

Chapter 12

Triggering and High-Level Data Selection



W. H. Smith

12.1 Level-1 Trigger

12.1.1 Introduction

The data taken by a particle physics collider detector consists of events, which are snapshots of the detector data at specific intervals in time. Usually these snapshots are taken at the frequency of the crossing of the colliding beams. For HERA this was 96 ns, for the Tevatron Run II this was 396 ns and for the LHC at design luminosity this is 25 ns. An individual bunch crossing may contain either no, one or many interactions between the particles in the colliding beams. The time during which beam collisions take place during a beam crossing is 1–2 ns. Even if there are multiple collisions in a single crossing the detector elements will make only one recording and the events will be superimposed. Therefore, each bunch crossing is individually evaluated. Not all of the detector data from an individual crossing is available immediately. Some may be stored as charge and need digitization. Other digital detector data may be inaccessible until further detector processing is complete.

The selection of bunch crossings is a highly complex function that involves a series of levels which take increasing amounts of time, process increasing amounts of data, use increasingly complex algorithms and make increasingly more precise determinations to reject increasing numbers of crossings. The first level(s) of the series usually involve(s) specific custom high-speed electronics. The subsequent level(s) involve more general CPU farms that run code similar to that found in the offline reconstruction. Due to this structure, the first level of trigger decision

W. H. Smith (✉)
University of Wisconsin-Madison, Madison, WI, USA
e-mail: wsmith@hep.wisc.edu

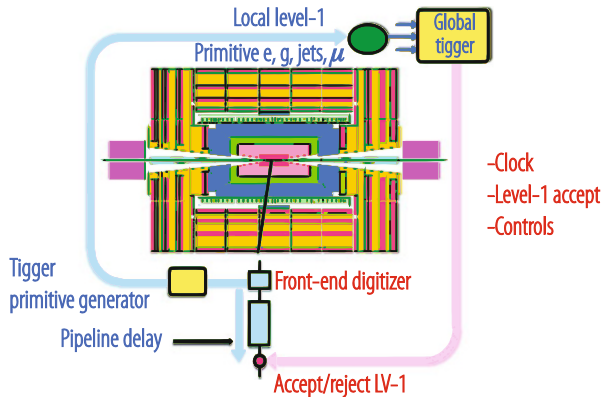


Fig. 12.1 Layout of the elements of the L1T

is based on particle identification (e.g. muon, electron, etc.) from local pattern recognition and energy evaluation. The higher trigger levels start by identifying the particle signature (e.g. Z, W, etc.), calculating kinematics for effective mass and event topology cuts and performing track reconstruction and detector matching (e.g. muon and tracking or calorimeter and tracking). The highest-level triggers perform identification of the physics process detected using event reconstruction and analysis. As shown schematically in Fig. 12.1 the Level-1 trigger¹ (L1T) inspects a subset of the detector information for each bunch crossing and provides the first in a series of decisions to either keep or discard it. The L1T system generally uses coarsely segmented data from calorimeter and muon detectors and in a few cases some rudimentary tracking detector information, while holding all the high-resolution data in pipeline memories in the front-end electronics. During the L1T decision time that is typically a few μ secs, all of the data from all crossings are stored. Usually a good fraction of the L1T time is used in transmission of the L1T data from the detector front ends to a central location where trigger processing is performed and transmission of the L1T decision back to the front ends, leaving a fraction of the L1T decision time available for the trigger processing.

The need to process each new crossing of data requires that the L1T function in a pipelined mode, e.g. be composed of a series of steps each of which processes its input and produces its output result at the crossing frequency. As noted above this can range from 396 ns at the Tevatron to 96 ns at HERA to 25 ns at the LHC. In order to avoid dead time, the trigger electronics must itself be pipelined: every process in the trigger must be repeated at the beam-crossing rate. This has important consequences for the requirements on the structure of the trigger system. The fact that each piece of logic must accept new data at the beam-crossing rate means that no

¹This is commonly called Level-1 but in such experiments as ALICE and LHCb this corresponds to the Level-0.

piece of individual data processing can take more than this time. This prohibits the use of iterative algorithms, such as jet finding based on finding a seed tower and then adding the surrounding towers to make a jet energy sum. This pipelined structure means that each step in the LIT logic must be completed within the time of the crossing frequency and the results output so that this step in the logic is available to process the data from the next crossing. The LIT logic therefore consists of a number of pipelined steps equal to the processing time multiplied by the crossing frequency.

The tight timing structure of the LIT presents a couple of challenges. Generally, the detector calorimeters have long pulse shapes that exceed the time between beam-crossings. This implies that particles produced in different bunch crossings can produce significant pulse-height in the bunch crossing of interest. Therefore, the detector systems that calculate the input information for the trigger need to correctly identify the energy associated with the correct bunch crossing, usually against a background of additional energy deposits from other bunch crossings. Typically, these systems use peak-finding algorithms and finite input response filters to perform this determination. The gaseous tracking detectors used in the muon systems can also have drift times or pulse widths exceeding the time between bunch crossings. These systems are also required to not only detect the passage of the charged track but also to identify the crossing that produced the track. Often this is resolved by combining and comparing the hits found in adjacent planes of chambers. Another challenge is that the physical extent of large HEP detectors produces times of flight to traverse them that exceed the time between bunch crossings. Therefore, at any particular point in time, the particles from interactions of more than one bunch crossing are present in the detector at different locations. This requires tight timing and synchronization of the detector trigger and readout systems.

The trigger is the start of the physics event selection process. A decision to retain an event for further consideration has to be made at the crossing frequency. This decision is based on the event's suitability for inclusion in one of the various data sets to be used for analysis. The data sets to be taken are determined by the experiment's physics priorities as a whole. Examples of data sets used in LHC experiments include di-lepton and multi-lepton data sets for top and Higgs studies, lepton plus jet data sets for top physics, and inclusive electron data sets for calorimeter calibrations. In addition, other samples are necessary for measuring efficiencies in event selection and studying backgrounds. The trigger has to select these samples in real time along with the main data samples.

The LIT is based on the identification of physics objects such as muons, electrons, photons, jets, taus and missing transverse energy. Each of these objects is typically tested against several p_T or E_T thresholds. The efficiency of a trigger is determined by dividing the number of events that pass the trigger by the number of actual events that would populate the final physics results plots if all of them passed the trigger. The trigger must have a sufficiently high and understood efficiency at a sufficiently low threshold to ensure a high yield of events in the final physics plots to provide enough statistics and a high enough efficiency for these events so that the correction for this efficiency does not add appreciably to the systematic error of the

measurement. The efficiency of the trigger is evaluated with respect to benchmark physics processes derived from the physics goals of the experiment. The criteria are a sharp turn-on curve of the efficiency at its threshold and an asymptote as close to 100% as possible. The L1T thresholds should be somewhat smaller than the offline physics analysis cuts. The reason for such a requirement is that the efficiency turn-on curves for the L1T will be somewhat softer than can be achieved with a full analysis including the best resolutions and calibration corrections.

Much of the logic in contemporary L1T systems is contained in custom Application Specific Integrated Circuits (ASICs), semi-custom or gate-array ASICs, Field Programmable Gate Arrays (FPGAs), Programmable Logic Devices (PLD), or discrete logic such as Random-Access Memories (RAM) that are used for memory Look-Up Tables (LUT). Given the remarkable progress in FPGA technology, both in speed and number of gates, the technology of many trigger systems has mostly moved towards full implementation in FPGAs.

