# ATLAS TDAQ upgrades for Phase-2

## F.Pastore[1] on behalf of ATLAS Collaboration

[1]Royal Holloway, University of London (United Kingdom)

E-mail: francesca.pastore@cern.ch

**Abstract.** The ATLAS experiment at CERN will be upgraded for the High-Luminosity LHC, with collisions due to start in 2029. In order to deliver an order of magnitude more data than previous LHC runs, 14 TeV protons will collide with an instantaneous luminosity of up to $7.5 \times 10^{34}$ cm$^{-2}$s$^{-1}$, resulting in higher pileup and data rates. This increase brings new requirements and challenges for the trigger and data acquisition system (TDAQ), as well as for the detector and computing systems.

The design of the TDAQ upgrade comprises the data acquisition, which combines custom readout with commodity hardware and networking to deal with 4.6 TB/s input, and the trigger system. The trigger is split into a hardware-based low-latency real-time trigger operating at 40 MHz, and the software trigger, called Event Filter, running at 1 MHz, which combines offline-like algorithms on a large commodity computing service, with the potential to be augmented by commercial accelerators. Commodity servers and networks are used as far as possible, with custom ATCA boards, high speed links and powerful FPGAs deployed in the low-latency parts of the system. Offline-style clustering and jet-finding in FPGAs, as well as accelerated track reconstruction are designed to combat pileup in the Trigger and Event Filter respectively.

This contribution will report recent progress on the design, technology and construction of the system. The physics motivation and expected performance will be shown for key physics processes.

## 1 ATLAS and the High-Luminosity LHC

The ATLAS [1] experiment [1] has planned a series of upgrades in view of the High-Luminosity LHC (HL-LHC) [2], for both detectors and the trigger and data-acquisition system (TDAQ) [3], the so-called Phase-2 upgrade. Planned to start proton-proton collisions at 14 TeV in 2029 in the so-called Run 4, the LHC will provide three times the current Run 3 instantaneous luminosity (up to 7.5 cm$^{-2}$ s$^{-1}$), to reach 4000 fb$^{-1}$ integrated luminosity, with an average 200 p-p interactions per bunch-crossing (pile-up). The harsher pile-up conditions, compared to the current pile-up of about 60 in Run 3, will require substantial changes in some detectors to maintain and possibly improve the selectivity: in particular two new detectors, a full-silicon tracker (ITk) [4, 5] and a High-Granularity Timing Detector (HGTD) [6], will be installed during Phase-2.

ATLAS will exploit the higher LHC luminosity both by increasing the data collection rate and by improving the data selection quality. The hardware trigger (Level-0, L0) running at 40 MHz will be improved and its decision rate, already a limit in Run 3, will increase from the current 100 kHz to 1 MHz. A full event-building at the L0 rate is foreseen, and the software trigger (Event Filter) will reduce the final data collection rate to 10 kHz (which is about five times the Run 3 output rate). The increased detector readout rate will force a renewal of all the Front-End electronics. The total DAQ throughput will increase from 0.2 TB/s to 4.6 TB/s, due also to the larger event size (from 2 to 4.6 MB/event), and

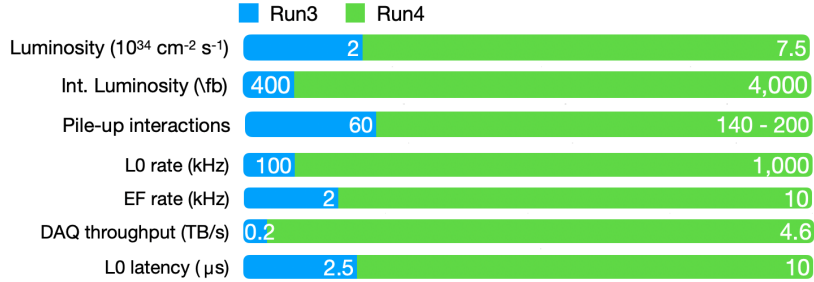| | Run3 | Run4 |
|---|---|---|
| Luminosity ($10^{34}$ cm$^{-2}$ s$^{-1}$) | 2 | 7.5 |
| Int. Luminosity (\fb) | 400 | 4,000 |
| Pile-up interactions | 60 | 140 - 200 |
| L0 rate (kHz) | 100 | 1,000 |
| EF rate (kHz) | 2 | 10 |
| DAQ throughput (TB/s) | 0.2 | 4.6 |
| L0 latency (μs) | 2.5 | 10 |

