

ATLAS Hardware-Based Endcap Muon Trigger for Future Upgrades

Yuya Mino^{ID}

Abstract—The large Hadron collider (LHC) is expected to increase its center-of-mass energy from 13 to 14 TeV for Run 3 scheduled from 2022 to 2024. After Run 3, upgrades for the high-luminosity-LHC (HL-LHC) program are planned and the operation will start in 2027, increasing the instantaneous luminosity to 5.0–7.5 times its nominal luminosity. Continuous upgrades of the ATLAS trigger system are planned to cope with the high event rate and to keep the physics acceptance. During the long shutdown period before Run 3, new detectors will be installed to improve the trigger performance. A new trigger logic, combining information from detectors located outside of the magnetic field and new detectors installed inside the magnetic field, will be introduced for Run 3 to reduce the trigger rate. In order to handle data from the various detectors, a new trigger processor board has been developed and the design is presented. During the upgrade for HL-LHC, the trigger and readout systems of the first level hardware-based part are planned to be upgraded. Full-granularity information will be transferred to the trigger processor board which enables more OFF-line like track reconstruction in the hardware-based system. To handle the full-granularity information and perform the hardware-based track reconstruction, the trigger processor board will be equipped with a field programmable gate array (FPGA) with hundreds of transceivers and a large memory. Expected performance for the hardware-based endcap muon trigger in Run 3 and HL-LHC will also be presented.

Index Terms—Data acquisition, field programmable gate arrays (FPGAs), high energy physics instrumentation, trigger circuits.

I. INTRODUCTION

THE large Hadron collider (LHC) [1] is the world's largest accelerator, colliding protons with a center-of-mass energy of $\sqrt{s} = 13$ TeV and peak instantaneous luminosity of $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. By the end of 2018, data corresponding to a total integrated luminosity of 140 fb^{-1} were collected. Future upgrades are planned for beyond standard model (BSM) searches and for Standard Model precision studies with higher energy and luminosity as shown in Fig. 1. Run 3 is planned to run from 2022 to 2024 and collect data at an increased center-of-mass energy of 14 TeV to extend the parameter space for various new physics models. The HL-LHC

Manuscript received October 26, 2020; revised February 12, 2021 and March 30, 2021; accepted May 23, 2021. Date of publication May 27, 2021; date of current version August 16, 2021. This work was supported by the ATLAS Collaboration.

The author is with the Department of Physics, Graduate School of Science, University of Kyoto, Kyoto 606-8502, Japan (e-mail: mino.yuya.33u@st.kyoto-u.ac.jp).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TNS.2021.3084098>.

Digital Object Identifier 10.1109/TNS.2021.3084098

TABLE I
TRIGGER RATE AND LATENCY

	Run 2	Run 3	HL-LHC
Trigger rate [kHz]	100	100	1000
Latency [μs]	2.5	2.5	10

The trigger rate and latency in each phase is listed. The trigger latency and rate will not be extended for Run 3, while there will be large improvement for the HL-LHC due to upgrades of the trigger and readout systems.

is planned to increase the peak instantaneous luminosity to $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and collect data for ten years, corresponding to a total integrated luminosity of 3000 fb^{-1} .

The ATLAS detector [3] is a general-purpose detector at the LHC, investigating a wide range of physics processes using data from proton-proton collisions at 40 MHz frequency. ATLAS uses a right-handed coordinate system with the z -axis pointing along the beam direction and x -axis pointing toward the center of the LHC ring. The two ends of the ATLAS detector are called the A-side and C-side, corresponding to the direction along the positive z -axis and negative z -axis respectively. Since the final recording rate of data from physics events is limited to approximately 1 kHz on average, ATLAS uses a two-level online event selection system to select events from interesting physics processes. The ATLAS trigger system consists of a hardware-based Level-1 (L1) trigger and a software-based high-level trigger (HLT). The L1 trigger uses a subset of information from the detector to make decisions and reduce the event rate to 100 kHz. The decision is made in $2.5 \mu\text{s}$ (called the L1 latency). The HLT trigger receives event candidates from the L1 trigger and refines the decision using the full detector information.

The ATLAS detector will include new hardware-based trigger capabilities for Run 3 and the HL-LHC. The trigger rate and latency for each phase are shown in Table I.

As shown in Table I, the reduction of the L1 trigger rate to 100 kHz will be kept for Run 3. New subdetectors will be installed to improve performance and reduce the trigger rate of current triggers. This period is called the Phase-I upgrade.

In order to cope with the higher luminosity of the HL-LHC, the trigger and readout systems of the hardware-based trigger are planned to be further upgraded. As shown in Table I, the trigger latency and rate will be increased to $10 \mu\text{s}$ and 1 MHz, respectively, by replacing the current system with high-bandwidth readout electronics. The increased latency will enable the use of much more sophisticated algorithms to

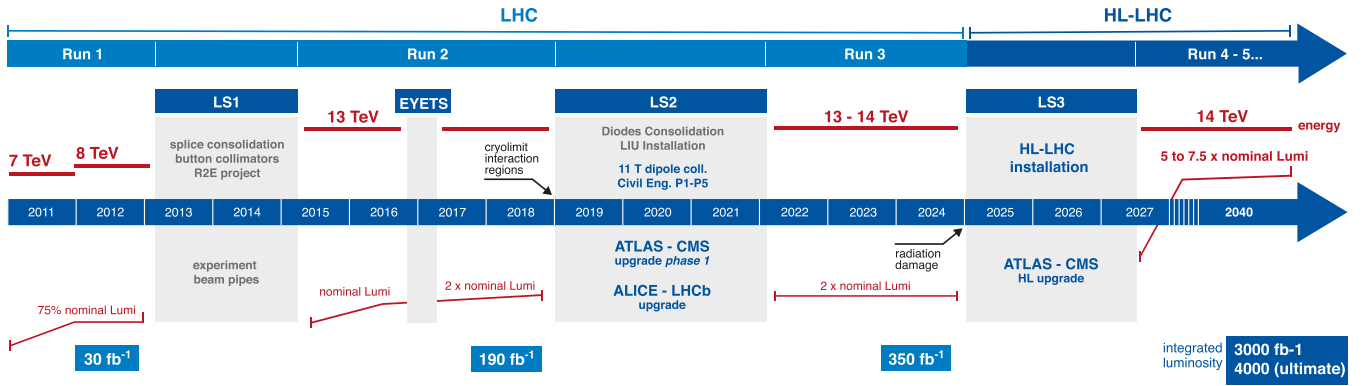


Fig. 1. LHC and HL-LHC upgrade plans [2]. Run 3 will start in 2022 and increase the center-of-mass energy to $\sqrt{s} = 14$ TeV during the run. The HL-LHC will start in 2027 and increase the luminosity to 5.0–7.5 times the nominal luminosity.

