

Particle Flow reconstruction in the CMS Level-1 Trigger for the HL-LHC

Giovanni Petrucciani^{1,*} on behalf of the CMS Trigger Project

¹CERN

Abstract. With the planned addition of the tracking information in the Level-1 trigger in CMS for the High-Luminosity Large Hadron Collider (HL-LHC), the algorithms for the Level-1 trigger can be completely reconceptualized. Following the example for offline reconstruction in CMS to use complementary subsystem information and mitigate pileup, we explore the feasibility of using Particle Flow-like and pileup-per-particle identification techniques at the hardware trigger level. We present the challenges of adapting these algorithm to the timing and resource constraints of the Level-1 trigger, the first prototype implementations, and the expected performance on physics object reconstruction.

1 Introduction

The Particle Flow (PF) approach aims to reconstruct and identify individually all particles produced in CMS [1] by combining information from all subdetectors [2].

Since LHC Run 1, PF has been widely used in the CMS offline and High Level Trigger (HLT) event reconstruction. Three areas where benefits from PF were most evident are the hadronic jet and missing transverse momentum (p_T^{miss}) performance, the reconstruction of hadronically-decaying tau leptons, and its use as input to pile-up mitigation strategies. Benefits have been especially important at low transverse momenta (p_T) relevant for many physics channels of interest at the High-Luminosity LHC (HL-LHC), e.g. top quark physics, ttH, compressed supersymmetry.

Two necessary ingredients for PF are an efficient reconstruction of charged particles in the inner tracker, and a fine granularity calorimetry to resolve the contributions from neighbouring particles. Neither is available in the current Level-1 trigger system [3, 4], limited to coarse granularity calorimetric information ($\Delta\eta \times \Delta\phi \sim 0.1 \times 0.1$)¹, and reconstruction in the muon spectrometer.

The CMS detector upgrade for HL-LHC [5] will bring key improvements that will make PF possible at the Level-1 Trigger for the first time:

- A new inner tracker supporting the readout of stubs, pairs of correlated hits in consecutive layers from high p_T particles, at the collision rate of 40 MHz, and a backend track finder system to reconstruct all tracks with $p_T > 2$ GeV, $|\eta| < 2.4$ within a latency of 4 μs .
- An updated backend for the electromagnetic barrel calorimeter (EB) supporting a full granularity readout at the Level-1 ($\Delta\eta \times \Delta\phi \sim 0.02 \times 0.02$).

*e-mail: Giovanni.Petrucciani@cern.ch

¹CMS uses a cylindrical coordinate system with the z axis along by the beam line. The pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$, with θ being the polar angle.

- A new endcap calorimeter (EC) detector with fine transverse and longitudinal segmentation.
- A new Level-1 trigger system [6] with improved processing power, a global correlator layer allowing complex algorithms using inputs from all trigger subsystems, and deeper buffers in all subsystems to allow a total latency up to $12.5 \mu\text{s}$.

2 Level-1 Algorithm

The PF algorithm was redesigned from first principles, as the algorithm used offline and at the HLT was not suitable to the Level-1 environment. In particular, the Level-1 PF algorithm must process events at the 40 MHz collision rate, and output the reconstructed particles after a fixed latency below $1 \mu\text{s}$. For comparison, the current HLT PF algorithm is run at a rate of about 20 kHz, and has a latency of the order of 100 ms.

Also, from the implementation point of view, the algorithm has to run on FPGA hardware instead of conventional CPUs. FPGAs provide a very large number of simple processing components that all operate in parallel, but to work efficiently the data flow and tasks have to be statically defined, and thus the logic of the algorithm needs to be simpler and more streamlined.

2.1 Inputs

A first prototype of the Level-1 PF algorithm has been developed using inputs from a preliminary simulation of the trigger primitives and the local trigger algorithms from the various CMS subsystems. Four classes of input objects are used: tracks, electromagnetic clusters, hadronic clusters, and muons.

Tracks are obtained from an early implementation of the Level-1 Track Finder, expected to reflect well the physics performance of the final system even if the algorithm implementation may be different.

Electromagnetic clusters are built from the EB and the electromagnetic layers of the EC, with fine granularity to allow the reconstruction and identification of non-isolated photons and electrons, e.g. from π^0 decays in hadronic jets and photon conversions in the detector material.

Hadronic clusters are built from the EB and the hadronic barrel calorimeter (HB) and from the full EC, with coarser granularity and larger size to ensure that showers from individual hadrons are contained in a single cluster. A simple clustering algorithm is used, not yet relying on information from the longitudinal shower profile.

Muon tracks in the muon spectrometer are derived from the current Phase 1 muon trigger emulation, not yet including the prospected Phase 2 upgrades in the algorithm and detector.

