Track-triggering at CMS for the High-Luminosity LHC
LHC
2808 bunches of $\sim 10^{11}$ protons crossing frequency: 40 MHz operating @ 13 TeV

CMS
Compact Muon Solenoid

~17 miles

silicon trackers (pixel + microstrips)
3.8 T solenoid
muon chambers
preshower
forward calorimeter
crystal electromagnetic calorimeter
hadron calorimeter
Luminosity

CMS Peak Luminosity Per Day, pp

Data included from 2010-03-30 11:22 to 2016-10-27 14:12 UTC

LHC design inst. luminosity exceeded this year!

CMS Integrated Luminosity, pp

Data included from 2010-03-30 11:22 to 2016-10-27 14:12 UTC

\[ N = \sigma \times \int L \, dt \]

number of events

cross section

instantaneous luminosity

integrated luminosity
LHC → High Luminosity LHC

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

LHC / HL-LHC Plan

TODAY
~40 fb^{-1} @ 13 \text{ TeV}

Integrated luminosity:
~3000 fb^{-1}
Motivation

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

![Graph showing Higgs boson couplings](image)

**Figure 1.10**: Observed and projected precision on Higgs boson couplings as a function of boson or fermion masses. Compared to a projection for the measurement of Higgs boson couplings in a dataset of 3 ab⁻¹, percent-level precision can be reached for most coupling measurements. The coupling to the second-generation fermions will be probed for the first time measuring the Higgs boson decay to two muons. Measurements of the Higgs boson couplings, probing of its tensor structure, and the search for Higgs boson pair production. The dominant Higgs boson pair production mode at LHC is through vector boson scattering processes. These measurements could also be sensitive to new physics beyond the Standard Model (BSM).
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

![Summary of CMS SUSY Projections with SMS](image-url)
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

![CMS Average Pileup, pp, 2016, $\sqrt{s} = 13$ TeV](image)

**CMS Average Pileup, pp, 2016, $\sqrt{s} = 13$ TeV**

$<\mu> = 24$

- **PILEUP**: number of overlapping interactions (expected average $\sim 200$)
- **Particularly challenging for trigger system!**

LHC: superb performance
- Thanks to the accelerator teams of CERN, the LHC has exceeded even the most optimistic performance estimates

The estimate prior to the start of the 2016 campaign were to achieve something similar to the previous best (2012)

Luminosity used for the results presented today: $12.9 \, \text{fb}^{-1}$

$\text{Peak Lumi} = 1.2 \times 10^{34} \, \text{Hz/cm}^2$
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

![Diagram of CMS trigger system]

- **L1 trigger decision** in ~2.5 (4) µs for ATLAS (CMS)
- **L1 output**: 100 kHz
- **40 MHz**
- **~4µs “latency”**
- **L1 output**: 100 kHz
- **Switching network**
- **Processor farms**
- **Detectors**
- **Front end pipelines**
- **Readout buffers**

- **HLT**
- **HLT output**: 100 Hz

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

Figure 2.4: Sketch of one quarter of the Tracker Layout. Outer Tracker: blue lines correspond to PS modules, red lines to 2S modules (see text). The Pixel Detector, with forward extension, is shown in green.

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 cm²
- 2x5 cm long strips, 90 µm pitch

**Stubs:** Correlated pairs of clusters, consistent with $\geq 2$ 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

![pT distribution for minbias tracks](image)

2.7% above 2 GeV
0.8% above 3 GeV
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 ~4 µs**

- 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

- Parallelization
  - Divide tracker in segments in $\phi / z$
  - 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

Time multiplexed architecture

- N independent processing boards
- Track Finder
- regional readout boards

Associative memories pattern matching

- Roads

Tracklet method

- fitted track
- stub pair

Anders Ryd, Cornell University
Emulation of Barrel+Endcap L1 Tracking
May 7, 2014
Page: 3/27
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
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 $\chi^2$ fit
- Gives final track parameters
  - $p_T$, $\eta$, $\phi_0$, $z_0$
  - 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 $\eta$ for single particles ($e/\mu/\pi$)
- High efficiencies achieved
- Minimal impact from truncation
Tracking performance

- $\sigma(z_0) \sim 1\,\text{mm}$ for wide range of $\eta$ thanks to PS modules
- $\sigma(p_T)/p_T \sim 1\%$ at central $\eta$ 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 $\phi$ 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 $\phi$ 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 $\phi$ 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 $\phi$ 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

“Virtual modules”

Layer 1: $3\phi \times 8z$ (2 shown) subregions
Layer 2: $4\phi \times 8z$ (2 shown) subregions

$\rightarrow$ 768 pairs, but only 96 can form valid tracklets
Project overview

8 processing + 2 transmission steps implements algorithm
Project overview

8 processing + 2 transmission steps implements algorithm

- Stub organization
- Forming tracklets
Project overview

8 processing + 2 transmission steps implements algorithm

Submit organization

Forming tracklets

Match tracklet projections to stubs

Match transmission to neighbors

Projection transmission to neighbors

8 processing + 2 transmission steps implements algorithm
Project overview

8 processing + 2 transmission steps implements algorithm

- Stub organization
- Forming tracklets
- Projection transmission to neighbors
- Organize tracklet projections
Project overview

Overview of Project (¼ of Barrel)

Eight processing steps + two transmission (red) implements the algorithm

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

8 processing + 2 transmission steps implements algorithm

Stub organization
Forming tracklets
Projection transmission to neighbors
Organize tracklet projections
Match tracklet projections to stubs
Match transmission
Project overview

8 processing + 2 transmission steps implements algorithm
Project overview

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
    - \( \phi \) sector for barrel vs hybrid+disk projects

- Process many simulated events in sequence

---

**L1 Tracking**

- DTC emulation (stub input)
- Total of 4 boards (one shared for input & output)
Demonstrator hardware

- Sector boards for demonstrator -- µTCA boards
  - Xilinx 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
Demonstrator results

✓ 100% agreement between board output & Vivado firmware simulation

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

![Graphs showing comparison between Emulation and Vivado simulation for different variables: $p_T$, $\phi$, $z_0$, and $\eta$.](image-url)
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 = \textbf{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*