Figure 1: Summary of the evolution of the LHC parameters and of the ATLAS TDAQ requirements, showing the expected changes due to the HL-LHC upgrade from Run 3 to Run 4.

this will require substantial changes in the readout system and the DAQ architecture. The evolution of the LHC and the ATLAS DAQ parameters are summarized in Figure 1.

The goal is also to improve the selectivity at trigger level, in order to extend the ATLAS physics acceptance at the electro-weak scale both by maintaining low energy thresholds (possibly even lower than previous runs), as shown in Figure 2(left) for some Standard Model and SUSY signals compared to Run 1, and by improving the robustness of algorithms against pile-up. The event complexity and the large combinatorics will require longer trigger processing time, which do not always grow linearly with luminosity. An example is shown in Figure 2(right) for the processing time of a tracking trigger algorithm which increases faster than linear with the expected pile-up. For this reason the latency of the L0 trigger will be extended up to 10 $\mu$s and the software trigger capability will be largely expanded. The ATLAS TDAQ architecture for Run 4, shown in Figure 3, is set to achieve the above-mentioned
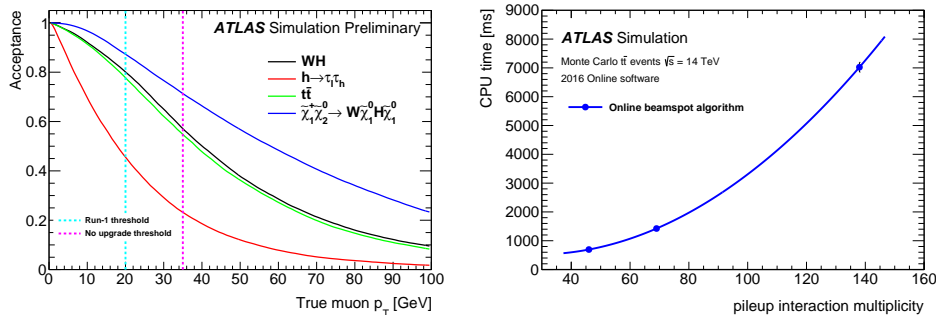


Figure 2: Left: signal acceptance as a function of generated muon $p_T$ for various signal samples at 14 TeV with one muon in the final state. The current estimated threshold for run 4 is 20 GeV. Generated with Pythia8 (WH), Powheg+Pythia8 (H$\to \tau\tau$), Powheg+Pythia6 (tt), MadGraph+Pythia8 (SUSY) [7]. Right: The trigger track reconstruction time for the beam-spot trigger for 14 TeV tt Monte Carlo events, measured on a 2.4 GHz Intel Xeon CPU. The software version is from 2016 online trigger system [8].

target parameters. It is described in the Technical Design Report (TDR) [9] and revisited with a TDR Amendment in 2022 [10], with a re-design of the Event Filter architecture. In the following sections more details are given for each component, showing their requirements and the status and plans for the Phase-2 upgrade.

