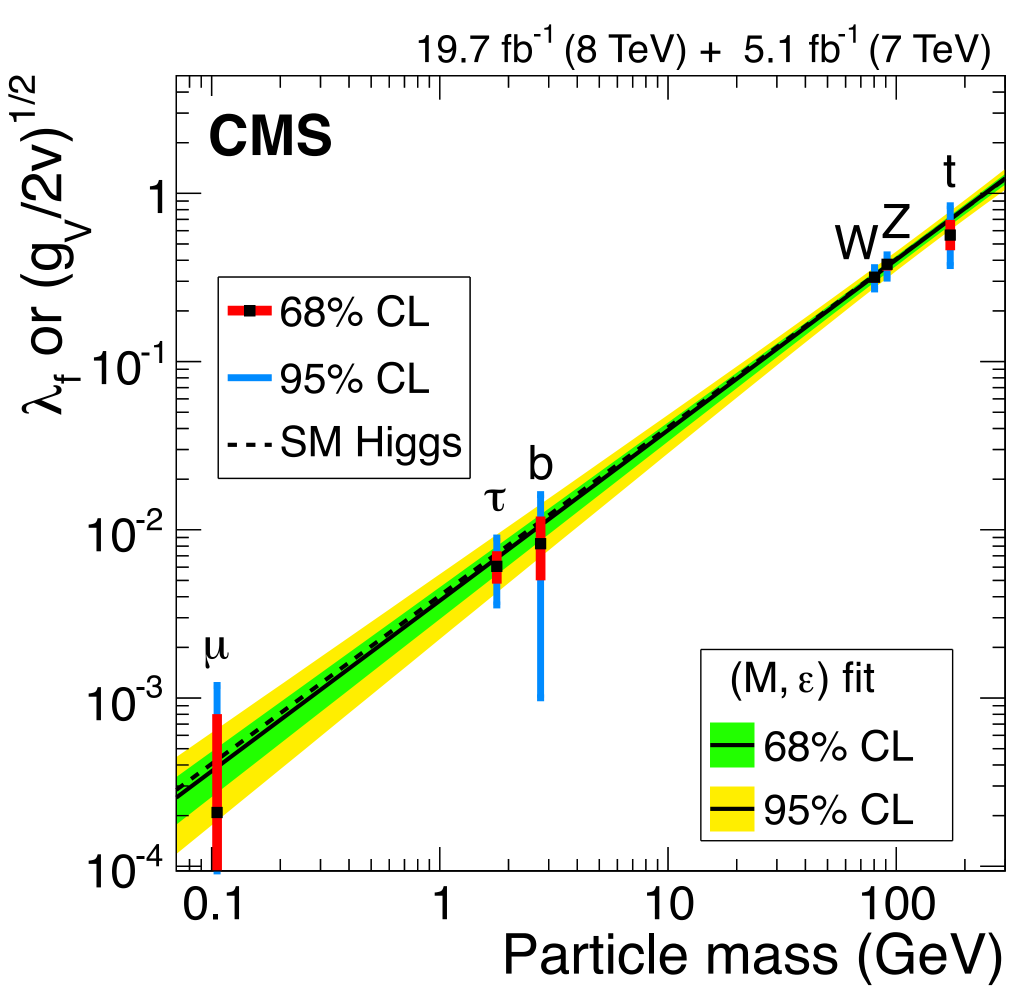
~40 fb^{-1} @ 13 TeV

Integrated luminosity:
~3000 fb^{-1}

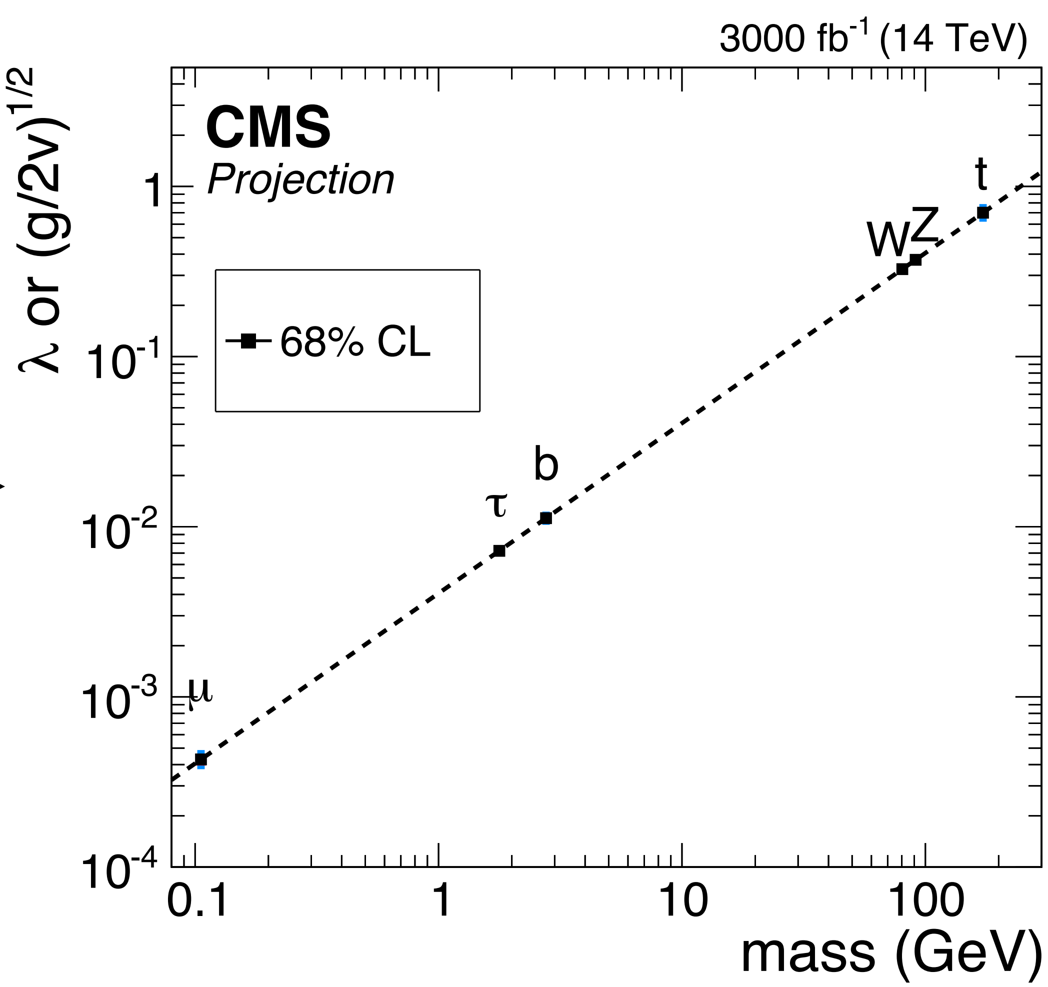
Motivation

- **Higgs boson**
  - Precision measurements of properties & couplings
  - Rare decays
  - Di-Higgs searches to measure Higgs self-coupling



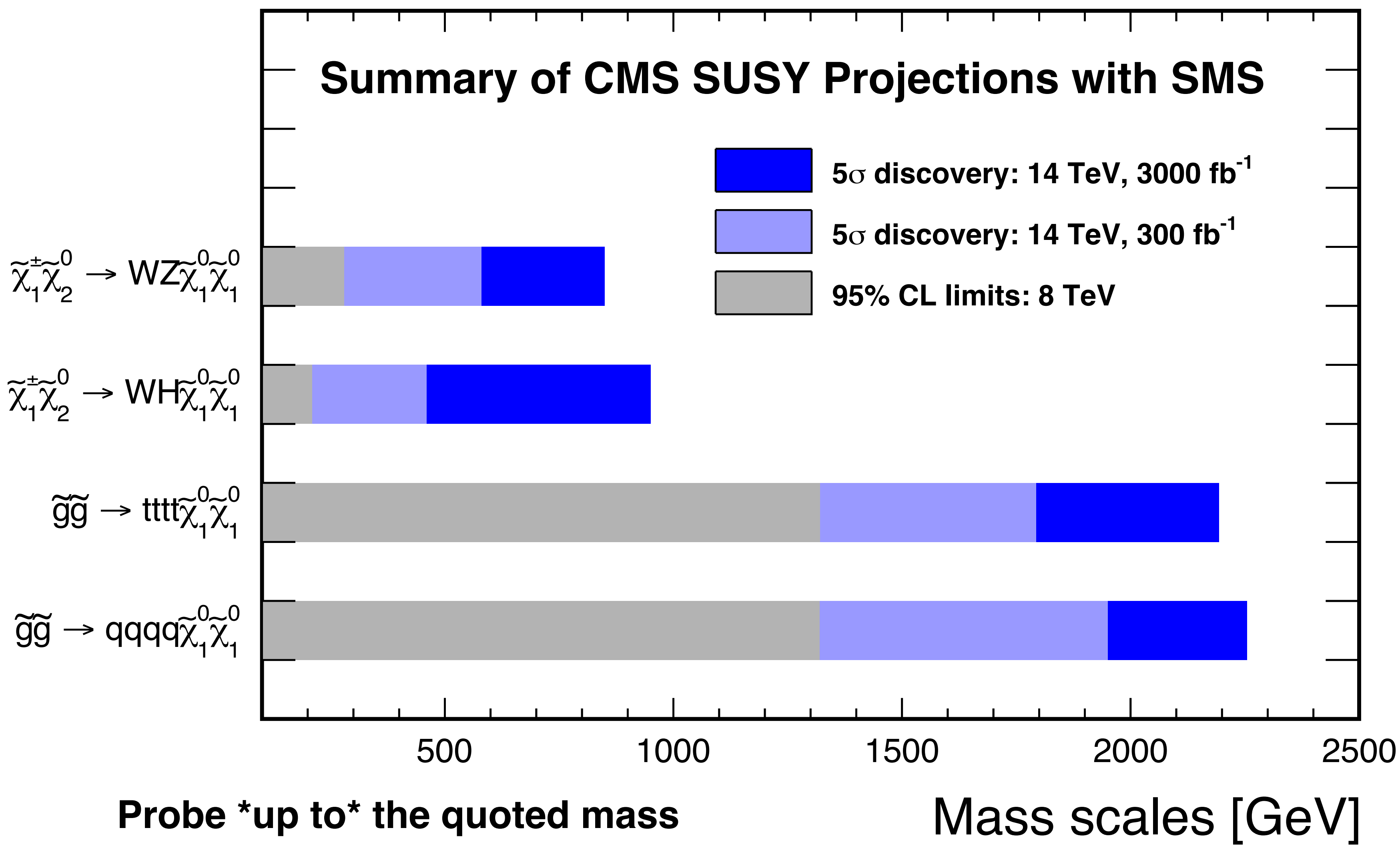


%-level precision for Higgs coupling measurements



Motivation

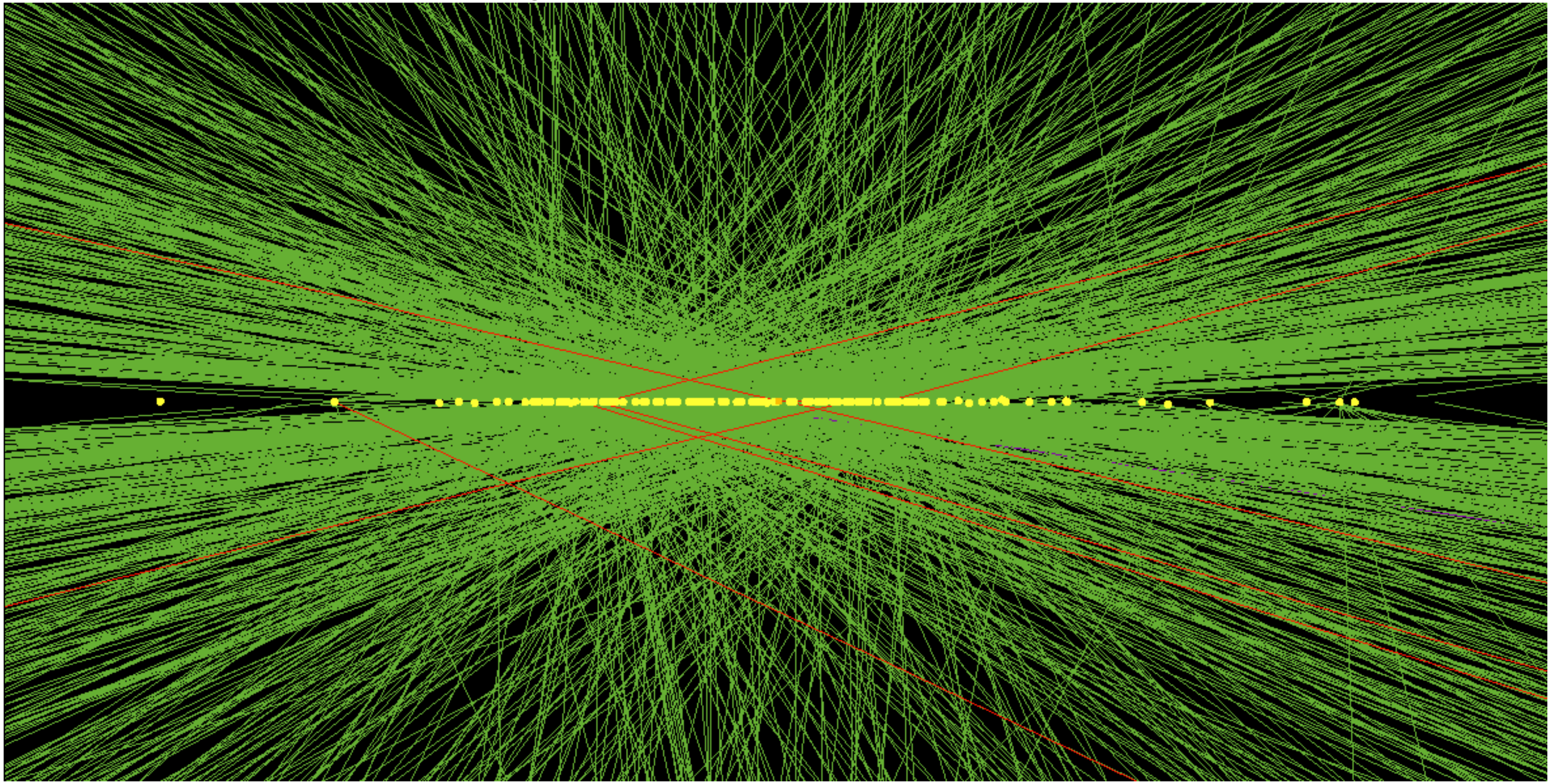
- Detailed studies of possible discovered new particles at the LHC
- Extend discovery reach in searches for SUSY & other BSM scenarios
- Search for rare SM processes, possibly enhanced by BSM physics



Probe *up to* the quoted mass

The price for high luminosity

Simulated event display with average pileup of 140

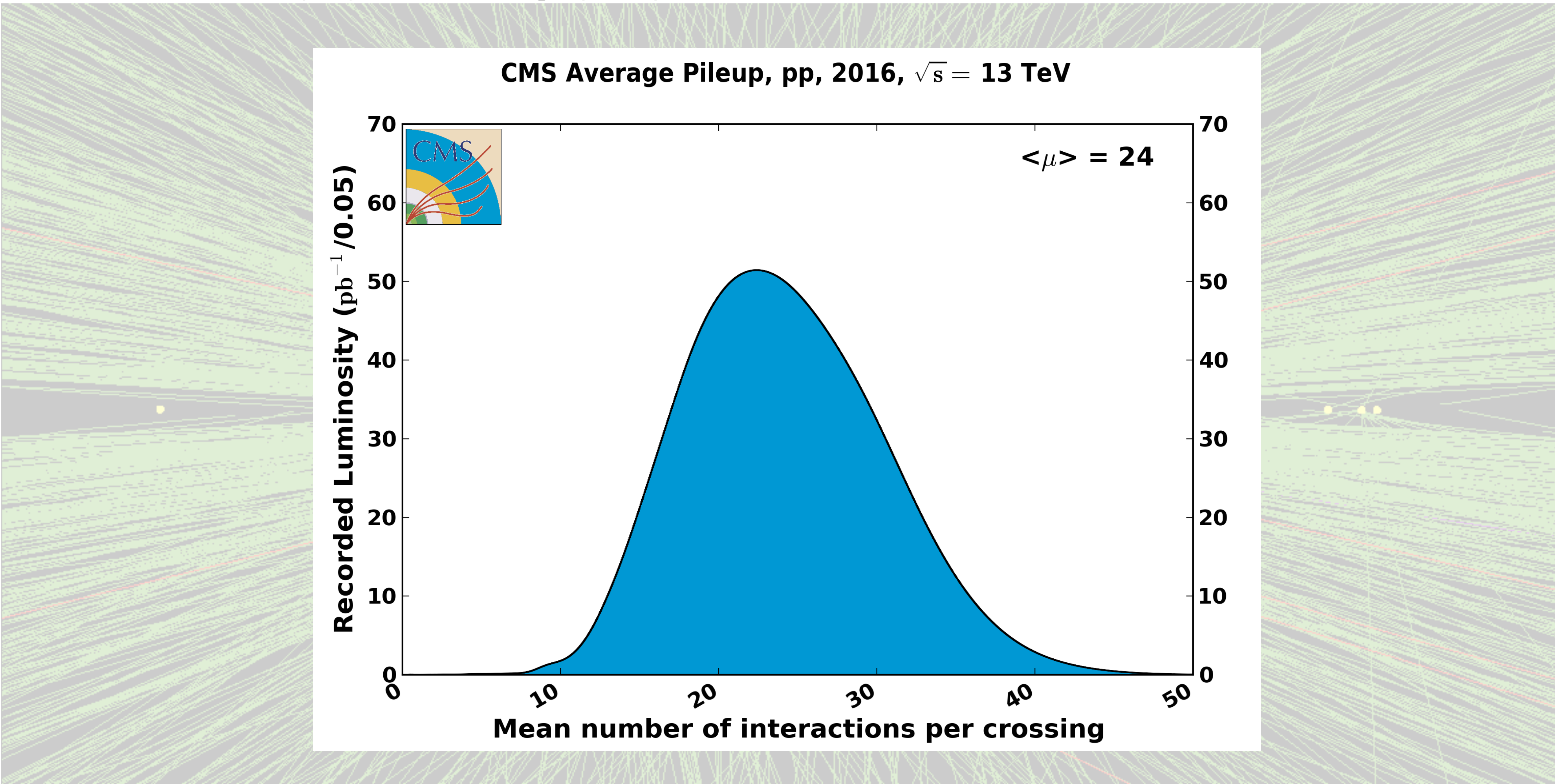


PILEUP: number of overlapping interactions (expected average ~200)

Particularly challenging for trigger system!

The price for high luminosity

Simulated event display with average pileup of 140



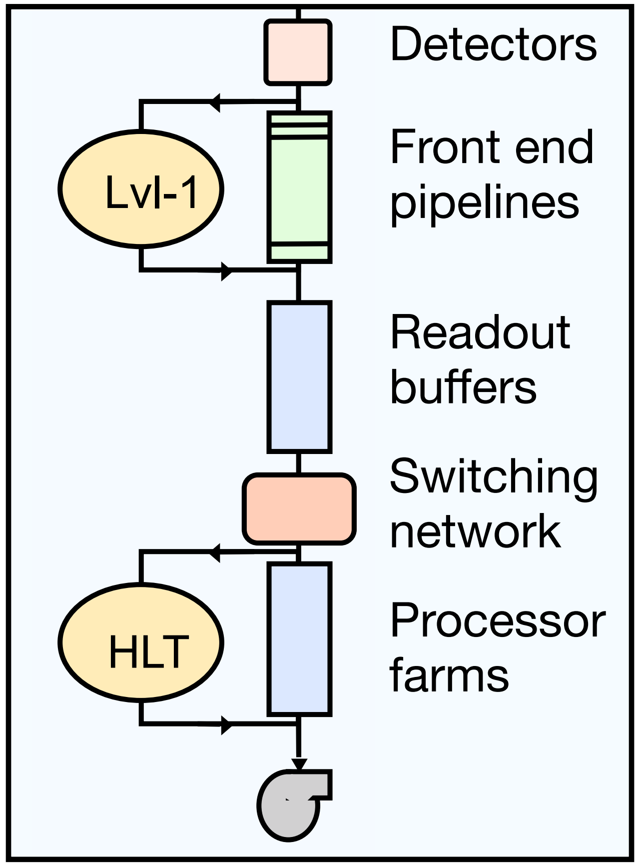
PILEUP: number of overlapping interactions (expected average ~200)

Particularly challenging for trigger system!

CMS trigger system

Which collision events to read out & store for offline analysis?

- **L1 trigger**
  - Hardware-based, implemented in custom-built electronics
  - Muon & calorimeter information with reduced granularity
- **High-Level Trigger (HLT)**
  - Software-based, executed on large computing farms
  - Tracking & full detector granularity



40 MHz

~4μs "latency"

L1 output: 100 kHz

HLT output: 100 Hz

Why tracking @ L1?