In the forward region, $|\eta| > 2.4$, tracks and muons are not available at Level-1, and a simpler algorithm is used based only on calorimetric information from the EC and the forward hadron calorimeter (HF).

2.2 Algorithm

The data flow in the Level-1 PF algorithm is shown in Figure 1.

1. For each input muon the best-matching track in the inner tracker by $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$ and p_T , is tagged as a muon and excluded from further processing in the PF algorithm.

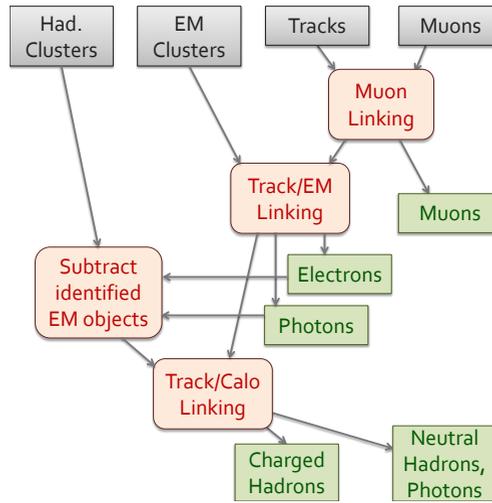


Figure 1. Data flow in the current prototype Level-1 PF algorithm. Gray boxes indicate the inputs from the different subsystems, tan boxes are the individual building blocks of the algorithm, and green boxes correspond to the output streams of PF objects (some of which also serve as input for further PF processing steps).

- Each track is linked to the closest electromagnetic cluster (EM) if any is found within $\Delta R < 0.04$, and the association is used to compute for all EM clusters the $\sum p_T$ of all associated tracks.

Clusters with no associated tracks are tagged as photons, while if $p_T^{\text{cluster}} \geq \sum p_T^{\text{track}}$ within uncertainties, tracks are tagged as electrons, and any significant p_T excess is promoted as photon.

If instead $\sum p_T^{\text{track}}$ significantly exceeds the cluster p_T , the cluster is discarded as it is likely originating from a hadronic shower starting in the electromagnetic calorimeter.

- Surviving EM clusters are linked to the closest hadronic cluster, and the energy of the hadronic clusters is updated subtracting that of all linked EM clusters. Hadronic clusters whose energy is reduced below about 10% of their original energy after this subtraction are discarded entirely.
- Tracks are linked to the closest hadronic cluster in ΔR^2 and p_T , with an additional requirement $p_T^{\text{calo}} \geq p_T^{\text{track}} - 2\sigma(p_T^{\text{calo}})$ in the linking, to reject fake high p_T tracks. Tracks are promoted to charged hadrons if linked to a cluster, or if $p_T^{\text{track}} < 10$ (20) GeV if the track passes loose (tight) quality criteria.

Conversely, calorimetric clusters can give rise to a neutral hadron or photon if p_T^{calo} significantly exceeds $\sum p_T^{\text{track}}$, depending on whether the excess is mainly electromagnetic or hadronic; otherwise, the cluster is discarded.

- Calorimeter clusters outside the $|\eta|$ coverage of the Level-1 track finder are also promoted to neutral hadrons or photons, depending on the fraction of energy in the electromagnetic calorimeter.

2.3 Pileup rejection

PF reconstructs all the particles produced in the collision, also from pileup interactions. To select only the contributions from the the leading interaction, once the primary vertex (PV) is identified the Pile Up Per Particle Identification [7] (PUPPI) algorithm is used.

Charged hadrons are selected if the track coordinate along the beam line is compatible with the PV.

For all neutral hadrons, a probability (w) for the particle to belong to the PV is estimated as function of $\sum p_T^2/\Delta R^2$, computed from all selected charged hadrons within $\Delta R < 0.4$, and the particle momenta is scaled by that probability. Particles with probability w below 1% or whose scaled p_T is below a threshold dependent on the pileup level are discarded entirely.

Outside the tracker coverage, the summation is performed using all particles, and higher thresholds are used.

3 Firmware implementation

Xilinx Vivado High Level Synthesis [8] (HLS) is used to compile the PF algorithm, coded in a subset of C++ with annotations, into a reusable firmware block (IP core).

To best profit from the FPGA capabilities, all the computation is implemented using integers instead of floating-point, and the mathematics is kept simple: mainly additions and comparisons, few multiplications, only one division (implemented as a lookup table), and no square roots or special functions.

To increase the throughput the entire algorithm is pipelined to accept new inputs every 1 or 2 clock cycles. All combinatorial loops, e.g. on object pairs in the linking, are unrolled to compute all values in parallel, which also reduces the latency.