## 2   Level-0 trigger for Run 4

The hardware trigger in Run 4 is still based on the identification of muons and calorimeter trigger objects (TOBs) in local regions called Region-of-interest. Both systems will be upgraded, with either new hardware or new firmware or both. A completely new component, called Global Trigger, is added
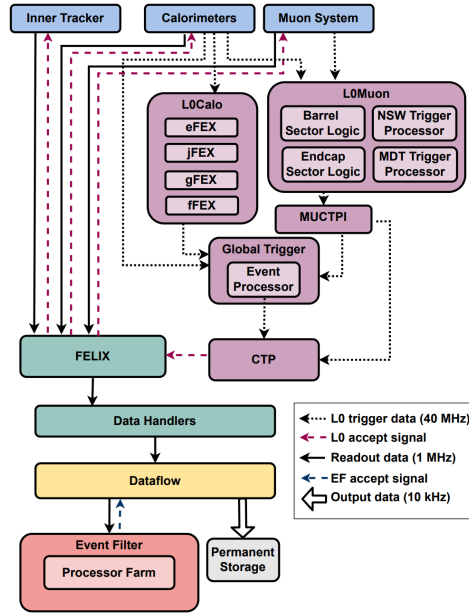
---

Figure 3: The ATLAS TDAQ architecture for Phase-2, showing the main components with different colors: the Level-0 hardware trigger (purple), the Event Filter software trigger (orange), the readout (green) and the dataflow (yellow). The flow of data and signals is also shown at the expected rates [10].

which performs quasi-offline reconstruction to improve the resolution of the selected TOBs and later combine them, exploiting the topological information of the event. A new Central Trigger Processor (CTP), which provides the final hardware trigger decision, sent to the Readout and the Event Filter, will be also built to accomplish the new requirements. The majority of the components of the L0 system are custom build, mostly using ATCA-based architectures [12] for internal data exchanges and modern powerful FPGAs for the trigger processing. The data transmission is mostly provided by optical links at 2.5-25 Gb/s.

The L0 calorimeter system is made of dedicated boards, able to identify calorimeter objects, with improved resolution: the coarse trigger towers of Run 2 will be replaced by 10-times finer granularity digital information of both Liquid-Argon electromagnetic calorimeter (already upgraded for Run 3) and Tile hadronic calorimeters. The new system, with the expected modifications of the upgrade, is shown in Figure 4(left). Four different boards execute the calorimeter Feature Extraction (FEX). The e(lectron)FEX, j(et)FEX, g(lobal)FEX are already installed and commissioned in Run 3. The hardware will be retained for Run 4, with significantly updated firmware to interface to the Global Trigger system and include the full-granularity hadronic calorimeter information. The f(orward)FEX is a new board, designed to identify forward electrons ($|\eta| > 2.5$) and forward jets ($|\eta| > 3.3$). The schematic capture and the board layout is now completed, so that a new prototype will be soon available. New optical connections are under design to communicate with the Global Trigger, which are custom designed, with the final design approval underway this year.