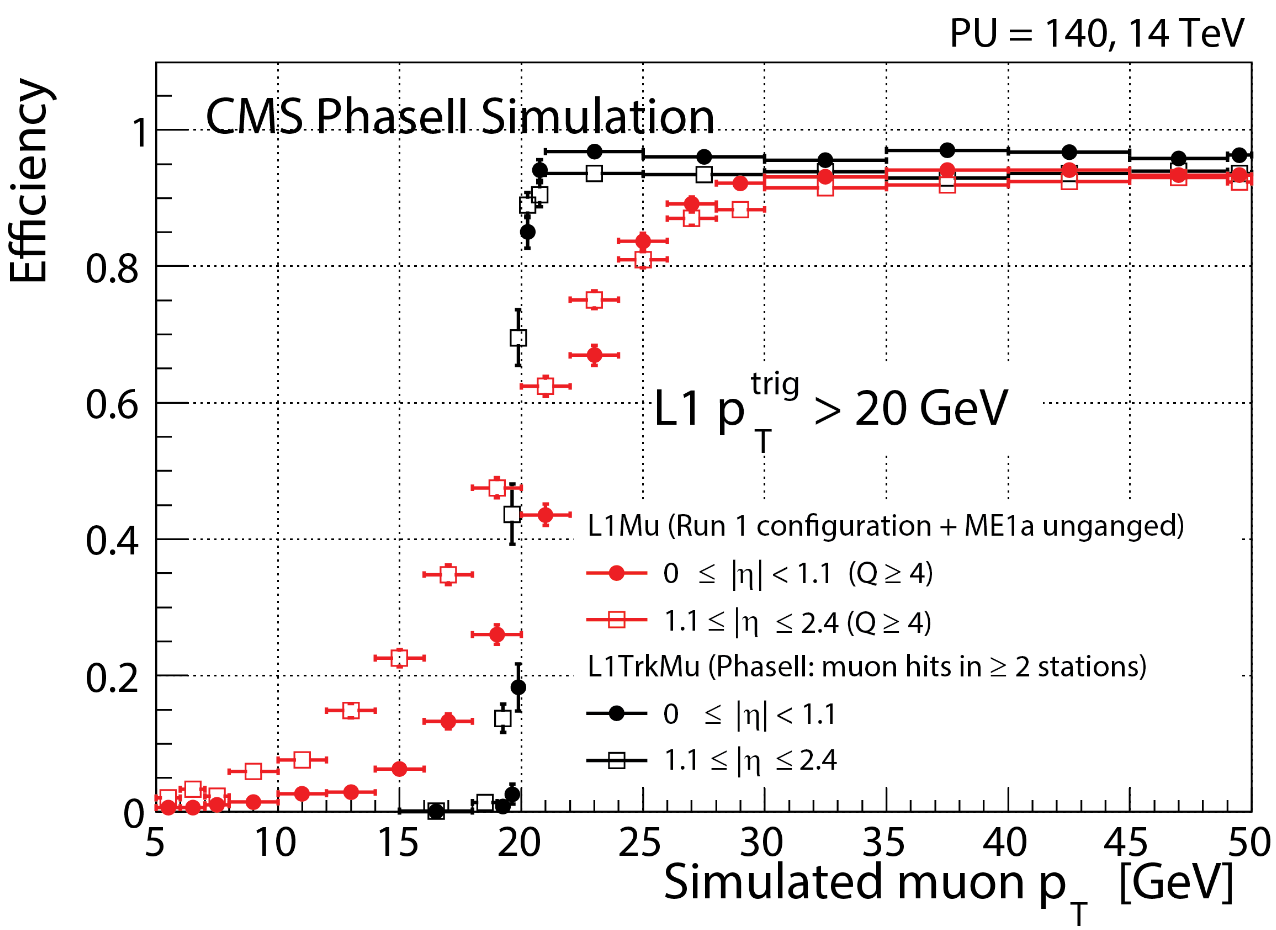
- With HL-LHC, event rates would exceed what can be read out at L1
- *Physics goals* rely on excellent detector performance & trigger capabilities
  - Must allow triggering on objects at electroweak scale!
- Typical handle to control event rates at trigger level -- momentum thresholds

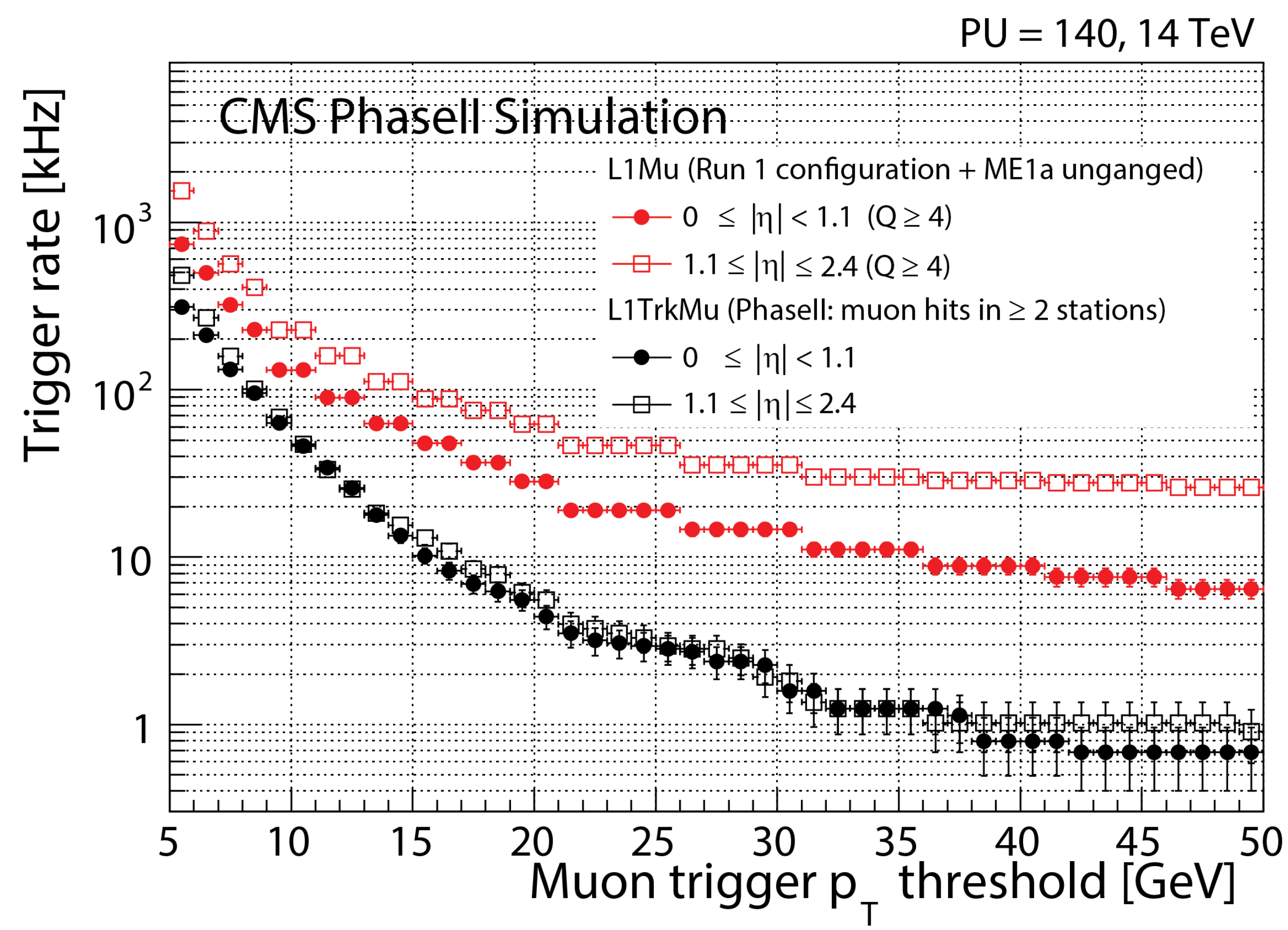
Increasing thresholds limits physics potential + alone insufficient!

⇒ Tracking @ L1

Using tracking @ L1

Example 1: Muons -- combine track with L1 muon object

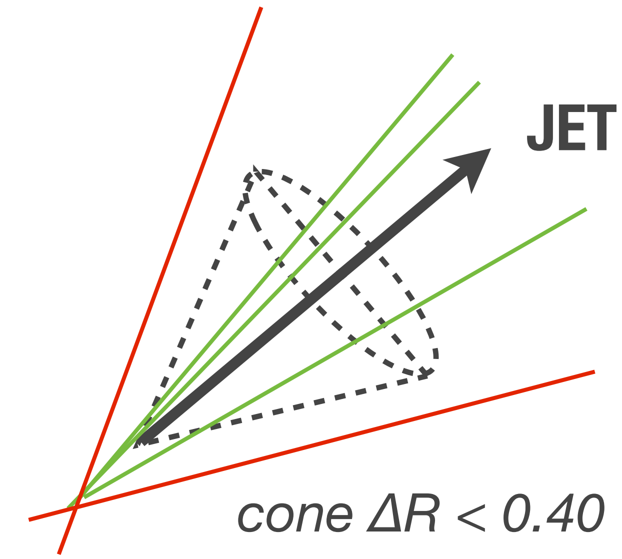


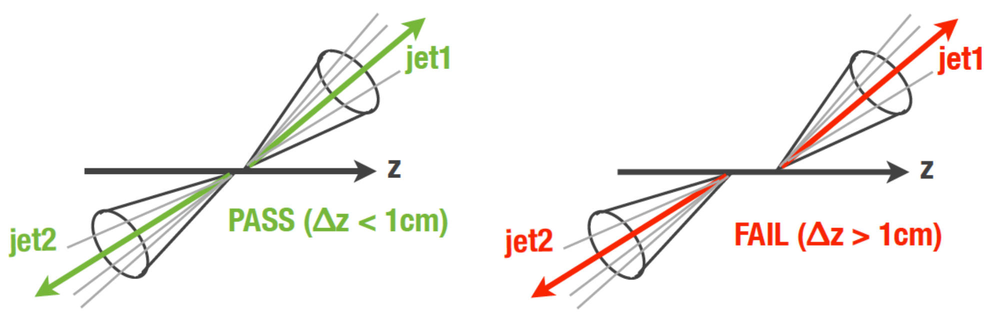


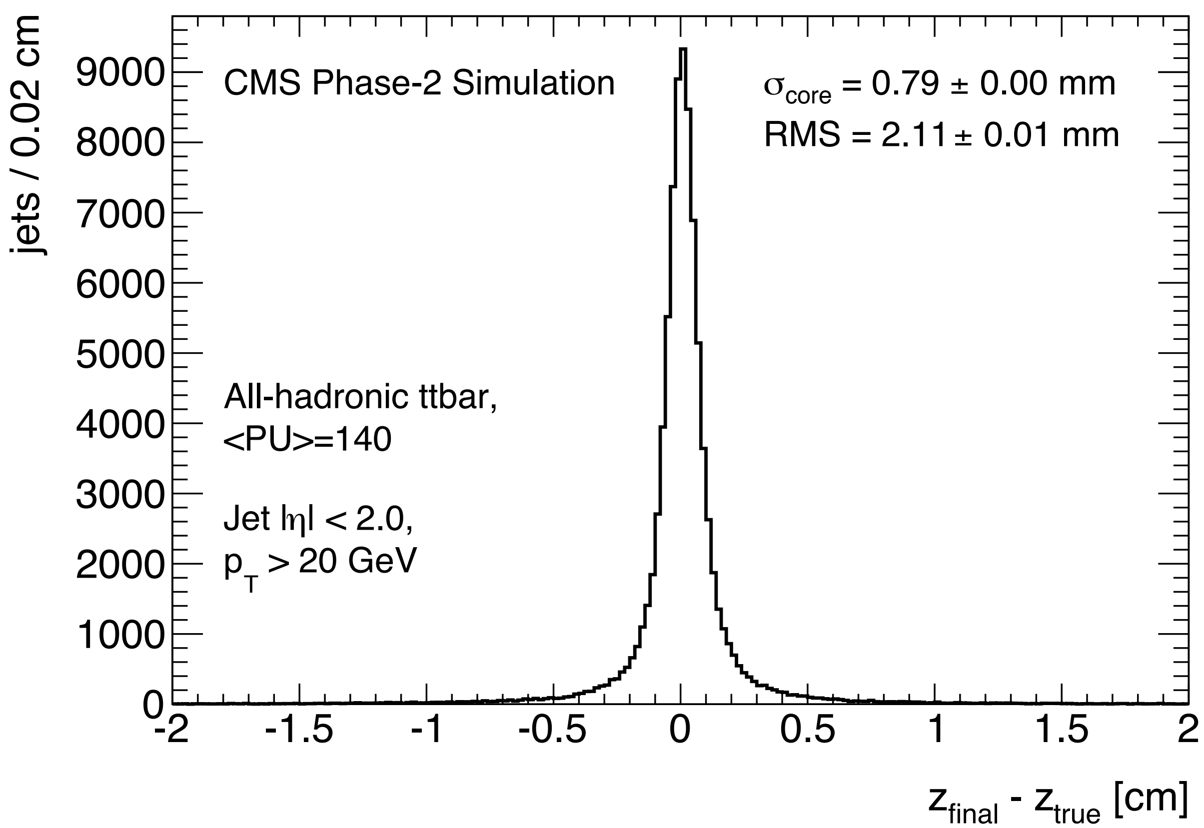
Sharpened p_T threshold → significant rate reductions

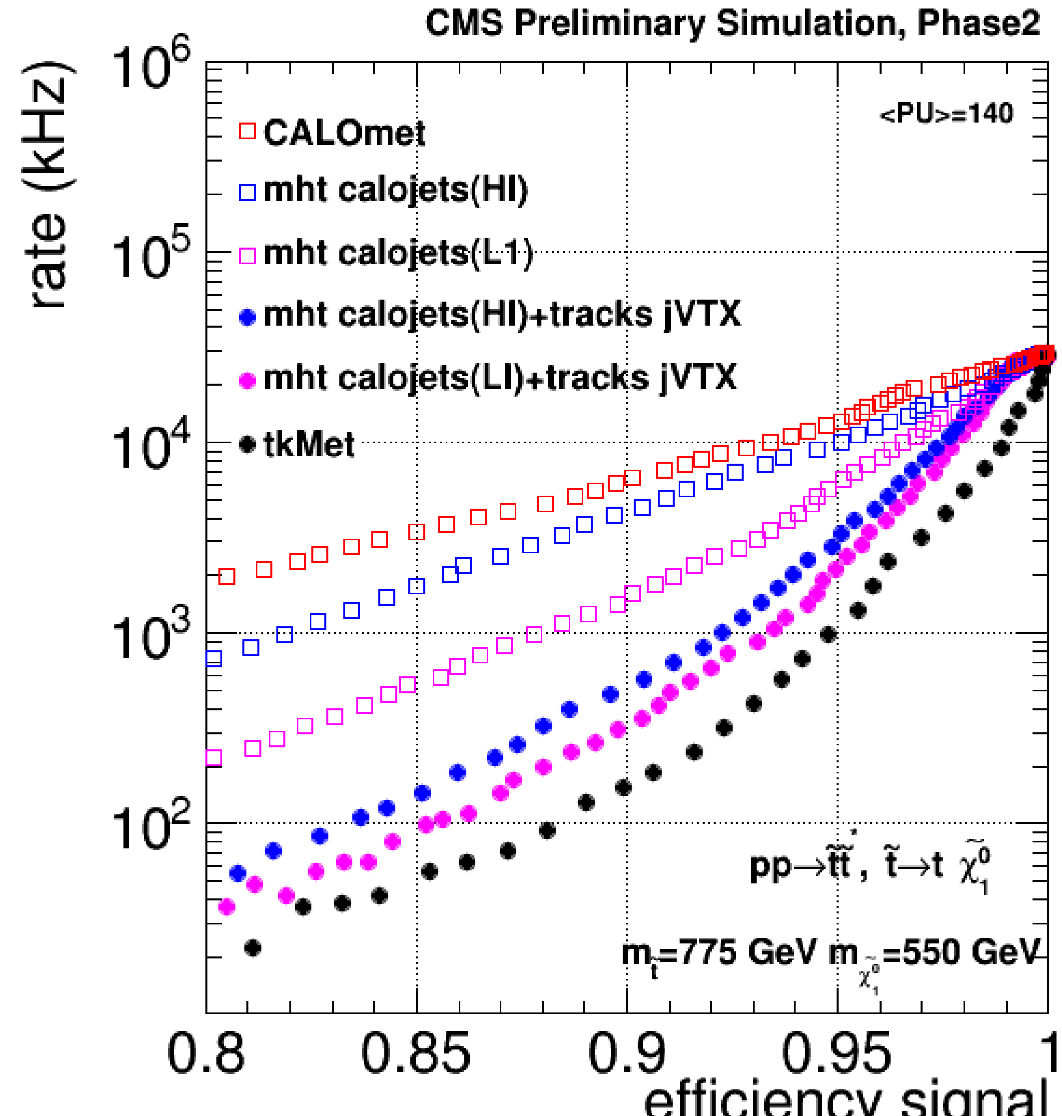
Using tracking @ L1

Example 2: Jets -- use nearby tracks to identify vertex position







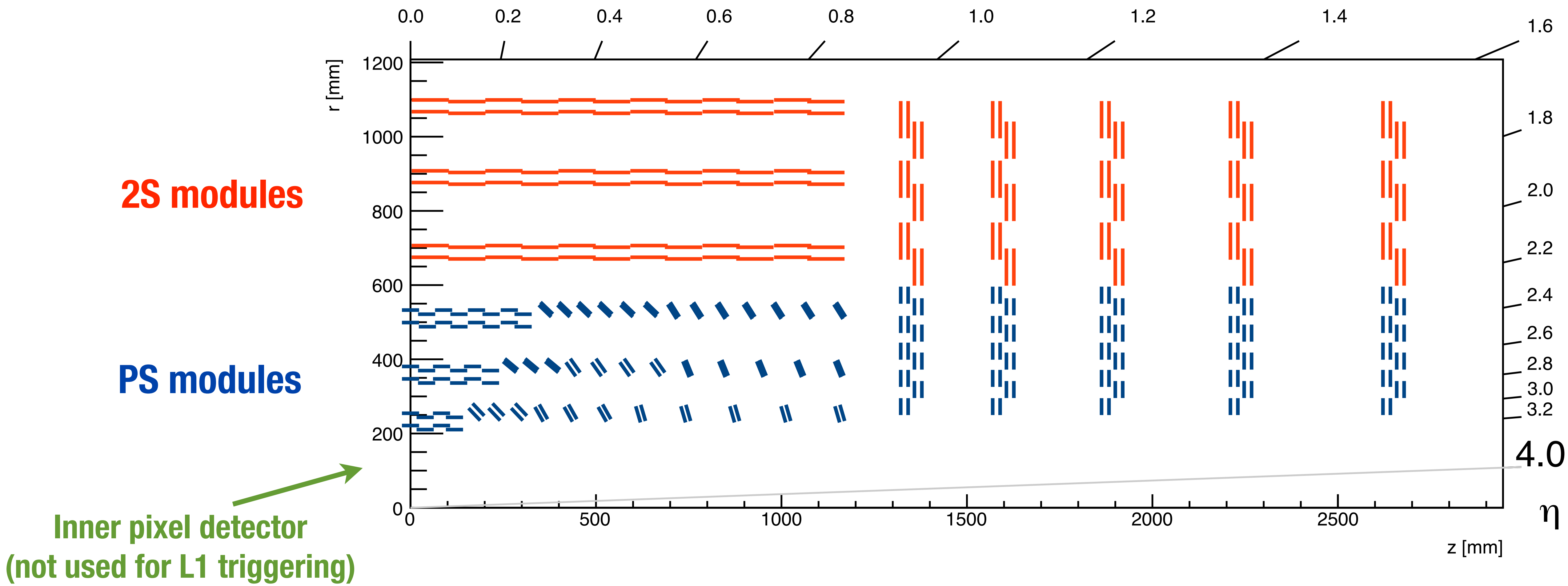


Jet vertex position → reject pileup → significant rate reductions

... how?

