The ATLAS Fast Tracker and Tracking at the High-Luminosity LHC

N. Ilic on behalf of the ATLAS collaboration

Stanford University
450 Serra Mall, Stanford, CA 94305, U.S.A.

E-mail: Nikolina.Ilic@cern.ch

ABSTRACT: The increase in centre-of-mass energy and luminosity of the Large Hadron Collider makes controlling trigger rates with high efficiency challenging. The ATLAS Fast TracKer is a hardware processor built to reconstruct tracks at a rate of up to 100 kHz and provide them to the high level trigger. The tracker reconstructs tracks by matching incoming detector hits with pre-defined track patterns stored in associative memory on custom ASICs. Inner detector hits are fitted to these track patterns using modern FPGAs. This proceeding describe the electronics system used for the massive parallelization performed by the Fast TracKer. An overview of the installation, commissioning and running of the system is given. The ATLAS upgrades planned to enable tracking at the High-Luminosity Large Hadron Collider are also discussed.

KEYWORDS: Data processing methods; Particle tracking detectors; Trigger concepts and systems (hardware and software)
1 Introduction

ATLAS [1] is one of two multi-purpose detectors at the Large Hadron Collider (LHC) built for a wide range of physics. Throughout its operation the ATLAS trigger system has efficiently selected interesting events, contributing to important results, including the discovery of the Higgs boson. Starting in 2015, the LHC is colliding protons at a higher center of mass energy and luminosity resulting in an increase in the mean number of simultaneous proton-proton interactions. This increase in luminosity and mean number of interactions during LHC operation is shown in figure 1. During the 2021 to 2023 period of LHC operation, the luminosity and number of simultaneous interactions will further increase. In these environments, controlling the trigger rates and selecting interesting physics events with a high efficiency will be challenging. The ATLAS Fast Tracker (FTK), discussed in section 2, is being installed and commissioned in order to cope with the challenge of selecting interesting physics events in the increasingly difficult environment. Another set of tracking and electronics upgrades are being developed in order to deal with even higher luminosity that will be delivered at the High-Luminosity LHC in 2026. The ATLAS tracking at High-Luminosity LHC is discussed in section 3.

2 The ATLAS Fast Tracker

The ATLAS trigger system consists of two levels: a hardware-based Level 1 (L1) trigger and a software-based High-Level trigger (HLT) [3]. At each level uninteresting events are rejected in order to efficiently reduce the amount of data in the final storage. The L1 trigger uses information from the muon tracks and electromagnetic and hadronic clusters to identify interesting regions of the detector containing high energy deposits, called regions of interest (RoI). The HLT further rejects events by utilizing tracks in the RoI from the inner detector, which consists of a silicon based pixel detector (Pixel), silicon semiconducting strip tracker (SCT) and transition radiation tracker (TRT) [4]. In the final step, the HLT decides which events to retain by reconstructing events using the energy deposits and tracks from the entire detector. The L1 trigger receives data from
In the current trigger system, tracks from the entire detector are not considered until the last stage of the software-based HLT processing, after many events have already been discarded. The Fast TracKer (FTK) is a new hardware system which will perform full track reconstruction after the L1 trigger, enabling the HLT to have access to tracks in the entire silicon detector at an earlier event selection stage [6]. By having access to all of the tracks earlier in the event selection stage, the trigger system can more efficiently reject uninteresting events. This greatly benefits physics analyses containing states identified using detailed tracking information, such as tau leptons and b-quarks.

There are many examples of physics analyses which will benefit from the tracking improvements the FTK will bring. It is important to measure Higgs properties and determine whether they match Standard Model (SM) expectations. For example, to uncover whether the fermionic couplings are SM-like, final states containing b-quarks and tau leptons are important but have large backgrounds from light quark and gluon jets. Since b-jet and tau identification largely relies on accurate tracking, the FTK can improve these SM measurements as well as SUSY analyses containing these final states. In addition, the FTK can allow for the possibility of a track-based Missing Transverse Energy trigger, which is less sensitive to overlapping interactions and can benefit exotic and other analyses [7].

### 2.1 Operational principle

The FTK receives hits from the twelve layers of the ATLAS Pixel detector and SCT and groups them into clusters of nearby hits in order to reduce the data size. The clustered hits are then sorted into 64 regions and sent downstream for parallel processing. Upon entering the processing units the hits are stored in full resolution, meaning they contain all of their original cluster information. While the full resolution hits are stored on input FPGAs in the processing units (PU), the remaining FPGAs and ASIC chips process the hits by grouping them in coarser resolution segments. During this
For eight of the twelve layers, the processing units identify which super-strips the clustered hits belong to. The same eight layers are always used. The selected super-strips are then compared to Monte Carlo (MC) track patterns stored in associative memories. In standard computer memory, once a user supplies a memory address the RAM returns the data word in that address. By contrast, in associative memory the user inputs a data word and the entire memory is searched in order to find that word. Thus when the associative memory finds an MC track pattern that matches the selected super-stripe, it returns it and calls it a “road”. For each road the full resolution hits stored on the PU input FPGA are retrieved and a goodness of fit is determined using a $\chi^2$ test for eight layers. The eight-layer tracks are then combined with the hits from the remaining four layers and a full twelve-layer track fit is performed. The twelve-layer tracks are then sent to the HLT. Figure 2 shows a visual representation of the process.