The L0 muon trigger will have extended functionalities, for example by including all muon detectors: in addition to the legacy trigger detectors running in Run 3, the MDT precision chambers will also be used in the trigger, to improve the momentum resolution on muons identified by RPC and TGC detectors as shown in Figure 4(right). The system will be enriched with multiple new boards: a new trigger processor for the end-cap New Small Wheel (NSW-TP), a new trigger processor for MDT precision chambers (MDT-TP) and a new off-detector trigger logic board called Sector Logic [11], common for all trigger detectors. The main difference with the current system will be that the data is streamed out from the Front-End boards and the entire trigger logic will be moved off-detector, thanks to the increased speed of the existing commercial links. While the final design is completed for the MDT-TP, with the recent start of pre-production, a first prototype of the Sector-Logic is already in place for a series of integration tests happening this year, mainly devoted to connectivity tests with all the different detector technologies and the new readout system. A second prototype, already being lay-down, will be realized as a result of these
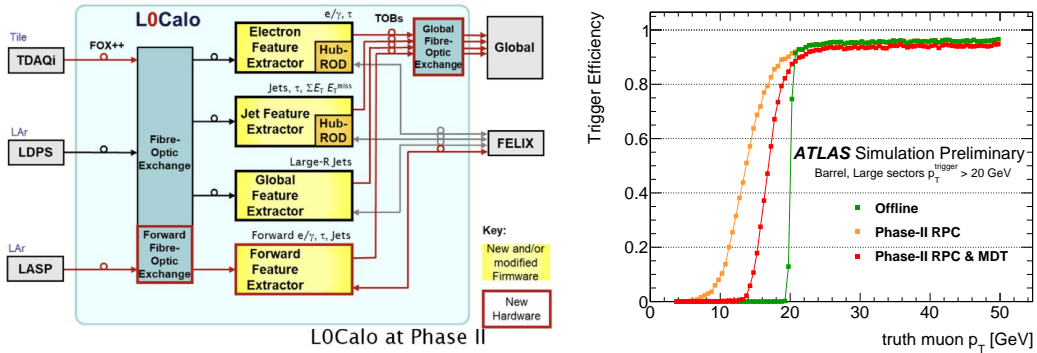
Figure 4: Left: The schema of the upgrades of the L0 calorimeter trigger, showing in yellow the new or modified components, and in red the new hardware components, together with all the connections to the Global Trigger and the new Readout system (FELIX). Right: Expected trigger efficiency for the Level-0 muon trigger for $p_T$ trigger threshold of 20 GeV, with and without the inclusions of MDT information, compared with offline. The efficiency is obtained for muon tracks in the Large sectors in the barrel. The values are obtained from single muon MC samples with no pile-up [13].

tests.

The Level-0 Global Trigger will process the L0 TOBs and additional high-granularity calorimeter information, using offline-like algorithms to refine identification of muons, calorimeter information (with topological-clusters), jets (with anti-kT algorithm) and execute pile-up subtraction. It will also extend the topological functionality already implemented in Run 3. It will consist of a farm of boards sharing the same hardware platform (called Global Common Module), and with different functionalities implemented in the firmware: data aggregation and time-multiplexing per bunch-crossing (MUX), algorithm execution per event in Global Event Processors (GEP), propagation of GEP outputs to the Central Trigger Processor (gCTPi). This system is a completely new component with very stringent requirements, in particular for the processing resources. A preliminary design review of the board has been completed last year and early tests on the new prototype are completed. The firmware development started well in advance and is now progressing, having passed the first reviews for the most critical algorithms (in particular the tau and pile-up suppression algorithms). The plan for this year is to integrate all the components in a slice test, once the single functionality tests are completed.

The Level-0 Central Trigger [2] takes the final L0 decision, possibly applying vetos and pre-scales factors, and distributes the timing signals synchronized with the LHC clock. A new Central Trigger Processor (CTP) board is under design for Run 4, which will be able to handle more trigger inputs (from 512 to 1024), sustain more bandwidth and apply more complex criteria. A preliminary design of the board is expected soon. The Muon-to-CTP-Interface (MuCTPI) board removes the overlaps between muons and calculate their multiplicities. The same board used in Run 3 will be reused, with upgraded firmware. The Trigger, Timing and Control (TTC) system network will be distributed via a new Local Trigger Interface (LTI) module, whose preliminary design is completed and a new prototype underway.

## 3 The Readout system for Run 4

To handle the increased data throughput, a new common interface between the detectors and the DAQ is being prepared. Replacing the current custom-based readout boards, it will be based on commodity servers and switched network. It will be composed of a FELIX (Front-End LInk eXchange) board, to collect data fragments from the detector front-end and sending them on the network, and the Data Handler, which performs more data collection, possibly organizes the data fragments for detector-specific preparation. FELIX [14] is a PCI-express card with a single FPGA, custom optical links for front-end communication and a custom TTC interface able to manage timing signals for the detectors. The first prototype of this board is already running in Run 3 for the new detectors and the new trigger components. For Run 4, other prototypes are being prepared: a second prototype (FLX-182) - Xilinx VM1802, PCIe Gen4, with 24 links up to 25 Gb/s - has been produced for integration tests, while a third prototype with

---

[2]More details on this system are presented by A. Koulouris at this conference

a more powerful FPGA (FLX-155) - Xilinx VP1552, PCIe Gen5, with up to 48 links - has been approved for production. The firmware is mature and expandable, and the full system is now starting integration tests with some front-end prototypes towards the final design approval.