CMS tracker for HL-LHC

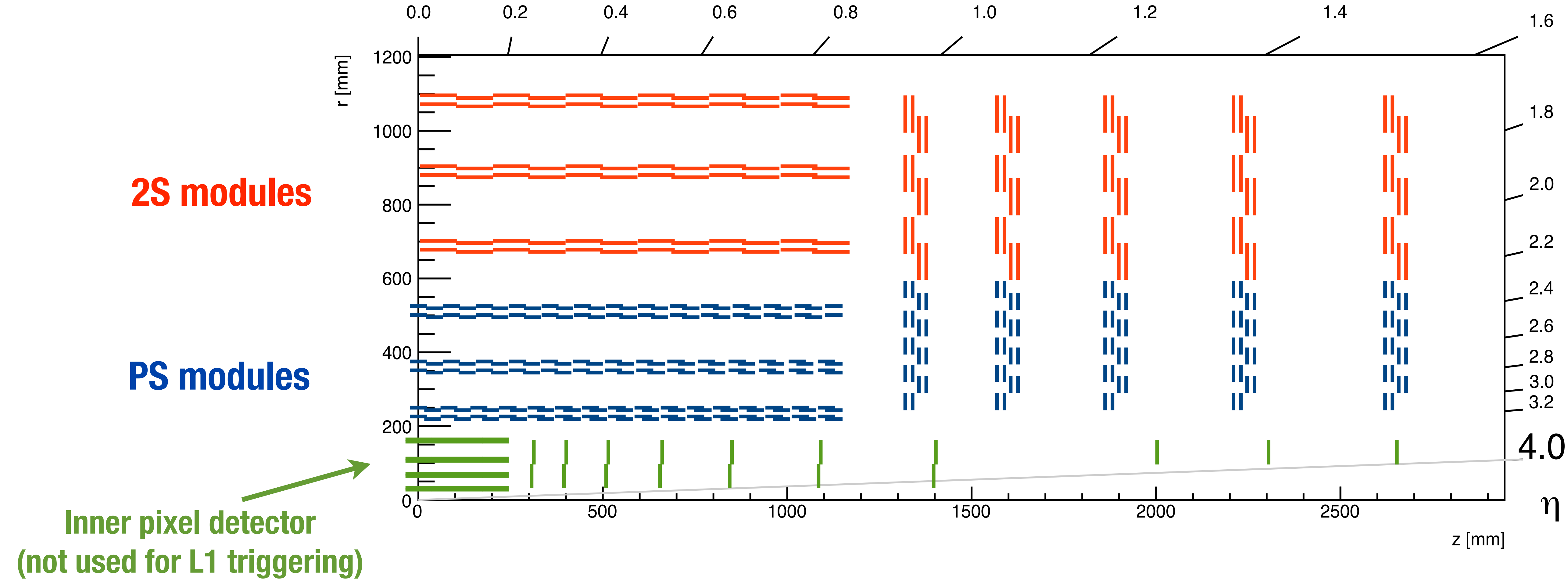
- New all silicon outer tracker + inner pixel detector
  - Increased granularity for HL-LHC occupancies
  - Tracking in hardware trigger → *Reconstruct trajectories of charged particles with $p_T > 2$ (3) GeV*



CMS tracker for HL-LHC

- New all silicon outer tracker + inner pixel detector
  - Increased granularity for HL-LHC occupancies
  - Tracking in hardware trigger

Reconstruct trajectories of charged particles with p_T > 2 (3) GeV



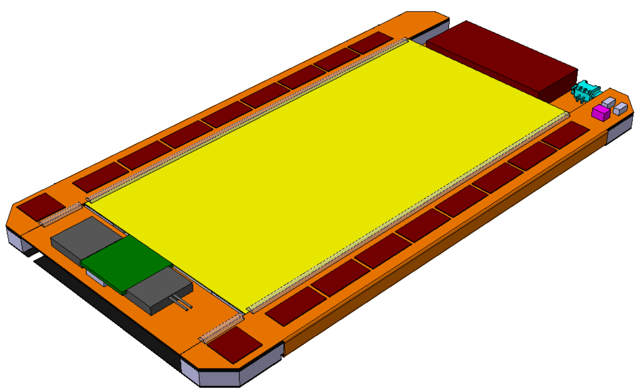
results shown today based on earlier version of geometry with flat barrel

p_T modules

- Modules provide p_T discrimination in FE electronics through hit correlations between closely spaced sensors

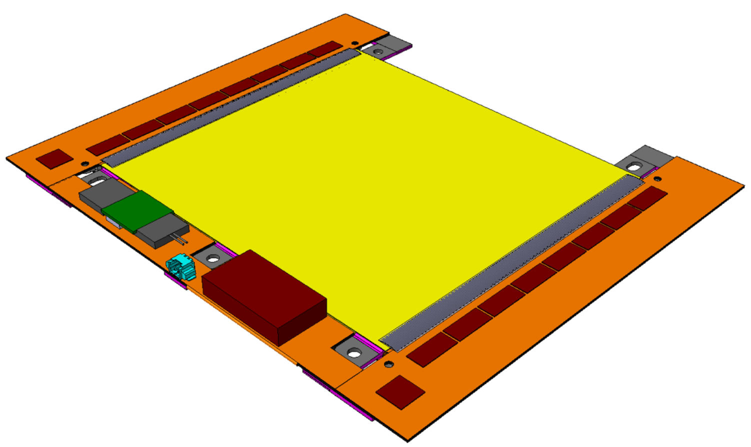
PS modules (pixel-strip)

- Top sensor: 2x2.5 cm strips, 100 μm pitch
- Bottom sensor: 1.5 mm x 100 μm pixels

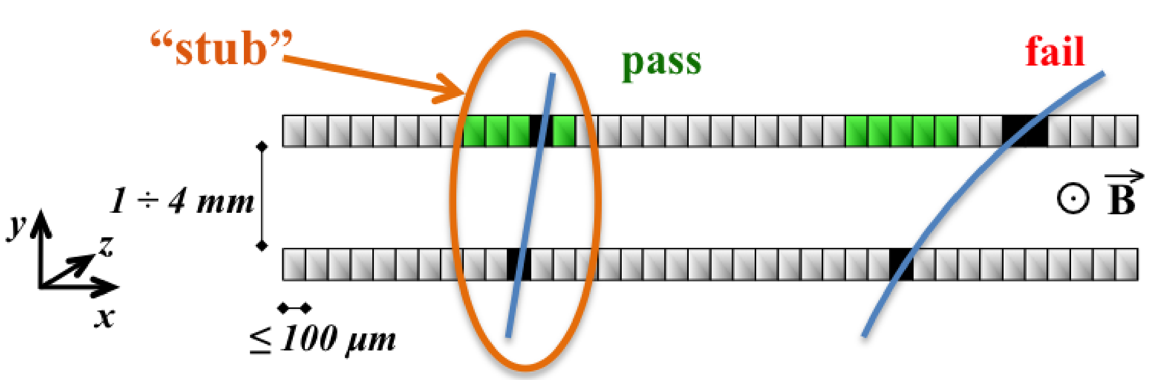


2S modules (strip-strip)

- Strip sensors 10x10 cm^2
- 2x5 cm long strips, 90 μm pitch

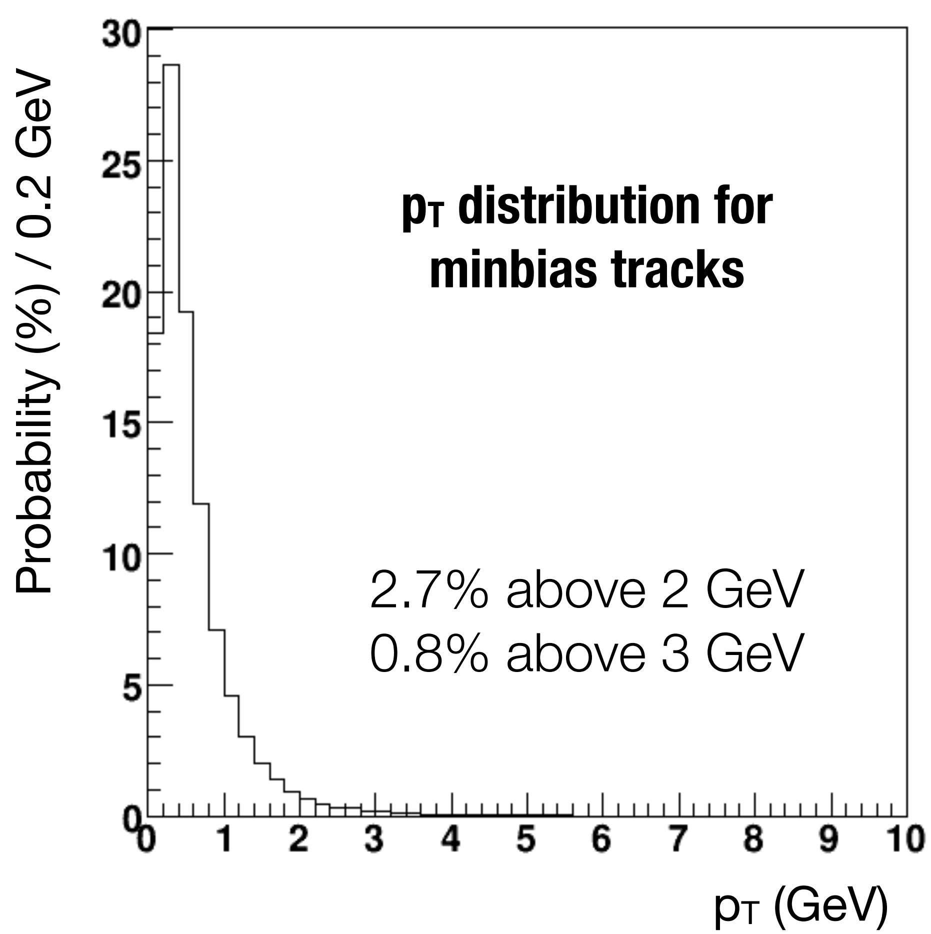


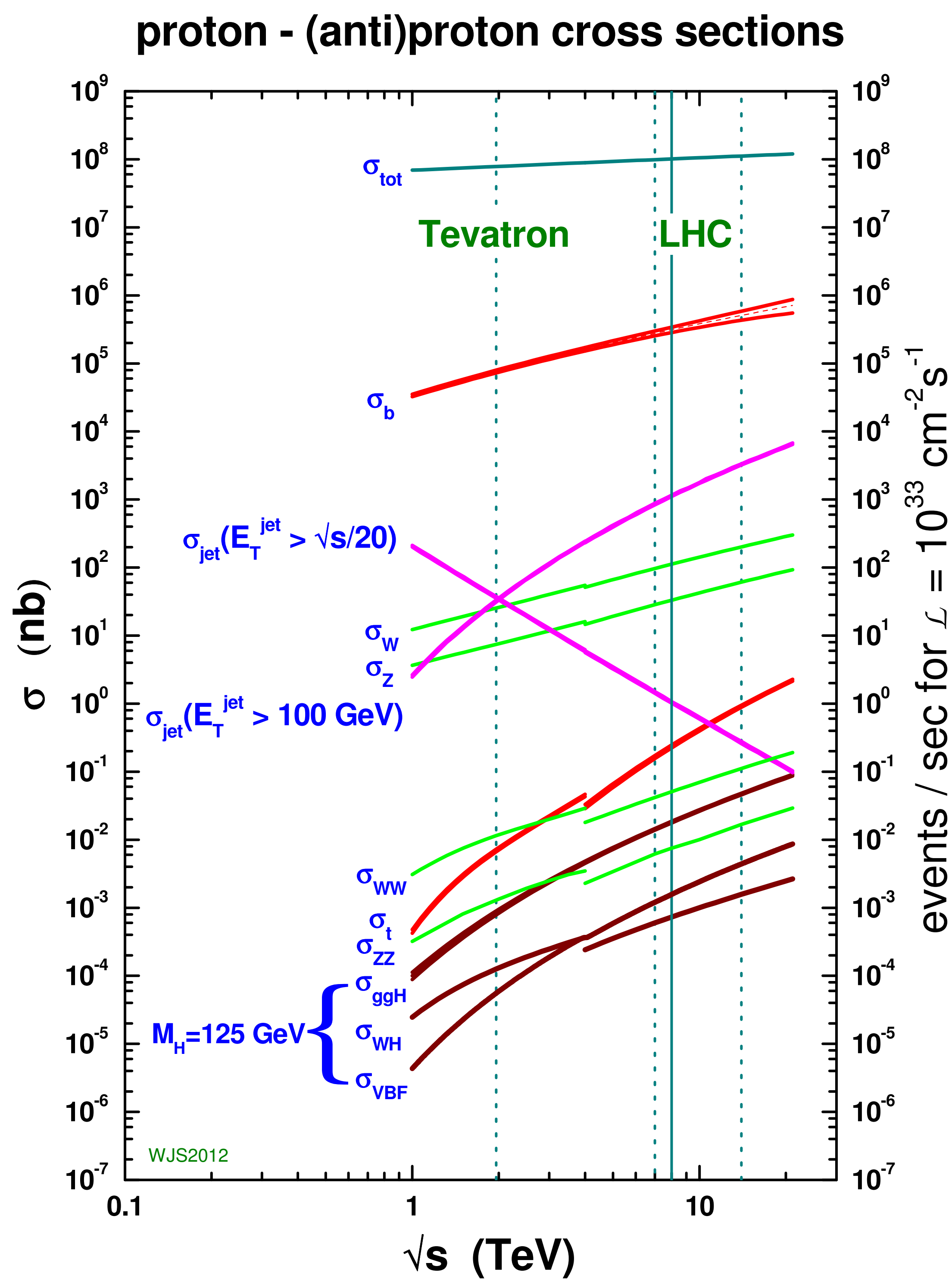
- **Stubs:** Correlated pairs of clusters, consistent with ≥ 2 GeV track
  - Data reduction at trigger readout
  - *Stubs form input to track finding*



HL-LHC conditions

- 40 million bunch crossings / second, each on average 200 interactions
- ~33 charged particles from minbias events @ 14 TeV
  - 6600 charged particles / bunch crossing!
  - ~180 tracks with p_T > 2 GeV per event





Challenges

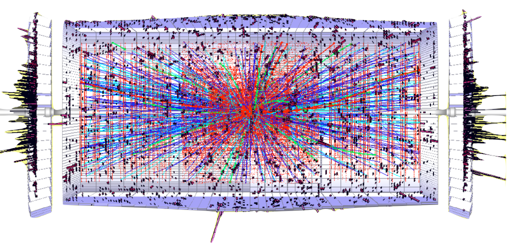
- **Combinatorics** ⇒ 15-20K input stubs / BX
- **Data volumes** ⇒ up to ~50 Tbits/s
- L1 trigger decision within 12.5 μs **(*)** ⇒ **time available for track finding ~<u>4 μs</u>**

- A track-trigger operating at 40 MHz with <10 μs latency has never been built!
  - **CDF:** L2 with lower input rate & less dense environment
  - **ATLAS FTK:** After L1 with lower input rate & longer latency

(*) *increased beyond current CMS trigger system, also increasing L1 output rate 100 kHz ⇒ 750 kHz*

Track trigger strategy

Data transfer (~50 Tbs)



Δt = 0 → Data formatting → Pattern recognition → Track fitting + duplicate removal → Δt = 4 μs

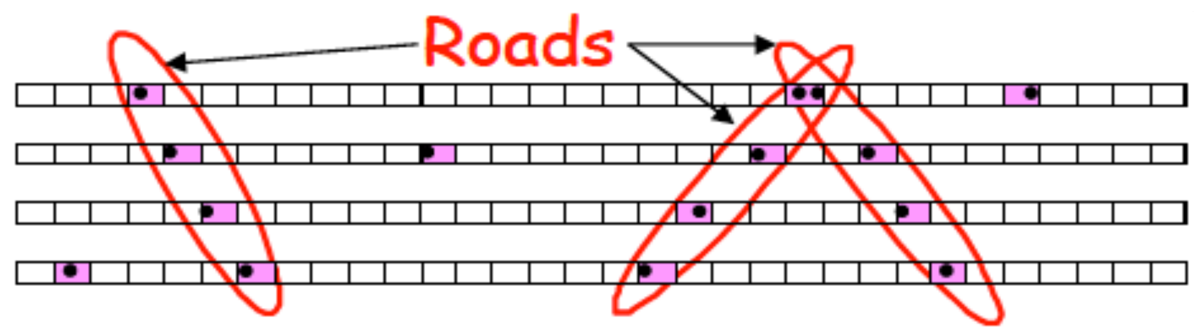
Output tracks to L1 trigger

- Parallelization
  - Divide tracker in segments in ϕ / z
  - Time-multiplexed systems -- process several BX simultaneously
- Different approaches to attack combinatorics & occupancies

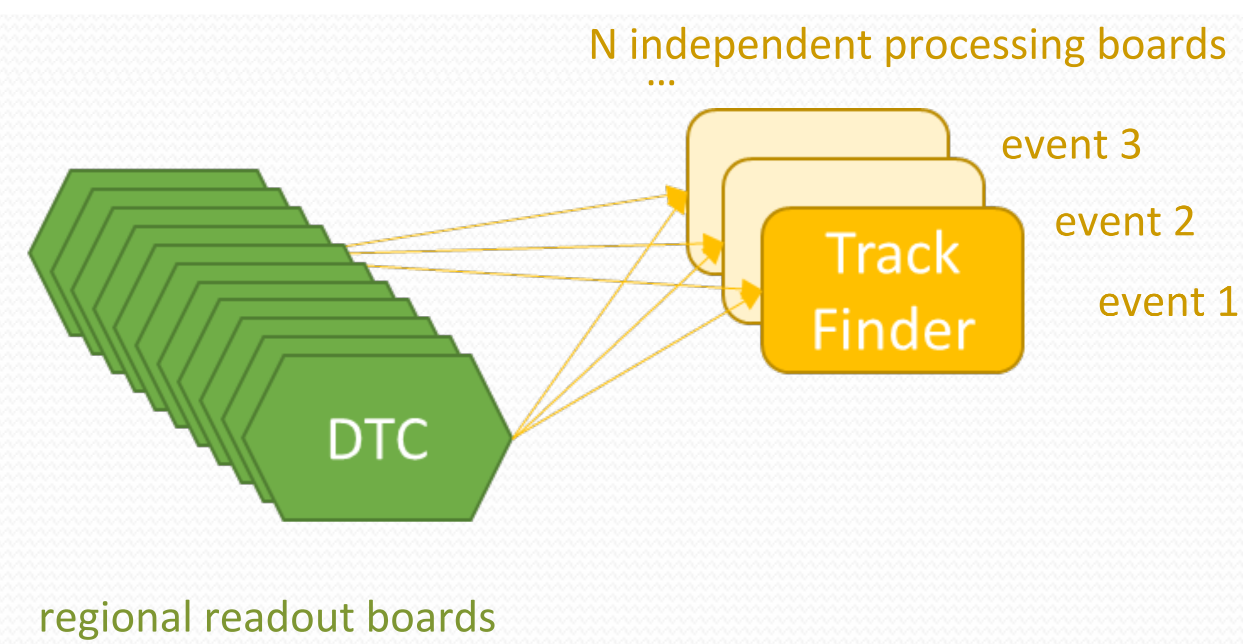
CMS track triggering

R&D efforts ongoing --
different approaches for handling occupancies & combinatorics