2.2 Hardware components

An overview of the different hardware components is shown in figure 3. Data are received from the Pixel and SCT detectors by the ATLAS readout drivers (RODs). Dual output high-speed optical links split the incoming data which can then take two paths: through the current Raw Data ROBs to HLT processing, or through the FTK system and to HLT processing.

The input mezzanines (IM) are 128 boards that receive the inner detector data at the 100 kHz input event rate and cluster them. Using a combination of Spartan-7 and Artix-7 FPGAs, the IMs calculate the cluster centroid, shape and size and information that contain the charge of the particle.
Figure 3. A diagram showing how different hardware components within the FTK interact.

The FPGAs for all boards were chosen based on the design requirements of the system. The IM boards sit on the data formatter (DF) boards which sort the data into 64 regions using Xilinx Virtex-7 FPGAs. The final system will contain 32 DF boards in Advanced Telecommunications Computing Architecture (ATCA) shelves, distributing data to one another using the ATCA full-mesh backplane and inter-shelf fibres. A custom microcontroller is used for board powering, sensor control and remote FPGA programming. The top left of figure 4 shows the DF board with four IMs on it.

The processing units shown in the green background of figure 3 consist of associative memory boards (AMB), and auxiliary boards (AUX). The processing units will be inside VME crates which are already used in other ATLAS subsystems. Since the FTK has developed new custom technologies used for the processing units, VME crates were chosen in order to minimize the amount of new technological changes introduced at once.

The auxiliary (AUX) boards use Altera Arria-5 FPGAs which have been programmed to contain four functional blocks: the data organizer (DO), track fitter (TF), and hit warrior (HW). The DO receives hits from the data formatter for 8 of the 12 inner detector layers. It stores the hits preserving their full information, and also compares them to map of the detector segmented in super-strips. It matches hits to their corresponding super-strip identification (SSID) and sends the SSID to the AMBs.

The AMBs perform the SSID to MC track pattern matching for eight inner detector silicon layers. Each AMB contains four local associative memory boards (LAMB) which each hold 16 associative memory chips [6]. The chips store $10^9$ patterns in HEP-specific content addressable memory (CAM), and perform $10^{14}$ word comparisons per second. The LAMBs are able to send data to all chips at every clock at a very large bandwidth. Approximately 750 serial links bring the hits to
the AM chips, with a total bandwidth of 200 GB/s. The pattern matching functionality uses a total bandwidth of 25 TB/s in the whole AMB system, which is composed of 128 AMBs. The buses can be viewed as 4 32-bit words for each pattern, leading to 128 thousand 4-word comparisons at each clock cycle. This is approximately 500 thousand comparisons per clock cycle, or $50 \times 10^6$ million comparison instructions per second (MIP). For 128 AMB boards, this means $40 \times 10^{10}$ MIPs for the whole system. This large number of instructions per second cannot be accomplished in any current CPU.

The chips also allows for matches with widths that vary layer by layer using ternary CAM cells. The LSB in the incoming data can use up to 6 ternary bits, each bit allowing for a value of 0, 1 or “X” meaning “don’t care” (DC) [8]. Using DC bits allows for efficient balancing of the match precision: low resolution patterns allow smaller bank sizes leading to less chips and lower cost but higher probability of random roads, while high resolution patterns lead to less fake roads but require more computing power. The DC bit allows for more precise patterns where this helps signal over noise separation, while less precise patterns can be used where this is not useful. One of the final 128 AMB boards with its four LAMBs is shown in the top middle of figure 4.

The AUX boards receive roads from the AMB, and retrieve the full resolution hits stored by the DO. The TF then does track fitting in eight layers using the full resolution hits. The TF board rejects tracks below a certain $\chi^2$ in order to reduce data to the next board. It also reconstructs tracks that do not have hits in certain layers using majority fit logic. The HW removes any duplicate tracks that might have been reconstructed before sending them to the second stage board (SSB). The final FTK system will consist of 128 AUX boards in VME crates. An AUX card is shown in the top right of figure 4.

The SSB receives eight-layer tracks from the AUX and hits from the remaining four layers from the DF. It uses Xilinx Kintex-7 FPGAs to extrapolate the eight layer tracks to the additional

---

**Figure 4.** Top Left: DF with four IMs on it. Middle Left: AMB with 4 LAMBs on it and 16 chips per LAMB. Top Right: AUX Card. Bottom Left: SSB. Bottom Right: fully installed ATCA shelf.
layers and does the full twelve-layer track fitting. The final system will consist of 32 SSB boards, one of which is shown in the bottom right of figure 4.

