The AMchip04 and the Processing Unit Prototype for the FastTracker

A. Andreani[a,d], **A. Annovi**[b], **M. Beretta**[b], **M. Bogdan**[c], **M. Citterio**[d], **F. Alberti**[a,d], **P. Giannetti**[e], **A. Lanza**[f], **D. Magalotti**[g], **M. Piendibene**[h,e*], **M. Shochet**[c], **A. Stabile**[a,d], **J. Tang**[c], **L. Tompkins**[c] and **G. Volpi**[b]

[a] *Department of Physics, University of Milan, Via G. Celoria, 16 20133 Milan, Italy*

[b] *Laboratori Nazionali, INFN, Via E. Fermi, 40 00044 Frascati, Roma, Italy*

[c] *Department of Physics and Enrico Fermi Institute, University, 5640 South Ellis Avenue. Chicago, IL 60637, USA*

[d] *Sezione di Milano INFN, Via Giovanni Celoria, 16 20133 Milan, Italy*

[e] *Sezione di Pisa INFN, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy*

[f] *Sezione di Pavia INFN, Via Agostino Bassi, 6 - 27100 Pavia, Italy*

[g] *Sezione di Perugia INFN, Via A. Pascoli, 06123 Perugia, Italy*

[h] *Department of Physics, University of Pisa, Largo B. Pontecorvo, 3 56127 Pisa, Italy*

E-mail: marco.piendibene@pi.infn.it

ABSTRACT: Modern experiments search for extremely rare processes hidden in much larger background levels. As the experiment's complexity, the accelerator backgrounds and luminosity increase we need increasingly complex and exclusive event selection. We present the first prototype of a new Processing Unit (PU), the core of the FastTracker processor (FTK). FTK is a real time tracking device for the ATLAS experiment's trigger upgrade. The computing power of the PU is such that a few hundred of them will be able to reconstruct all the tracks with transverse momentum above 1 GeV/c in ATLAS events up to Phase II instantaneous luminosities (3×10^{34} cm^{-2} s^{-1}) with an event input rate of 100 kHz and a latency below a hundred microseconds. The PU provides massive computing power to minimize the online execution time of complex tracking algorithms. The time consuming pattern recognition problem, generally referred to as the "combinatorial challenge", is solved by the Associative Memory (AM) technology exploiting parallelism to the maximum extent; it compares the event to all pre-calculated "expectations" or "patterns" (pattern matching) simultaneously, looking for candidate tracks called "roads". This approach reduces to a linear behavior the typical exponential complexity of the CPU based algorithms. Pattern recognition is completed by the time data are loaded into the AM devices. We report on the design of the first Processing Unit prototypes. The design had to address the most challenging aspects of this technology: a huge number of detector clusters ("hits") must be distributed at high rate with very large fan-out to all patterns (10 Million patterns will be located on 128 chips placed on a single board) and a huge number of roads must be collected and sent back to the FTK post-pattern-recognition functions. A network of high speed serial links is used to solve the data distribution problem.

KEYWORDS: Trigger concepts and systems (hardware and software); Pattern recognition; VLSI circuits.

Contents

1.   Introduction

The trigger system at hadron colliders must maintain high trigger efficiencies for the physics we are most interested in, while suppressing the enormous QCD backgrounds. A multi-level trigger [1] is an effective solution for this task. The ATLAS trigger system [2] consists of three levels. The hardware Level-1 Trigger quickly locates the regions of interest in the calorimeter and the muon system, operating at rates up to 100 kHz. The subsequent trigger levels, Level-2 and the Event Filter (EF), are collectively known as the high-level trigger (HLT). They consist of software algorithms running on a farm of commercial CPUs. The Level-2 algorithms may request track information in a Level-1 region of interest while the EF has access to information throughout the entire detector. The final EF output rate is limited to 200 Hz.

The trigger event selection requires massive computing power to minimize the online execution time of complex algorithms. The online track reconstruction is the most demanding task. The FastTracker processor (FTK) [3] will be an important element in triggering at CERN's Large Hadron Collider (LHC), and even more so after the planned luminosity upgrade.

The FTK is highly parallel, with the detector segmented into η-φ towers, each with its own tracking processor. Each processor covers one sixteenth of the detector in φ, 22.5°, plus 10° overlap to maintain high efficiency. The η range of each region is divided into four overlapping intervals, for a total of 64 η-φ towers. Consequently, a tower receives only a fraction of the silicon hits, and the Processing Units (PUs) executing track reconstruction have substantially fewer candidates to process. Within each tower, we distribute the high luminosity data on 12 parallel buses at the full 100 kHz Level-1 Trigger rate.