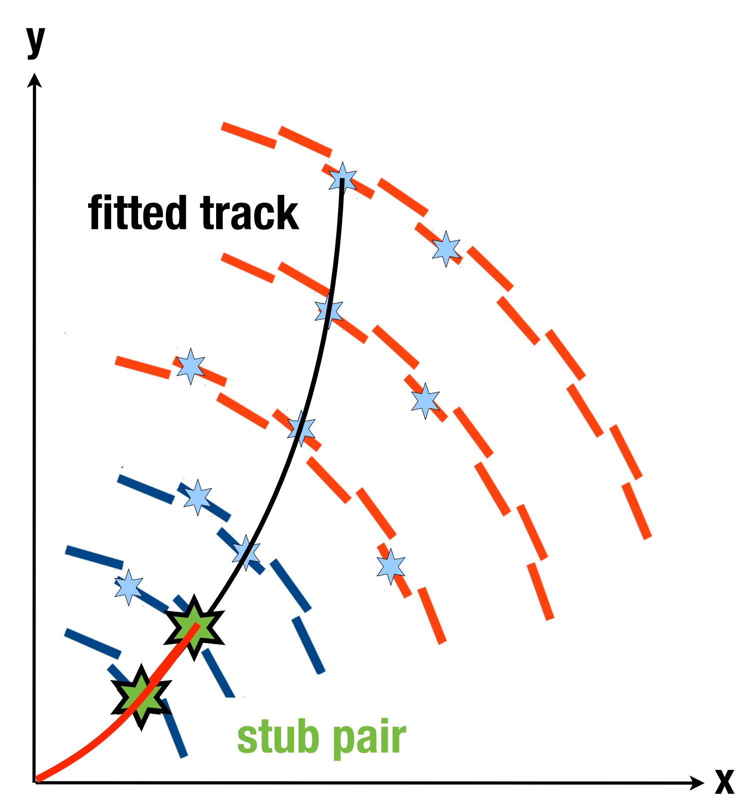
Associative memories pattern matching



Time multiplexed architecture



Tracklet method

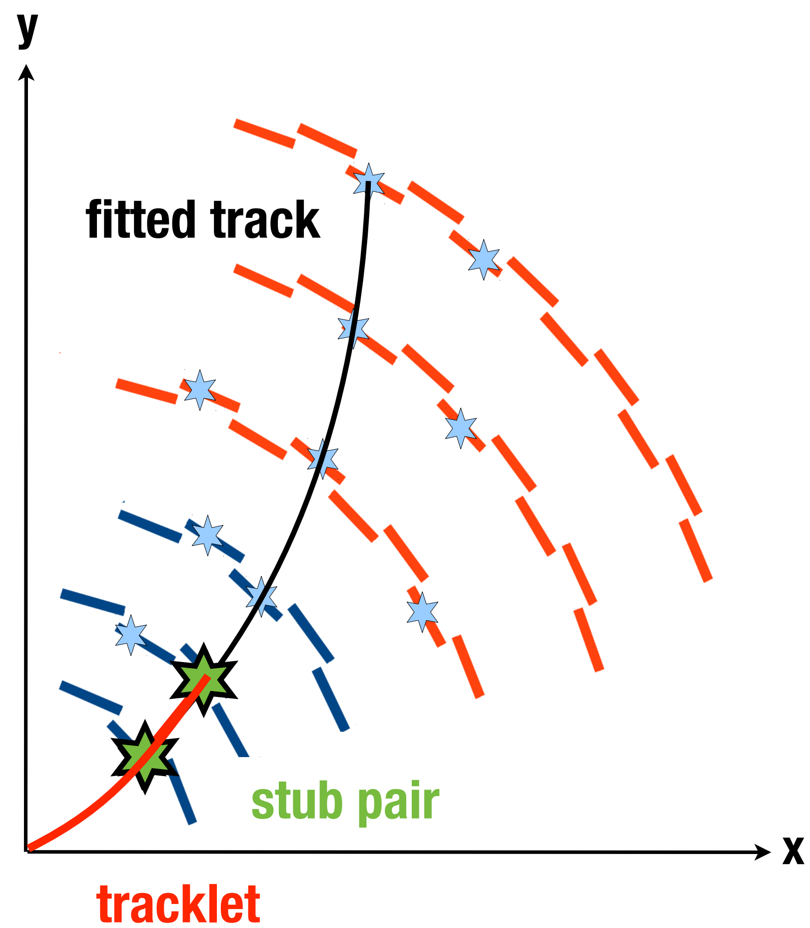


Tracklet method

Tracklet approach

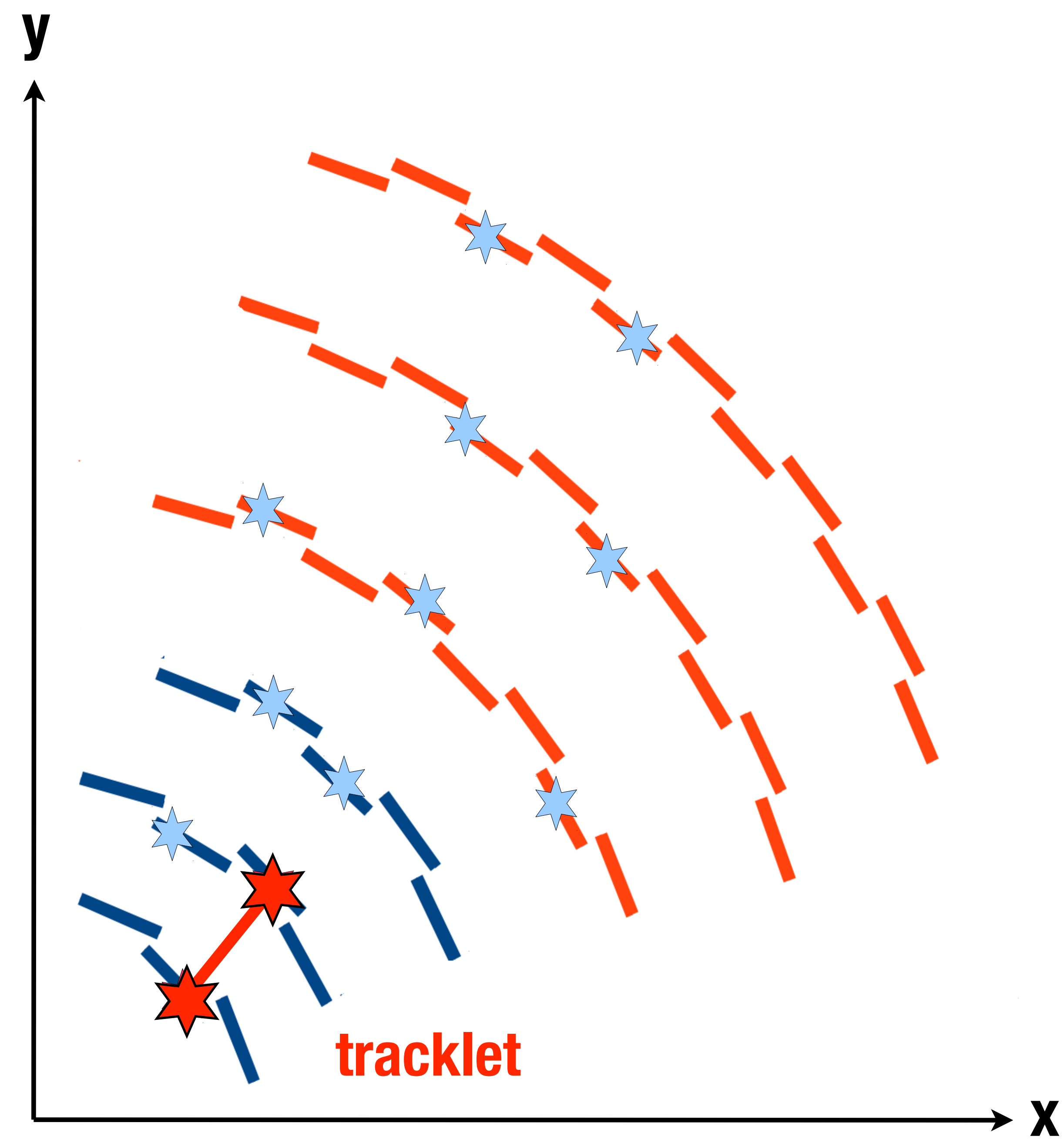
- Minimal hardware system based on commercial **FPGAs**
  - Off-the-shelf hardware
  - Ever-increasing capability + programming flexibility → **ideal for fast track finding**

- Tracklet algorithm
  - **Road search algorithm**
  - Few (simple) calculations
  - Parallelized processing in time & space
  - Naturally pipelined implementation
  - Operates at a fixed latency -- truncate if necessary



Tracklet algorithm: **Seeding**

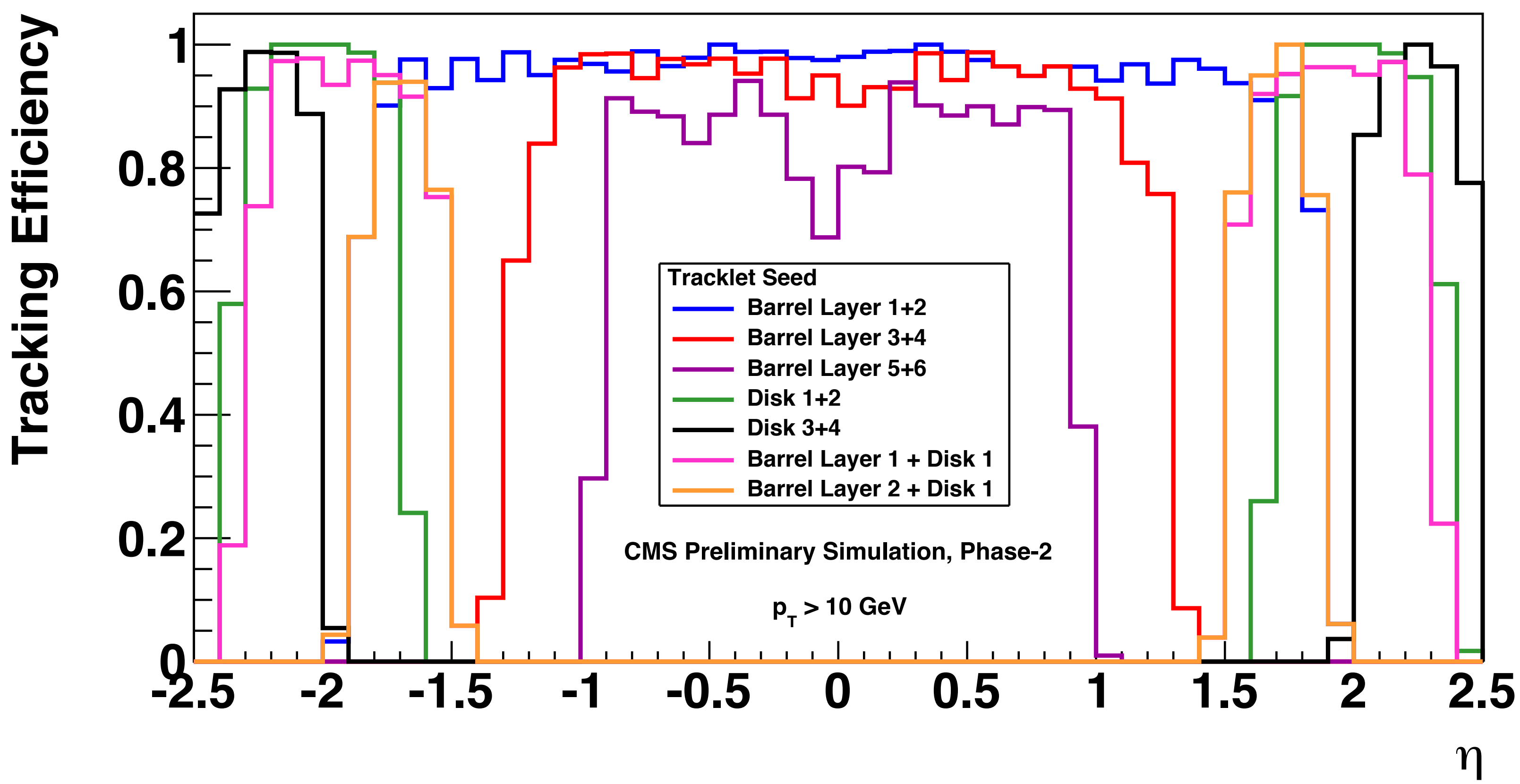
- **Seed** by forming tracklets
  - Pairs of stubs in adjacent layers/disks
  - Initial tracklet parameters from stubs + beamspot constraint
  - Consistent with $p_T > 2$ GeV



Tracklet algorithm: **Seeding**

- **Seed** by forming tracklets
  - Pairs of stubs in adjacent layers/disks
  - Initial tracklet parameters from stubs + beamspot constraint
  - Consistent with $p_T > 2$ GeV

Seed multiple times in *parallel* to ensure good coverage & redundancy

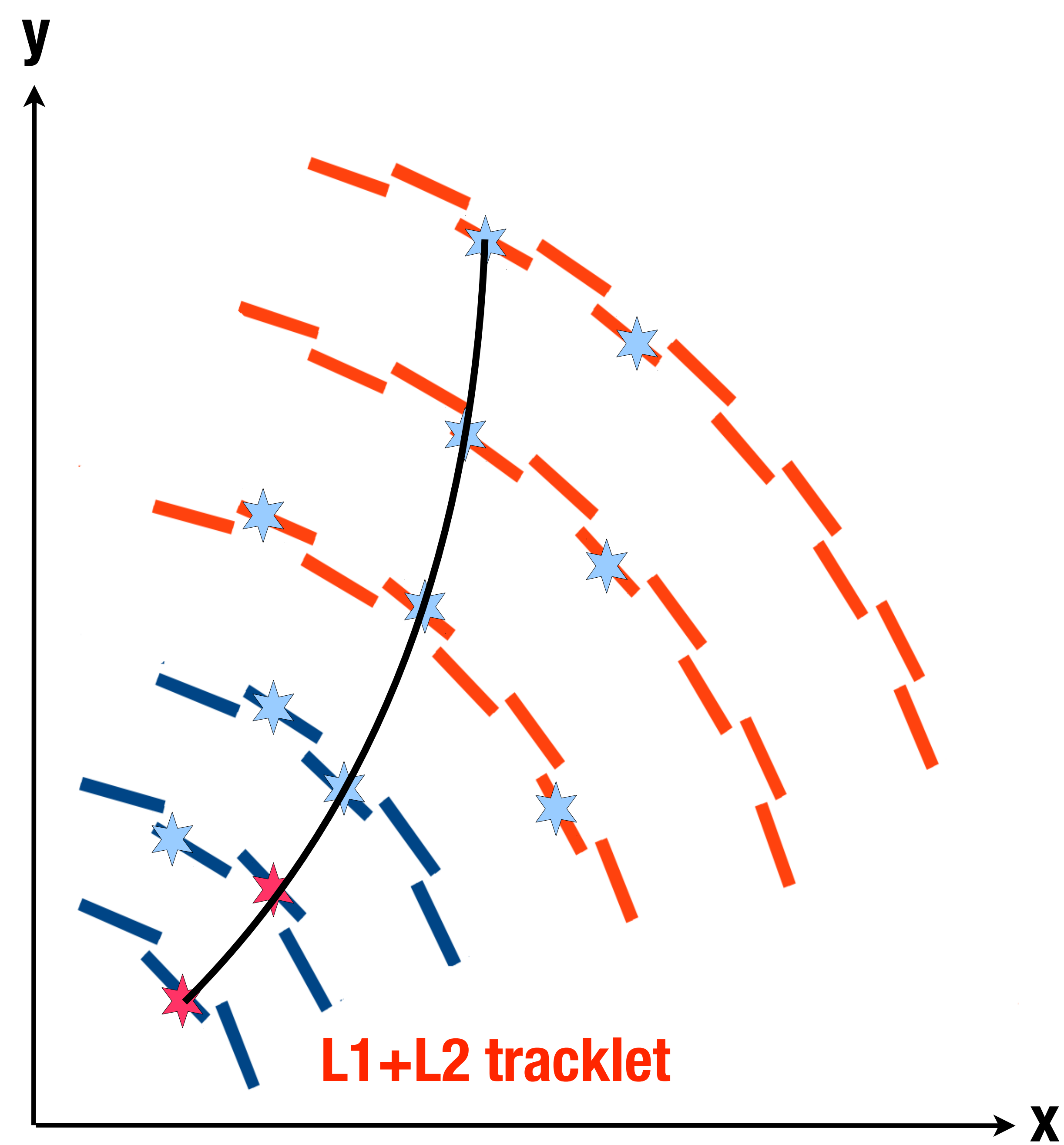


Efficiency for single muons (using firmware emulation)

Tracklet algorithm: **Project**

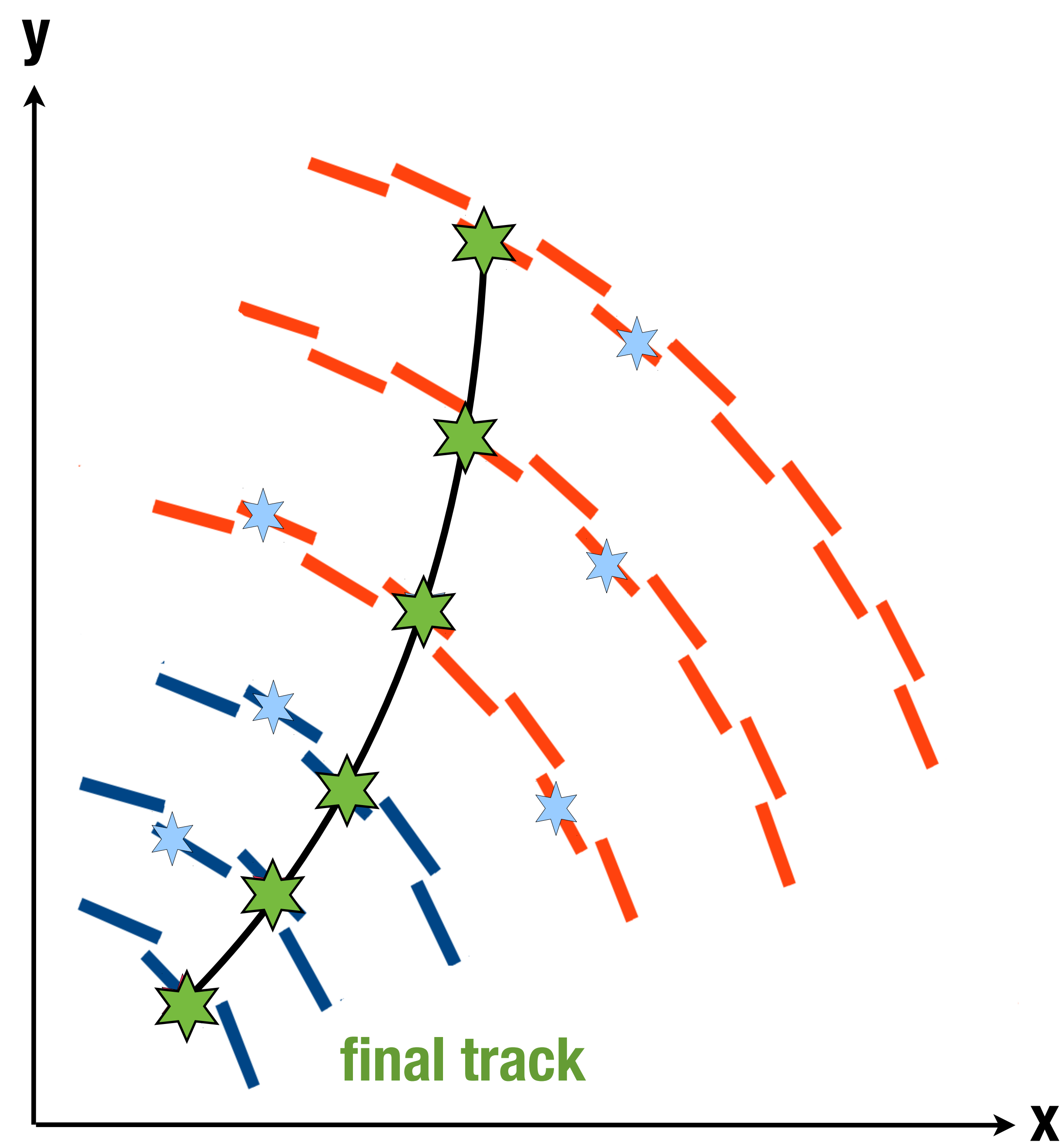
- Project tracklets to other layers & disks to search for matching stubs
- Use predefined search windows
- Project both inside-out & outside-in

projections to different layers/disks done in parallel!



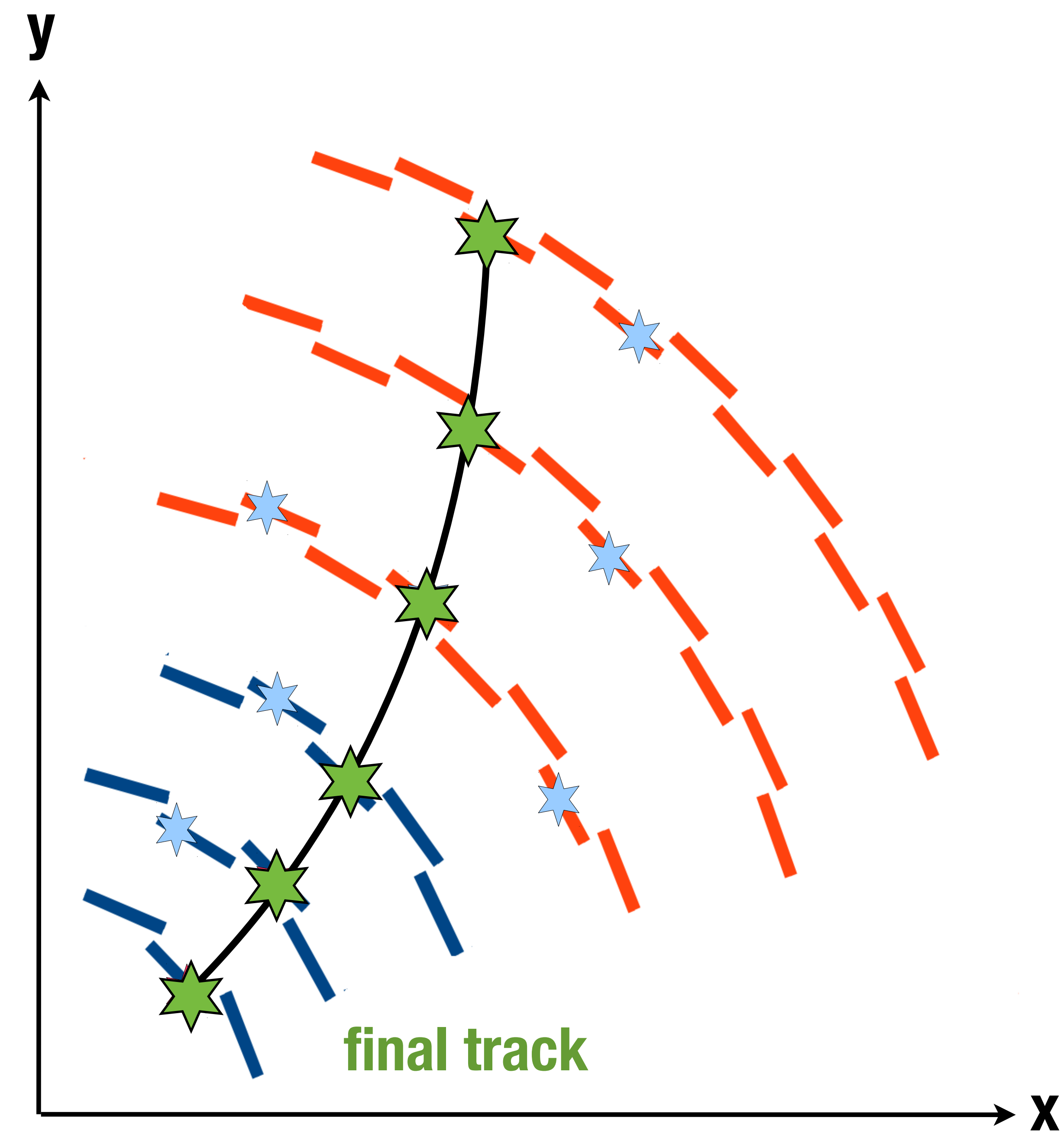
Tracklet algorithm: **Fit**

- Perform track fit of stubs matched to trajectory
- Linearized χ^2 fit
- Gives final track parameters
  - p_T, η, ϕ_0, z_0
  - Optionally d_0



