

The design of a fast Level-1 track trigger for the High Luminosity Upgrade of ATLAS

Joakim Gradin*

On behalf of the ATLAS collaboration

Uppsala Universitet, Université Grenoble Alpes

E-mail: per.olv.joakim.gradin@cern.ch

The high-luminosity upgrade of the LHC will increase the rate of the proton-proton collisions by approximately a factor of 5 with respect to the initial LHC design. The ATLAS experiment will upgrade consequently, increasing its robustness and selectivity in the expected high radiation environment. In particular, the earliest, hardware based, ATLAS trigger stage ("Level 1") will require higher rejection power, still maintaining efficient selection on many and various physics signatures. The key ingredient is the possibility of extracting tracking information from the brand new full-silicon detector and use it for the decision process. While fascinating, this solution poses a big challenge in the choice of the architecture, due to the reduced latency available at this trigger level (few tens of micro-seconds) and the high expected working rates (order of MHz). In this paper, we review the design possibilities of such a system in a potential new trigger and readout architecture, and present the performance resulting from a detailed simulation of possible hardware-based algorithms, to be implemented in the context of Associative Memories and FPGA technologies, as foreseen by R&D plans on these devices.

Fourth Annual Large Hadron Collider Physics

13-18 June 2016

Lund, Sweden

*Speaker.

1. Introduction

The ATLAS experiment at the Large Hadron Collider (LHC) at CERN is a multipurpose experiment with the aim of investigating the phenomena of high energy proton-proton collisions. The physics program includes both measurements of parameters of the Standard Model of particle physics as well as searches for new, unobserved physics. The ATLAS detector consists of three subdetectors; the inner detector, the electromagnetic (EM) and hadronic calorimeters, and the muon detector. The LHC provides the experiment with collisions at a very high rate to be able to observe rare events. To handle the large amount of data from these collisions ATLAS uses a two tiered trigger system, a first (L1) hardware based trigger and a second software based trigger. With the High Luminosity LHC upgrade, planned to start in 2026, the peak luminosity is expected to exceed five times the design value ($5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$). This will push the rate of the single lepton triggers beyond readout capability if the trigger thresholds are preserved as the current Run-2 values. Simply increasing the thresholds would reduce the physics potential of the ATLAS experiment. The proposed solution is to take tracking information into account already at L1, with the so-called L1Track trigger. Previous studies have shown that a trigger of this kind could reduce the rates to an acceptable level [1].

The current baseline design foresees a L1Track trigger seeded by Regions of Interest (RoI), corresponding to about 10% of the event data volume, from the new L0 trigger (similar to current L1). The L0 single lepton triggers react on signatures in the EM calorimeter and muon detectors to select events with potential high p_T leptons at a rate of 1 MHz and the L1 trigger will bring it down to 200 kHz. A new full-silicon inner detector, called the Inner Tracker (ITk), is currently under design, a few layouts are under study for validating the L1Track trigger, all including outer strip and inner pixel layers. L1Track will process hits from the ITk in an RoI ($\Delta\eta \times \Delta\phi = 0.2 \times 0.2$) and cluster them in "super strips" to form coarse granularity patterns.¹ These patterns are then compared to precomputed patterns within Associative Memory (AM) chips [2], the efficiency of this pattern matching is limited by the number of patterns each chip can store. The size of the super strips is chosen to be large enough to reduce number of patterns stored and small enough to limit the random combinations in background events. This step will reduce the large number of hit combinations and allow to perform the fit to extrapolate the track parameters as the number of fits per event is limited by the processing latency. The fitting procedure can be implemented in FPGAs which are expected to be able to perform one fit in 0.25 ns, the track parameters are then propagated to the L1Global processor for a final L1 trigger decision. By design the system should be able to satisfy the goals of a five times reduction of the L1 single lepton trigger rates with a signal efficiency above 95% with respect to offline reconstructed tracks, with at least 10 mm resolution on the track position along the beam axis, and have a short latency of few tens of μs . All these steps in the track reconstruction have been simulated in detail, with different layouts of the silicon detectors and at the expected luminosity, which correspond to an average of 200 interactions per collision (called pileup or $\langle\mu\rangle$).

