

FTK: The hardware Fast Tracker of the ATLAS experiment at CERN

Ioannis Maznas^{1,a} on behalf of the ATLAS Collaboration

¹ *Aristotle University of Thessaloniki*

Abstract.

In the ever increasing pile-up environment of the Large Hadron Collider, trigger systems of the experiments must use more sophisticated techniques in order to increase purity of signal physics processes with respect to background processes. The Fast Tracker (FTK) is a track finding system implemented in custom hardware that is designed to deliver full-scan tracks with p_T above 1 GeV to the ATLAS trigger system for every Level-1 (L1) accept (at a maximum rate of 100 kHz). To accomplish this, FTK is a highly parallel system which is currently being installed in ATLAS. It will first provide the trigger system with tracks in the central region of the ATLAS detector, and next year it is expected that it will cover the whole detector. The system is based on pattern matching between hits coming from the silicon trackers of the ATLAS detector and one billion simulated patterns stored in specially designed ASIC Associative Memory chips. This document will provide an overview of the FTK system architecture, its design and information about its expected performance.

Introduction

The ATLAS experiment [1], in order to select events for future analysis, uses a two-level trigger system. The first level or L1 trigger is a custom hardware based system that selects events using the energy depositions in the electromagnetic and hadronic calorimeters and on the momentum of muons detected in the muon spectrometer. After L1 the trigger rate is reduced from 40 MHz to 100 kHz. The second level is the High Level Trigger (HLT), running in software on a commodity server farm, that receives L1 accepted events and reduces the rate even further to 1.5 kHz.

During 2015-2016 the collision centre-of-mass energy has reached $\sqrt{s} = 13$ TeV[2] and the delivered luminosity increased from 4.2 fb^{-1} in 2015 to more than 38 fb^{-1} in 2016 (Figure 1b). Additionally, there was an increase in the number of interactions per bunch crossing (Figure 1a). In Run 3, when FTK will be fully integrated, it is expected in total a delivered luminosity greater than 300 fb^{-1} , LHC will deliver more than 50 fb^{-1} per year. At each bunch crossing, the number of overlapping proton-proton collisions (pile-up) will escalate to an average of 60 pile-up events. The trigger rates of many processes increase dramatically with additional pile-up. In order to cope with the high pile-up rates, the HLT exploits information coming from the ATLAS Inner Detector (ID) [3] that proves to be crucial as the fine resolution and high granularity of the detector allow the system to

^ae-mail: imaznas@cern.ch

decide whether an L1 accepted event is interesting using charged particle tracks. However, as the HLT is a CPU farm, extensive tracking becomes impossible given limited computing resources, as the detector occupancy increases dramatically with luminosity. As a counterweight, the FTK[4] will provide to the HLT reconstruction information by processing the full readout of the ATLAS semiconductor ID.

The FTK is designed to help the HLT by providing it with all tracks above $p_T > 1$ GeV within $100 \mu\text{s}$. Thus the FTK will allow the HLT to overcome hardware limitations and to leave time for more sophisticated event selection. The FTK is a highly parallel system of electronics that performs the tracking for all L1 accepted events, at 100 kHz, by using the readout of all twelve layers of the ATLAS ID, i.e. the four Pixel layers, including the newly installed Insertable B-Layer (IBL) [5], plus the eight Strip (SCT) layers.

In this document, the architecture of the system is presented, together with a description of the design and development of the FTK components, a report on the simulated performance of the whole system, and on the integration with the HLT.