The key to a good trigger system is flexibility. Not only should all thresholds be programmable, but also as mentioned above, algorithms are either implemented in FPGAs or LUTs. Reprogramming the FPGAs or downloading new LUT contents allows for revisions of the trigger algorithms. The only fairly fixed aspect of the trigger system is which data is brought to which point for processing. However, this is determined by the detector elements, size of showers and curvature of tracks, which are well known and basic features of the detectors and physics signals. There are new technologies being developed that are expected to provide flexibility in data routing, including backplanes and cards that use programmable cross-point switches.

The L1T system sustains a large dataflow. This is either carried on optical fibres, copper cables, or on backplanes within crates. At the LHC, the data carried by these means may be sent in parallel at either 40 MHz, or a higher multiple of this frequency, or converted from parallel to serial and transmitted at a higher rate on a single lines or pair of lines. Serial data transmission has the advantage of transmitting more data per cable wire or backplane pin but the disadvantage of extra latency for the parallel to serial and serial to parallel operations plus the risk of data errors involved with the encoding, high frequency transmission and link synchronization. In many cases this requires the overhead of monitoring and error detection bits. Copper cables in general avoid the necessity for optical drivers with their cost, size and power requirements, but have limited length capability, take up more volume and use more material.

12.1.2 L1T Requirements

The L1T has to be inclusive, local, measurably efficient, and fill the DAQ bandwidth with a high purity stream. The local philosophy of the trigger implies an initial trigger selection of electrons, photons, muons and jets that relies on local information tied directly to their distinctive signatures, rather than on global topologies. For

example, electron showers are small and extremely well defined in the transverse and longitudinal planes. Information from a few Electromagnetic and Hadronic calorimeter towers at the L1T, the corresponding elements of the preshower detector, and a small region of the tracking volume (at higher trigger levels) are sufficient for electron identification. The only global entities are neutrinos (from a global sum of missing E_T).

For the trigger to be measurably efficient the tools to measure lepton and jet efficiencies must be built into the trigger architecture from the start. One such tool is overlapping programmable triggers so that multiple triggers with different thresholds and cuts that can run in parallel. A second tool is pre-scaled (e.g. random selection of a fraction) triggers of lower threshold or weaker criteria that run in parallel with the stricter triggers. A third tool is pre-scaling of a particular trigger with one of its cuts removed.

The requirement on the use of DAQ bandwidth implies two conditions. First, each level of the trigger attempts to identify leptons and jets as efficiently as possible, while keeping the output bandwidth within requirements. The selected event sample should include all events that would be found by the full offline reconstruction. Hence, the selection criteria in the trigger must be consistent with those of the offline. Second, since the bandwidth to permanent storage media is limited, events must be selected with care at the final trigger level.

The measurement of trigger efficiency requires the flexibility to have overlapping triggers so that efficiencies can be measured from the data. The overlaps include different thresholds, relaxed individual criteria, prescaled samples with one criterion missing, and overlapping physics signatures. For example, measurement of the inclusive jet spectrum uses several triggers of successively higher thresholds, with the lower thresholds prescaled by factors that allow a reasonable rate to storage. These triggers overlap in jet energy all the way down to minimum bias events so that the full spectrum can be reconstructed accurately. The efficiency and bias of each higher threshold can be measured from the data sets of lower threshold. A requirement for understanding the trigger efficiency is that the data used as input to the L1T system is also transmitted via the DAQ for storage along with the event readout data. In addition, all trigger objects found, whether they were responsible for the L1 trigger or not should also be sent.

The L1T accept rate is limited by the speed of the detector electronics readout and the rate at which the data can be harvested by the data acquisition system. Since it is pipelined and deadtimeless, the L1T renders a decision on every bunch crossing. The maximum L1T accept rate is set by the average time to read information for processing by the Higher Level Triggers (HLT) and the average time for completion of processing steps in the HLT logic.

The high operational speed and pipelined architecture also requires that specific data is brought to specific points in the trigger system for processing and that there cannot be fetching of data based on analysis of other data in an event. The data must flow synchronously across the trigger logic in a deterministic manner in the same way for each crossing. At any moment there are many crossings being processed in sequence in the various stages of the trigger logic. The consequence is that most

of the L1T operations are either simple arithmetic operations or functions using memory lookup tables where an address of data produces a result previously written into the memory.

The L1T requirements evolve with the experiment luminosity, energy and event pile-up (number of p–p collisions per beam crossing). For example for the LHC trigger systems [20], algorithms used by the ATLAS [13] and CMS [14] experiments at the LHC during the period before 2014 (Run-1) were optimized for 7–8 TeV center-of-mass energy, PU up to 40 due to the 20 MHz beam crossing frequency and luminosities up to $7 \times 10^{-33} \text{ cm}^2 \text{ s}^{-1}$, whereas afterwards (Run-2) these were optimized for a 13 TeV center-of-mass energy and PU above 50 due to the 40 MHz beam crossing frequency and luminosities exceeding $15 \times 10^{-33} \text{ cm}^2 \text{ s}^{-1}$ [15, 16].

12.1.3 Muon Triggers

The design of L1T muon trigger logic depends on the detectors being used to generate the trigger information. These detectors include those with timing resolution and prompt signals that are generally less than the time between bunch crossings such as Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs). They also include special signal handling of detectors with individual signals and resolution greater than the bunch crossing time, such as Cathode Strip Chambers (CSCs) and Drift Tube Chambers (DTs). For these detectors, offset detector planes, front-end logic that processes over the drift time, and combinations of planes provide identification of the bunch crossing associated with the muon passage. Another important feature in muon trigger design is whether the muon chamber measuring stations are placed in a magnetic field in air or embedded in iron. In the former case, the muon momentum resolution is usually sufficient to provide an efficient threshold up to relatively high p_T . In the latter case, information from the tracking detectors is needed to provide a sufficiently sharp threshold.

L1T muon algorithms depend on comparison of tracks of hits with predefined geometrical patterns such as roads. For example, the ATLAS muon trigger employs RPCs and TGCs in an air-core magnetic field and the trigger algorithm uses Coincidence Windows that start with a hit in a central “pivot plane” and searches for time-correlated hits within an η – ϕ window in a “confirm plane” [1]. Different “confirm planes” are used for low and high p_T muons, as is shown in Fig. 12.2. The RPC barrel algorithm extrapolates hits in the middle RPC 2 station to a point and coincidence window in the innermost RPC 1 station along a straight line to the nominal interaction point. The size of this coincidence window depends on the muon’s bend in the magnetic field. A low- p_T candidate is found if there is one hit in this window and hits in both views and planes of either RPC 1 or RPC 2. If there is also a hit in RPC 3, then a high- p_T candidate has been found. For Run 2 ATLAS commissioned a fourth layer of barrel RPCs that improved the acceptance and added new trigger logic to the end-cap requiring additional coincidences with the TGC’s

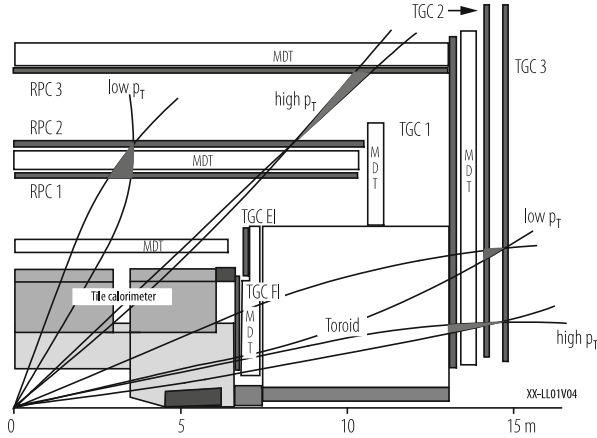


Fig. 12.2 ATLAS muon trigger algorithms

or the Tile hadronic calorimeter to reject particles not originating at the interaction point [15].