¹The ATLAS coordinate system uses \hat{z} along the beam axis, ϕ as the azimuthal angle and the pseudorapidity $\eta = -\ln(\tan(\theta/2))$. The particle tracks are described using perigee parameters; q/p_T the charge over transverse momentum, ϕ , η , d_0 and z_0 the distance to and position on the z -axis at the point of closest approach respectively.

2. Pattern generation

The patterns stored in the AM chips are so-called pattern banks. They are generated from single muon events with p_T greater than 4 GeV/c and the expected LHC beam spot size (250 mm) along the beam axis. Each AM chip is expected to store 0.5M patterns and cover a 0.1×0.1 region, so a maximum of 1M patterns can be used for one RoI. A subset of the ITk layers are used to construct a pattern from each event. The number of layers, eight in our studies, has been chosen to reduce the number of superstrip combinations, keeping the pattern banks small while having enough information to effectively identify high p_T tracks among pileup. The simulations were made with the layout from the Phase-II upgrade Letter of Intent [1] which has four pixel layers closest to the interaction point and five double layers of strip detectors at larger radii. Less hit combinations are found using only the outer strip layers, which have lower occupancy and less granularity in z than the pixel layers. The latter on the other hand can improve the quality of the tracks with the longer lever arm and the greater precision in z . This is crucial for high η regions in particular. By not using all the available detector layers, geometrical inefficiencies due to unavoidable dead regions between the silicon modules may become relevant. To recover the efficiency, a “wild card” scheme has been applied in the generation of the banks. This means that if a single muon track has hits in all but one or two of the designated layers, the missing layers gets marked as wild cards. In the pattern matching procedure a wild card is always considered as hit, thereby the match is true even when a track has a missing hit. The number of patterns needed to ensure a high efficiency is large, often larger than the amount foreseen as feasible to store in the AM chips. An effective way to reduce the number of patterns without losing much efficiency is to use “don’t care” (DC) bits [2]. Technically, this means leaving out the least significant bits in the matching of pattern hit words. The effect is that patterns differing only by one or two super strip IDs, i.e. adjacent super strips, can be stored as one pattern with slightly coarser resolution. The number of allowed DC bits in each layer can be different. A good strategy is to allow more on the outer layers as tracks close in p_T will “fan out” at larger radii, as illustrated in Figure 1.

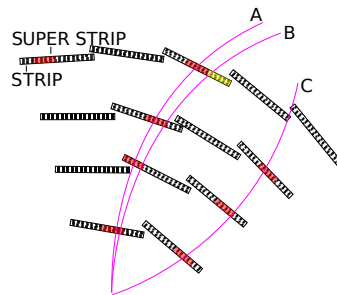


Figure 1: Cartoon representation of pattern generation. Tracks A and B have hits in all layers, they only differ in the outermost layer and can be combined into one pattern with a DC bit. Track C is missing a hit in one layer, a pattern with a wildcard on the outermost layer will be stored.

3. Tracking performance of the simulated L1Track trigger

A linear fit of the track parameters is made from the full granularity hits within the matched patterns. The fit is extracted from precomputed fit constants relating cluster positions to track

parameters; in a similar way the χ^2 of the fit is computed. A configuration using only strip layers has been proven to be sufficient to provide good quality tracks, ensuring at least 95% efficiency on reconstructed single leptons (with $p_T \geq 20$ GeV/c for muons and $p_T \geq 25$ GeV/c for electrons) with a five times rejection on corresponding backgrounds. This is shown in Figure 2 where the efficiency on single leptons is compared to that on the corresponding main backgrounds, semileptonically decaying b-jets and jets respectively, overlaid with a pileup of $\langle \mu \rangle = 200$. The background lepton momentum spectrum has been weighted to the expected shape from the L0 triggers. The L0 muon triggers requires coincidence hits in multiple layers of the muon detectors and a $p_T \geq 20$ GeV/c (L0 MU20) while the L0 electron triggers require a deposited $E_T \geq 18$ GeV (L0 EM18) in neighbouring EM calorimeter cells.