improve the trigger performance. Along with this upgrade, the L1 trigger will be renamed to Level-0 (L0) trigger after Run 3. This period is called the Phase-II Upgrade.

II. PHASE-I UPGRADE OF THE ATLAS LEVEL-1 MUON TRIGGER

The rate of the single muon trigger with the highest p_T threshold (“primary muon trigger”) in Run 3 is required to be reduced to 15 kHz considering other trigger and physics requirements [4]. With the current trigger scheme in Run 2, the trigger rate could not be reduced below 20 kHz at an instantaneous luminosity of $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Raising the p_T threshold would reduce the trigger rate but would also reduce the physics acceptance. Therefore higher rejection in the muon trigger is required to reduce the trigger rate. For example, Higgs-strahlung from W bosons is one of the production processes to determine the Higgs boson coupling to the gauge bosons and fermions at high precision. Leptons from the W boson are used to trigger this process and 93% of the muons have p_T larger than 20 GeV. If the p_T threshold is raised to 40 GeV, more than 30% of events from the Higgs-strahlung process would be lost [5]. Thus, the upgrade of the current trigger scheme is mandatory to keep the physics acceptance.

During Run 2, the L1 muon trigger rate was dominated by low p_T muons below the p_T threshold and charged particles emerging from the endcap toroidal magnets (“fake” muons). Fig. 2 shows the η distribution of trigger candidates in the L1 single muon trigger at p_T 20 GeV in Run 2 (L1_MU20). Charged particles from the endcap toroidal magnets are bent by the magnetic field, so only the positively charged particles point in the direction of the interaction point, for the A-side. As shown in Fig. 2, the number of track candidates from fake muons is larger in the A-side compared to the C-side because most of the fake muons originate from protons with positive charge.

Approximately 80% of the trigger candidates are from the endcap region ($|\eta| > 1.05$). In order to reduce triggers from low p_T and fake muons in the endcap region, new algorithms are implemented using information from the new detectors installed for Run 3.

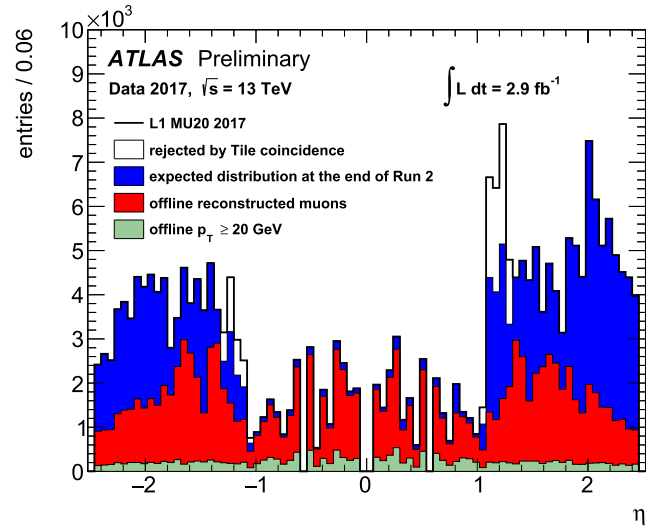


Fig. 2. η distribution of trigger candidates from L1_MU20 [6]. The blue and red regions show track candidates from fake and low p_T muons, respectively.

Fig. 3 shows the detectors used in the L1 endcap muon trigger for Run 3. Muons passing the toroidal magnetic field are bent and their position information measured by thin-gap chambers (TGC) [7] is used to measure their p_T . The TGC chambers are aligned in a disk-shaped structure called the TGC big wheel (TGC-BW) and the BWs are placed on both sides of the ATLAS detector. In the current trigger scheme, coincidence is taken on the front-end boards before sending hit information to the trigger processor board.

The fake muons leave hits in the TGC-BW which imitate hits from high p_T muons from the interaction point, as shown in Fig. 3. However, the fake muons do not leave hits in the detectors inside the magnetic field since they emerge directly from the toroidal magnet. Thus, the fake muons are reduced by combining hit information from the TGC-BW and detectors inside the magnetic field. Various detectors are placed inside the magnetic field and used in the L1 endcap muon trigger in Run 3: new small wheel (NSW) [5], TGC in the endcap inner station (TGC EI) [7], resistive plate chambers in the barrel inner station (RPC BIS78) [8], and Tile hadronic

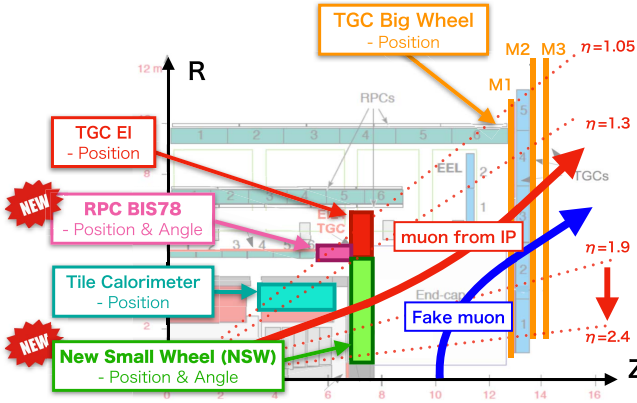


Fig. 3. Layout of the ATLAS muon spectrometer in Run 3. In the current system, detector coverage in the magnetic field is limited to $|\eta| = 1.9$. By installing the NSW, the detector coverage will be extended to $|\eta| = 2.4$.

calorimeter (TileCal) [9]. The NSW and RPCBIS78 are new detectors installed during the Phase-I upgrade. The NSW consists of eight layers consisting of sTGC (small-strip TGC) and micromegas [10] detectors. Combining hit information from the multiple layers enables reconstruction of tracks with an angular resolution of 1 mrad in the θ direction. Resolutions of the NSW are 0.005 and 10 mrad in the η and ϕ directions, which are much better than the resolution of the current detectors, 0.15 and 65 mrad, respectively. Coverage of the detectors inside the magnetic field will be extended from $|\eta| = 1.9$ to $|\eta| = 2.4$, which will result in a further reduction of the fake muons.