The CMS Detector uses Drift Tubes (DT), Cathode Strip Chambers (CSC) and overlapping Resistive Plate Chambers (RPC) for muon triggering in iron. The RPC readout strips are connected to pattern logic, which is projective in η and ϕ and connected to segment processors that find the tracks and calculate the p_T . As shown in Fig. 12.3 the CSC logic forms Local Charged Tracks (LCT) from the charge distributions in the CSC planes, which are combined with the Anode wire information for bunch crossing identification and assignment of p_T and “quality”, which is an indicator of the number of planes hit. The CSC Track Finder combines the LCTs into full muon tracks and assigns p_T values to them. As is also shown in Fig. 12.3 the DTs are equipped with Bunch and Track Identifier (BTI) electronics that finds track segments from coincidences of aligned hits in four layers of one drift tube superlayer. The DT Track Finder combines the segments from different stations into full muon tracks and assigns p_T values to them. In Run 1, the Global Muon Trigger sorted and then correlated the RPC, DT and CSC muon tracks. In Run 2, the RPC, DT and CSC information were combined earlier, in the track-finding stage [12].

The LHCb Level-0 muon trigger searches for candidates in the quadrants of five stations of Multi-Wire Proportional Chambers separated by iron and sends the two highest p_T candidates from each quadrant to the Level-0 Decision Unit (L0DU) [2]. The ALICE dimuon trigger system is based on two stations of 18 RPCs each read out on both sides of the gas gap by X - Y orthogonal strips with high resolution front-end electronics which feed local trigger electronics modules that find tracks in 3 out of the 4 detector planes in both X and Y [3]. The track is found and the magnetic deviation is calculated to enable a cut on a p_T threshold using memory Look-Up Tables (LUTs). Two unlike-sign muons are then required in the LIT.

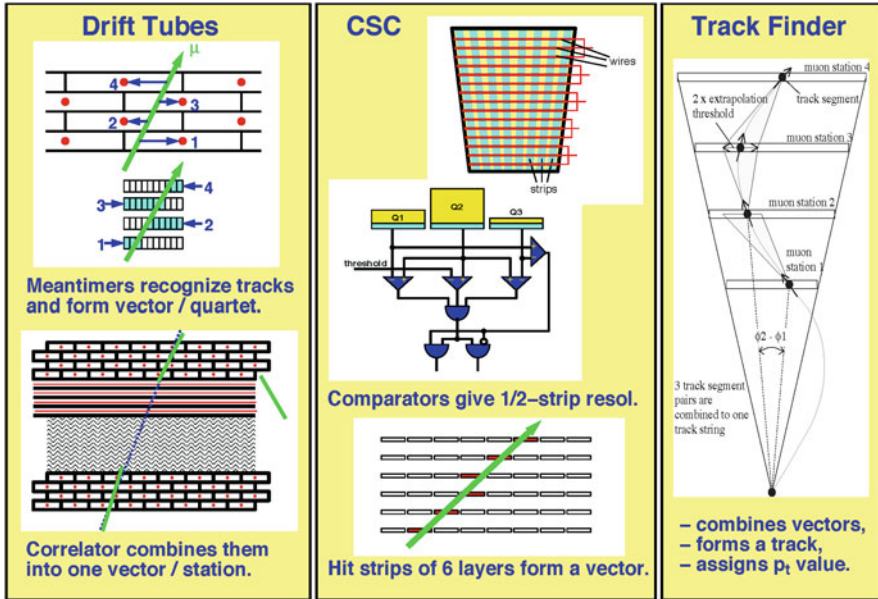


Fig. 12.3 CMS muon chamber trigger algorithms

12.1.4 Calorimeter Electron and Photon Triggers

The calorimeter trigger begins with trigger tower energy sums formed by the detector electromagnetic calorimeter (ECAL) hadronic calorimeter (HCAL) and forward calorimeter. Experiments vary on whether these sums are performed by analog methods before digitization or by digital summation after an initial ADC.

For the ATLAS experiment, the calorimeter trigger begins with a Preprocessor (PPr) which sums analog pulses into 0.1×0.1 ($\eta \times \phi$) trigger towers, assigns their bunch crossing and adjusts for calibration. The Cluster Processor then identifies and counts electron/photon and tau candidates based on the energies and patterns of energy isolation found in overlapping windows of 4×4 ECAL and HCAL trigger towers as shown in Fig. 12.4. For Run 2, the PPr was upgraded to provide improved Finite Input Response (FIR) filtering and dynamic bunch by bunch pedestal correction [15]. New cluster merging modules (CMX) were added that transmitted the location and energy of trigger objects, rather than the threshold multiplicities used in Run 1.

The CMS Calorimeter trigger algorithm for electron and photon candidates uses a 3×3 trigger tower sliding window centered on all ECAL/HCAL trigger towers. A diagram of this electromagnetic algorithm is shown in Fig. 12.5. Two types of electromagnetic objects are defined. The non-isolated electron/photon identification is based on a large energy deposit in one or two adjacent ECAL 5-cell ϕ strips in the trigger tower, the lateral shower profile in the central tower comparing maximum E_T

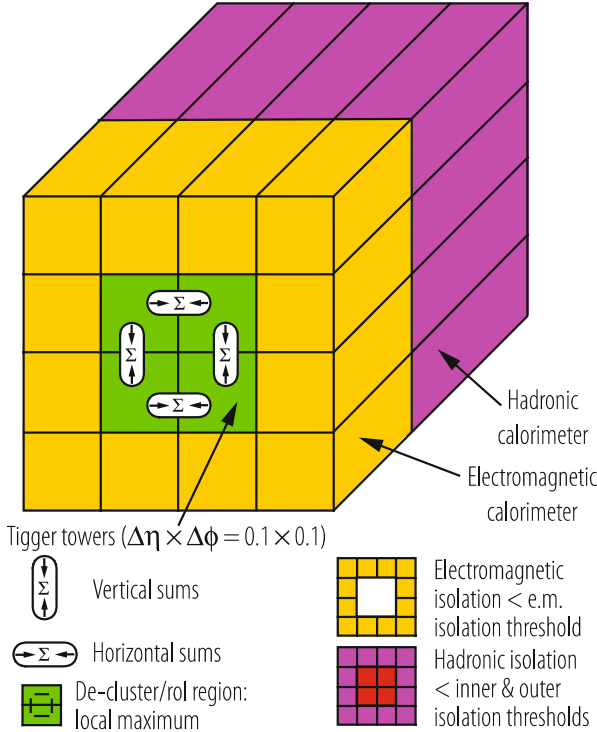


Fig. 12.4 ATLAS calorimeter electron/photon trigger algorithm

of each of four pairs of strips of 5 cells to the total tower level E_T of all 25 crystals (this “Fine Grain” veto uses a strip due to electron bending in the magnetic field), and the longitudinal shower profile defined by the ratio of E_T deposits in the HCAL and ECAL portions of the calorimeter (H/E veto). The isolated electron/photon has two additional requirements: the ECAL E_T deposited in one of the five trigger towers surrounding the central tower is below a programmable E_T threshold and the eight trigger towers surrounding the central tower in the 3×3 region have passed the Fine Grain and H/E vetoes. For Run-2, the CMS Calorimeter Trigger hardware was upgraded so that more complex algorithms could be deployed [21]. The e/γ and τ candidates started with a local maximum around which the trigger towers were dynamically clustered.