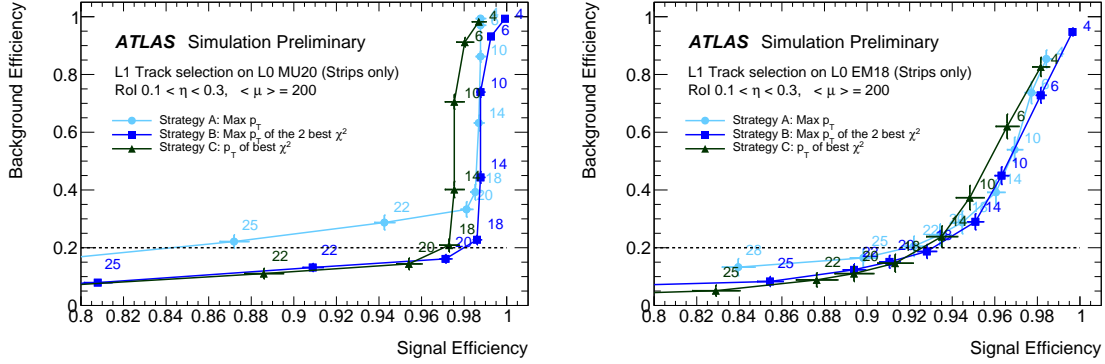


Figure 2: Single muon (left) and electron (right) efficiencies vs. background efficiencies for three track selection strategies as functions of a track p_T cut in the region of interest $0.1 \leq \eta \leq 0.3$, $0.3 \leq \phi \leq 0.5$. The number next to each marker signifies the p_T cut applied to one of the track candidates resulting from our fit, the candidate was selected either by highest p_T (light blue), highest p_T of the two candidates with best χ^2 (dark blue) or the candidate with the best χ^2 (black). The dashed line at background efficiency of 0.2 represents the target rate reduction. [3]

Table 1: Summary of the pattern recognition and track fitting performance on single muon and minimum bias events in the barrel region $0.1 \leq \eta \leq 0.3$, $0.3 \leq \phi \leq 0.5$, for two layer configurations: one with strip layers only and one where the innermost strip layer has been replaced by a pixel layer. The pattern matching efficiency, $\epsilon_{pattern}$, is defined as the fraction of single muon events with a matched pattern; the track fitting efficiency, ϵ_{fit} , is defined as the fraction of those events where at least one track fit is successful and has a $\chi^2 \leq 40$; and $\langle N_{fits} \rangle$ is the average number of fits in minimum bias events at a $\langle \mu \rangle = 200$ level of pileup interactions. [3]

Detector layers	single muon		min. bias		
	$\epsilon_{pattern}$	ϵ_{fit}	$\epsilon_{pattern}$	ϵ_{fit}	$\langle N_{fits} \rangle$
Strip layers only	99.4%	99.5%	92.5%	7.5%	114
Strip + 1 pixel layer	99.5%	99.7%	99.0%	61.5%	331

Table 2: The resolutions of the track parameters from the fit for single muon events in the barrel region $0.1 \leq \eta \leq 0.3$, $0.3 \leq \phi \leq 0.5$, for two layer configurations: one with strip layers only and one where the innermost strip layer has been replaced by a pixel layer. [3]

Detector layers	q/p_T [e/GeV]	ϕ [rad]	η	d_0 [mm]	z_0 [mm]
Strip layers only	0.003	0.001	0.002	0.3	1.7
Strip + 1 pixel layer	0.003	0.001	0.001	0.2	0.3

Including a pixel layer greatly improves the track parameter resolutions, especially z_0 , details are presented in Tables 1 and 2. These results are promising for the prospect of triggering on more complex topologies such as τ decays, multi-jet events etc.