## 4 Dataflow, network and online software

The dataflow system includes all components needed to aggregate the data fragments towards a full-event-building, to buffer the events for the software trigger (Event Filter) and send the selected events to permanent storage. The software prototype for Run 4 is based on the Run 3 system, including optimisations to scale the event building rate to 1 MHz. After a preliminary design completed in 2023, large scale tests have been performed using the Run 3 system, which enables testing the software and the distributed system at a scale comparable with the one expected for Run 4. Studies have been done with different network simulation models, to check the ability to expand the networking capabilities [15], focusing on traffic-shaping techniques for the network control and quantifying the necessary switch buffers. The online software framework glues the whole TDAQ system together for common configuration, control and monitoring. It will also undergo to a major upgrade. A prototype based on Kubernetes [16] as farm orchestrator has already been proven successful on large scale tests that included the full Run 3 farm, with more than 2600 nodes. This choice has great advantages being an open-source platform to automate deployment, management and scaling with containerised applications. The scaling of the Kubernetes cluster size is still a research topic [17] and these tests represent one important goal for its application.

## 5 The software trigger evolution for Run 4: Event Filter

The software trigger will run at the full event-building rate (1 MHz) and will not include the regional readout and reconstruction currently adopted in Run 3. The Event Building and the Event Filter (EF) will run on the same server within a farm of heterogeneous commodity processors. This trigger will make use of offline-like algorithms for reconstruction and selection. To reduce the size of the farm it could make use of accelerators, like GPUs and/or FPGAs, helping the CPU processors. Studies have been performed to reduce the processing time of the most resource-demanding algorithms and to demonstrate the feasibility of a heterogeneous processing with single algorithms running on different platforms. For example a fast tracking algorithm with performance very close to offline has shown an eight-fold speed-up on CPU [10], while running topological calorimeter cell clustering on GPUs showed a factor 12 speed-up compared to the offline version [18]. Many Neural Network (NN) approaches (like GNN, CNN, RNN) are being studied on multiple platforms and look promising, with vivid interest on the implementation on FPGAs, also in conjunction with other LHC experiments. The choice of the technology to adopt for the Event Filter is still not defined, and the final decision is expected for next year. A development phase is starting for building demonstrators to investigate the use of these accelerators, focusing in particular on the tracking algorithms which are the main resource consumers.

A project, called EF Tracking, is dedicated to the assessment of the best technology choice for Event Filter in the tracking realm. It collects all of the most recent developments of tracking algorithms on accelerators, comparing CPU, GPU and FPGA performance in terms of tracking efficiency and resolution on one side and timing, power, cost on the other. It allows to coordinate the developments towards common FPGA family (AMD/Xilinx) and GPU language/API for the benefit of the full ATLAS experiment. Options being explored include using commodity boards with multiple technologies (e.g. CPU and GPU or CPU and FPGA), each adapted to execute different steps of the tracking workflow (split into track seeding, pattern recognition, track fitting and ambiguity removal). Neural Network options are also investigated, in particular the Graphical NN has shown good results for many use-cases. The use of High Level Synthesis (HLS) language is also being explored as side developments of modern technology toolkits. All these developments will be included in a common interface to the ATLAS software, called ACTS (A Common Tracking Software) [19], which is an experiment-independent toolkit dedicated to tracking.

