L1 Track Finding for the CMS HL-LHC Upgrade

Kristian Hahn – Northwestern University
on behalf of the CMS Track Trigger Working Group

Vertex 2017 – Las Caldas, Asturias, Spain
September 15, 2017
High luminosity is crucial for future CMS physics goals

- Instantaneous: $5-8\times10^{34}$ cm$^2$s$^{-1}$
- Integrated: 3000 fb$^{-1}$ (10 yr)

Achieve precision knowledge of Higgs & rare-SM processes

- Improvement in couplings
- Di-Higgs becomes accessible

Extend reach of searches for new physics

- SUSY, Dark Matter, Long-lived ...
An incredibly challenging physics environment

- 140-200 <pile-up> interactions
- In and out of time ...

Confusing the L1 trigger

- Additional energy in the calorimeters
  - Impacts efficiency EG triggers
  - EG rate @ 25 GeV > 100 kHz
  - Isolation becomes less effective

- Muon pT resolution degrades
  - Causing rate to plateau as a function of threshold

- Total L1 rate > 1.5 MHz

Must be overcome to maintain sensitivity to “low-mass” physics

- Thresholds as loose as possible

Need a fundamentally new handle in the L1 Trigger : tracking!

- Calorimeter & muon used exclusively in the present CMS L1 trigger
- The HL-LHC upgrade introduces L1 tracking to temper impact of pileup

Tracks combined with calo & muon info in the “Correlator”

- Improves trigger object reconstruction, provides access to new quantities
CMS L1 Tracking

Tracking lowers rates & thresholds across the HL-LHC L1 trigger menu!
L1 tracking enabled by new Outer Tracker design

- 6 Barrel + 5 Endcap layers
  - Strip-strip outer modules
  - Pixel-strip inner modules
- Pixel not part of L1Tk
- Trigger info sent off detector via VTRx+ & LpGBT
  - ~80% of OT data is for trigger!

pT (bend) measurements at the detector front-end

- Double-sided modules correlate pairs of hits into “stubs”
- Send only data that correspond to high momentum tracks
  - Enable L1 tracking for pT > 2 GeV
- Self-seeding tracking: immediate rate reduction by a factor of >10!

More module details in talk from J. Luetic, Tues. session: https://indico.cern.ch/event/627245/contributions/2675779
Data from front-ends received at Data, Trigger Control Board (DTC)

- 50-70 modules per DTC
- FPGA emulation of GBT links to the front-ends
- Approximately 600 Gbps outputs to L1Tk system

**L1Tk system builds and delivers track primitive to the L1 trigger**

- Tracks combined with calo, muon, etc in the L1 “correlator”
- Presently anticipating $O(130)$ links, TMUXed
Formidable technical challenges!

- Total data rates $> 50$-$100$ Tbps
- Occupancy & combinatorics: $O(10^4 \text{ hits/BX})$ @ $200 <PU>$
- Latency: just $4+1 \mu s$ for tracking

No precedent for real-time tracking trigger at this scale

CMS launched 3 R&D programs to confront the challenge

- 2 FPGA-only: tracklet, TMTT
- 1 ASIC-assisted: AM+FPGA
- Primarily differing in how pattern recognition is performed
- Each performed technical demonstrations in 2016 to establish feasibility (w/ present-day technology)
Time Multiplexed Track Trigger

- NB: all three approaches are TMUXed ...

Detector segmented into octants, with ~32 DTCs / octant

Processing octants shifted by ½ DTC octant

- Data from at most 2 DTC octants needed for tracks >2 GeV
- DTCs duplicate data near processing octant boundaries
- Within each processing octant, \( N_{\text{TMUX}} = 36 \) Track Finding Processors
- Pattern recognition via Hough Transform, fitting via Kalman Filter
HT: pattern recognition in the r-φ plane

- Circular trajectories described by two parameters: $q/p_T$, $\phi_{65}$
- Create 2D histogram of $q/p_T$, $\phi_{65}$
- For each stub, fill those bins for which a track with corresponding $q/p_T$, $\phi_{65}$ passes through the stub
  - Individual stubs mapped to lines
  - Where lines cross: a track!