4. Latency estimation

The latency of the L1Track trigger is comprised of several components, the main ones being the readout of the Front-End chips and the trigger processing time. The readout latency of the strip modules has been evaluated using a discrete event simulation of the detector and ABC130 ASICs [1], the readout chips for the silicon strip modules, linked to Hybrid Chip Controllers in a star layout with a stave link bandwidth of 320 Mbps [3]. These studies have shown that with a L0-priority, i.e. the modules in the RoI defined by the L0 trigger are given priority to be read out first, 99% of all regional readout requests can be completed within the 6 μ s latency, as shown in Figure 3. The equivalent studies for the pixel layers have not yet been performed. The pattern matching and track fitting step have not been evaluated with the discrete event simulation since a detailed hardware description is unknown at this time. The latency of this step can be approximated considering that modern FPGAs can perform a single track fit in 0.25 ns, meaning that a single chip can fit tracks from several RoI's with a latency of a few μ s. The AM chips used for the pattern matching are expected to have an input capacity of 200 MHz and the highest occupancy strip layer has 250 clusters on average, which puts the latency of the pattern matching step at a few μ s as well, furthermore the pattern matching will run in parallel on the data from the priority request while the full detector data is read out.

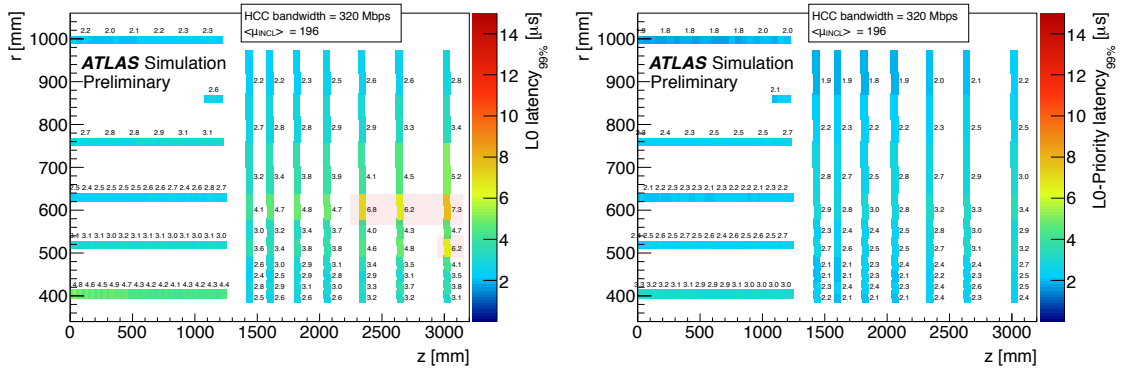


Figure 3: Detector map showing the latency within which 99% of all (left) non-prioritised full detector readout requests and (right) L0-Priority Regional Readout Requests (R3) can be completed. The full system delivers a rate of 1MHz of full detector data of which 10% are L0-Priority requests corresponding an R3 request. The chip hit occupancies correspond to a mean inclusive pileup interaction multiplicity, $\langle \mu_{INCL} \rangle$ of 196 interactions per bunch crossing, the upper limit expected with a bunch separation of 25 ns at an instantaneous luminosity $7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. [3]

References

- [1] ATLAS Collaboration, *Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment*, CERN, Geneva, CERN-LHCC-2012-022. LHCC-I-023,2012.
- [2] A. Annovi et al. *A new variable-resolution Associative Memory for High Energy Physics, Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA), 2011 2nd International Conference on*, pp 1-6, 2011
- [3] ATLAS Collaboration, Approved plots for the L1Track Trigger project, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L1TrackPublicResults>