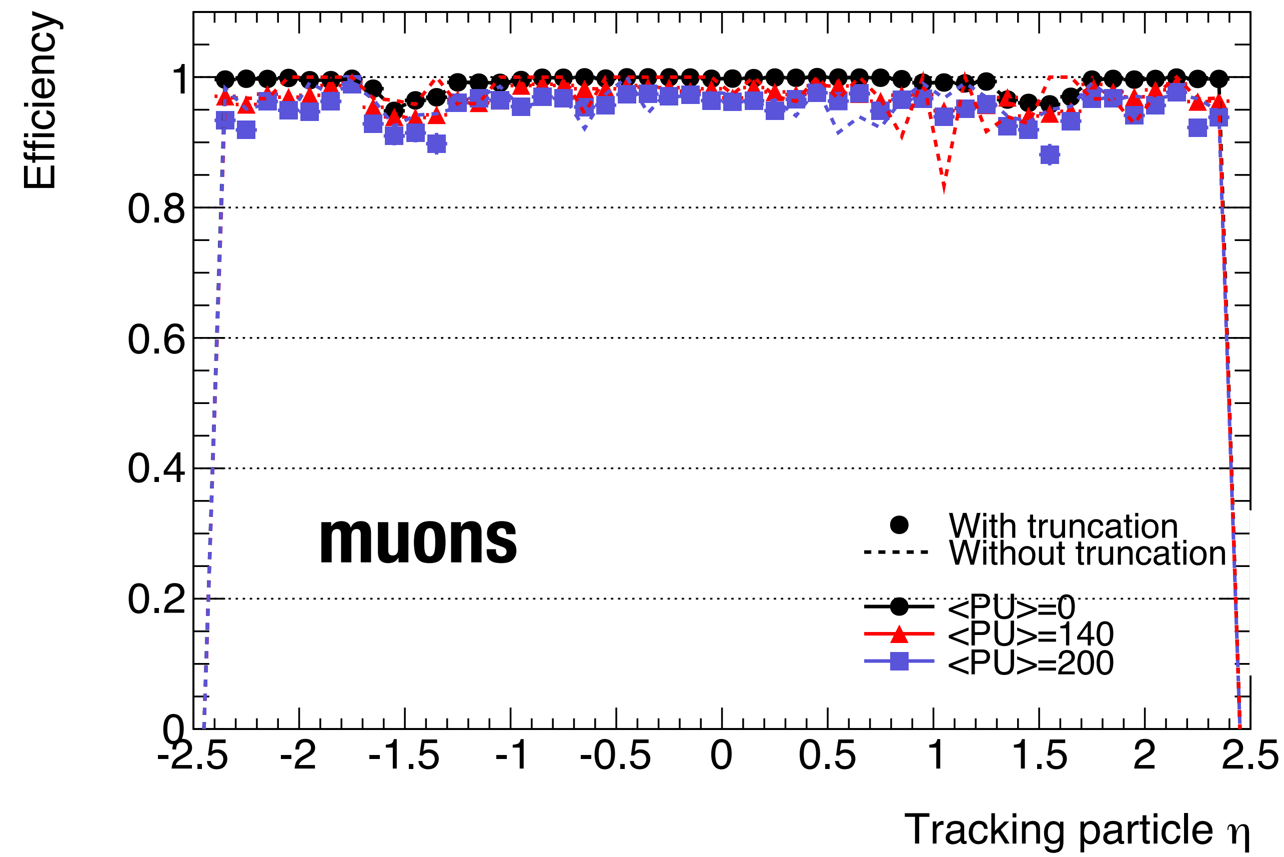
Tracklet algorithm: **Duplicate Removal**

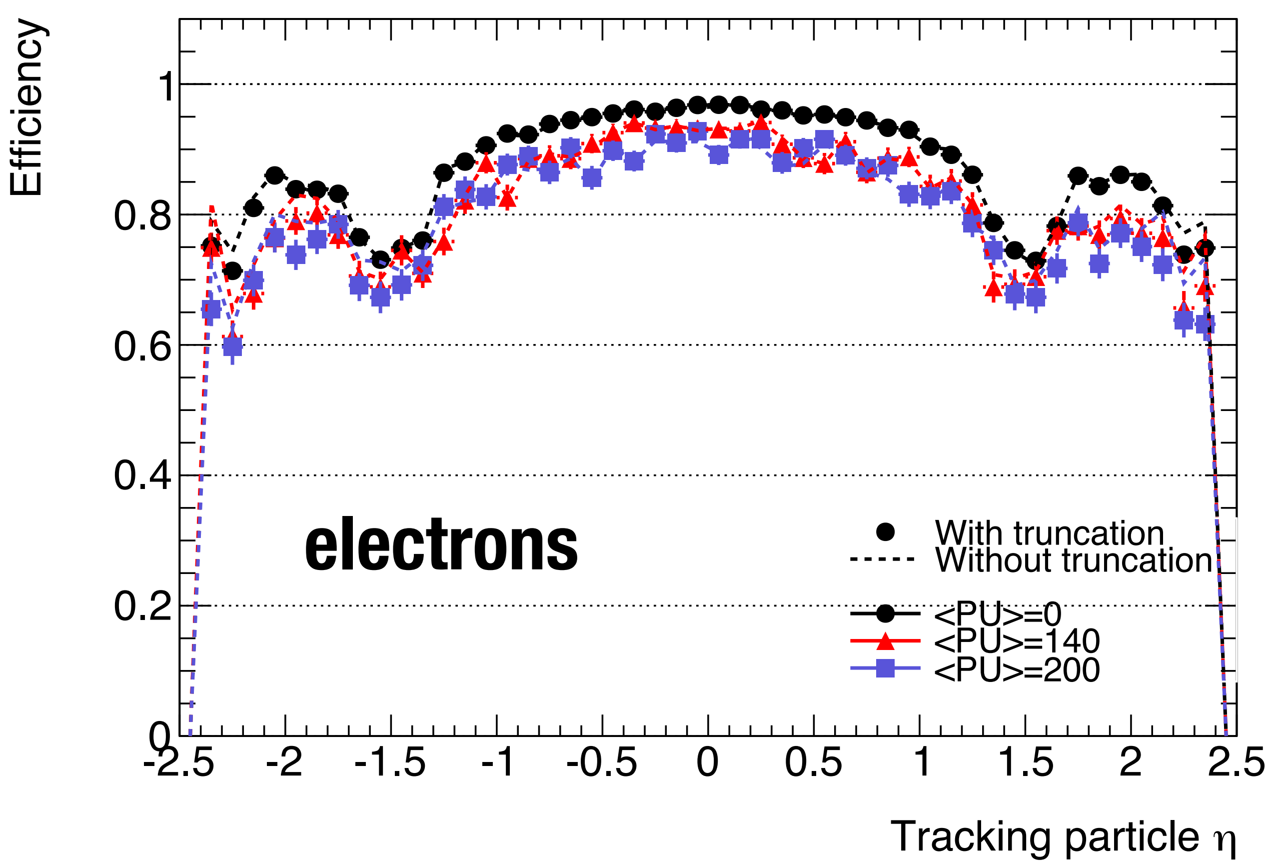
- A given track can be found many times due to seeding in multiple pairs of layers
  - Ensures high efficiency
- Remove duplicates based on shared stubs
  - Compare pairs of tracks & count # independent / shared stubs

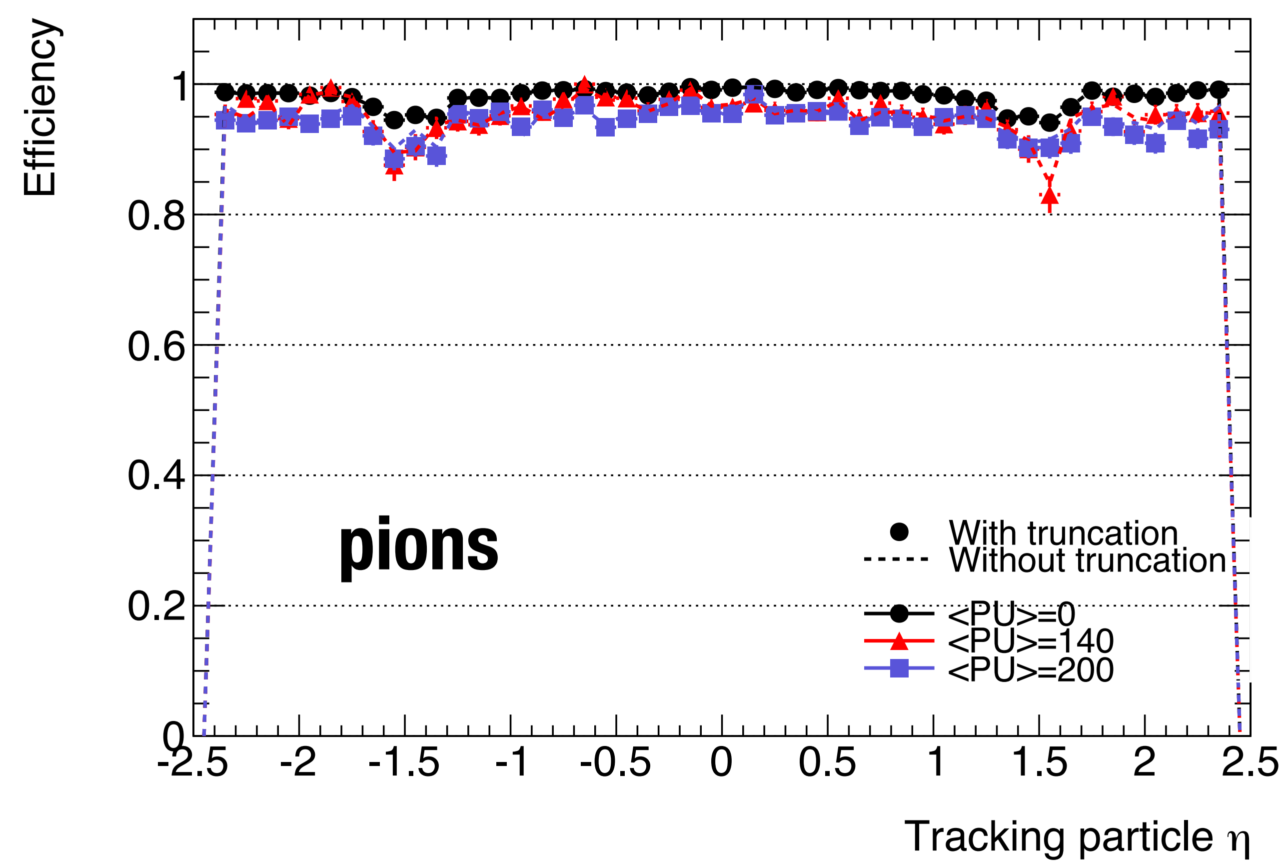


Tracking performance

- Efficiency as function of η for single particles (e/μ/π)
- High efficiencies achieved
- Minimal impact from truncation

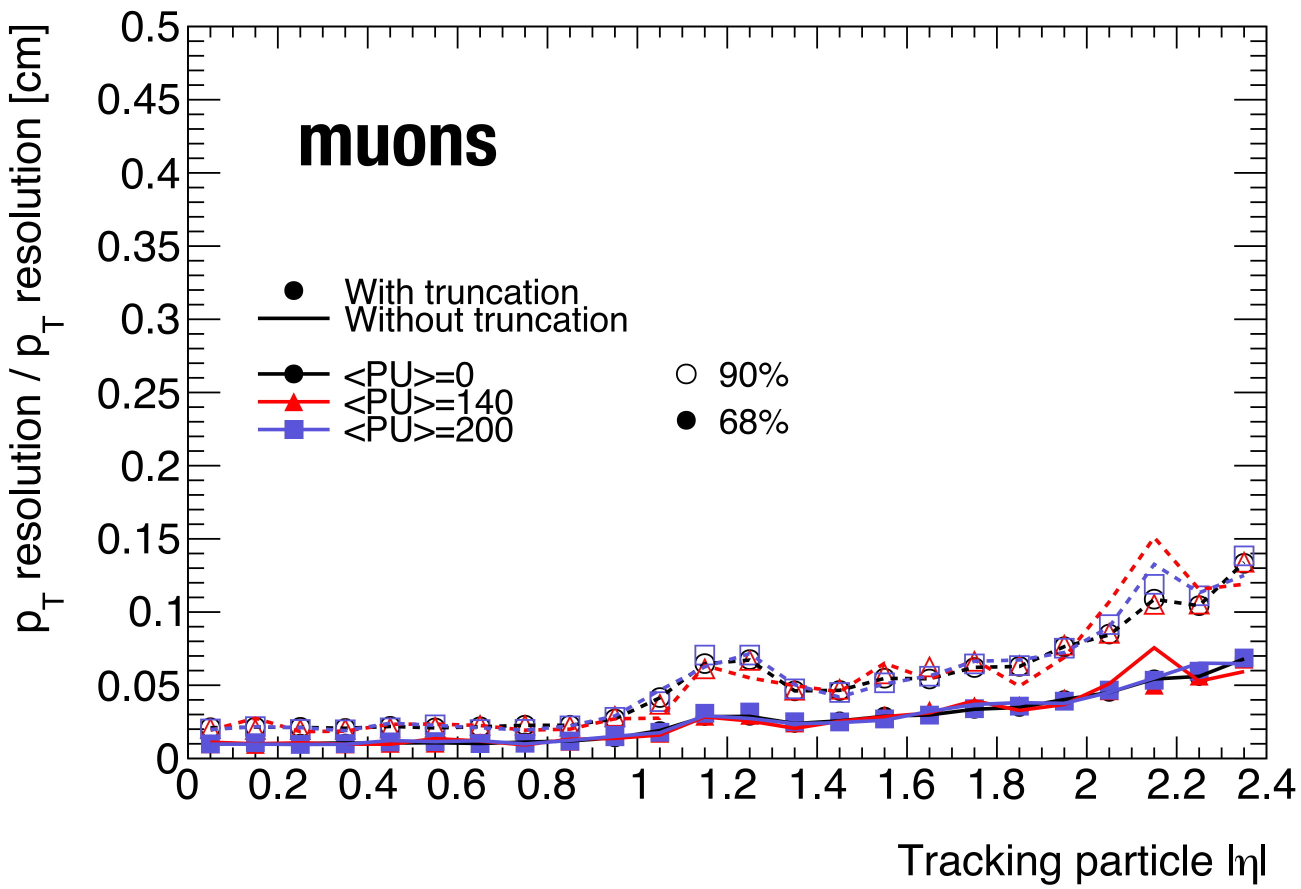


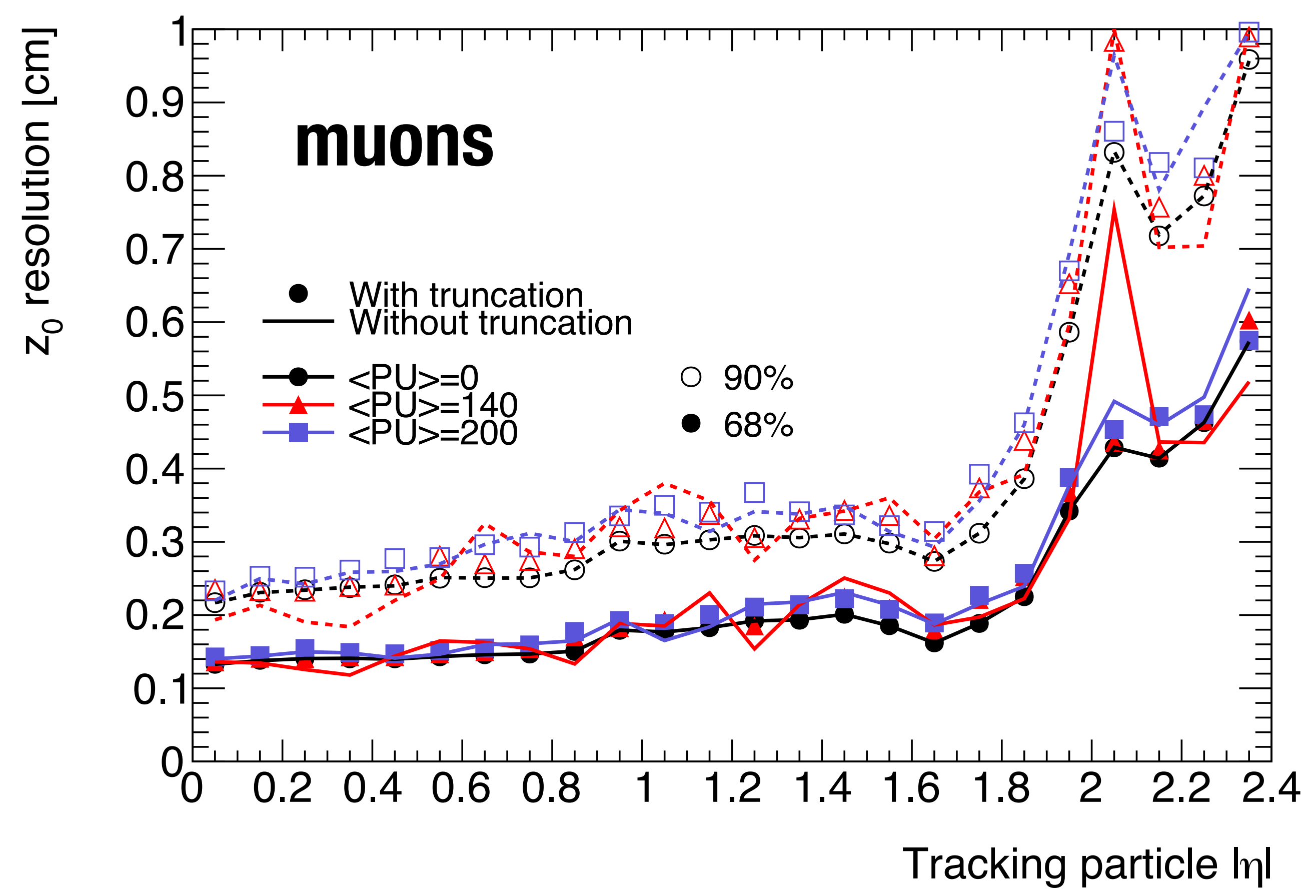




Tracking performance

- **σ(z_0) ~ 1 mm** for wide range of η thanks to PS modules
- **σ(p_T)/p_T ~ 1%** at central η for high-p_T track





- Already good enough resolution for trigger

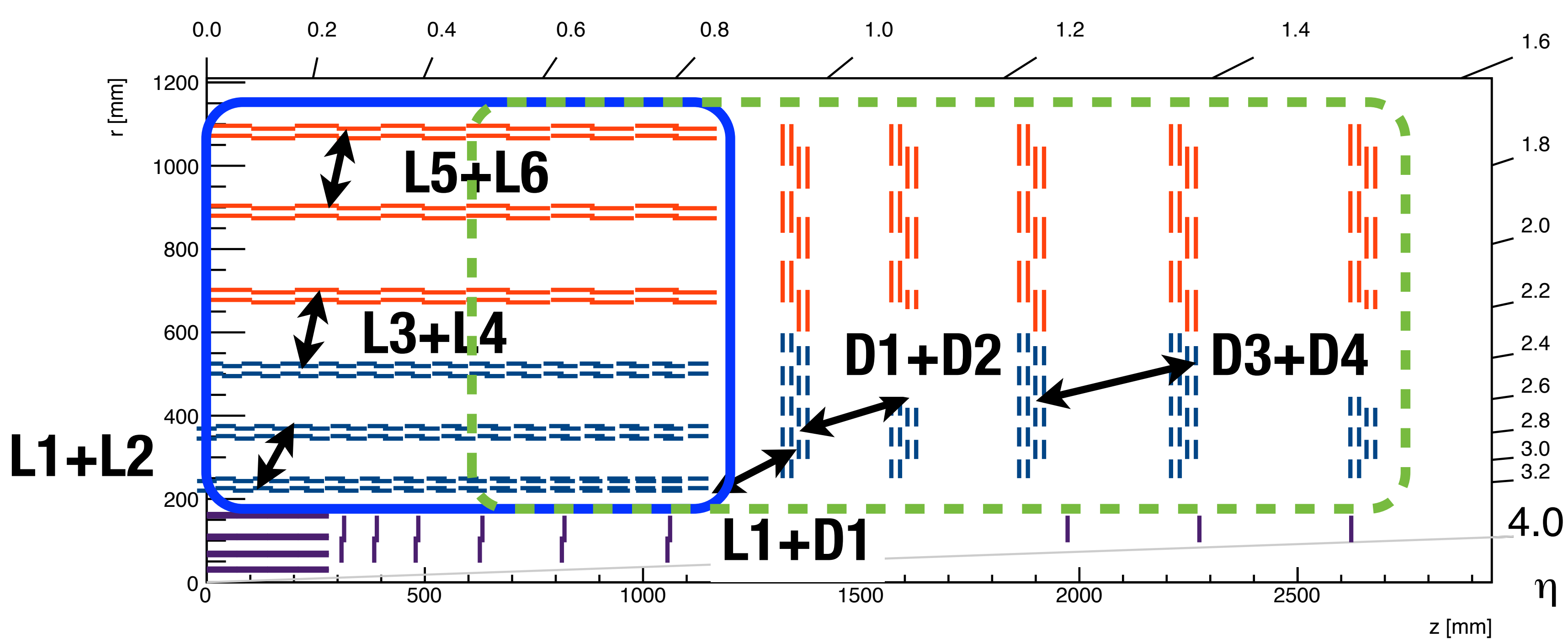
- Known degradation from using too few bits in certain points of calculations, can be corrected

... how to implement this?