Track fitting via Kalman Filter

- Equivalent to offline tracking algorithm
  - Implemented via Java/HDL hybrid Max-J
- Iterative, data-dependent ...
  - Impose fixed latency cutoff
  - Output states with lowest $\chi^2 / \text{DOF}$
Demo uses Imperial MP7 (μTCA)

- Virtex 7 690T, 0.94 Tbps FP I/O
- 6x12 channel miniPOD, 11.3 Gbps / link
- Expertise from CMS Phase-1 Calo trigger

Implement 1 geometric / time slice of the full system

- Mock-up future, larger FPGAs w/ >1 daisy-chained boards

http://www.hep.ph.ic.ac.uk/mp7/
The Tracklet Approach

Stubs in pair of layers form a seed (tracklet)
- Seeds in layer pairs 1+2, 3+4, 5+6
- Use tracklet + IP to project into other layers
- Look for stubs near the projected track
- Refit final track parameters using all stubs

To handle combinatorics, split detector into 28 phi sectors (spanning $\eta$)
- 1 sector primarily processed by 1 board
- 2 GeV tracks will be in $\leq 2$ $\varphi$-sectors
- Neighboring board I/O required

Timing model
- Factor of 6 TMUX (150 ns)
- Fixed-time processing model, truncate if a time allocated for a given step is exceeded
Tracklet : Implementation

Tracklet diagram for 1/4 barrel in a $\phi$ sector.

- Firmware and emulation generated together with SW configuration tool!
- Memories
  - Processing modules
- Stub organization
- Forming tracklets
- Projection transmission to neighbors
- Organize tracklet projections
- Match tracklet projections to stubs
- Match transmission
- Track fit
Using Wisconsin CTP7 (μTCA) boards for the demonstration

- Used in the CMS Phase-1 Calo trigger
- Virtex-7 690 FPGA
- Zynq SoC with dual Cortex-A9 ARM
- GTH: 80 RX & 61 TX

3 CTP7 boards are used as sector boards

- 1 central sector
- +/- phi sectors for neighbor communication
- 1 CTP7 handles sending input stubs and receiving output tracks
- Demonstrates 1 TMUX period

AMC13 card is used for central clock distribution
Parallelize computationally expensive PR tasks with AMs

- CAM cells + majority logic
  - CAMs match hits in layers to those stored
  - Hits=“SuperStrips”, coarse groupings of stubs
  - ML associates hits to across layers, matching patterns
  - “Roads” = coarse, multi-layer hit patterns
- Completes as soon as all hits arrive

Simplifies downstream track fitting!
- Avoids power-law dependence of execution time on occupancy
- Already employed to this end in L2 applications (SVT, FTK)

Finer grained PR and param estimation from FPGA track fitting
Split detector into trigger “towers”, nominally 48 (6ξ x 8φ)

- 3 types of towers (barrel, endcap, hybrid)

Utilize modern, high-speed ATCA backplane

- Nominally processing via 1 shelf per tower, 1 mezzanine per BX
- Data delivery to a single processing board via backplane
Trigger tower demo with 2 ATCA shelves

- One to emulate front-end modules
  - 400 modules/fibers, no requirements on the DTC (pass through)
- Other for Pattern Recognition Boards
  - Receive, format and deliver data to Pattern Recognition Mezzanines (PRM)
  - X10 (x2) TMUX via BP (PRM)

FNAL Pulsar2b used as DS & PRB

- Virtex 7 690T
- 80 GTH @ 10 Gbps
  - 40 to RTM (DS ↔ PRB)
  - 28 to BP (PRB ↔ PRB)
  - 12 to FMC (PRB ↔ PRM)

FNAL & INFN Kintex Ultrascale PRMs

- AM emulation on a 2nd US FPGA
  - Realistic latency, but limited pattern density
- x12 AM06, PRM06 mezzanine
  - Realistic pattern density, but high latency
A collective vision of the goals of the demonstrations developed within the CMS L1Tk group

- Common analysis framework
- Performance Metrics
  - Basic tracking
  - “L1-oriented”
- Hardware demonstration (latency)
- System extrapolation & scaling

### Demonstration Specifications

**Truth**

| TrackingParticles in the low and high-$p_T$ regimes with ≥ 4 stubs in ≥ 4 layers with $|\eta^{TP}| \leq 2.4$ with $|z^{TP}| \leq 30$ cm with $p_T > 3$ GeV with $d_{xy} < 1$ cm |