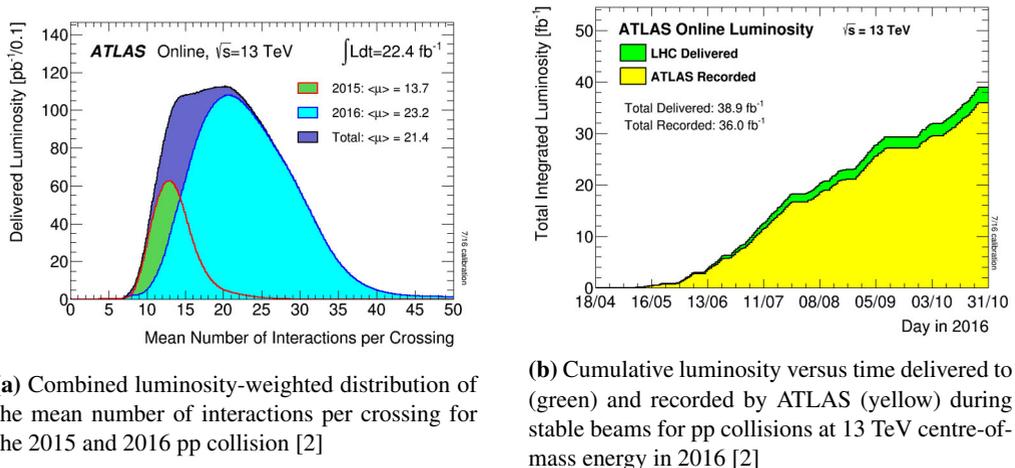


Figure 1: ATLAS luminosity in 2016

1 Architecture of the FTK System

Figure 2 depicts a sketch of the FTK architecture. FTK uses the Associative Memory (AM) chip, which is a custom designed ASIC, to perform pattern matching, along with numerous FPGAs for the detailed track fitting and other needed functionalities. Data from the ID Read-Out-Drivers (RODs) are transmitted to the FTK Input composed of Input Mezzanines (IM) and the mother Data Formatter (DF) boards. The IMs receive the data and perform one- or two-dimensional cluster finding in the SCT or the Pixel/IBL detectors, respectively. The DFs reorganize the ID data into projective $\eta - \phi$ towers and distribute the cluster centroids to the respective Processing Units, where pattern matching and track fitting take place. The division of the FTK into $\eta - \phi$ towers is done to achieve maximum parallelization. The inefficiencies generated by this segmentation at the tower boundaries are cancelled by an overlap in η and ϕ that takes into account the curvature of the particle tracks and the

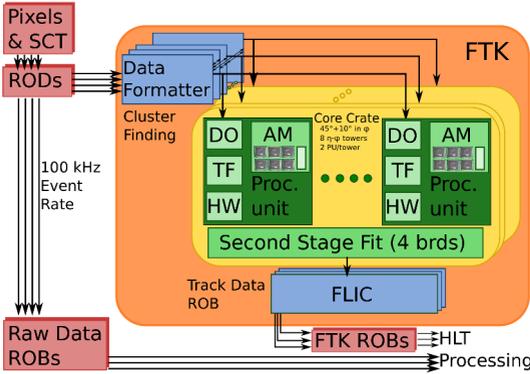


Figure 2. Functional diagram of the FTK components. DF reorganizes hits coming from ID Read-Out Drivers and propagates it to the DO. The DO organizes hits for pattern matching and sends it to the AMB. Track Fitter calculates track parameters for patterns matched. Hit Warrior removes duplicate tracks. Second Stage Fit (on SSB) extrapolates tracks for twelve ID layers. FLIC re-formats FTK output to ATLAS global coordinates and sends it to Read Out Buffers [4]

beam's luminous region in z . In total, 64 $\eta - \phi$ towers are used; 16 in ϕ (azimuthal angle) and 4 in η (pseudorapidity $|\eta| < 2.5$). There are 8 $\eta - \phi$ towers in a crate. The overlap between neighbouring crates is 10° in ϕ but there is also overlap between boards of the same crate.