The LHCb Level-0 calorimeter trigger system combines the E_T measurement in clusters of 2×2 cells in the electromagnetic (ECAL) and hadronic calorimeters (HCAL), as well as information from the Scintillator Pad Detector (SPD) and a Preshower (Prs) to indicate the charged and electromagnetic nature of the clusters. The calorimeter trigger system sends the highest E_T hadron, electron, photon and π^0 candidates and the total HCAL E_T and SPD multiplicity to the Level 0 Decision Unit (LODU).

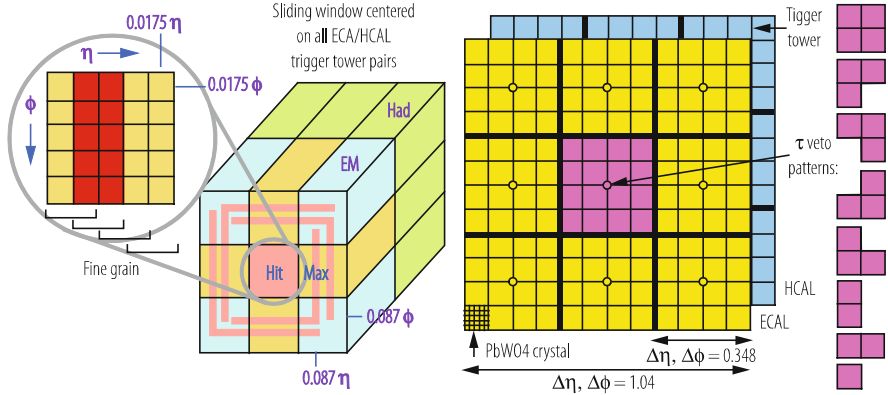


Fig. 12.5 CMS calorimeter electron/photon and jet/tau trigger algorithms

12.1.5 Calorimeter Jet and Missing Energy Triggers

The Level-1 Calorimeter Jet trigger needs to approximate the offline and higher-level trigger iterative jet finding in cones around seed towers with rectangular sliding windows of trigger towers. As shown in Fig. 12.5, the CMS jet trigger algorithms are based on sums of 3×3 calorimeter regions. This corresponds to 12×12 trigger towers in the barrel and endcap where a region corresponds to 4×4 trigger towers. The algorithm uses a 3×3 sliding window technique that uses the complete (η, ϕ) coverage of the CMS calorimeter. The E_T of the central region is required to be higher than that of the eight neighbours. The central jet or τ -tagged jet is defined by the 12×12 trigger tower E_T sum. In the case of τ -tagged jets, none of the nine 4×4 regions are allowed to have energy deposited outside the patterns of ECAL or HCAL towers (i.e. above a programmable threshold). For Run-2 the upgraded CMS calorimeter trigger formed Jet candidates by grouping the trigger towers around a local maximum in a 9×9 tower region in $\eta \times \phi$ with a PU subtraction estimated using four surrounding 3×9 tower regions. The ATLAS Jet and Energy L1T algorithm is based on a sliding window of 4×4 sums of trigger towers. It operates on a 4×8 matrix of core towers as shown in Fig. 12.6. In order to perform its calculations, it also needs the energy deposited in the “environment” of 7×11 towers. The execution of this algorithm depends on the duplication and distribution of energies in order to supply the needed information to perform these sums.

12.1.6 Tracking Information in Level-1 Triggers

Tracking information is very effective in reducing backgrounds to level-1 electron triggers from π^0 s. It improves tau triggers by identifying isolated tracks and it

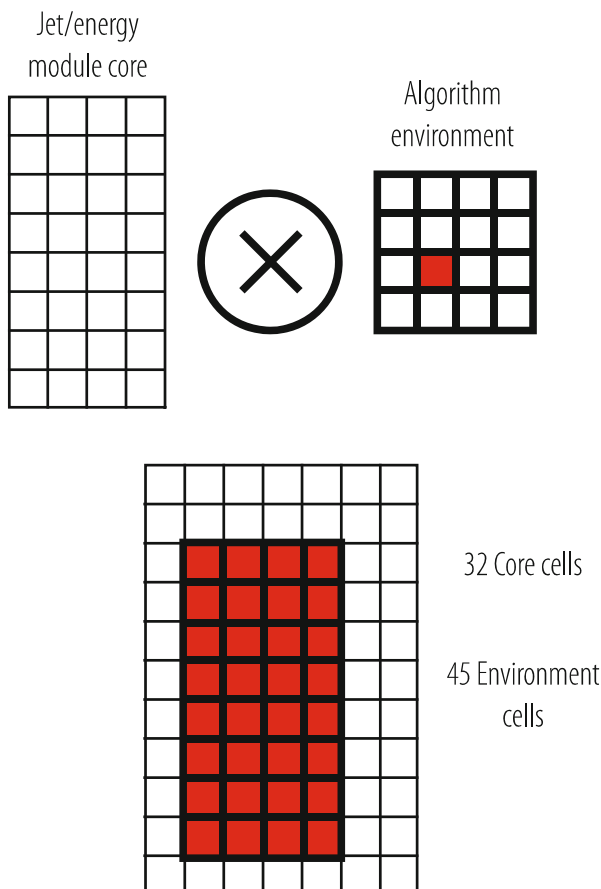


Fig. 12.6 Organization of the ATLAS jet trigger system

refines the muon trigger with a sharper momentum threshold that is not affected by the backgrounds in the muon chambers. It also can be used to identify heavy flavour candidates. Both Tevatron experiments CDF and DØ employed level-1 tracking triggers. CDF used signals from the Central Outer Tracker (COT) open-cell drift chamber in the eXtremely Fast Tracker (XFT) to perform charged track reconstruction in the r - ϕ plane for the L1T [4]. Track segments were found by comparing hit patterns in a COT superlayer to a list of valid patterns or “masks”. These masks contained specific patterns of prompt and delayed hits on the 12 wire layers of an axial COT superlayer. Tracks were found by comparing track segment patterns in all four layers to a list of valid segment patterns or “roads”. The XFT had an efficiency $>90\%$ for tracks with $p_T > 1.5$ GeV/c, transverse momentum resolution of $\delta p_T/p_T = 0.002 p_T$ and pointing resolution of $\delta\phi = 0.002$ radians with respect to the beam line [5]. The XFT reported the highest p_T track in each of 288 azimuthal

segments (1.25° each) to the XFT “Linker system” modules which cover 15° each and are matched to the segmentation of the trigger signals from the muon and calorimeter systems. The results from the linker system were passed to the Track Extrapolation System (XTRP), which sent one or more bits in 2.5° segmentation to the muon trigger systems set according to the calculated p_T , ϕ and multiple scattering. The XTRP also sent a set of 4 bits (for four momentum thresholds) for each 15° calorimeter wedge to the Level-1 calorimeter trigger. Finally the XTRP created a Level-2 tracking trigger based on the number of tracks and their p_T and ϕ information.