**GoodTracks**

| TTTtracks corresponding to fit tracks from simulation with $|\eta^{fit}| \leq 2.5$ with $|z^{fit}| \leq 30$ cm that satisfy all other approach-specific selections/constraints (e.g. $\chi^2$) |

**MatchedTracks**

| GoodTracks output from simulation matched to Truth using the trackToTrackingParticleMap |

### Specifications for the CMS L1 Tracking Demonstration

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Versus</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon$</td>
<td>$p_T$ $\eta$ nPU</td>
<td>single muon/electron + 0,140,200 PU $t\bar{t} + 0, 140, 200$ PU</td>
</tr>
<tr>
<td>$\sigma_p$</td>
<td>$p_T$ $\eta$ nPU</td>
<td>single muon/electron + 0,140,200 PU $t\bar{t} + 0,140,200$ PU</td>
</tr>
<tr>
<td>$f$</td>
<td>nPU</td>
<td>single muon + 140,200 PU $t\bar{t} + 140,200$ PU</td>
</tr>
</tbody>
</table>

Compiled by:

Kristian Hahn, Peter Wittich

*Northwestern University, Evanston, Illinois, USA
Cornell University, Ithaca, New York, USA
E-mail: kristian.hahn@cern.ch, peter.wittich@cern.ch
Excellent (& comparable) performance from each of the demonstrators!

- Latency < 4 us with ~95% efficiency
- $p_T (z_0)$ resolution of ~2% (~1 mm)
- Significant (~x100) rate reduction
Demonstration results detailed in Tracker TDR

2 Overview of the Phase-2 Tracker Upgrade
  2.3.1 Tracker input to the L1 trigger

3 The Outer Tracker
  3.5 L1 track finder
    3.5.1 L1 track finder approaches
    3.5.2 The way forward

6 Expected Performance
  6.3 Tracking performance
    6.3.1 Level-1 tracking performance

9 The Outer Tracker: Additional Information
  9.4 L1 track finder
    9.4.1 Associative memory plus FPGA approach
    9.4.2 FPGA-based Hough transform approach
    9.4.3 FPGA-based tracklet approach

12 Additional Information on Material Budget, Local Reconstruction, and Tracking
  12.3 Tracking performance
    12.3.1 Level-1 tracking performance

July '17 : CMS-TDR-17-001, CERN-LHCC-2017-009
Outcomes

CMS has converged to an all-FPGA baseline for L1Tk
- AM+FPGA R&D will continue as a backup option

Optical “flow-forward” ATCA architecture for the DTC & TF
- Stub routing / duplication performed in the DTC
  - Provides flexibility in terms of TF algorithms
- System based on high-speed (eg: 16, 25) Gbps links
- Algorithm agnostic

Continued algorithm development
- Firmware-based → schedule permits further R&D
- Possibly combining aspects of the three demonstrated approaches
- Possibly expanding to include additional L1 primitives
  - Eg: vertex information
A sequence of system development milestones must now be met

- Ensuring coherent prototyping / integration together w/ the Phase-2 OT

Near-term : finalize specifications of the DTC & TF systems

- Additional R&D needed to hash out details, collapse remaining options
  - Means of slow control, link speed(s), etc.
  - Share/reuse as much as possible between the DTC & TF
Current R&D

Development of modular ATCA carrier w/ PCIe interconnect
Ultrascale Kintex development card
Exploring COM Express for board control

Extensive use of COTS - Commercial Off The Shelf

Service Card – ATCA System Builder

PCIe Card (e.g. MP-Ultra)
Could be any development card
Xilinx or Altera
Design is manufacturer agnostic

COM Express (COTS)
High Level Control

IPMC from CERN (COTS)
Low Level Control

25Gb/s Test Structure w/wo Retimers
(PCB & Backplane)
Current R&D

YUGE: \( \mu \)TCA Ultrascale (Kintex/Virtex) card for 25 Gbps evaluation
Exploring Zynq for board control
Summary

L1 Tracking is the key-enabler of CMS HL-LHC physics goals

- A truly novel system, with clear physics impact and potential for exploiting new physics signatures

Three CMS R&D programs have successfully and independently demonstrated its feasibility

An all-FPGA TF & optical flow-forward architecture now serve as the reference platform

- R&D underway to finalize a comprehensive back-end system design