Data Organizers (DO) on the AUX receive the hits coming from DFs and convert them to coarser resolution hits (the so-called Super Strips) which are appropriate for the pattern matching performed by the AM chips. The Associative Memory boards (AMB) contain roughly 8.2 million patterns each in the AM chips, corresponding to the possible combinations of Super Strips in the eight silicon layers. The AM is a massively parallel system in which all patterns see each silicon hit nearly simultaneously. As a result, pattern recognition in FTK is complete shortly after the last hit has been transmitted from the ID RODs. The matched patterns are referred to as "roads" containing track candidates. When a pattern has seven or eight layers matched, the AM sends the road identifier back to the DO which fetches the associated full resolution hits and sends them, together with the road identifier, to the Track Fitter (TF). The TF has access to a set of constants for each detector sector, which consists of eight physical silicon modules, one for each layer. The TF calculates track parameters linearly based on these pre-calculated constants using the formula:

$$p_i = \sum_j c_{ij} x_j + q_j, \quad \forall i = 1 \dots 5, \quad (1)$$

where p_i represents the five track parameters, x_j are the cluster local coordinates and c_{ij}, q_j are the constants used in the linearized formula. Following fitting in each road, duplicate track removal (the Hit Warrior or HW function) is carried out among those tracks which pass the χ^2 cut.

When a track passes the first stage criteria, the road identifier and hits are sent to the Second Stage Boards (SSB). The SSBs receive from the DFs the cluster centroids in the four layers not used in the previous stage (Figure 3). The track is extrapolated into the four additional layers, nearby hits are found, and a full 12-layer fit is carried out. A minimum of three hits in these four layers is required. Duplicate track removal is again applied to the tracks that pass the χ^2 test, but now tracks in all roads are used in the comparison. The SSBs also share track data with each other for $\eta - \phi$ overlap removal. There are three main functions running on different FPGAs: the *Extrapolator* that computes likely positions of hits in the other four SCT layers, the *Track Fitter* that determines the helix parameters and the track quality χ^2 for road hits of all 12 silicon layers, and the *Hit Warrior*. SSB output tracks consisting of the hits on the track, the χ^2 , the helix parameters, and a track quality word that includes the layers with a hit are sent to the FTK-to-Level2 Interface Crate (FLIC). The FLIC performs the look-up for the local FTK to global ATLAS hit identifiers and sends them to the Readout System (ROS), sending them to the High Level Trigger via the ATLAS ReadOut System.

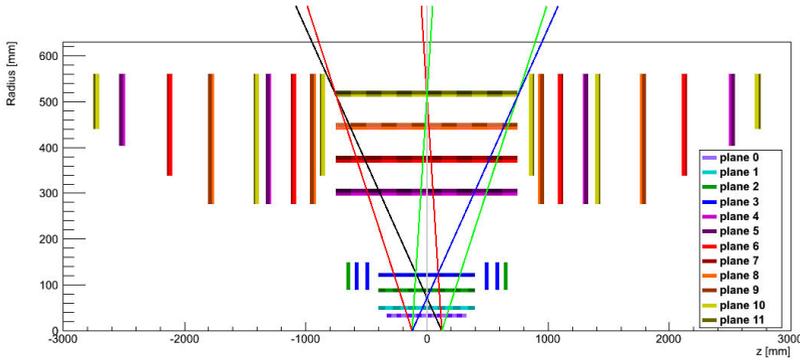


Figure 3: The assignment of barrel layers and end-cap disks to FTK logical layers. Layers 0-3 are pixels; the rest are SCT. The division into 4 η regions with the appropriate overlap is indicated by the thin colored lines: two endcap regions, one to the left of the black line and the other to the right of the blue line; and two barrel regions, one between the red lines and the other between the green lines.[4]

2 System Design and Implementation

FTK as configured for a luminosity of $3 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ consists of 13 crates; 8 VME "core crates" (Figure 4) and 5 ATCA shelves. The DF boards occupy 4 crates. These are ATCA shelves with full-mesh backplanes to facilitate the massive reorganization of the silicon hits from the readout optical fibres to the $\eta - \phi$ tower organization of FTK. As the AMBs need a full 9U height to hold all of the AM chips and because the AUX behind each AMB is 280 mm deep in order to hold the DOs and first-stage track fitters, the core crates are 9U VME crates. A final crate contains the two FLIC cards; this is also an ATCA shelf with a full mesh backplane.