A. New Coincidence Logic in Run 3

The new trigger algorithms using new detectors in the L1 endcap muon trigger system in Run 3 are introduced in this section.

1) *Position Matching*: The p_T resolution of a track candidate found by the TGC-BW is limited by the detector granularity as shown in Fig. 4. The p_T resolution can be improved by refining the p_T using the position difference between the TGC-BW and new detectors in the η and ϕ directions because the new detectors have finer granularity. The rate of low p_T muons, which could not be reduced by the hit information in the TGC-BW, is reduced by requiring the position difference to have an appropriate value. The position difference information is handed over to a lookup table (LUT) implemented on a field programmable gate array (FPGA) and the corresponding p_T is immediately returned to refine the p_T decision of the TGC-BW. LUTs are defined depending on the position of the trigger candidate since the toroidal magnetic field is nonuniform and the correlation between p_T and position difference differs.

2) *Angle Matching*: Angular information from the new detectors enables further reduction of the rate of low p_T muons in addition to the position matching algorithm. The $d\theta$ information from the new detectors is defined by the angular difference between the angle of the reconstructed track and the angle of the straight line connecting the nominal interaction point and the measured track position. The $d\theta$ of the track

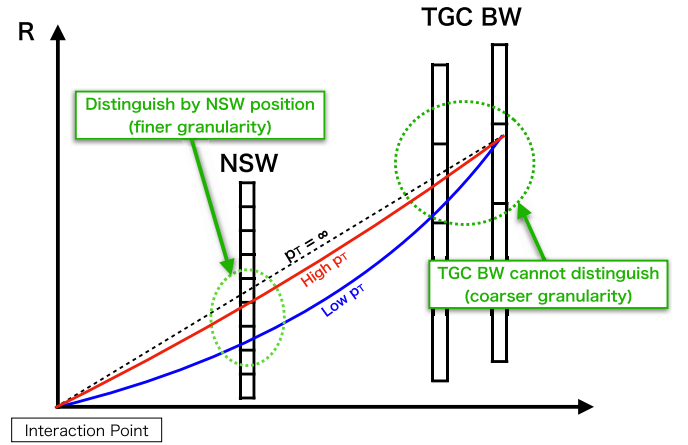


Fig. 4. Schematic of position matching algorithm. Low p_T and high p_T muons can be distinguished using position information of the new fine granularity detectors.

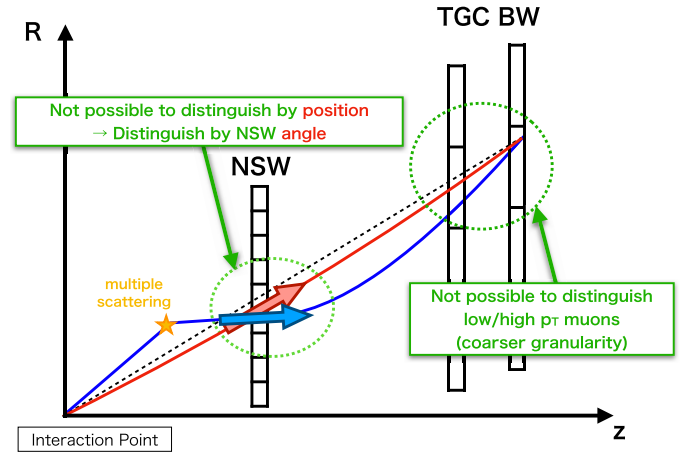


Fig. 5. Schematic of angle matching algorithm. A combination of the angle information and position difference enables further reduction of the remaining low p_T muons from the beam spot size and multiple scattering in detector materials.

should have a value near zero when the muon is produced at the detector center and enters the detector head-on. However, the interaction point can differ from the detector center within the beam spot size in the z -direction of approximately 10 cm. In addition, when multiple scattering occurs inside detector materials, especially in the calorimeter, there is a shift in the direction from the initial path. Low p_T muons imitating high p_T muons for these two reasons cannot be reduced by applying position matching as shown in Fig. 5. In this case, the $d\theta$ from the new detectors is different for low p_T and high p_T muons even with the same position difference. By combining the $d\theta$ information with the position difference, a further reduction of the rate of low p_T muons is achieved.

B. Hardware Design of the Sector Logic Board

The trigger processor board used in the muon trigger system is called the sector logic (SL) board. The endcap muon trigger system is grouped into 144 sectors and each SL board handles information from two sectors. As shown in Fig. 6, a VME 9U

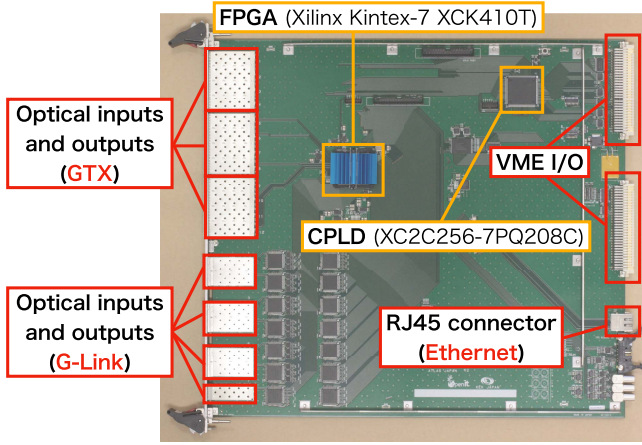


Fig. 6. Picture of the endcap SL board for Run 3. Xilinx Kintex-7 FPGA is implemented on a VME 9U board as the main processor.

board is used in Run 3 to control the registers on the FPGA by the VME bus.