The $D\phi$ experiment Central Tracking Trigger (CTT) used information from the Central Fiber Tracker (CFT) and the Central Preshower System (CPS). Hit information from each of the 80 axial sectors of the CFT/CPS detectors was fed through boards programmed with 16,000 Boolean equations that identified patterns of hits likely to be produced by a charged particle. A list of tracks in four momentum ranges between 1.5 and 10 GeV/c was then sent to the L1 muon trigger system [6]. The $D\phi$ L1 CTT also identified the number of tracks in each event for each of these four momentum ranges, whether a coincident CPS hit had been found, and whether the track was isolated. This information was also used in the $D\phi$ L1T decision. The $D\phi$ CTT had an efficiency of $97.3 \pm 0.1\%$ for tracks with $p_T > 10$ GeV/c [4].

Although both ATLAS and CMS are planning the use of Tracking information at Level-1 in their designs for the High Luminosity LHC (HL-LHC) project [22], this information was not included in Run-1 or Run-2.

12.1.7 Global Triggers

An experiment Global Trigger accepts muon, calorimeter and tracking (if available) trigger information, synchronizes matching sub-system data arriving at different times and communicates the Level-1 decision to the timing, trigger and control system for distribution to the sub-systems to initiate the readout. The global trigger decision is made using logical combinations of the input trigger data. Besides handling physics triggers, the Global Trigger provides for test and calibration runs, not necessarily in phase with the machine, and for prescaled triggers, as this is an essential requirement for checking trigger efficiencies and recording samples of large cross section data.

The ATLAS Level-1 Global trigger is called the Central Trigger Processor (CTP). It combines information on the multiplicities of calorimeter and muon trigger objects which have sufficiently high momentum. These are electrons/photons, taus, jets, and muons. These are also the “seeds” for the Level-2 trigger that are sent to the Region of Interest Builder (RoIB). In addition, threshold information on the global transverse energy and missing energy sums is also used in the Level-1 decision. In Run-1, the CTP discriminated the delivered multiplicities of the trigger objects against multiplicity conditions and then combined these conditions to form more complex triggers when multiple object triggers are needed. In Run-2, the ATLAS

L1 Global trigger added a topological trigger (L1Topo) to allow geometrical or kinematic association between trigger objects received from the L1 Calorimeter or Muon Triggers [23].

The CMS L1 Global Trigger sorts ranked trigger objects, rather than histogramming objects over a fixed threshold. This allows all trigger criteria to be applied and varied at the Global Trigger level rather than earlier in the trigger processing. All trigger objects are accompanied by their coordinates in (η, ϕ) space. This allows the Global Trigger to vary thresholds based on the location of the trigger objects. It also allows the Global Trigger to require trigger objects to be close or opposite from each other. In addition, the presence of the trigger object coordinate data in the trigger data, which is read out first by the DAQ after a Level-1 trigger accept (L1A), permits a quick determination of the regions of interest where the more detailed HLT analyses should focus. The Global L1 Trigger transmits a decision to either accept (L1A) or reject each bunch crossing. This decision is transmitted through the Trigger Throttle System (TTS) to the Timing Trigger and Control system (TTC). The TTS allows the reduction by prescaling or blocking L1A signals in case the detector readout or DAQ buffers are at risk of overflow. For Run-1, the Global L1 Trigger allowed up to 128 algorithms to contribute to the overall trigger decision. For Run-2, this was upgraded to a modular design capable of up to about 500 algorithms that was typically running about 300 [16].

12.2 Higher-Level Selection

12.2.1 Introduction

The design of the Higher-Level Selection of events after Level-1 takes place in a number of “trigger levels”. Generally, collider experiments use at least two additional trigger levels, referred to as the Level-2 and Level-3 trigger. Some experiments have a Level-4 trigger. The higher the number the more general purpose (or commercial) the implementation, with the Level-3 and Level-4 triggers being composed of farms of standard commodity computers. The physical implementation of the Level-2 trigger varies substantially between experiments from inclusion in the Level-3 farm of processors to an independent farm of processors to customized dedicated processing hardware. The Level-2 trigger has to operate at the output rate of the Level-1 trigger, generally with a subset of the higher resolution and full-granularity available to the full reconstruction code available at Level-3 and higher. Typically, the Level-1 output rate ranges between 1 and 100 kHz depending on the experiment. The Level-2 trigger is generally limited in execution time so that the full event data cannot be unpacked and processed. Instead, the higher resolution and full-granularity data is unpacked in “regions of interest” determined by the Level-1 trigger data.

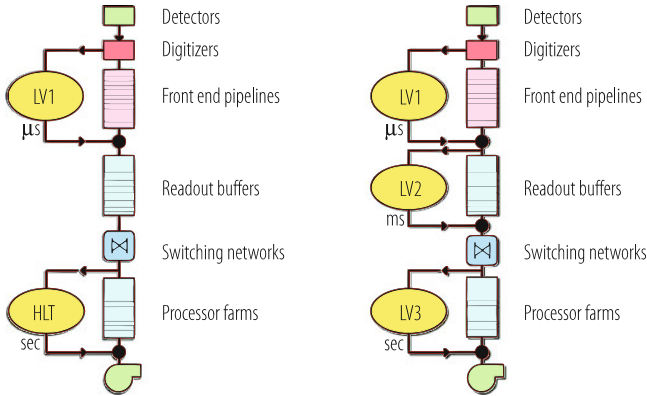


Fig. 12.7 Common architectures for collider detector trigger and data acquisition systems. Left: two physical levels. Right: three physical levels

The architectures of Level-2 trigger systems vary depending on the rejection factor required, the information provided as input, the interconnections with the front-end electronics, Level-1 and Level-2. Examples of two types of architecture presently employed by general-purpose collider detectors are shown in Fig. 12.7. Including Level-1, experiments such as H1, ZEUS, CDF, DØ and ATLAS have three physical levels of processing [18]. For Run-2, the ATLAS Higher Level Trigger Layers (HLT) were combined [15]. CMS has two layers of physical processing [19]. LHCb has three levels of processing, but the first level (Level-0) output trigger rate is 1.1 MHz, an order of magnitude higher than other collider experiments [7, 17]. The subsequent levels, HLT1 and HLT2, are software-based, running on the Event Filter Farm. In Run-2, HLT1 and HLT2 became two independent asynchronous processes on the same node and HLT2 was able to run a full reconstruction on real-time aligned and calibrated data [17].