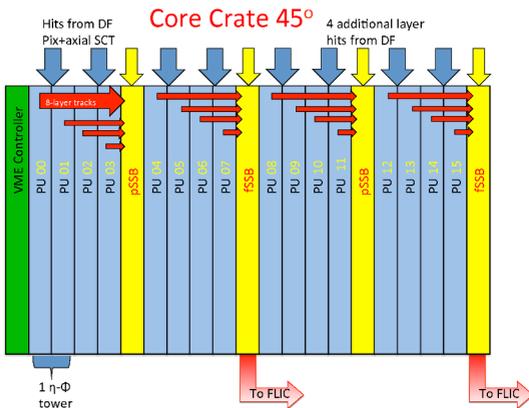


Figure 4. Layout of an FTK core crate and the interboard and intercrate data flow. There are 16 Processing Units (PU), 2 per FTK $\eta - \phi$ tower, each consisting of an Associative Memory board with an AUX card behind it. Four PUs send their 8-layer tracks to a common Second Stage Board. Twelve-layer tracks from each pair of SSBs are sent to the FLIC crate.[4]

2.1 Hardware Description

The FTK IM functions are implemented in a mezzanine card that is connected to the DF main board with a High Pin Count (HPC) FMC connector. Data from the ID are received by the four SFP+ connectors of the FTK IM via four S-Link channels and are directly connected to two FPGAs that

process data independently and transmit the output via the FMC connector to the DF. As shown in Figure 5 data is fed into the FPGA on the DF board. The FPGA shares data with other DF boards on the same shelves through the ATCA backplane and on other shelves via optical fibres using the Rear Transition Module (RTM). The RTM is also used to send the data downstream to the FTK core crates.

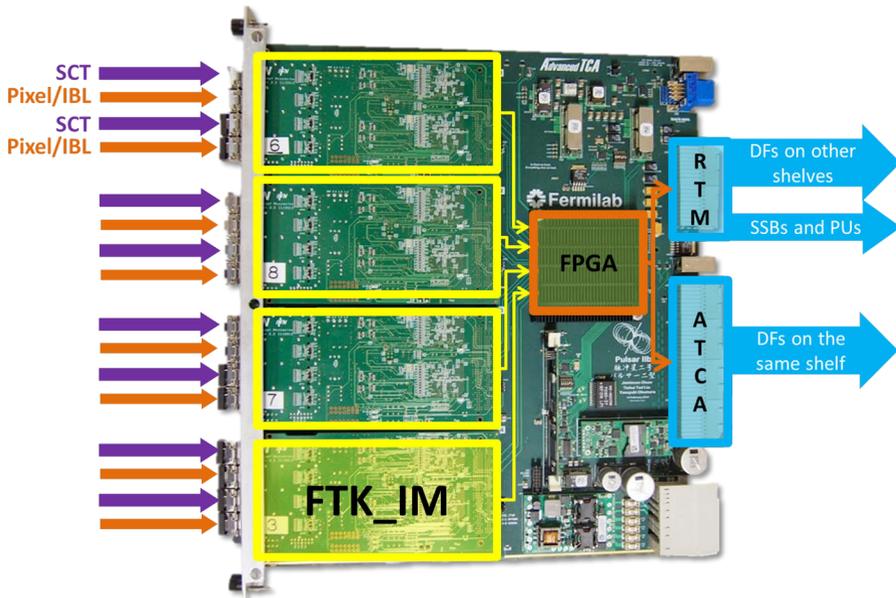


Figure 5: The Data Formatter data flow[4]

Reorganized data by FTK $\eta - \phi$ towers are sent to the Processing Units (PUs). Each PU is a pair of boards; one Auxiliary Card (AUX) and one AMB. Each $\eta - \phi$ tower is handled by two PUs. As seen in Figure 6 the hits from the DFs enter through two QSFP+ connectors with one silicon layer per fibre. Within the input FPGAs, the Super Strip number for each hit is generated and sent through the VME J3 connector to the Associative Memory Board. The Super Strip number and the full-resolution hit coordinates are sent to the four Processor FPGAs, each of which contains a DO and TF to handle a quarter of the AMB (one LAMB mezzanine card). Pattern addresses for roads containing at least seven hit layers are sent back from the AMB through the J3 connector to the processor FPGAs on the AUX. Processed data are sent after that to the SSBs through an SFP+ fibre connector.