In order to handle data from various detectors, the endcap SL board is required to have enough I/O ports. Due to the Phase-I upgrade, data from the NSW, RPC BIS78, and TGC EI will be received by eight GTX [12] transceivers. GTX is a multi-gigabit transceiver for Xilinx Kintex-7 FPGAs, supporting line rates up to 12.5 Gbps. Thirteen G-Link [11] connections, which require a large amount of I/O ports, are also used to receive data from the TGC-BW and TileCal. G-link has been used to transfer data during Run 2 and will continue to be used because the TGC-BW and TileCal will not be upgraded during the Phase-I upgrade. Two hundred and seventy-three I/Os are required for each G-Link connection because 21 I/Os are used for each G-Link channel. A large amount of memory is also required to implement new coincidence logic with new detectors.

Xilinx Kintex-7 FPGA [13] (XC7K410T-1FFG900) is selected as the main processor of the endcap SL board which meets these requirements. The XC7K410T has 500 I/O pins and 16 GTX [12] transceivers which is enough to handle data from all detectors. The XC7K410T has 795 Block RAMs (BRAMs) [14], which is about 20 times the memory compared to the FPGA used in Run 2. BRAM is a RAM module which provides storage for up to 36 Kb of data. LUTs for the coincidence logic will be implemented on BRAMs. Thus, the larger BRAM leads to an improved trigger performance.

A complex programmable logic device (CPLD) is also placed on the SL board to control the VME bus. Nonvolatile memory on the CPLD enables the configuration of FPGA at power up. The FPGA configuration by a byte peripheral interface (BPI) memory, containing data of the firmware design, is also controlled by the CPLD.

C. Firmware Implementation for the New Coincidence Logic

The new coincidence logic will be implemented on the FPGA with a fixed latency. The SL is required to send trigger candidates at 53 LHC clocks after the bunch crossing. Considering the arrival time of the track information from

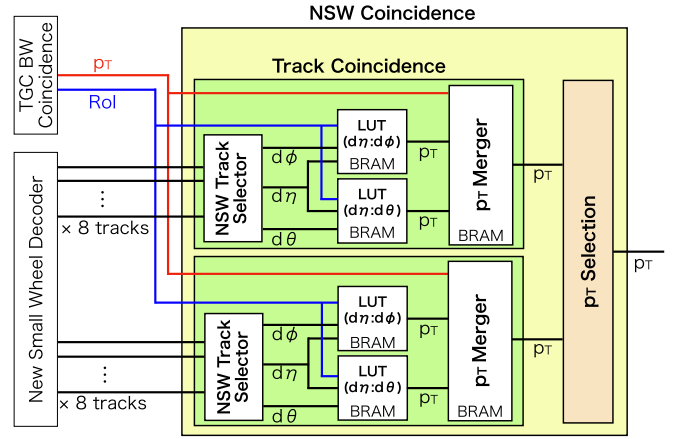


Fig. 7. Block diagram for the new coincidence logic. The track coincidence module and p_T selection module operate at 320-MHz clock to calculate the p_T for 16 track candidates in two LHC clocks.

the NSW and serialization, data transfer, and trigger logic before sending information to the subsequent boards, the new coincidence logic is required to be finished within two LHC clocks.

One track candidate from the TGC-BW will be compared with several track candidates from the new detectors. For example, the NSW will send 16 track candidates to the endcap SL at the maximum. Calculating the p_T using the NSW track candidates in parallel by placing 16 identical LUTs is the simplest implementation to achieve a short latency. However, this implementation makes the memory usage 16 times as large. To minimize the memory usage, processing the NSW track candidates in serial would be an alternative implementation. Still, the latency would be 16 times longer which would not meet the requirements.

To reduce the latency and minimization of memory usage, the firmware is designed as in Fig. 7. This firmware design consists of two modules operated at 320-MHz clock.

1) *Track Coincidence*: Two track coincidence modules are placed in parallel with identical LUTs and receive eight NSW track candidates per module. The NSW track selector receives eight NSW track candidates and sends track information to the LUTs for position and angle matching in serial. Sixteen tracks can be processed in one LHC clock by two track coincidence modules using the 320-MHz clock, which is eight times faster than the LHC clock. After the p_T extraction from the LUTs is finished for each track, the p_T merger module is used to choose which p_T to take as the final decision. Basically, the highest p_T is chosen to keep as many muons as possible for precise measurement by the software-based trigger. The decision could be changed to choose the lowest p_T when the trigger rate is higher than expected. The p_T merger is implemented by a BRAM to cope with various conditions.

2) *p_T Selection*: The p_T selection module receives two candidates per clock tick. To select the highest p_T among the 16 candidates, the p_T selection module selects the highest p_T among three candidates per clock tick: the two candidates from the track coincidence module and the candidate with

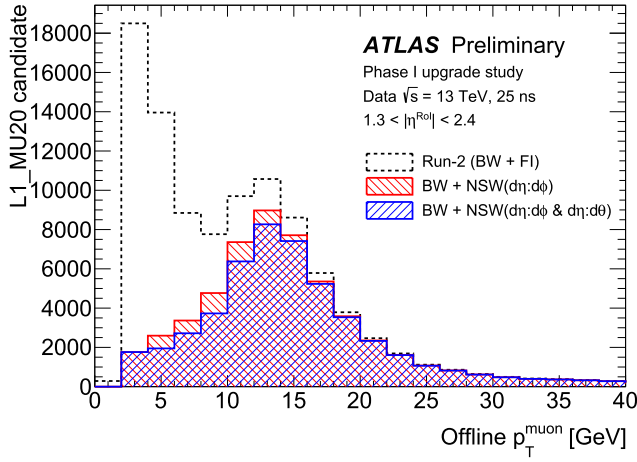


Fig. 8. p_T distribution of the offline reconstructed muons matched to the L1_MU20 for $1.3 < |\eta| < 2.4$ [6]. The dotted line indicates the p_T distribution of the Run 2 L1_MU20 candidates. The red (blue) line indicates the p_T distribution including the position (position and angle) matching algorithm.