There are more substantial differences in trigger architecture for experiments such as ALICE that is designed to study heavy ion collisions with a bunch spacing of 125 ns at a lower luminosity than the LHC experiments ATLAS and CMS. However, each Pb–Pb collision produces much higher multiplicities of secondary particles than a p–p collision, resulting in a much higher event size. Since the detectors in ALICE have different readout times, there are three parallel trigger systems, allowing readout from the faster detectors while slower detectors are occupied with reading out the data from earlier events [8]. The first decision is made 1.2 μs after the event (Level-0), the Level-1 decision comes after 6.5 μs , and the Level-2 trigger is issued after 88 μs . The Level-1 and Level-2 decisions can veto trigger signals from Level-0. The ALICE Central Trigger Processor also checks the events for pile-up from events in a programmable time interval before and after the interaction at all three levels. For Run-2, an earlier L0 trigger decision time of 525 ns provides a pre-trigger for the TRD [24].

The algorithms deployed in the HLT are dynamic, reflecting continuing improvements in the offline reconstruction that represent the functions that the HLT is attempting to approach within the constraints of processing time. The descriptions below of algorithms in the LHC experiments represent a snapshot at the time of Run-1 processing. These considerably evolved during Run-1 into Run-2, although the general techniques shown continue to be applied.

12.2.2 Tracking in Higher Level Triggers

The principal new information in the higher level triggers is tracking information. Either it is introduced for the first time in the event selection process or it is greatly refined over rudimentary tracking used in the Level-1 trigger. There are two major sources of tracking information. A pixel detector provides the most inner tracking and some vertex information. Outside of the pixel detector, silicon strip and then in some cases drift chambers, fibers, or straw tube detectors provide additional information at larger radius. For example, ATLAS [9] uses space points found in the pixels and silicon central tracker (SCT) to find the z -vertex location, fit tracks into the Transition Radiation Tracker (TRT) and measure the ϕ and p_T of the track above a p_T of 0.5 GeV/c. In the latter part of Run-2, ATLAS commissioned the Fast TracKer (FTK), a dedicated Associative-Memory hardware processor which delivers tracks with $p_T > 1$ GeV/c for every L1A to the HLT within 100 μ s [25]. In CMS, two types of tracking are employed. Charged particle tracks are first quickly reconstructed using pixel hits and then more laboriously but more accurately reconstructed with additional hits from the silicon strip tracker. Generally, tracking is “seeded” by the confirmed Higher Level Trigger objects, which themselves are “seeded” by Level-1 trigger objects.

12.2.3 Selection of Muons

The first algorithms executed in Level-2 on Level-1 selected muons are refinements of the reconstruction of the tracks in the muon chambers. In the case of ATLAS, where only the RPC (Barrel) and TGT (Forward) chambers provide information for the L1T, the precision hit information from the Monitored Drift Tubes (MDTs) is added to the RPC and TGC determined candidates. This provides good track reconstruction in the muon spectrometer. Since the ATLAS Muon Chambers are mostly in air, there is little multiple coulomb scattering. The found tracks are extrapolated for combination with tracks found in the Inner Detector. Matching between muon tracks measured independently in the muon system and those in the Inner Detector selects prompt muons and rejects fake and secondary muons. The isolated muon triggers also use information from the calorimeter towers surrounding the found muon track.

In CMS, all of the muon chamber systems participate in the L1T. The L1T muon candidates are used to seed the reconstruction of tracks in the muon chambers in the Level-2 algorithm. First, an initial pattern recognition is performed on muon segments along the trajectory, then a second more precise fit using all hits on these segments is used to determine the muon parameters. Since the CMS chambers are surrounded by steel, the propagation of track parameters to adjacent muon stations must take into account material effects such as multiple Coulomb scattering, and energy losses due to ionization and bremsstrahlung in the muon chambers and the iron. To avoid excessive processing times, these are estimated from fast parameterizations. Muons passing this first reconstruction are then input to the Level-3 reconstruction that uses hits in the silicon tracker within a rectangular $\eta \times \phi$ region. Pairs or triplets of hits in the innermost layers of the tracker form trajectory seeds that are required to be compatible with the $\eta \times \phi$ region and the primary vertex constraints. These are then grown into tracks of about seven hits and optionally combined with the reconstructed hits from the Level-2 algorithm. In Level-2, the isolation variable is calculated from the weighted sums of energies deposited in the ECAL and in the HCAL in the region around the muon track. For the Level-3 isolation variable, only charged-particle tracks near the vertex of the candidate muon are selected for inclusion. This excludes tracks from pile-up of contributions from other pp collisions (which occur at another vertex location), making this isolation less sensitive to pile-up than calorimetric isolation.

12.2.4 Selection of Electrons and Photons

The first algorithms executed in Level-2 on Level-1 selected electrons and photons are refinements of the clustering algorithms. For example, in ATLAS [9], the energy deposited in windows of the electromagnetic LAr calorimeter cells and the energy-weighted position information, as well as the leakage energy into the hadronic calorimeter are calculated. CMS [10] also reconstructs energy in clusters of electromagnetic calorimeter cells corresponding to the Level-1 calorimeter triggers, adding a margin around the trigger region to ensure complete collection of energy. These clusters are then formed into “Super Clusters” which are groups of clusters along a road in the ϕ direction, chosen due to bending in the magnetic field. These clusters are then required to be isolated in the electromagnetic calorimeter. The hadronic calorimeter energies are then reconstructed and the energies in the hadronic tower behind the cluster and the adjacent towers are required to be small with respect to the electromagnetic cluster energy.

The second tier of algorithms performed on electrons and photons confirmed by the first algorithms are tracking algorithms. The first or more local steps of these are generally called Level 2.5 algorithms. This involves establishing track isolation around the electromagnetic cluster and for electron triggers, associating the electromagnetic cluster with a track. For CMS electron triggers, the energy and position of the Super Cluster is used to search for hits in the pixel detector. These

hits are reconstructed and the track p_T is checked for consistency with the Super Cluster energy. For both electron and photon triggers, tracks are seeded from pairs of hits in the pixel layers in a rectangular $\eta \times \phi$ region around the direction of the reconstructed electron or photon, where these seeds are required to be consistent with the nominal vertex spread (photons) or closest approach of the electron path to the beam line (electrons). Then for electrons a threshold is applied to the p_T sum of the tracks within a cone around the electron direction and on the number of tracks for the photon. In ATLAS [11] the electromagnetic cluster is identified as an electron by association with a track in the Inner Detector, which is found by independent searches in the SCT/Pixel and TRT detectors in the region identified by the L1T RoI. For electron candidates, matching in both position and momentum between the track and cluster is required.

12.2.5 Selection of Jets and Missing Energy

The primary processing of the jet candidates at Level-2 begins with the L1T jet candidates, which are used as seeds for the Level-2 jets. The first step is to recalculate the jet energy for these candidates using the full granularity and calorimeter energy resolution information, which is not available to the Level-1 jet energy calculation. In ATLAS, the Level-2 jet finding searches in the RoIs produced by the Level-1 calorimeter logic. In CMS, jets are reconstructed using an iterative cone algorithm with cone size $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.5$ that sums over all projected electromagnetic and hadronic calorimeter cells with energy greater than a threshold set above the level of noise (0.5 GeV). In addition, to be declared a jet, at least one seed tower must have $E_T > 1$ GeV. After summation, the jet energy is adjusted by an η -dependent correction for the calorimeter response.