The AMB is a 9U VME board on which four LAMBs are mounted. Figure 6b shows the AMB layout. The bus distribution on the AMB is done by a network of high speed serial links: twelve input serial links that carry the silicon hits from the J3 connector to the LAMBs, and sixteen output serial links that carry the road numbers from the LAMBs to J3. These buses are connected to the AUX card through the high frequency J3 connector. The data rate is up to 2 Gb/s on each serial link. Thus, the AMB has to handle a challenging data traffic rate: a huge number of silicon hits must be distributed at high rate with very large fan-out to all patterns (eight million patterns will be located on the AM chips

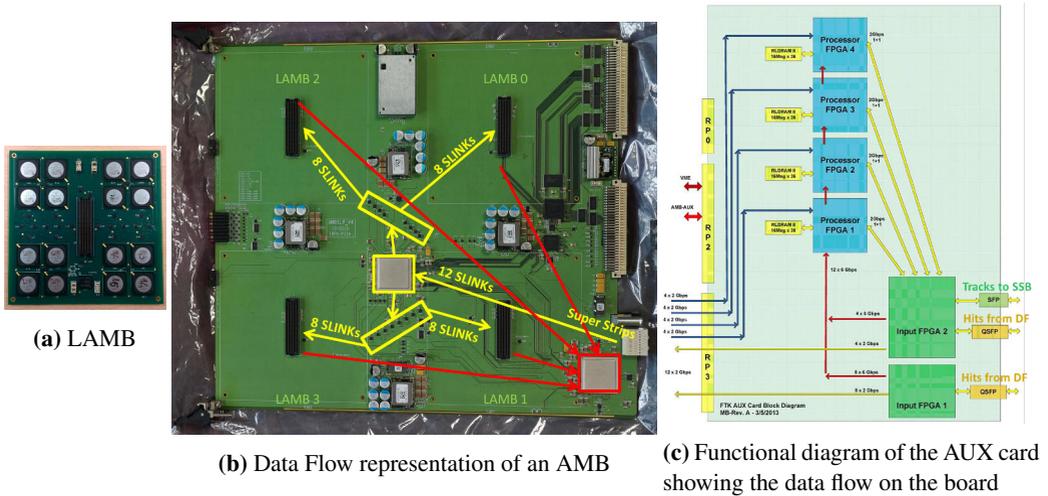


Figure 6: The FTK Processing Unit: 128 PUs in total distributed in 8 VME 9U bins[4]

of a single AMB) and, similarly, a large number of roads must be collected and sent back to the AUX. There is a dedicated FPGA handling the input data from the AUX and another one handling the output to it. One FPGA handles the communication via the VME bus, and a control FPGA is connected to all FPGAs in the AMB to control event processing and handle error conditions.

The SSBs receive through optical links the output from four AUX cards and the hits on the additional layers from the DF system for the two $\eta - \phi$ towers associated with those AUX cards.

The FLIC is implemented in a single separate ATCA shelf with full mesh backplane. Each FLIC main board receives the data coming from the SSBs of the core crate via eight SFP links and performs the look-up for the local FTK to global ATLAS hit identifiers. A passive Rear Transition Module holding eight SFP connectors sends the FTK tracking information to the ATLAS Readout System in ATLAS standard coordinate format via optical fibres using the S-LINK protocol. It is expected that the incoming data rate to the FLIC will be less than 1.0 Gb/s.

Each one of the aforementioned boards has been tested with simulated data and their installation in ATLAS has already started.