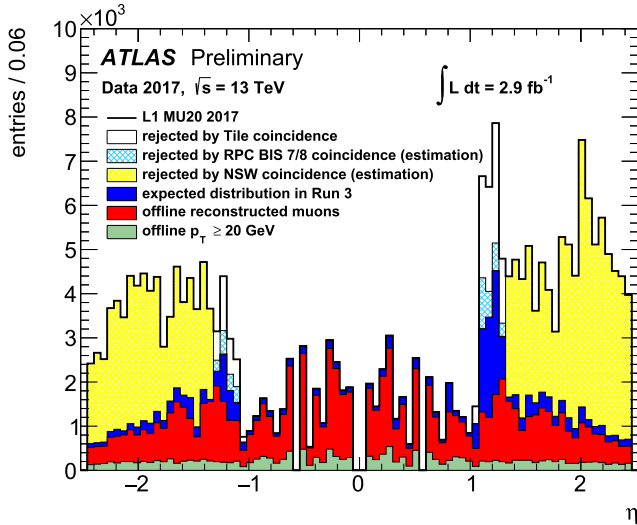


Fig. 9. η distribution of trigger candidates triggered by L1_MU20 expected in Run 3 [6]. The yellow and blue hatching area shows the number of reduced low p_T and fake muons by introducing the new coincidence logic with the NSW and RPCBIS78, respectively.

the highest p_T selected so far. A register is placed in the p_T selection module to keep the candidate selected so far. After receiving 16 candidates, the candidate kept in the register is sent back to the LHC clock domain.

D. Performance of Level-1 Endcap Muon Trigger

The rejection power for low p_T muons is estimated from a single muon MC simulation sample. A track reconstruction efficiency of the NSW is assumed to be 97% and included in the calculation. Higher reduction of the rate of low p_T muons relative to the Run 2 trigger is achieved by including the position and angle matching algorithm as shown in Fig. 8.

The rejection power for the fake muons is estimated from 2017 data since the fake muons cannot be modeled by MC simulations. Fig. 9 shows the η distribution of trigger

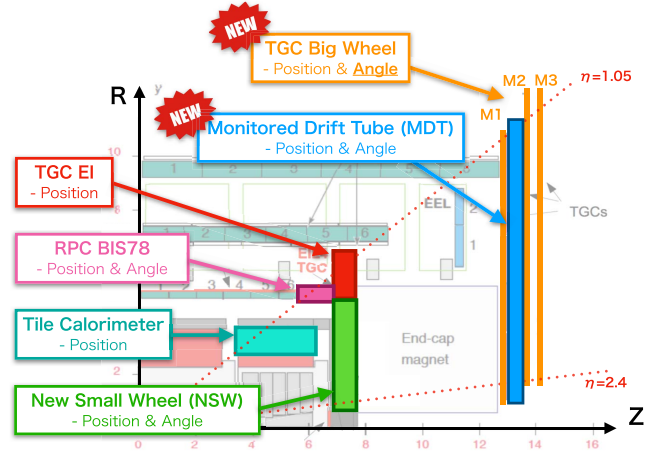


Fig. 10. Layout of the ATLAS muon spectrometer for the HL-LHC. Hardware-based track reconstruction will be implemented to obtain position and angle information with the TGC-BW. MDTS used in the current software-based trigger will be used in the hardware-based trigger for precise p_T determination.

candidates triggered by the L1 MU20 trigger expected in Run 3. 90% of the fake muons are reduced by the new inner muon detectors compared to the Run 2 logic. The expected trigger rate in Run 3 is 13 kHz, which meets the requirements for Run 3.

III. PHASE-II UPGRADE OF THE ATLAS LEVEL-0 MUON TRIGGER

Fig. 10 shows the detectors used in the L0 endcap muon trigger for the HL-LHC. In the new endcap muon trigger system, the SL will receive all TGC hit information from the new boards on the detector side. Hardware-based track reconstruction using all TGC hit information will be enabled to measure the p_T of track candidates with higher resolution.

During Run 2, monitored drift tubes (MDTs) were used for precise segment reconstruction only in the software-based trigger due to their long latency. In Phase-II, with its $10 \mu s$ latency, it will now be possible to include the slower MDTs into the L0, which was not possible in Phase-I because there the latency was not extended.

A. TGC Track Reconstruction

Tracks will be reconstructed with a pattern matching algorithm using all hits from the TGC-BW. The TGC-BW consists of three stations, M1, M2, and M3, which consist of three, two, and two TGC detector layers, respectively. Fig. 11 shows the main concept of the pattern matching algorithm. After receiving all TGC-BW hit information, a coincidence is taken to define a position in each station. A combination of the position in each station is compared with a predefined list of hit patterns, which has the position and angle information for high p_T muons.

In the current trigger scheme, at least two (one) hits in the M1 station and at least three (two) hits in the M2 and M3 stations together are required for wires (strips) on the front-end boards. Due to this requirement and assuming the

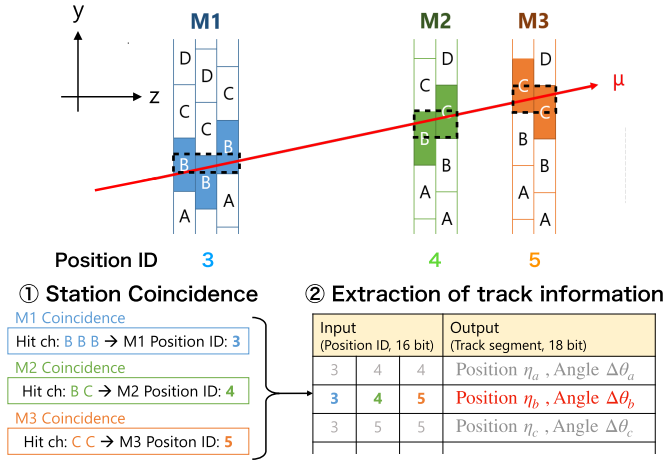


Fig. 11. Schematic drawing of the pattern matching algorithm. TGC tracks are extracted in two steps. First, the position ID is defined by taking the coincidence in each station. Second, hit patterns corresponding to the combination of the position ID are extracted from the pattern list.