Algorithm implementation

- Simulations of method
  - **Floating-point** simulation (C++)
  - **Integer emulation** of firmware (C++)
  - **FPGA firmware simulation** (Vivado)

- Hardware implementation
  - Currently implemented in firmware as two projects (**half barrel** vs **hybrid+disks**)



Hardware configuration

- System replicated for parallel data processing

- Divide detector in ϕ sectors
  - Tracks with $p_T > 2$ GeV span max. 2 sectors
  - Dedicated processing board for each sector

- System time multiplexed by factor 6
  - New event every 150 ns

- Tracklet formation within sector, projections to neighboring sectors sent there for stub matching

Hardware configuration

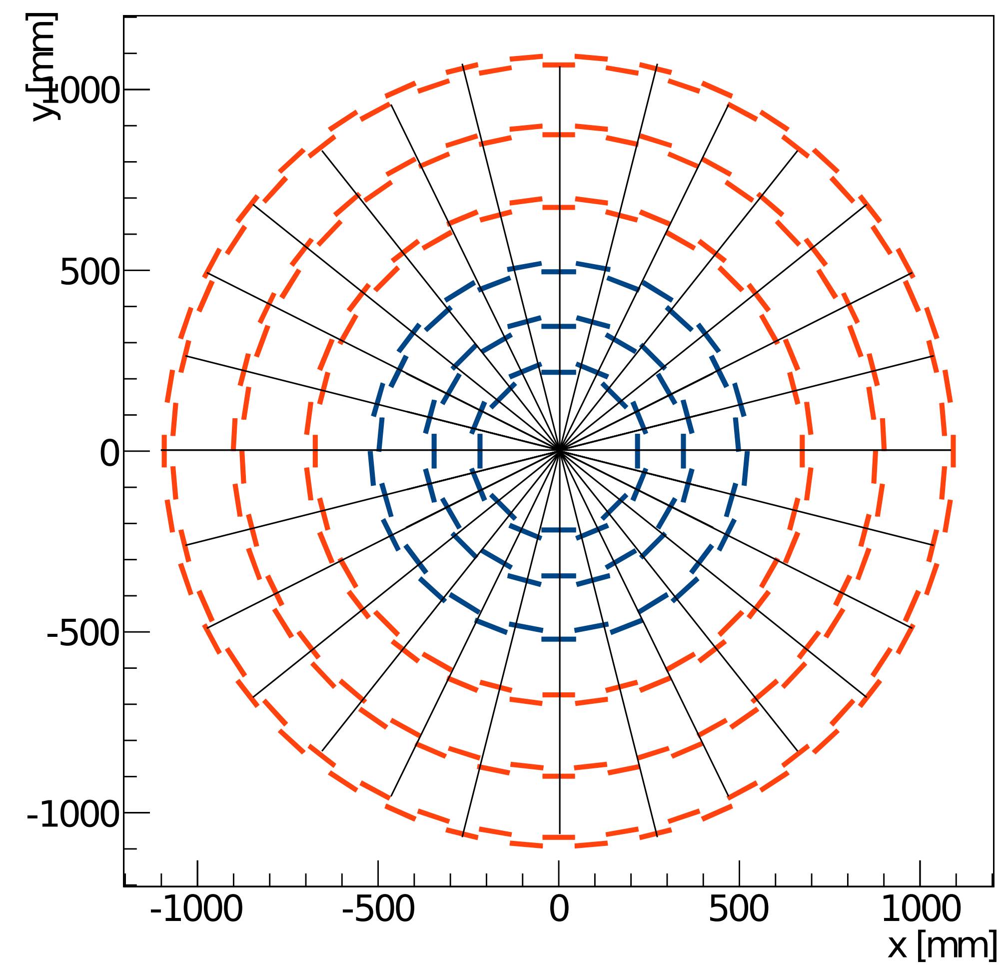
- System replicated for parallel data processing

- Divide detector in ϕ sectors
    - Tracks with $p_T > 2$ GeV span max. 2 sectors
    - Dedicated processing board for each sector

- System time multiplexed by factor 6
    - New event every 150 ns

- Tracklet formation within sector, projections to neighboring sectors sent there for stub matching

28 ϕ sectors



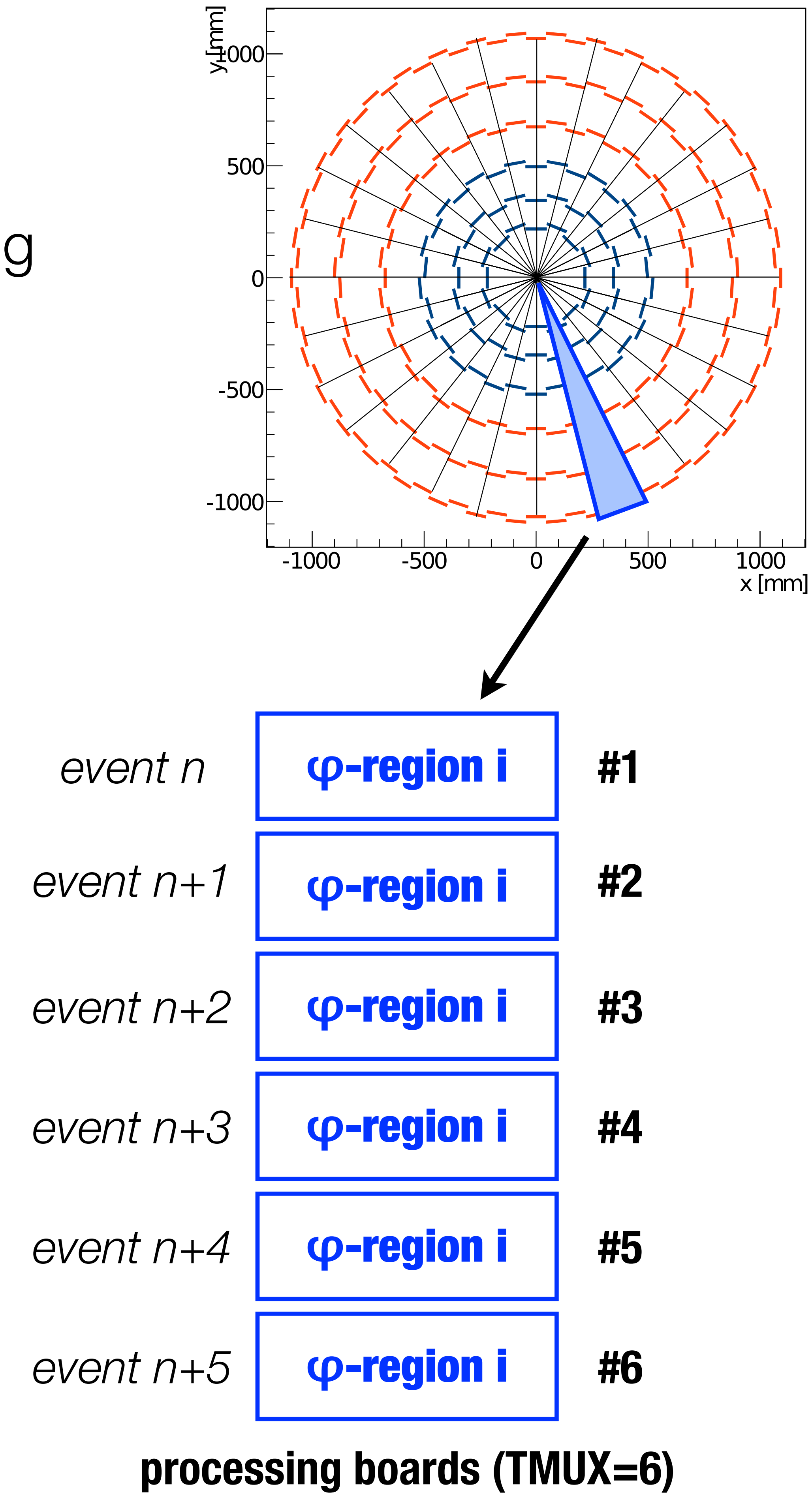
Hardware configuration

- System replicated for parallel data processing

- Divide detector in ϕ sectors
  - Tracks with $p_T > 2$ GeV span max. 2 sectors
  - Dedicated processing board for each sector

- System time multiplexed by factor 4-8
  - New event every 100-200 ns

- Tracklet formation within sector, projections to neighboring sectors sent there for stub matching



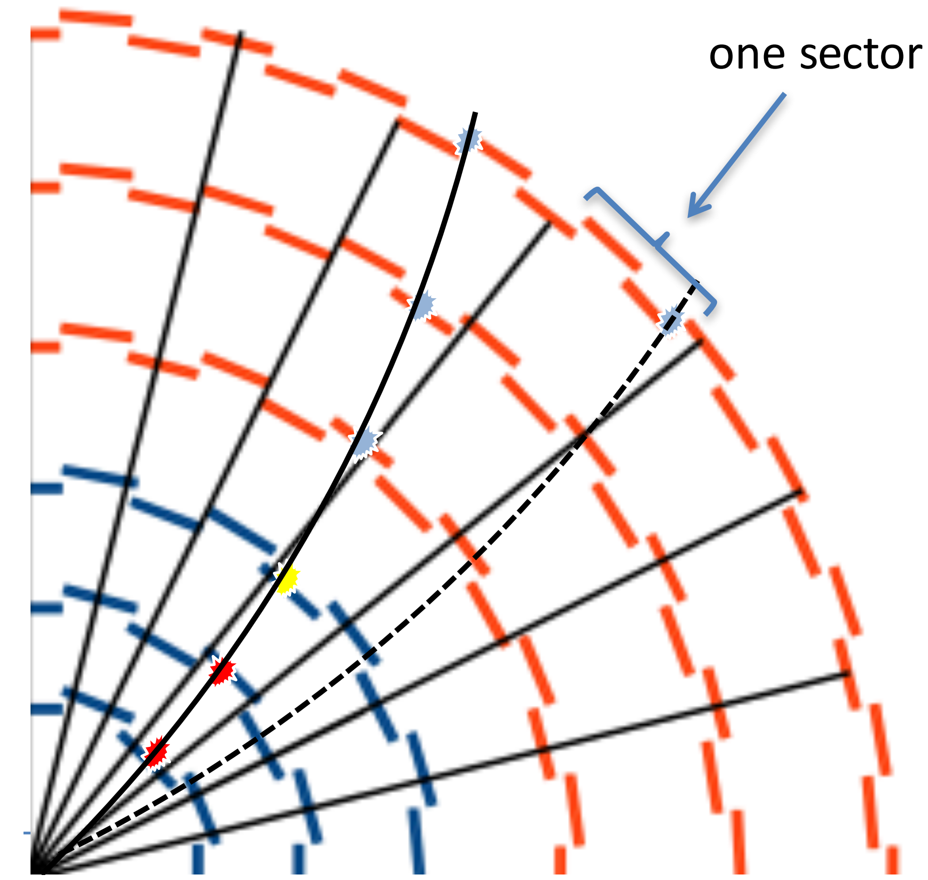
Hardware configuration

- System replicated for parallel data processing

- Divide detector in ϕ sectors
  - Tracks with p_T > 2 GeV span max. 2 sectors
  - Dedicated processing board for each sector

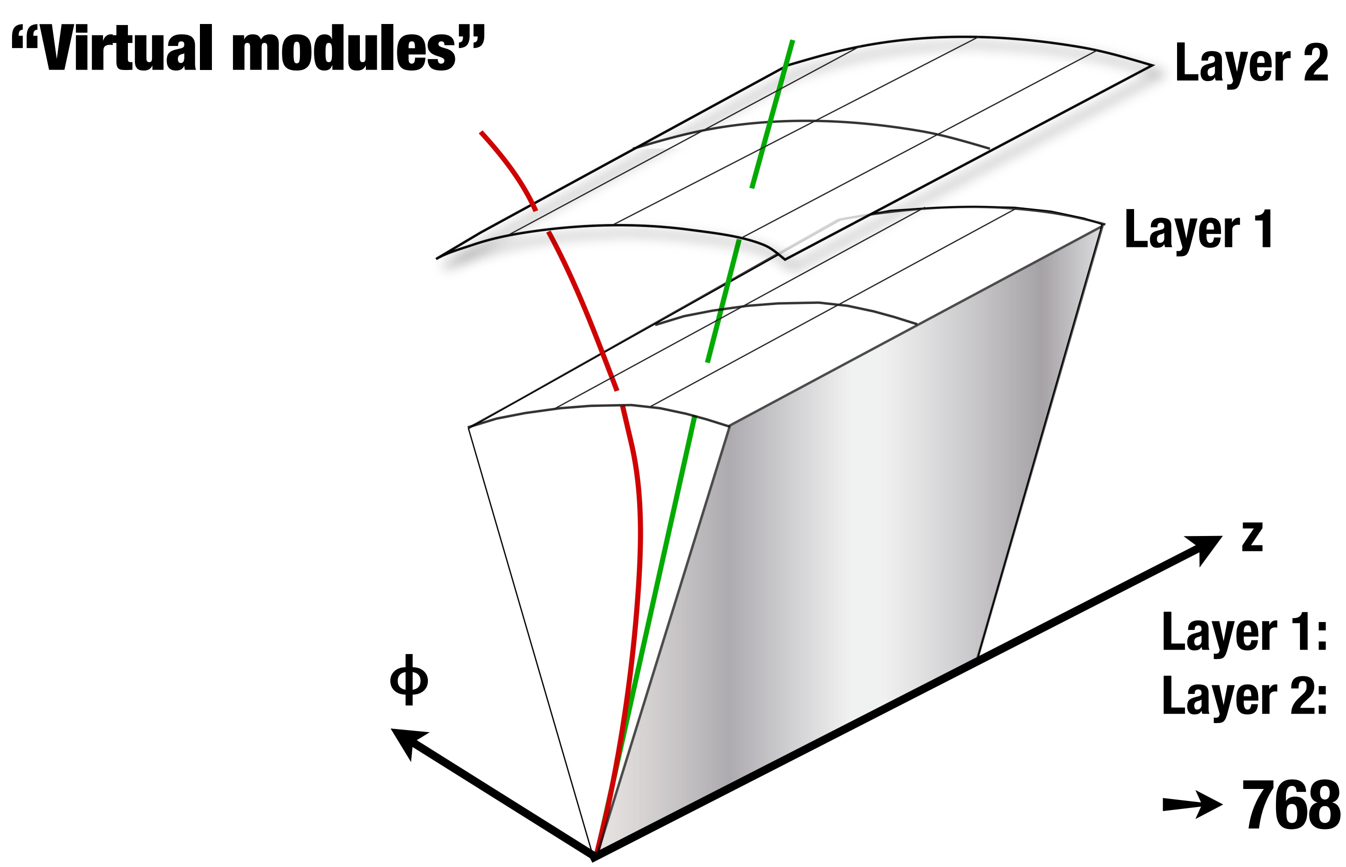
- System time multiplexed by factor 6
  - New event every 150 ns

- Tracklet formation within sector, projections to neighboring sectors sent there for stub matching



Challenge of combinatorics

- Main challenge -- combinatorics in forming tracklets & matching projections
- Subdivide layers & sector into smaller units to allow parallel processing



Layer 1: 3ϕ x 8z (2 shown) subregions
Layer 2: 4ϕ x 8z (2 shown) subregions

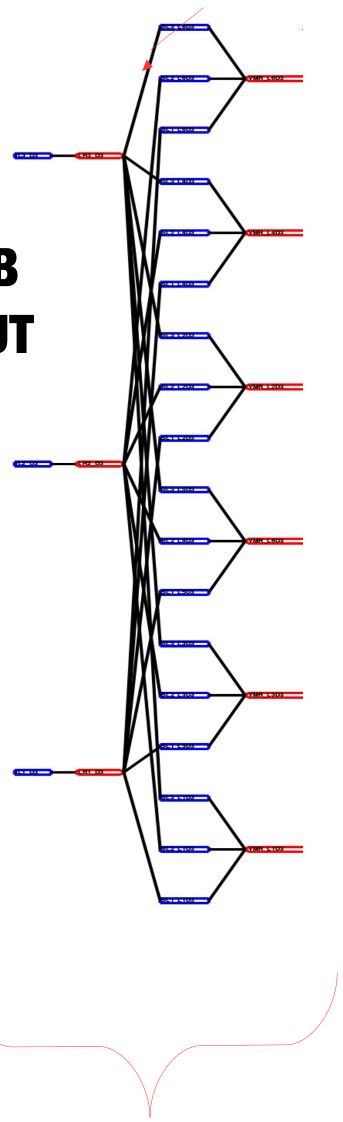
→ 768 pairs, but only 96 can form valid tracklets

Project overview

memories
processing modules

8 processing + 2 transmission steps implements algorithm

STUB
INPUT

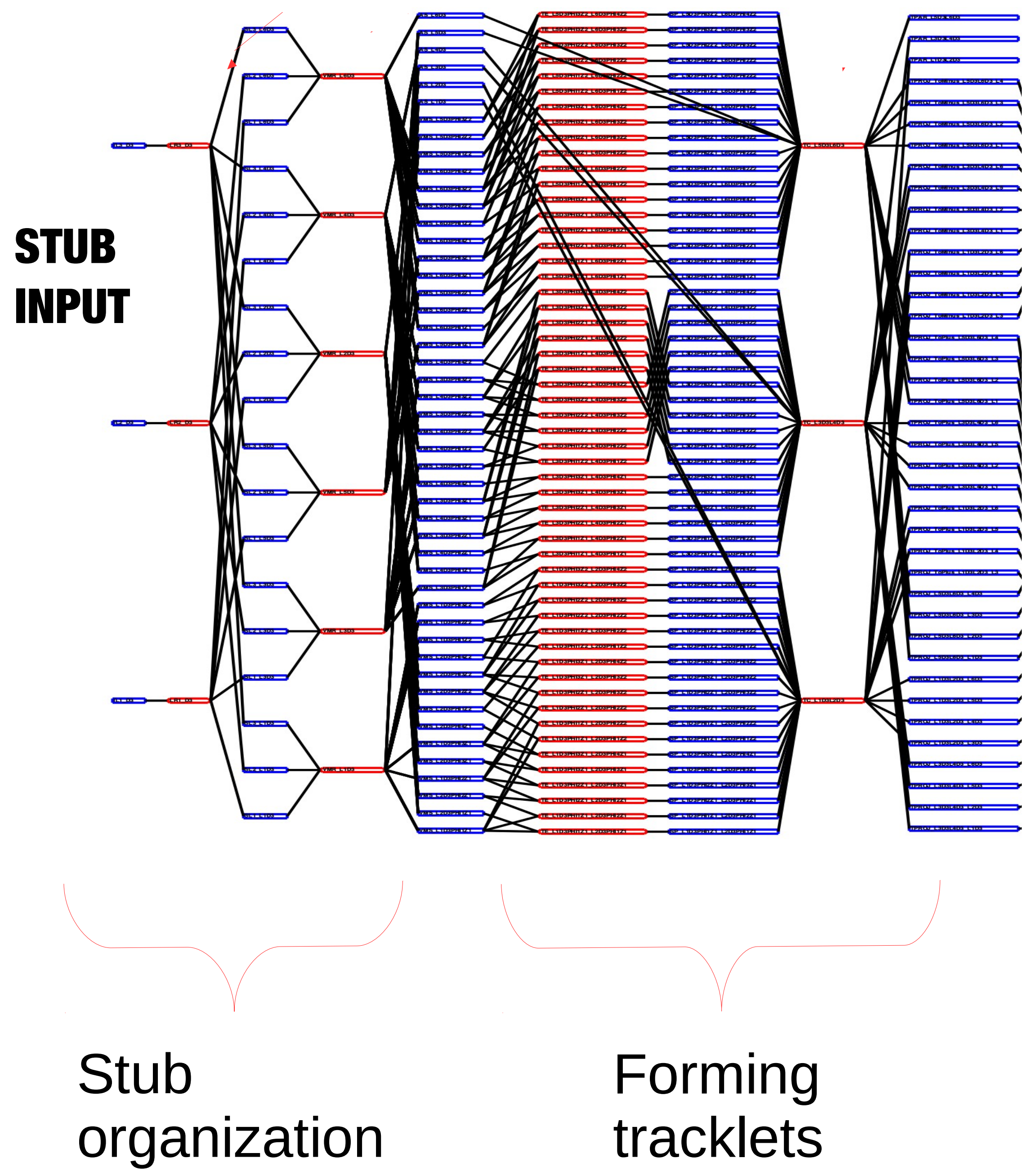


Stub
organization

Project overview

memories
processing modules

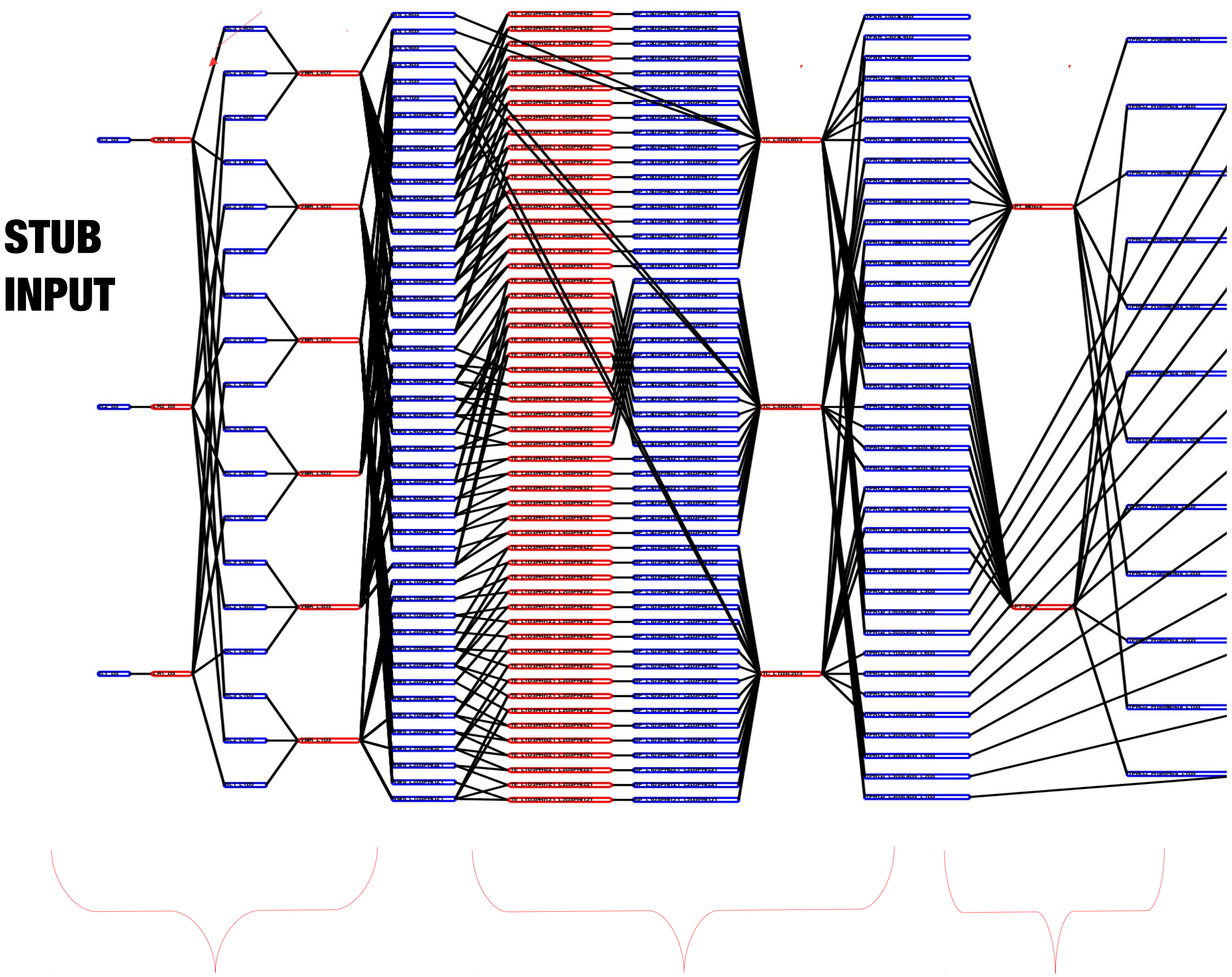
8 processing + 2 transmission steps implements algorithm