3 FTK Performance

3.1 Timing Performance

A simulation tool for calculating FTK latency has been developed in order to tune the system architecture and to ensure that FTK can handle a 100 kHz L1 trigger rate at high luminosity. The system has been divided into functional blocks: DF, DO (read and write mode), AMB, TF, HW, SSB. Emphasis has been put on the time-consuming tasks from DO write mode through TF, as seen in Figure 7a. The DF seems to have a small impact on the overall latency, due to its high speed, and the fact that its timing overlaps almost completely with the DO's write mode. DO's write mode, during which DO receives data from the DF and propagates it to the AMB, might take up to 8 μ sec for an event. The AMs start the pattern matching some nanoseconds after DO's receiving data from DF which means

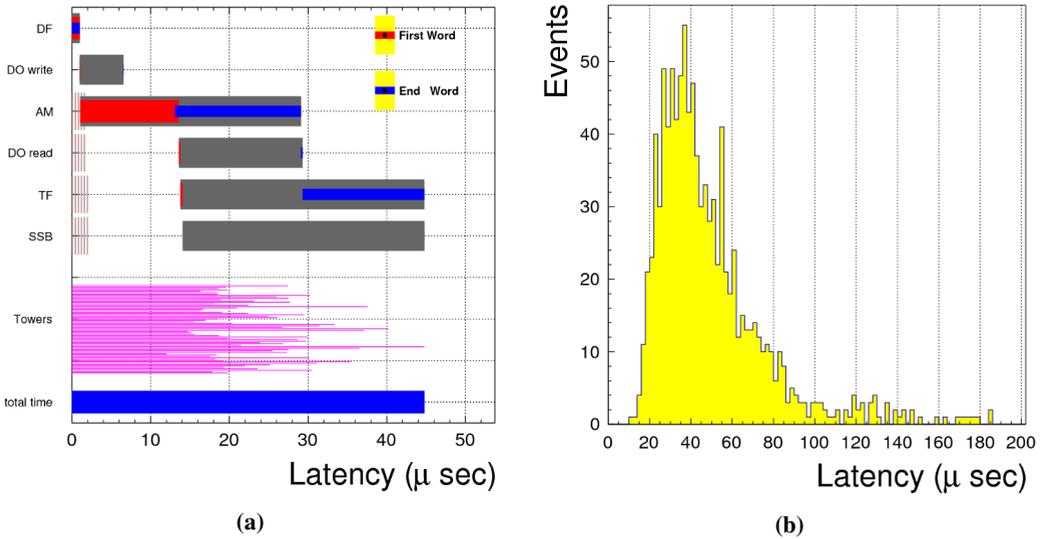


Figure 7: FTK latency for $Z \rightarrow \mu\mu$ events with 69 pileup.

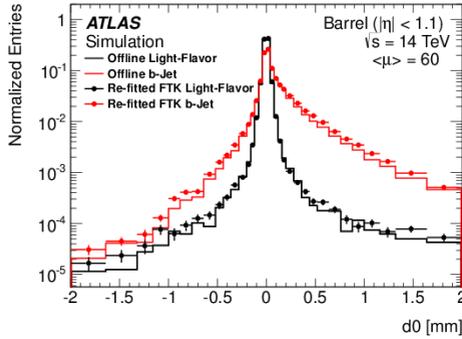
a) The timing of the functional blocks is given for the core crate (region) that takes the most time. The time for each of the 64 regions is shown below that, with the total latency shown in the bottom bar.

b) Latency histogram for $Z \rightarrow \mu\mu$ events with 69 pileup. For each event, the execution time starts when the event is available (10 μ s after the previous event, corresponding to a 100 kHz level-1 trigger rate) and ends when the FTK has completed analyzing that event.