muon single layer efficiency of 92.7% and 92.1% respectively for the TGC wires and strips, the muon efficiency of the BW-TGC wires and strips is limited to 95.6% and 96.9%, respectively. Combining these two efficiencies, the muon efficiency of the TGC-BW is estimated as 91.8% in the current system. In the HL-LHC trigger scheme, at least five (four) hits in the M1, M2, and M3 stations together are required. The muon efficiency of TGC wires and strips is improved to 98.2% and 97.8%, respectively with a looser coincidence. Combining these two efficiencies, the muon efficiency of the TGC-BW is estimated as 96.0%. The new trigger scheme is expected to improve the trigger efficiency by 4.2% compared to the current trigger scheme.

B. Hardware Design of the Sector Logic Board

Fig. 12 shows the hardware design of the SL board for the HL-LHC. In order to handle data from every TGC channel, the endcap SL is required to have enough I/O ports. A few hundred Mbits of memory is also required to implement pattern lists for track reconstruction. The Xilinx Virtex UltraScale+ FPGA [15] (XCVU9P-1FLGA2577E) is selected as the main processor of the endcap SL which meets these requirements. XCVU9P has 448 I/O pins and 120 GTY [16] transceivers which is enough to handle data from all detectors. GTY is a multi-gigabit transceiver for Xilinx UltraScale FPGAs, supporting line rates up to 32.75 Gbps in UltraScale+ FPGAs. XCVU9P has 2160 BRAMs and 960 UltraRAMs [17], which is about ten times the memory resource compared to the FPGA used in Run 3. UltraRAM is a new memory block with a large capacity, up to 288 Kb implemented in UltraScale+ families.

FireFly [18] and QSFP+ modules will be placed on the SL board to manage the data transfer with optical connections. FireFly modules are capable of handling 12 channels up to 16 Gbps per channel. QSFP+ modules are capable of handling four channels up to 10 Gbps per channel. QSFP+ modules are used to communicate with the Muon-to-Central-Trigger Processor Interface (MUCTPI) and FrontEnd

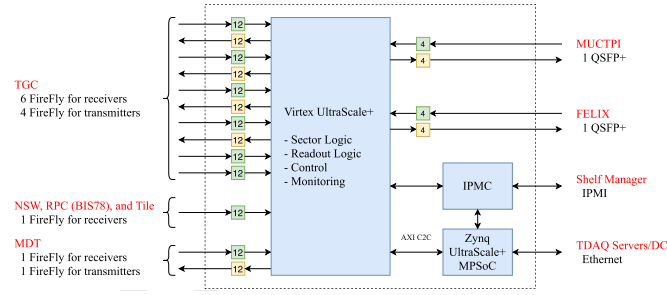


Fig. 12. Block diagram of the endcap SL for the HL-LHC. A Xilinx Virtex UltraScale+ FPGA will be implemented on a ATCA blade as the main processor.

Link eXchange (FELIX) because these modules are developed for Run 3 and will continue to be used for the HL-LHC.

The CERN-developed intelligent platform management controller (IPMC) for ATCA blades will be implemented for power monitoring, temperature monitoring, and for power up through the ATCA shelf manager.

A multiprocessor System-on-Chip (MPSoC) [19] device implemented on the SL board will be the interface for the central ATLAS run control, configuration, and monitoring of the status registers of the FPGA.

C. Firmware Implementation for the TGC Track Reconstruction

The TGC track reconstruction logic will be implemented on the FPGA with a fixed latency. The arrival of the TGC hit signals to the endcap SL is estimated to be $0.888 \mu\text{s}$ after the bunch crossing. The estimated latency of the TGC track reconstruction is $1.013 \mu\text{s}$ after the bunch crossing. Thus, the firmware design is required to reconstruct tracks in $0.125 \mu\text{s}$.

This algorithm is parallelly processed in subdivided small regions (“unit”) to reduce redundant pattern lists as shown in Fig. 13. Units are defined to include TGC hits from muons (both μ^+ and μ^-) with p_T as low as 4 GeV. This leads to a triangular-shaped region which consists of eight wire channels per layer in M3, 16 wire channels per layer in M2 and 32 wire channels per layer in M1. These units are subdivided into four regions (“subunit”) and one URAM block is allocated for each subunit to store pattern lists. A subunit is defined by dividing the eight wire channels consisting of a unit into four regions, two wire channels per layer in M3.

A firmware block diagram for the TGC track reconstruction is shown in Fig. 14. After receiving hit information from the TGC-BW for every bunch crossing at 40 MHz, a 160-MHz clock synchronous to the 40-MHz LHC clock is used to process the hits. The firmware block for TGC track reconstruction consists of three modules.

1) *Station Coincidence*: Coincidence of the TGC hits is taken in each station of the TGC-BW and coincidence results (“position IDs”) are output for each station. There are seven types of modules for the station coincidence, corresponding to the number of hits required in each station. For example, the “M1 2/3 coincidence” block requires hits in two layers with

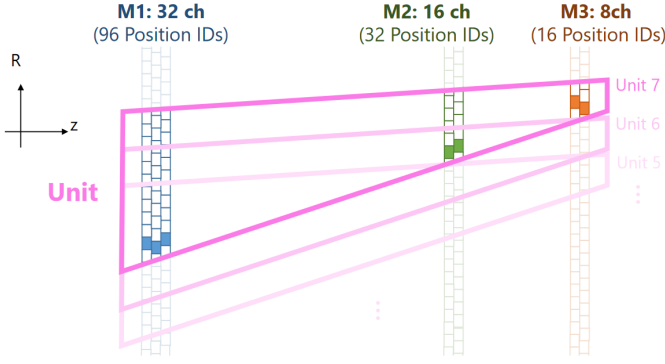


Fig. 13. Definition of the subdivided region ("unit") of the TGC-BW.