Project overview

memories
processing modules

8 processing + 2 transmission steps implements algorithm



Stub organization

Forming tracklets

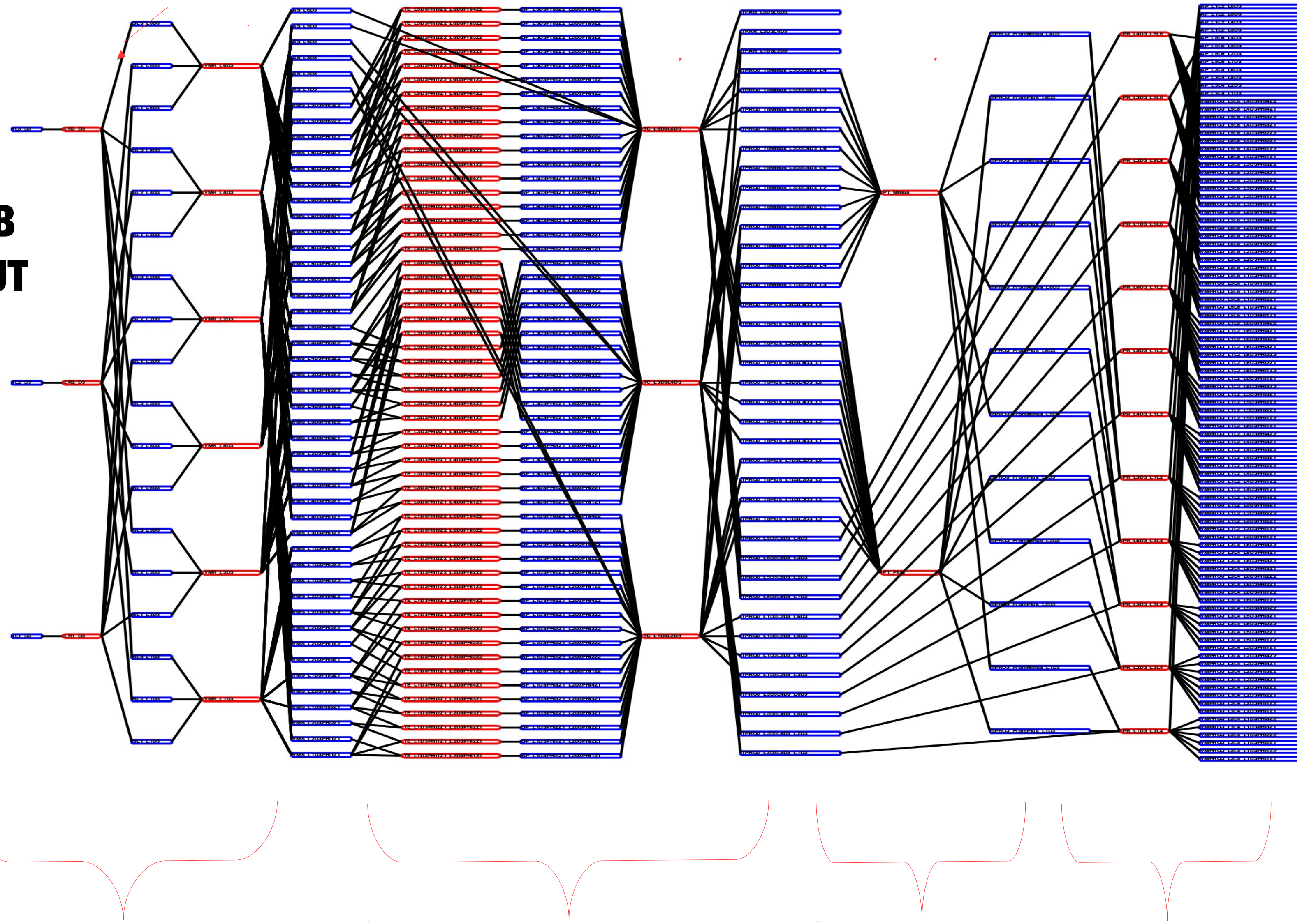
Projection transmission to neighbors

Project overview

memories
processing modules

8 processing + 2 transmission steps implements algorithm

STUB INPUT



Stub organization

Forming tracklets

Projection transmission to neighbors

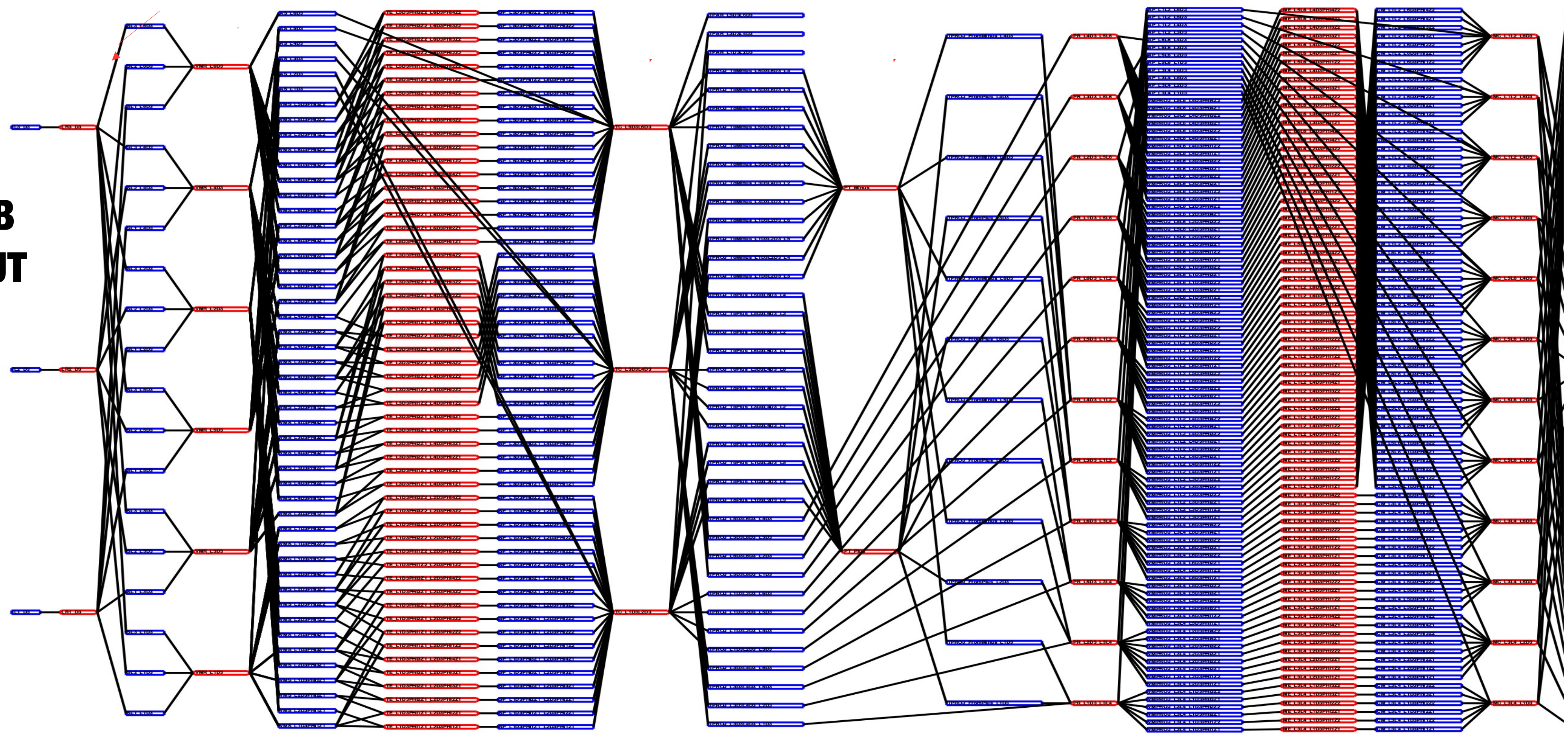
Organize tracklet projections

Project overview

memories
processing modules

8 processing + 2 transmission steps implements algorithm

STUB
INPUT



Stub
organization

Forming
tracklets

Projection
transmission
to neighbors

Organize
tracklet
projections

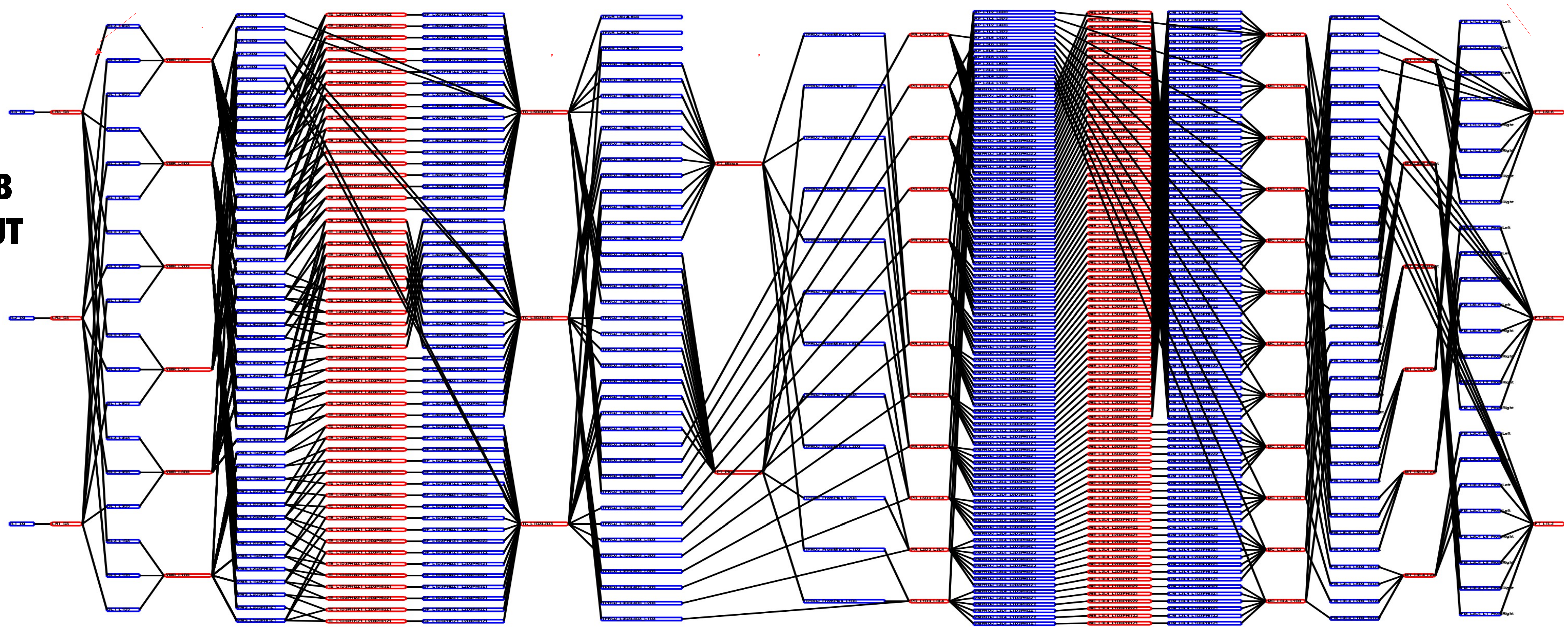
Match tracklet
projections to
stubs

Project overview

memories
processing modules

8 processing + 2 transmission steps implements algorithm

STUB
INPUT



Stub organization

Forming tracklets

Projection transmission to neighbors

Organize tracklet projections

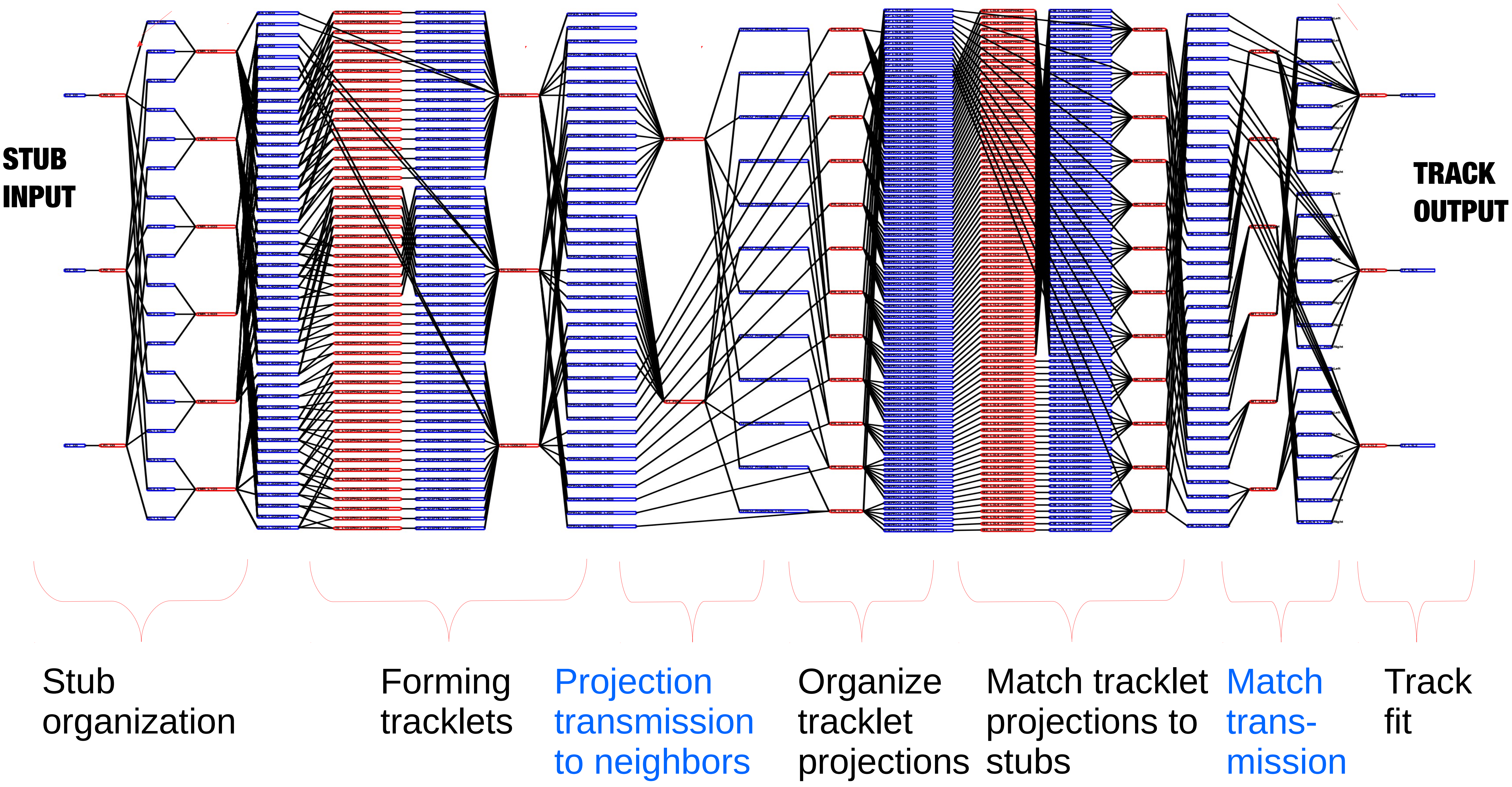
Match tracklet projections to stubs

Match trans- mission

Project overview

memories
processing modules

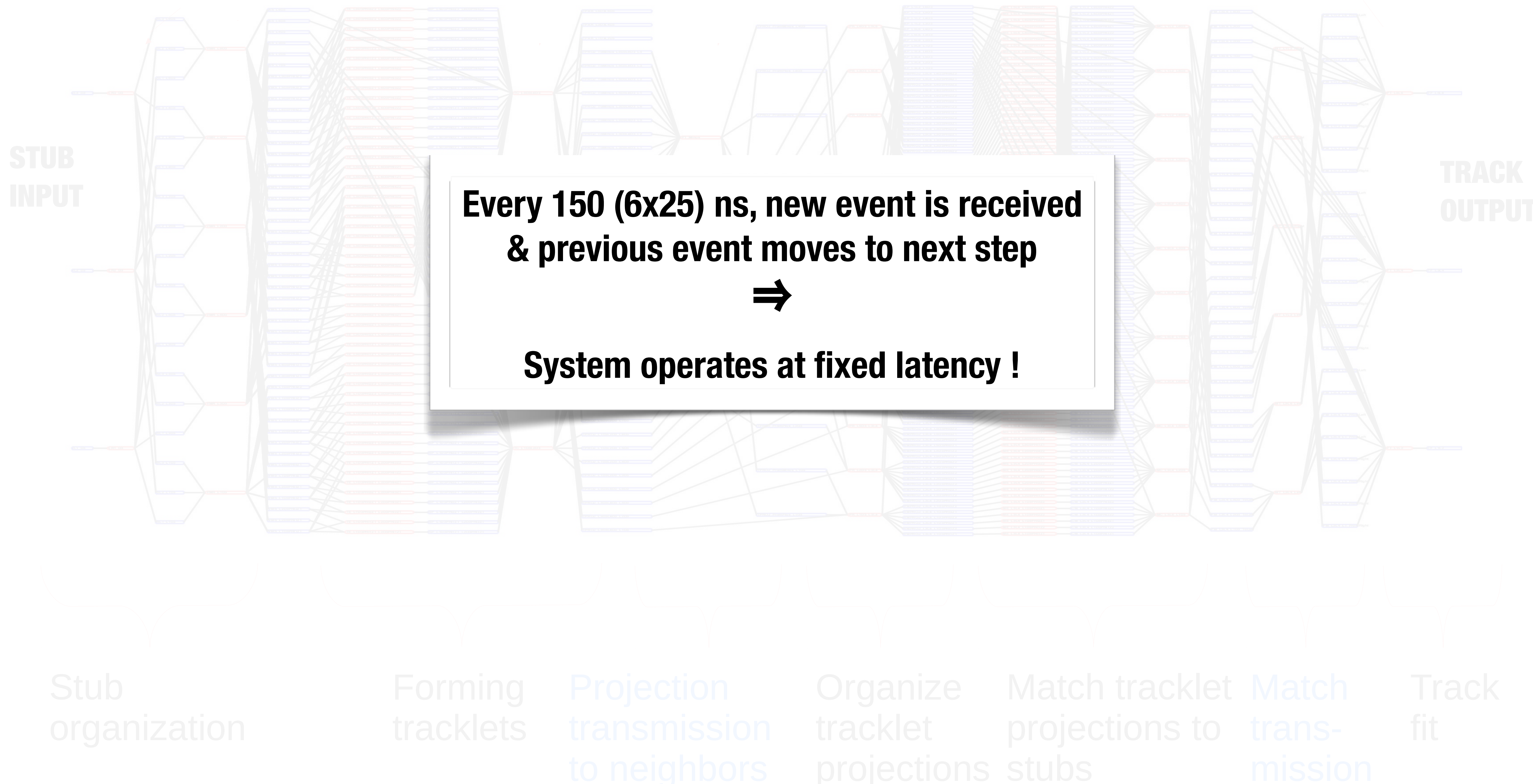
8 processing + 2 transmission steps implements algorithm



Project overview

memories
processing modules

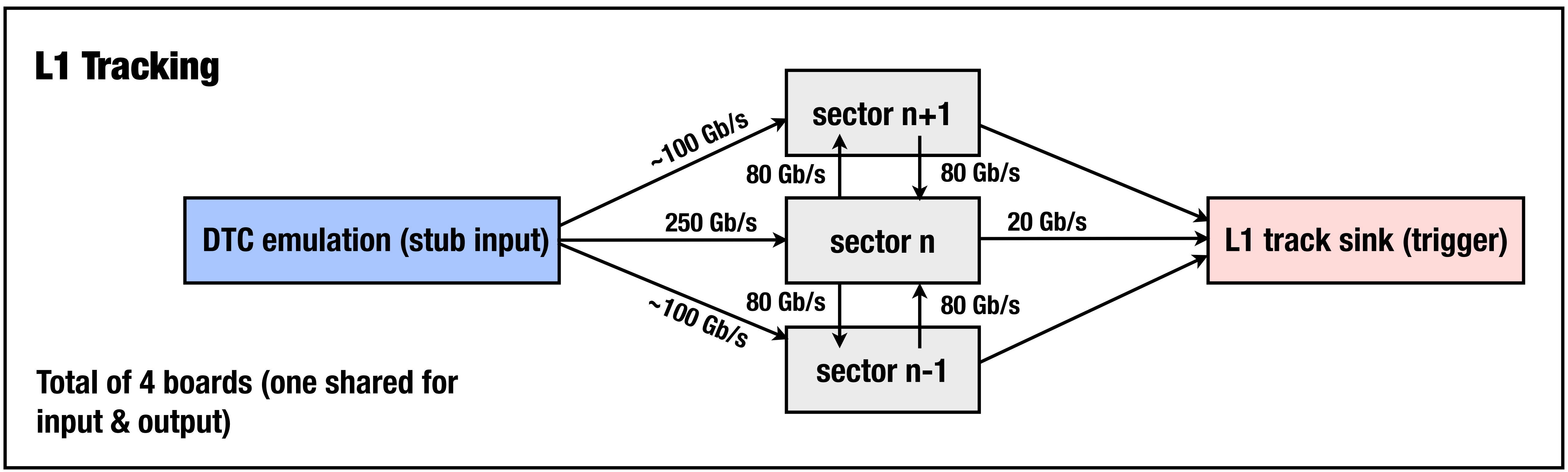
8 processing + 2 transmission steps implements algorithm



Every 150 (6x25) ns, new event is received
& previous event moves to next step
⇒
System operates at fixed latency !