The PF relies only on local information, so different detector regions can be processed independently. The complexity and FPGA resource use depend on the maximum allowed number of input objects, determined by the size of the detector region.

Preliminary estimates show that 4 regions of size $\Delta\eta \times \Delta\phi \sim 0.62^2$ with 25 tracks and 20 clusters can be processed every 25 ns on a Xilinx VU9P FPGA with $\sim 40\%$ resource usage and $\sim 0.5 \mu\text{s}$ latency, in line with the requirements.

3.1 Test setups

The PF IP core has been successfully tested on current and early prototype trigger boards based on Virtex-7 FPGAs (for reduced object multiplicities), and on the VU9P FPGA in the VCU118 development kit and on Amazon AWS.

VHDL or C++ code is used to interface the core with the board infrastructure using IPbus [9] or AXI-PCIe to inject input patterns from the CMS detector simulation into the core, and the output is checked for bitwise identity with the expectations from HLS.

4 Physics performance

The performance of the Level-1 PF + PUPPI algorithm for physics objects has been evaluated using simulated events with the CMS Phase 2 detector and HL-LHC pile-up conditions.

A benchmark Level-1 trigger for hadronic signatures is defined based on H_T , the scalar sum of the calibrated p_T of all jets with $p_T > 30 \text{ GeV}$ and $|\eta| < 2.4$, built from calorimeters clusters, tracks associated to the PV, or PF+PUPPI particles. For a fixed Level-1 accept rate of 20 kHz at 140 pile-up, PF+PUPPI inputs yield a lower threshold and a sharper turn-on curve of the efficiency as function of the true p_T in simulated $t\bar{t}$ events (Figure 2, left).

For missing-energy signatures, the p_T^{miss} computed from PF+PUPPI particles is likewise found to yield a better trade-off between Level-1 rate and signal efficiency (Figure 2, right). As a reference signal for the study, $t\bar{t}$ events with true $p_T^{\text{miss}} > 100$ GeV from neutrinos were used.

The applications of PF reconstruction to other physics objects and the integration in the Level-1 trigger menu studies for Phase 2 is ongoing.

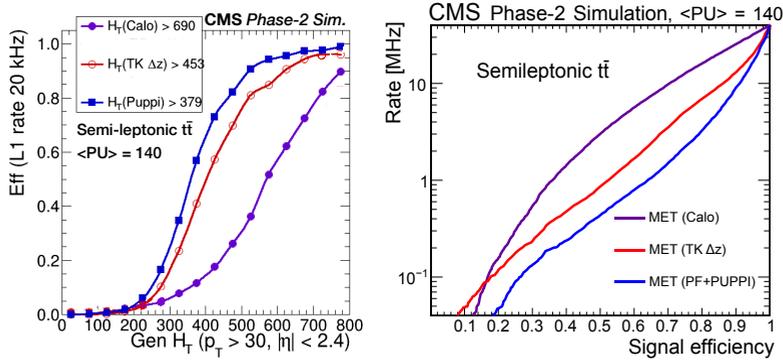


Figure 2. Left: turn-on for prototype H_T triggers using calorimeters, tracks and PF+PUPPI objects, for thresholds corresponding to a Level-1 rate of 20 kHz. Right: Level-1 Rate as function of signal efficiency for prototype p_T^{miss} triggers using calorimeters, tracks and PF+PUPPI objects.[6]

References

- [1] S. Chatrchyan et al. (CMS), JINST **3**, S08004 (2008)
- [2] A.M. Sirunyan et al. (CMS), JINST **12**, P10003 (2017), 1706.04965
- [3] V. Khachatryan et al. (CMS), JINST **12**, P01020 (2017), 1609.02366
- [4] CMS Collaboration, Tech. Rep. CERN-LHCC-2013-011. CMS-TDR-12 (2013), <https://cds.cern.ch/record/1556311>
- [5] CMS Collaboration, Tech. Rep. CERN-LHCC-2015-010, CMS-TDR-15-02 (2015), <http://cds.cern.ch/record/2020886>
- [6] CMS Collaboration, Tech. Rep. CERN-LHCC-2017-013. CMS-TDR-017 (2017), <https://cds.cern.ch/record/2283192>
- [7] D. Bertolini, P. Harris, M. Low, N. Tran, JHEP **10**, 059 (2014), 1407.6013
- [8] Xilinx, Vivado High Level Synthesis, <https://www.xilinx.com/products/design?tools/vivado/integration/esl?design.htm>
- [9] C. Ghabrous Larrea, K. Harder, D. Newbold, D. Sankey, A. Rose, A. Thea, T. Williams, JINST **10**, C02019 (2015)