## 6 Conclusions

The Phase-2 upgrade in the ATLAS Trigger and DAQ systems is a very active area of developments. The Level-0 systems are focusing on finalising their design and are progressing towards the integration of prototypes that already exist, with major effort in the firmware development. The readout and DAQ have developed prototypes that are progressively under scaling tests, still ready to adopt the latest products available on the market. The Event Filter software trigger is studying the best options for use of accelerators, with extensive R&D for both algorithms and hardware choices. In parallel, detailed simulation studies are on-going to set the ATLAS physics plans for the HL-LHC phase, and to ensure that

the performance goals are met with the designed TDAQ system, for example by enabling updated and new trigger algorithms and including them in the Run 4 trigger menu selections. A detailed integration plan is being formed, to coordinate all these efforts, starting now through to upcoming years of installation and commissioning.

## References

[1] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", 2008 JINST 3 S08003 1-407

[2] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont and L. Rossi, "High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report," doi:10.5170/CERN-2015-005

[3] ATLAS Collaboration, "ATLAS TDAQ Phase-II Upgrade: Technical Design Report", ATLAS-TDR-029; CERN-LHCC-2017-020, `https://cds.cern.ch/record/2285584`

[4] ATLAS Collaboration, "ATLAS Inner Tracker Strip Detector: Technical Design Report", ATLAS-TDR-025; CERN-LHCC-2017-005, `https://cds.cern.ch/record/2257755`

[5] ATLAS Collaboration, "ATLAS Inner Tracker Pixel Detector: Technical Design Report", ATLAS-TDR-030; CERN-LHCC-2017-021, `https://cds.cern.ch/record/2285585`

[6] ATLAS Collaboration, "A High-Granularity Timing Detector (HGTD) in ATLAS: Performance at the HL-LHC", ATL-LARG-PROC-2018-003, `https://cds.cern.ch/record/2302827`

[7] ATLAS Collaboration, "TDAQ Phase-II Simulation Plots", ATL-COM-DAQ-2016-059, `https://cds.cern.ch/record/2155664`

[8] ATLAS Collaboration, "Run 2 HLT tracking timing plots for approval", ATL-COM-DAQ-2015-148, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HLTTrackingPublicResults`

[9] ATLAS Collaboration. "ATLAS TDAQ Phase-II Upgrade: Technical Design Report", ATLAS-TDR-029, `https://cds.cern.ch/record/2285584`

[10] ATLAS Collaboration, "Technical Design Report for the Phase-II Upgrade of the ATLAS Trigger and Data Acquisition System - Event Filter Tracking Amendment", ATLAS-TDR-029-ADD-1, `https://cds.cern.ch/record/2802799`

[11] Y. Mitsumori on behalf of the ATLAS Collaboration, "Sector logic development for the ATLAS Level-0 muon trigger at HL-LHC", `https://dx.doi.org/10.1088/1748-0221/18/02/C02019`

[12] AdvancedTCA, `https://www.picmg.org/openstandards/advancedtca/`

[13] ATLAS Collaboration, "L0 Muon Trigger Public Results", `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L0MuonTriggerPublicResults`

[14] ATLAS Collaboration, "FELIX: the Detector Interface for the ATLAS Experiment at CERN", EPJ Web Conf.251 04006, `https://cds.cern.ch/record/2814356`

[15] E.Pozo Astigarraga et al., "Benchmarking Data Acquisition event building network performance for the ATLAS HL-LHC upgrade", ATL-DAQ-PROC-2023-009, `https://cds.cern.ch/record/2872107`

[16] Kubernetes, `https://kubernetes.io`

[17] OpenAI, "Scaling Kubernetes to 7500 nodes", `https://openai.com/index/scaling-kubernetes-to-7500-nodes/`

[18] ATLAS Collaboration, "GPU acceleration of the ATLAS calorimeter clustering algorithm", ATL-DAQ-PROC-2022-002, doi 10.1088/1742-6596/2438/1/012044, `https://cds.cern.ch/record/2802139`

[19] ACTS, `https://acts.readthedocs.io/en/v9.0.0/`