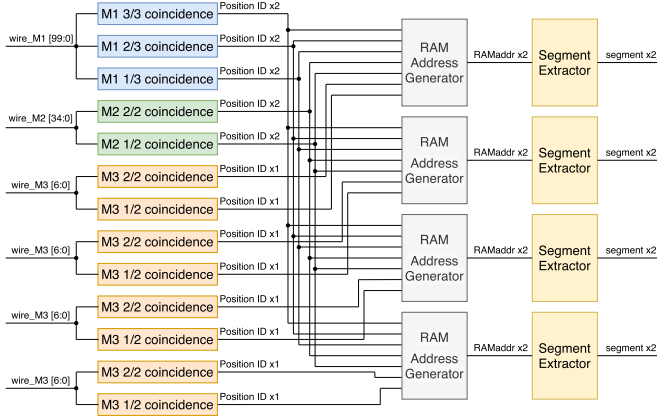


Fig. 14. Block diagram for the TGC track reconstruction.

no hit in the remaining one layer in the M1 station. Position IDs from the center of the unit region output by the M1 and M2 station coincidence modules are preferentially sent to the subsequent modules. For the M3 station coincidence modules, position IDs with smaller η regions are prioritized. This prioritization scheme is optimized to select track candidates with higher p_T . M1 and M2 station coincidence modules output two position IDs, and M3 station coincidence modules output one position ID.

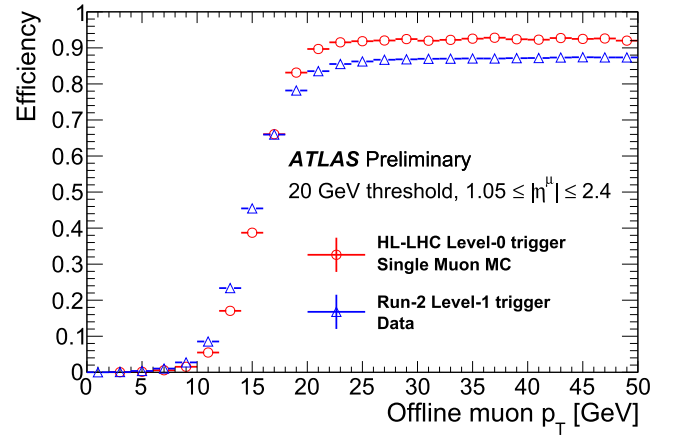
2) *RAM Address Generator*: The RAM address generator module receives position IDs from the station coincidence module and combines them to create a RAM address for the corresponding hit pattern. The RAM address is described in 12 bits using M1, M2, and M3 position IDs, which are described in 5, 5, and 2 bits, respectively. The considered patterns for the combination of position IDs are shown in Table II. Combinations of position IDs with a larger number of hits are prioritized and a maximum of eight hit pattern candidates are obtained per subunit. When the candidates have the same coincidence pattern, candidates with smaller η are preferentially selected.

3) *Segment Extractor*: The segment extractor module receives two RAM addresses per clock, using four clock ticks in total to receive eight RAM addresses per subunit. URAM blocks are configured in the true dual port mode, and track

TABLE II
DEFINITION OF COINCIDENCE PATTERNS

Coincidence pattern	Hit pattern			Fraction
	M1	M2	M3	
7/7	3/3	2/2	2/2	0.649
6/7A	2/3	2/2	2/2	0.124
6/7B	3/3	1/2	2/2	0.083
6/7C	3/3	2/2	1/2	0.083
5/7A	2/3	1/2	2/2	0.016
5/7B	2/3	2/2	1/2	0.016
5/7C	3/3	1/2	1/2	0.011
5/7D	1/3	2/2	2/2	0.008
Total				0.988

The coincidence patterns for TGC wire track reconstruction are listed in descending order of the fraction. The fraction of muons for each coincidence pattern is calculated by assuming that the hit efficiency for each TGC layer is 94%.

Fig. 15. p_T dependency of trigger efficiency for the L0 muon trigger with the HL-LHC scheme (red) and with the Run 2 scheme (blue) [20].

information corresponding to the RAM address is extracted from the RAM per clock tick.

D. Performance of the Level-0 Endcap Muon Trigger

The trigger performance of the L0 endcap muon trigger is evaluated. However, the precise information from the MDTs is not used in the evaluation. Further rate reduction is expected using the MDTs which improves the p_T resolution of the track candidates. Fig. 15 shows the expected efficiency of the new trigger algorithm with respect to offline muons in a single muon MC simulation sample. Compared to the Run 2 trigger scheme, higher efficiency in the plateau region and better rejection for the low p_T muons are obtained due to the looser coincidence and the improved angular resolution, respectively.

Fig. 16 shows the estimated trigger rate from Run 2 data taken with a random trigger to reproduce the higher luminosity expected in the HL-LHC. The trigger rate for a 20 GeV threshold is about 23 kHz, which constitutes only about 2.3% of the assumed total L0 trigger rate of 1 MHz.

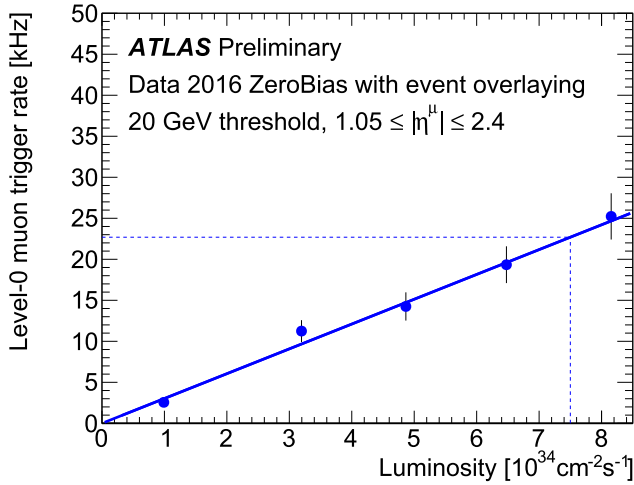


Fig. 16. Luminosity dependency of the estimated trigger rate of the L0 single muon trigger for a p_T threshold of 20 GeV [20]. The lowest luminosity point corresponds to the Run 2 data taken with a random trigger and the higher luminosity points are produced by overlaying the Run 2 events.