Missing energy is calculated by summing all towers with E_T above a noise threshold. For CMS, this threshold is 0.5 GeV. No energy corrections are applied to Missing E_T . Since Missing E_T is susceptible to noise because it is summing over many channels, an alternative is often considered. This is Missing H_T , which is Missing E_T calculated by summing over the jets in the event rather than the calorimeter cells. Since there are fewer cells involved in the computation of missing H_T , there is less noise included in this sum.

It is typical to ask for two or more jets in the HLT algorithms. It is also common to combine two or more jets with missing E_T or H_T . Also, topological constraints are often employed such as requiring forward jets or acoplanarity between multiple jets or jets and Missing E_T .

12.2.6 Selection of Hadronic Tau Decays

The Level-2 processing of tau jets relies only on calorimeter information. In ATLAS, the tau finding uses the same algorithms used for electron and photon candidates, but retuned for taus. The inputs are the Level-1 RoIs. A cluster summed over the full resolution data for the electromagnetic and hadronic cells is required to have $E_T > 20$ GeV with at least 10 GeV required individually in the electromagnetic and hadronic cells. The position of the candidate cluster is required to be consistent with the Level-1 tau-jet candidate. Then shower shape variables are used to discriminate tau jets from regular jets. An example of one such variable is R_{37} , defined as the ratio of E_T contained in a 3×7 cell cluster to the E_T contained in a 7×7 cell cluster centred on the same seed cell calculated for the second electromagnetic layer of the LAr calorimeter. In CMS, the Level-1 tau jets are used as seeds for the Level-2 tau-jet reconstruction that employs an iterative cone algorithm with a radius of $R = 0.5$. Level-2 tau candidates are then these jets which have $E_T > 15$ GeV and are tagged as isolated if the sum of electromagnetic calorimeter deposits in an annulus $0.13 < R < 0.4$ around the jet direction, $E_T < 5$ GeV.

The subsequent processing of tau candidates involves tracking. ATLAS requires a track formed from the pixel and SCT detector space points in the RoI to be within $\Delta R < 0.3$ of the Level-2 tau candidate cluster direction. At Level-3 a requirement is made that the number of tracks within $\Delta R < 0.3$ be either one or three. Additional detailed jet shape requirements also refine the identification. In CMS, at Level 2.5 (the higher level trigger processing following the initial Level-2 processing that uses calorimeter and muon information alone), tau selection is based on tracks with a $p_T > 5$ GeV/c that are reconstructed from seeds from the pixel hits found in a small rectangle ($\Delta\eta = \Delta\phi = 0.1$) around the tau-candidate direction. At Level 3 the rectangle is expanded to 0.5 and the p_T cut is reduced to 1 GeV/c. To save CPU time, these tracks are terminated when seven hits in the silicon strip tracker are acquired since the resolution with seven hits is close to final. Reconstructed tracks are associated with the tau-jet candidate if they are within a radius $R < 0.5$ and originate from the primary vertex as determined by the pixel tracks. Tracks within a radius $R < 0.1$ of the tau-jet candidate direction are classed as tau tracks. The leading tau track must have $p_T > 3$ GeV and there must be no reconstructed tracks within an annulus $0.07 < R < 0.3$ around this track.

12.2.7 Selection of b-Jets

The b-jet selection is based on track reconstruction to tag displaced vertices associated with the jet. In ATLAS at Level-2, b-tagging uses reconstructed tracks from the silicon tracker within the Level-1 jet RoI. For each of these tracks the significance of the transverse impact parameter is computed and its error is

parameterized as a function of p_T . A b-jet discriminator is constructed using the likelihood ratio method to determine for each track in the jet the ratio of probability densities for the track to come from a b-jet or a u-jet. In CMS, Level-2 starts with events with 1, 2, 3 or 4 jets passing various thresholds or a high total E_T for the whole event. At Level 2.5 tracks are reconstructed using only pixel hits (at least three required), which are used to reconstruct the primary vertex. The b-tag algorithm runs on the four highest E_T jets with $E_T > 35$ GeV and uses the pixel tracks and primary vertex to tag jets as b-jets if they have at least two tracks with a signed 3D impact parameter with large significance. Events pass Level 2.5 if they have at least one b-tagged jet. At Level 3, tracks of up to eight hits are reconstructed in a cone of size $\Delta R = 0.25$ around the b-tagged jets. The level-3 filter selects events where there is at least one jet having at least two tracks with large impact parameter significance.

12.3 Outlook

Trigger and DAQ requirements will further evolve in the next decade with large increases in luminosity and the associated pile-up. ALICE will continuously read out the majority of its detectors with different latencies, busy times and technologies, differently optimized for pp, pA and AA running scenarios [26]. Triggered readout will be used by some detectors and for commissioning and some calibration runs. LHCb will run trigger-free at 30 MHz, reading every bunch crossing with inelastic collisions [27].

A major upgrade to the LHC, the HL-LHC [28], is planned to start in the middle of this decade and deliver a luminosity of $5\text{--}7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ at the LHC design centre of mass energy of 14 TeV, which corresponds to a pile-up of 140–200 at 25 ns bunch spacing. Present link technologies operable in the radiation and magnetic field environments of their inner detectors do not allow ATLAS and CMS to adopt a “triggerless” architecture with an acceptable detector power and material budget for their tracking detectors. Therefore, at the HL-LHC, both ATLAS and CMS will retain architectures with Level 1 triggers.

In order to maintain Run-2 physics sensitivity at the HL-LHC, ATLAS and CMS will add L1 tracking triggers for identification of tracks associated with calorimeter and muon trigger objects and will also feature a significant increase of L1 rate, L1 latency and HLT output rate. Additionally, ATLAS and CMS are also studying the use of fast timing information in the L1T. The ATLAS experiment will divide its L1T into two stages [29]. A L0 trigger with a rate of 1 MHz and latency of $6 \mu\text{s}$ will use calorimeter and muon trigger information to produce seeds used with tracking and more fine-grained calorimeter and muon trigger information in the L1 trigger with an output rate of 400 kHz and latency of $30 \mu\text{s}$. This is processed by the HLT with an output storage rate of 5–10 kHz. The CMS L1T latency will increase to $12.5 \mu\text{s}$ with an output range of 500–750 kHz for pileup ranging between 140 and 200 [30]. It will use an un-seeded L1 Track trigger along with finer granularity

calorimeter and muon triggers. The CMS HLT output rate to storage will range between 5 and 7.5 kHz for pileup ranging between 140 and 200.

The hardware implementations of the HL-LHC ATLAS and CMS L1T will use high-bandwidth serial I/O links for data communication and large, modern field-programmable gate arrays (FPGAs) for sophisticated and fast algorithms. The development and synthesis of FPGA firmware incorporating these algorithms is significantly enhanced in reliability, accessibility and performance with Higher Level Synthesis (HLS) tools [31]. The latest developments and expectations for future FPGAs not only include significant increases in the number of logic gates available and high-speed serial links, but also increases in the number of high-bandwidth serial links per device, more sophisticated and fast DSPs, embedded Linux, and integration with high speed networking. Fast Tracking Trigger devices such as the ATLAS FTK [25] use Associative Memories. The hardware framework will be designed following standards deployed in industry, such as the Advanced Telecommunications Architecture (ATCA) for backplanes, which offers substantial backplane bandwidth and flexibility and provides for users to extend the backplane connectivity using the spare I/O available on each card. Further interconnectivity technology developments such as optical backplanes and wireless data transmission may provide additional opportunities.