The FTK Level-2 Interface (FLIC) organizes tracks from the SSB and sends them to the High-Level Trigger Readout systems. It uses two custom microcontrollers and Xilinx Virtex-6 FPGAs. The final system will have 2 FLIC boards in ATCA shelves, one of which is shown in the bottom middle of figure 4.

2.3 Expected performance

The FTK reconstruction efficiency with respect to offline reconstruction is greater than 90% as can be seen from figure 5. The left of figure 5 shows that the offline and FTK variance of the impact parameter ($d_0$) as a function of the particle charge over its momentum closely match. The right of figure 5 shows that the efficiency with respect to offline for a muon and pion as a function of their momentum is greater than 90%. The small differences are due to using only silicon detectors, using a simplified clustering model and considering only tracks with momentum greater than 1 GeV. The simulation was carried out by emulating each of the firmware components for every board and running Monte Carlo physics events through them. Since most of the firmware is meeting its specifications the simulation is an accurate representation of the system’s performance. Currently every aspect of the event processing in the hardware is being compared to the simulation to ensure desired performance is met.

2.4 FTK commissioning and plans

The FTK is currently being installed and commissioned in the electronics room beside the ATLAS caverns. Currently 65% of the IM+DF boards are installed in their ATCA shelves. An example of a full DF ATCA shelf is shown in the bottom right of figure 4. A slice containing IMs, a DF and AUX board is currently integrated with the ATLAS detector and regularly taking data. The hits from this slice are written into the ATLAS event data stream. The firmware for all of the systems is complete and undergoing debugging with simulated data as well as real LHC data. The communication between the different boards is being stress tested in a separate laboratory that mimics the setup of the electronics room beside the cavern. The data flowing through all of the boards are being

![Figure 5](image-url)

Figure 5. Left: comparison of offline to FTK reconstruction for the variance of $d_0$ vs charge over momentum [5]. Right: FTK efficiency with respect to offline for a muon and a pion versus its momentum [5].
compared to the data expected from simulation. By the end of 2016 a part of the FTK system will be stably taking data with the ATLAS detector. In 2017 the full detector is expected to be commissioned using half of the processing units (64 AUX+AMB pairs), and the ATLAS trigger plans to use the FTK. By 2018 it is planned that the full detector will be commissioned with all of the processing units.

3 Trigger upgrades at high-luminosity LHC

In 2026 the LHC will resume operation after it has been upgraded to operate at a higher luminosity. The number of interactions per bunch crossing at the High-Luminosity LHC is expected to rise to typical values of 140, and an upper value of 200. The current single lepton triggers are not able to deal with this high rate at their current momentum thresholds. Simply increasing the momentum threshold below which events are rejected would lead to a large loss in physics events. Thus several upgrades, including tracking information at the hardware-based L1 trigger are being developed in order to help ATLAS deal with the increase in luminosity.

A new all-silicon tracker, called the ATLAS Inner Tracker (ITK) will replace the inner detector. In the calorimeter front-end electronics will be replaced in order to enable sending of finer granularity information to the trigger at a 40 MHz rate. The muon spectrometers will also be upgraded to include new resistive plate chambers to allow for a larger acceptance, and the use of muon drift tubes in the first level of the muon trigger. Finally, FTK++ will be an upgraded version of the FTK with newer FPGAs and a larger number of patterns. The FTK++ will provide global track reconstruction for tracks above 1 GeV, when requested by the HLT.

The upgrade for the High-Luminosity LHC is currently being shaped and various schemes are being studied [8]. One possibility for a new system involves a new L0 trigger and changes to the L1 trigger. An L0 trigger would be introduced and operate at 1 MHz with a 6 µs latency. The L1 track trigger would reduce the rate to 400 kHz with a latency of 24 µs. The L1 trigger would consist of an L1 track trigger, complementary to FTK++, and would perform tracking in regions of interest for tracks with momentum above 4 GeV. The L1 track trigger would feed the L1 Global trigger which would process finer granularity calorimeter information to improve the electron, photon, and tau jets measurement.

References

[1] ATLAS collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, 2008 JINST 3 S08003.
[2] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.
[3] ATLAS collaboration, ATLAS detector and physics performance: Technical Design Report, CERN-LHCC-99-014.
[4] ATLAS collaboration, ATLAS inner detector: Technical Design Report, CERN-LHCC-97-016.
[5] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FTKPublicResults.
[6] ATLAS collaboration, Fast TracKer (FTK) Technical Design Report, CERN-LHCC-2013-007.
[7] ATLAS collaboration, Measurement of the missing transverse momentum based on tracks in proton-proton collisions at $\sqrt{s} = 900$ GeV centre-of-mass energy with the ATLAS detector, ATLAS-CONF-2010-020.

[8] K. Pagiamtzis and A. Sheikholeslami, Content-addressable memory (CAM) circuits and architectures: a tutorial and survey, IEEE J. Solid-State Circuits 41 (2006) 712.

[9] ATLAS collaboration, ATLAS Phase-II Upgrade Scoping Document, CERN-LHCC-2015-020.