IV. CONCLUSION

A step-by-step upgrade of the hardware-based (Level-1, -0) endcap muon trigger is planned for Run 3 and the HL-LHC to handle higher luminosity. For Run 3, new detectors with finer granularity track information will be installed inside the toroidal magnetic field. A new coincidence logic using the position and angle information of the new detectors was suggested. The estimated L1 endcap muon trigger rate for a 20 GeV threshold is about 13 kHz at an instantaneous luminosity of $2.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, which meets the requirements for Run 3. The new trigger processor board has been produced for Run 3 to handle data from various detectors and implements the new coincidence logic. For the HL-LHC, the trigger and readout electronics will be replaced to extend the L0 trigger rate and latency. Using full-granularity information from the TGC-BW, a fast track segment reconstruction will be implemented on the SL board in Run 3. The new L0 endcap muon trigger scheme shows about 4% higher efficiency compared to the current trigger system. In addition, the estimated trigger rate for a 20 GeV threshold is about 23 kHz at an instantaneous luminosity of $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, which constitutes only about 2.3% of the assumed total L0 trigger rate. Further rate reduction is expected by adding the MDTs in the L0 trigger. The design of the SL board for the HL-LHC has been introduced to implement the TGC track reconstruction algorithm.

REFERENCES

- [1] L. Evans and P. Bryant, "LHC machine," *J. Instrum.*, vol. 3, no. 8, Aug. 2008, Art. no. S08001.
- [2] *HL-LHC : Project Schedule*. Accessed: Jan. 12, 2012. [Online]. Available: <https://project-hl-lhc-industry.web.cern.ch/content/project-schedule>
- [3] A. Collaboration, "The ATLAS experiment at the CERN large hadron collider," *Jinst*, vol. 3, Aug. 2008, Art. no. S08003.
- [4] ATLAS Collaboration. (2013). *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*. [Online]. Available: <https://cds.cern.ch/record/1602235>
- [5] ATLAS Collaboration. (2013). *New Small Wheel Technical Design Report*. [Online]. Available: <https://cds.cern.ch/record/1552862>
- [6] ATLAS Collaboration. *L1 Muon Trigger Public Results*. Accessed: Mar. 20, 2021. [Online]. Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L1MuonTriggerPublicResults>
- [7] ATLAS Collaboration. (1997). *ATLAS Muon Spectrometer: Technical Design Report*. [Online]. Available: <https://cds.cern.ch/record/331068>
- [8] L. Massa, "Proposal of upgrade of the ATLAS muon trigger in the barrel-end cap transition region with RPCs," in *Proc. Technol. Instrum. Part. Phys. PoS (TIPP)*, Jul. 2015, p. 117.
- [9] ATLAS Collaboration. (1996). *ATLAS Tile Calorimeter: Technical Design Report*. [Online]. Available: <https://cds.cern.ch/record/331062>
- [10] Y. Giomataris, P. Rebougeard, J. P. Robert, and G. Charpak, "MICROMEGAS: A high-granularity position-sensitive gaseous detector for high particle-flux environments," *Nucl. Instrum. Methods Phys. Res. A, Accel. Spectrom. Detect. Assoc. Equip.*, vol. 376, no. 1, pp. 29–35, Jun. 1996.
- [11] Agilent *HDMP-1032/1034 Transmitter/Receiver Chip-Set Data Sheet*. Agilent Technologies. Accessed: Mar. 20, 2021. [Online]. Available: <https://www.asc.ohio-state.edu/physics/cms/cfeb/datasheets/hdmp1032.pdf>
- [12] *7 Series FPGAs GTX/GTH Transceivers User Guide*, Xilinx, San Jose, CA, USA, Aug. 2018. [Online]. Available: https://www.xilinx.com/support/documentation/user_guides/ug476_7Series_Transceivers.pdf
- [13] *7 Series FPGAs Data Sheet: Overview*, Xilinx, San Jose, CA, USA, Sep. 2020. [Online]. Available: https://www.xilinx.com/support/documentation/data_sheets/ds180_7Series_Overview.pdf
- [14] *7 Series FPGAs Memory Resources*, Xilinx, San Jose, CA, USA, Jul. 2019. [Online]. Available: https://www.xilinx.com/support/documentation/user_guides/ug473_7Series_Memory_Resources.pdf
- [15] *UltraScale Architecture and Product Data Sheet: Overview*, Xilinx, San Jose, CA, USA, Mar. 2021. [Online]. Available: https://www.xilinx.com/support/documentation/data_sheets/ds890-ultrascale-overview.pdf
- [16] *UltraScale Architecture GTY Transceivers User Guide*, Xilinx, San Jose, CA, USA, Sep. 2017. [Online]. Available: https://www.xilinx.com/support/documentation/user_guides/ug578-ultrascale-gty-transceivers.pdf
- [17] *UltraRAM: Breakthrough Embedded Memory Integration on UltraScale+ Devices*, Xilinx, San Jose, CA, USA, Jun. 2016. [Online]. Available: https://www.xilinx.com/support/documentation/white_papers/wp477-ultraram.pdf
- [18] *Firefly: Application Design Guide*, Xilinx, San Jose, CA, USA, May 2021. [Online]. Available: <http://suddendocs.samtec.com/ebrochures/firefly-brochure.pdf>
- [19] *Zynq UltraScale+MPSoC Data Sheet: Overview*, Xilinx, San Jose, CA, USA, May 2021. [Online]. Available: https://www.xilinx.com/support/documentation/data_sheets/ds891-zynq-ultrascale-plus-overview.pdf
- [20] ATLAS Collaboration. *Level-0 TGC Trigger Performance of Trigger Algorithms in Software and Firmware Implementations*. Accessed: Mar. 20, 2021. [Online]. Available: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L0MuonTriggerPublicResults>