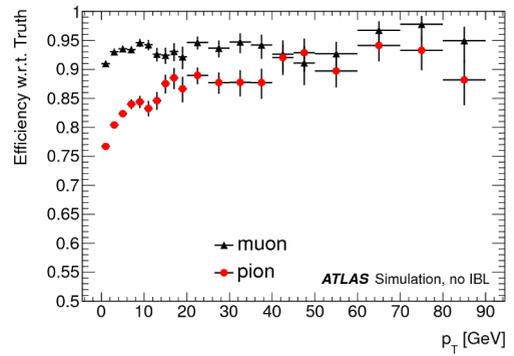
that the latency between the two functions is comparatively small. It has been calculated that the most time-consuming tasks are the pattern matching and the TF on the AUX. These functions take roughly 30 μ sec each for a single event and are the main contributors to the latency of the system. The Hit Warrior has a relatively short latency for each track that enters before the track is either sent to the output or discarded. Depending on the event complexity, the FTK latency fluctuates mostly between 30 and 100 μ sec as shown in Figure 7b. Some events take longer than others to do global tracking, but after such an event the latency quickly returns to the typical range. In conclusion FTK operates well for a 100 kHz L1 trigger rate at 69 pile-up events.

3.2 Physics Performance

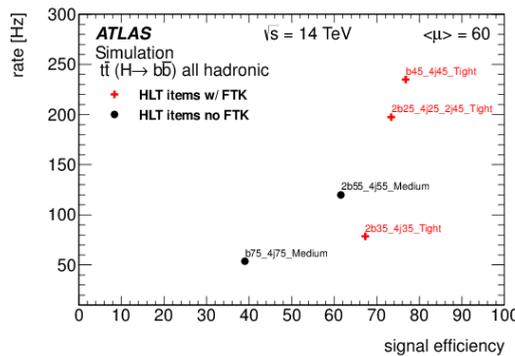
In order to perform studies for the FTK efficiency, the full ATLAS simulation framework and its geometry was used. Comparisons of the FTK tracks to both truth and offline tracks were made. Simulation results indicate that the FTK has an acceptable performance with respect to truth particles and for particles with $p_T > 10$ GeV the efficiency is above 90% (Figure 8b). With respect to the offline tracks, FTK has a general performance $> 90\%$ that in cases can reach efficiencies of 95% (Figure 8a). Simulations show also promising performance for b -tagging and τ which renders the HLT capable to apply event selection algorithms for such events with relaxed p_T criteria at high rates with performance close to the offline one (Figure 8a). The accurate tracks provided by FTK allow the HLT to use the transverse impact parameter to separate b -jets from light jets and thus keep control of



(a) The transverse impact parameter and its $d0$ significance are shown for tracks associated to light-flavor (black) and heavy-flavor (red) jets. [6]



(b) Absolute efficiency with respect to truth particles in muon and pion samples versus pT . [4]



(c) Absolute efficiency with respect to truth particles in muon and pion samples versus pT . [6]

Figure 8: FTK performance

the event rates without sacrificing the signal efficiency (see Figure 8c, where the lower jet P_t thresholds boost the efficiency for $H \rightarrow b\bar{b}$ above 65%, with a reduced trigger accept rate).

Conclusion

The FTK system will provide global tracking information to the ATLAS High Level Trigger with good performance and low pile-up dependency. Studies show that this will allow the HLT to use more complex algorithms. FTK integration has begun and the hardware and firmware commissioning are progressing well.

Acknowledgements

Participation in the conference was partly supported by the European Commission, call: FP7-PEOPLE-2012-IAPP, Grant No: 324318.

References

- [1] *The ATLAS Experiment at the CERN Large Hadron Collider*, ATLAS Collaboration, JINST **Vol 3** S080031
- [2] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>
- [3] *The ATLAS Inner Detector commissioning and calibration*, ATLAS Collaboration, Eur.Phys.J. **C70**, 787 (2010)
- [4] *Fast TracKer (FTK) Technical Design Report*, ATLAS Collaboration, CERN-LHCC-2013-007
- [5] *ATLAS Insertable B-Layer Technical Design Report*, ATLAS Collaboration, CERN-LHCC-2010-013
- [6] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FTKPublicResults>