The increase in L1 output rate from 100 kHz to possibly as high as 1 MHz requires higher bandwidth into the DAQ system and more CPU power in the HLT. The addition of a tracking trigger and more sophisticated algorithms at L1 increases the purity of the sample of events passing the L1 trigger, but requires a higher sophistication and complexity of algorithms used at the HLT. This implies a greater CPU power than scaling with the L1 output rate but is somewhat mitigated by the availability of the L1 Tracking Trigger primitives in the data immediately accessible by the HLT. Without a L1 tracking trigger, the opportunity to access most of the tracker information at the first levels of the HLT is limited by the CPU time to unpack and reconstruct the tracking data. This is significantly improved in the ATLAS FTK that provides quick access to tracking information in the HLT. For the HL-LHC, the addition of the L1 tracking trigger means that the results from the L1T track reconstruction can be immediately used without the overhead of tracking data unpacking and reconstruction.

The evolution of the computing market towards different computing platforms and co-processors offers an opportunity to achieve substantial gains in HLT processing power at the price of adapting code to the new hardware. Examples include Graphical Processor Units (GPUs), such as the NVIDIA Tesla and GeForce (used by ALICE [32]), ARM processors, FPGAs (e.g. the Xeon/FPGA used by LHCb [33]) and the Intel Xeon Phi coprocessor. Additional HLT processing power may result from improved code such as machine learning algorithms for track reconstruction [34].

References

1. M. J. Woudstra and the ATLAS Collaboration, Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV, 2014 J. Phys.: Conf. Ser. 513 012040.
2. E. Aslanides et al., (LHCb Collaboration), The Level-0 muon trigger for the LHCb experiment, Nucl. Instr. and Meth. A579 (2007) 989–1004.
3. B. Forestier, Nucl. Instrum. Meth. A 533 (2004) 22–26.
4. R. Downing et al., Nucl. Instrum. Meth. A 570 (2007) 36–50.
5. E.J. Thompson et al., *Online Track Processor for the CDF Upgrade*. IEEE Trans. Nucl. Sci. 49(3) (2002).
6. Y. Maravin et al., *First Results from the Central Tracking Trigger of the DØ Experiment*. Proc. 2003 IEEE Nuclear Science Symposium, Portland, OR.
7. R. Aaij et al., (LHCb collaboration), *The LHCb Trigger and its Performance in 2011*, 2013 JINST 8 P04022.
8. M. Krivda et al. (ALICE Collaboration), *The ALICE trigger system performance for p-p and Pb-Pb collisions* 2012 JINST 7 C0105.
9. I. A. Christidi and the ATLAS Collaboration, *The tracking performance of the ATLAS High Level Trigger in pp collisions at the LHC*, 2011 J. Phys.: Conf. Ser. 331, 032006.
10. The CMS Collaboration, *The Trigger and Data Acquisition project*, Vol. II: *Data Acquisition and the Higher Level Trigger*. CERN/LHCC 2007-021.
11. Will Buttinger (ATLAS collaboration) *The ATLAS Level-1 Trigger System*, 2012 J. Phys.: Conf. Ser. 396 012010.
12. D. Acosta et al. (CMS Collaboration), *CMS Trigger Improvements Towards Run II*, Nuclear and Particle Physics Proceedings 273–275 (2016) 1008–1013.
13. The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J.C 72, 1849 (2012).
14. V. Khachatryan et al., (CMS Collaboration) *The CMS Trigger*, 2017 JINST 12 P01020.
15. The ATLAS Collaboration, *Performance of the ATLAS Trigger system in 2015*, Eur. Phys. J. C (2017) 77: 317.
16. L. Cadamuro et al. (CMS Collaboration), *The CMS Level-1 trigger system for LHC Run II*, 2017 JINST 12 C03021.
17. B. Sciascia (LHCb Collaboration), *LHCb Run 2 trigger performance*, PoS BEAUTY2016 (2016) 029.
18. The ATLAS TDAQ Collaboration, *The ATLAS Data Acquisition and High Level Trigger system* 2016 JINST 11 P06008.
19. The CMS Collaboration, *The CMS Trigger*, 2017 JINST 12 P01020.
20. W. H. Smith, *Triggering at the LHC*, Ann. Rev. Nucl. Part. Sci. 66, 123 (2016).
21. CMS collaboration, *Triggering on electrons, jets and tau leptons with the CMS upgraded calorimeter trigger for the LHC RUN II*, 2016 JINST 11 C02008.
22. ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, CERN-LHCC-2015-020; CMS Collaboration, *Technical Proposal for the Phase-II Upgrade of the CMS Detector*, CERN-LHCC-2015-010.
23. M. zur Nedden (on behalf of the ATLAS Collaboration), *The LHC Run 2 ATLAS trigger system: design, performance and plans*, 2017 JINST 12 C03024.
24. M. Krivda et al. (The ALICE Collaboration), *The ALICE Central Trigger Processor (CTP) Upgrade* 2016 JINST 11 C03051.
25. Asbah, N. & ATLAS collaboration, *A hardware fast tracker for the ATLAS trigger*, Phys. Part. Nuclei Lett. (2016) 13: 527.
26. ALICE Collaboration, *Upgrade of the ALICE Readout & Trigger System*, CERN-LHCC-2013-019; M. Krivda and J. Pospisil for the ALICE Collaboration, *The ALICE Central Trigger Processor (CTP) Upgrade*, JINST 11, C03051 (2016); F Costa et al., *The detector read-out in ALICE during Run 3 and 4*, J. Phys.: Conf. Ser. 898 032011 (2017).

27. LHCb Collaboration, LHCb Trigger and Online Upgrade Technical Design Report, CERN-LHCC-2014-016; T. Szumlak, Real time analysis with the upgraded LHCb trigger in Run III, *J. Phys.: Conf. Ser.* 898 032051 (2017).
28. Apollinari, G. (ed.) et al., HL-LHC Preliminary Design Report, CERN-2015-005; P. Campana, M. Klute and P.S. Wells, *Physics Goals and Experimental Challenges of the Proton-Proton High-Luminosity Operation of the LHC*, *Ann. Rev. Nucl. Part. Sci.* 66, 273 (2016).
29. ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, CERN-LHCC-2015-20 (2015).
30. CMS Collaboration, The Phase-2 Upgrade of the CMS Level-2 Trigger, CERN-LHCC-2017-013 (2017); CMS Phase-II Upgrade Scope Document, CERN-LHCC-2015-19 (2015).
31. N.P. Ghanathe, et al., *Software and firmware co-development using high-level synthesis*, *JINST* 12 C01083 (2017).
32. S. Gorbunov et al., *ALICE HLT High Speed Tracking on GPU*, in *IEEE Transactions on Nuclear Science*, 58, no. 4, 1845 (2011).
33. C. Färber, R. Schwemmer, J. Machen and N. Neufeld, *Particle identification on an FPGA accelerated compute platform for the LHCb upgrade*, 2016 IEEE-NPSS Real Time Conference (RT), Padua, 2016, pp. 1–2.
34. S. Farrell et al., *The HEP.TrkX Project: deep neural networks for HL-LHC online and offline tracking*, *EPJ Web of Conferences* 150, 00003 (2017).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