Demonstrator

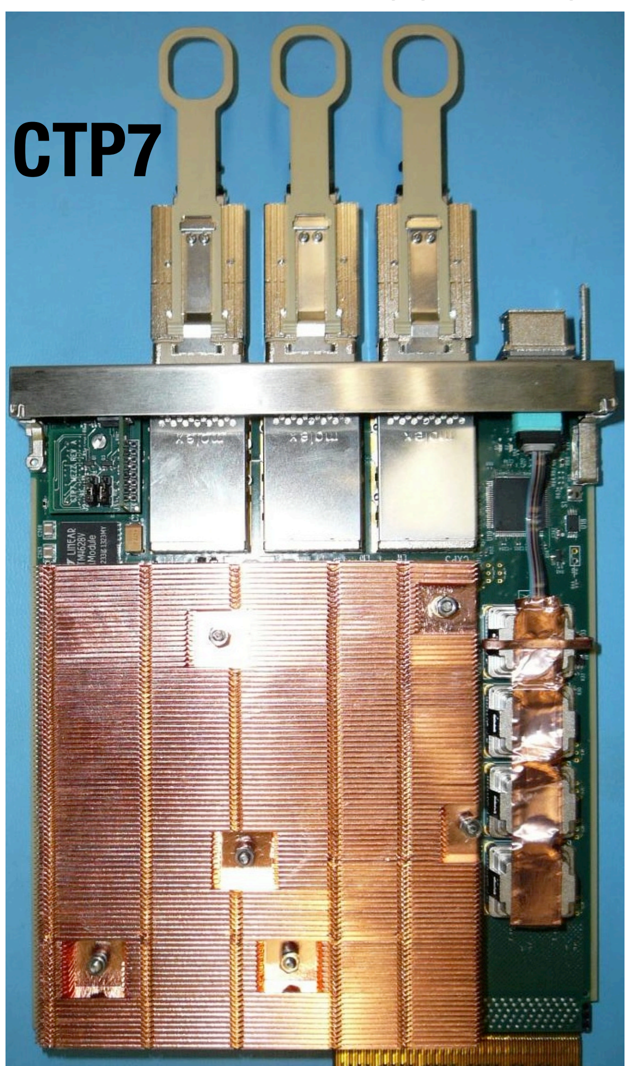
- Demonstrate that full tracking chain meets required performance within available latency
  - For final system process each sector with a single (future) FPGA
  - **2016 demonstrator**
    - *ϕ sector for barrel vs hybrid+disk projects*
- Process many simulated events in sequence



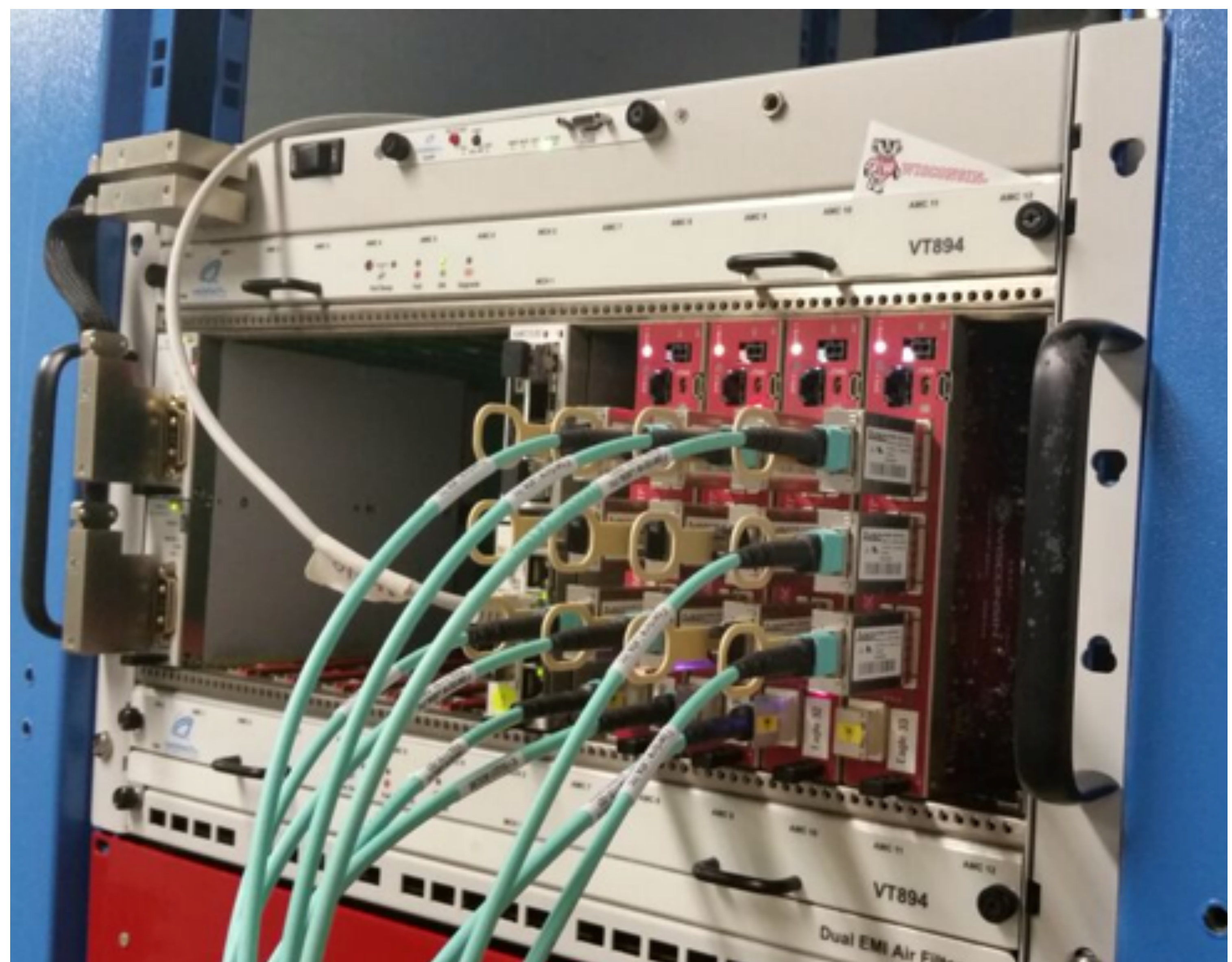
Demonstrator hardware

- Sector boards for demonstrator -- **µTCA boards**
    - Xilinix Virtex-7 FPGA + Zynq chip for outside communication
    - AMC13 card provides central clock distribution

Boards developed by University of Wisconsin for 2016 CMS L1 trigger upgrade



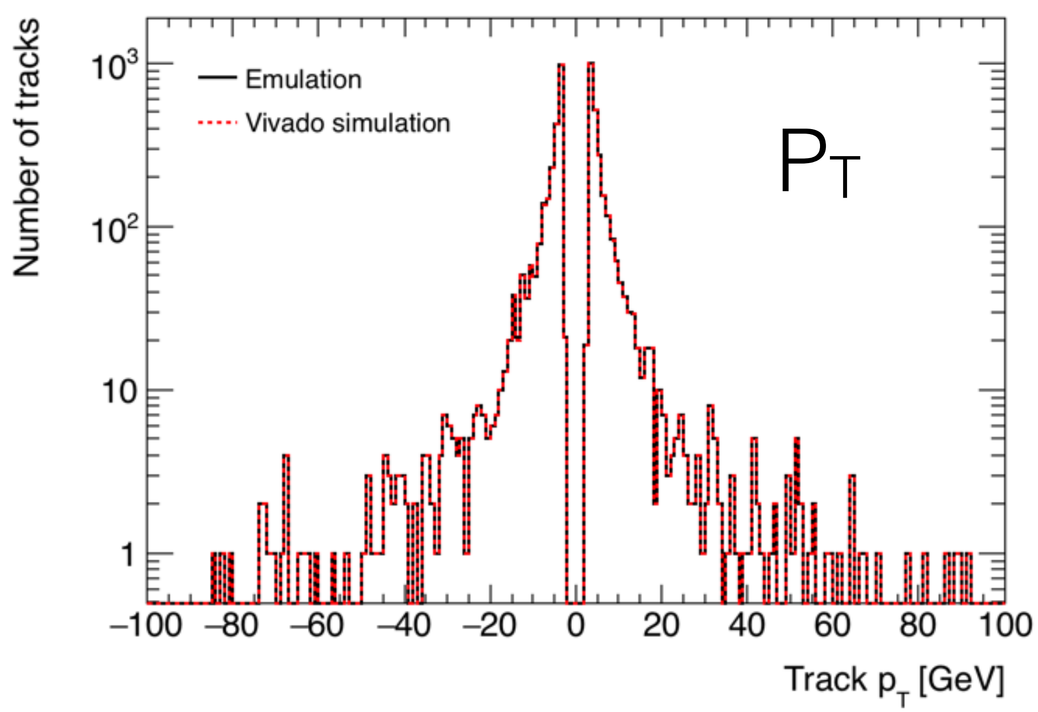
Test stand @ CERN

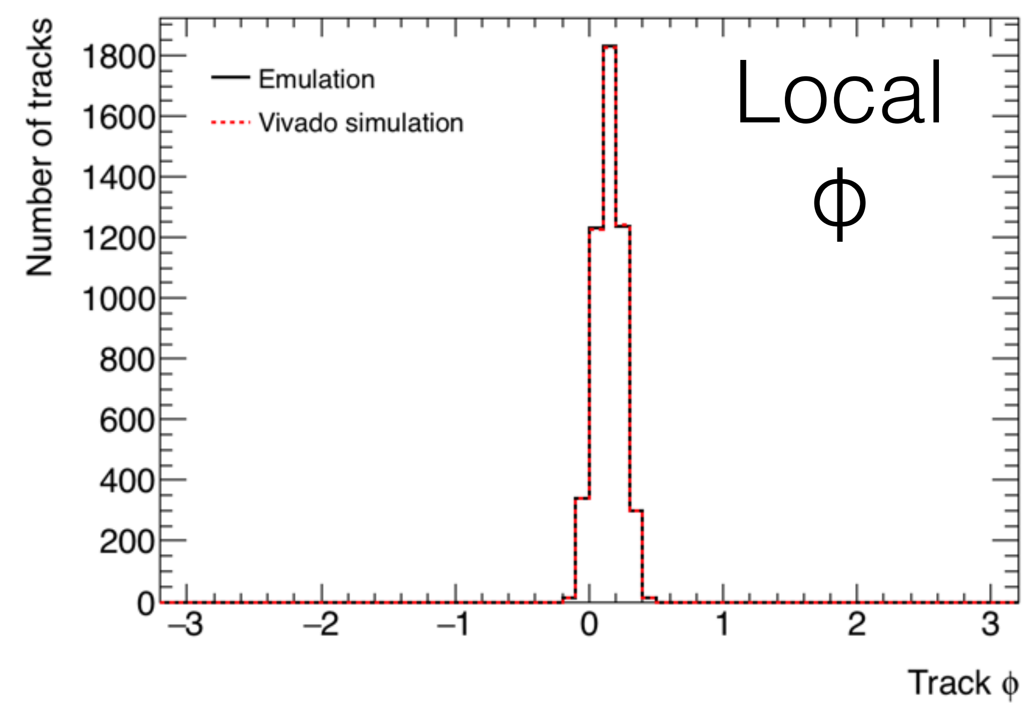


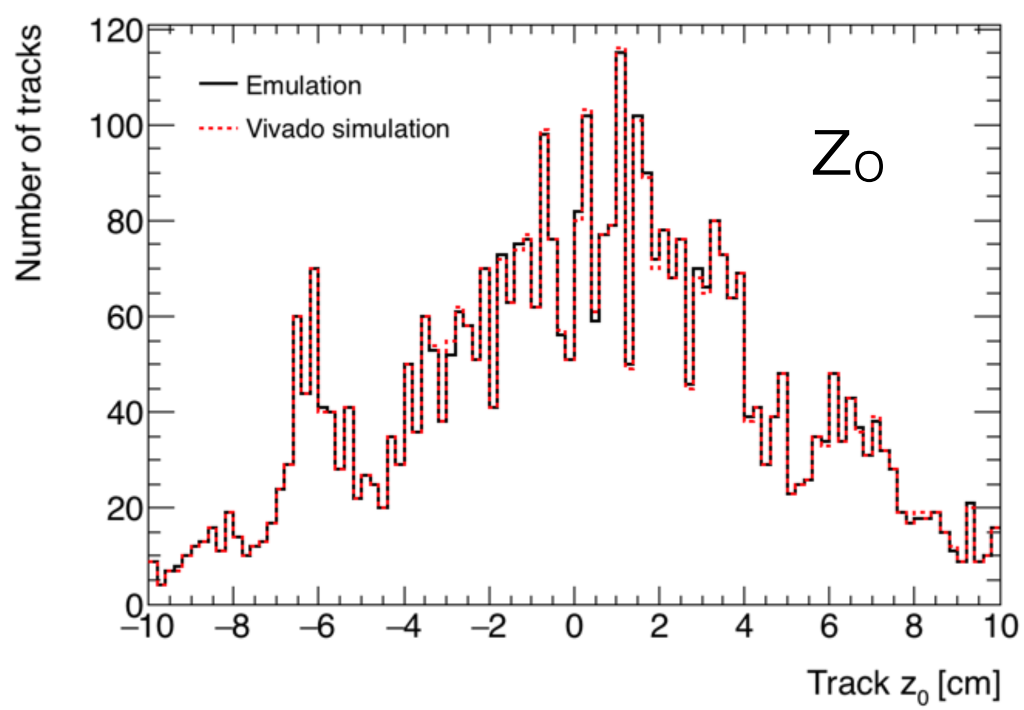
Demonstrator results

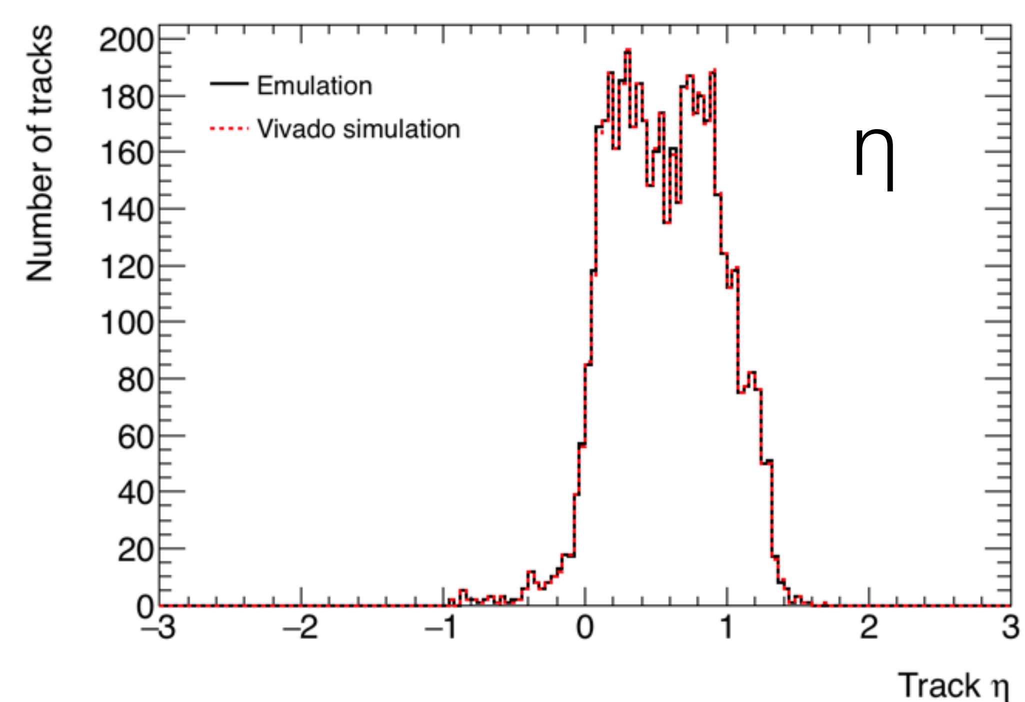
✓ 100% agreement between board output & Vivado firmware simulation

✓ C++ emulation vs firmware implementation:
- single μ: 100% agreement
- ttbar+PU=200: >99% agreement



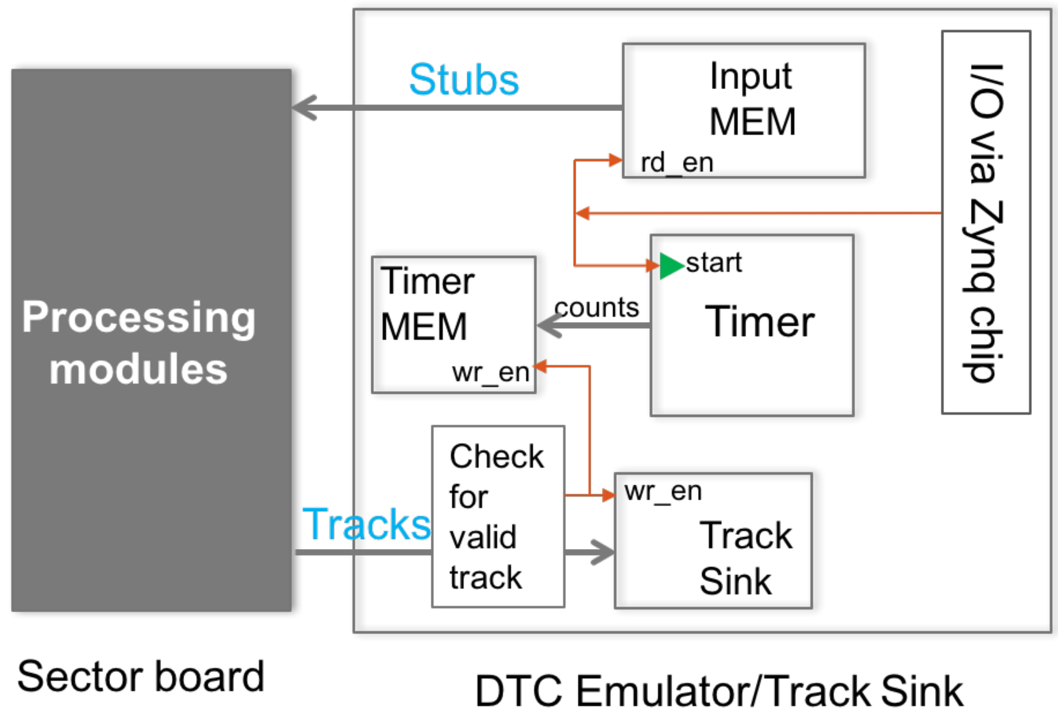






Latency measurement

- A full end-to-end latency measurement done using clock counter
  - 240 MHz clock (same as processing clock)
  - Implemented on input emulator board
- First track out latency: 800 clks = **3.33 μs**
- Well within budget (4μs)!



- Compare with latency model
  - Each processing step has fixed latency => 3.35 μs
  - In good agreement with measured latency (3 clks / 0.38% difference)

Summary

Conclusions

- Incorporating tracking in L1 trigger critical to achieve required rate reductions for CMS at HL-LHC

- Highly challenging -- track triggering on this scale never implemented before
  - Aggressive R&D efforts ongoing
  - System demonstrators in 2016 show feasibility of the systems

- One of these efforts: **tracklet approach**
  - Road search algorithm using commercial FPGAs
  - Manage data volume & combinatorics -- segmentation & parallel processing
  - Feasibility demonstrated!
    - *Implemented on Virtex-7 FPGAs with 3.33 μs latency*
  - **Ongoing work**
    - *Improvements to improve load balancing & reduce latency even further*
    - *Migrate to new tracker geometry*