The pattern recognition inside each detector tower is executed by two PUs working in parallel. Figure 1a shows one of the 8 FTK Core Crates [3]. Each core crate contains16 PUs, corresponding to 8 towers, all contiguous and contained in an azimuthal detector section of 45°.

The pixel and SCT data are transmitted from the Read-Out Drivers (RODs) on S-Link [4] fibers to the Data Formatters (DF) which perform cluster finding (see figure 1a). The DFs organize the detector data into the FTK η-φ tower structure, taking the needed overlap into account, for output to the core crates. The barrel layers and the forward disks are grouped into logical layers. The cluster centroids in each logical layer are sent on high-speed serial links to the PUs located in the Core Crates.
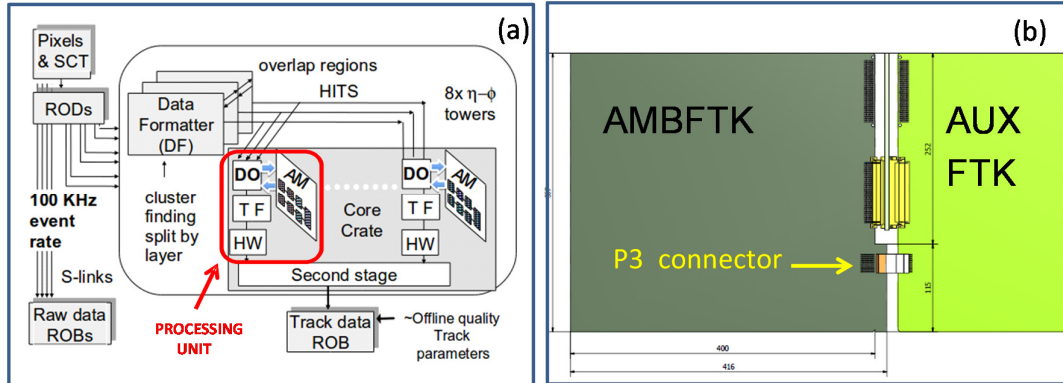

Figure 1: (a) The FastTracker architecture. The Core Crate, containing up to 16 PUs, is shown. (b) The PU, composed of a 9U VME board, the AMBFTK, and a rear card, the AUXFTK. Both are placed in the same slot of the VME Core Crate.

2. The Processing Unit Architecture

The PU algorithm consists of two sequential steps. In the first step, pattern recognition is carried out by a dedicated device called the Associative Memory (AM), which finds track candidates (coarse-resolution roads). When a road is found that has hits in at least seven of the eight silicon layers used for pattern recognition, the second step is carried out in which the full-resolution hits within the road are fit to determine the track helix parameters and a goodness of fit. Tracks that pass a χ^2 cut are kept. The ATLAS inner tracker geometry is described in [5], while the geometry for the Phase II upgrade is being defined.

The PU consists of a 9U VME board, called the AM Board or AMBFTK, along with an auxiliary card on the back of the crate (AUXFTK). A special P3 connector allows for communication between the front and rear boards placed in the same VME slot (see figure 1b). The AMBFTK has 128 AM chips organized into 4 large mezzanines called LAMBs. The AMBFTK finds the roads, while the AUXFTK refines the track finding by doing preliminary track fitting with full resolution detector information. The AUXFTK contains also the Data Organizer (DO), which is an important interface between the DF, the AMBFTK and the Track Fitter (TF).

The DO engines are smart databases where full resolution hits received from the DFs are stored in a format that allows rapid access based on the pattern recognition road ID. When the AM finds roads, the DO retrieves the requisite number of hits. In addition to storing hits at full

resolution, the DO also converts them to a coarser resolution, referred to as super-strips (SS), appropriate for pattern recognition in the AM. The AMBFTK contains a very large number of preloaded patterns, corresponding to the possible combinations of a SS in each silicon layer that a real track could produce. These are determined in advance from a full ATLAS simulation of single tracks using detector alignment extracted from real data. The AM is a massively parallel system in that each hit is compared with all patterns nearly simultaneously. When a pattern has been found with the requisite number of hit layers, it is then labelled as a road, and the AM sends the road number back to the DO. The DO immediately fetches the associated full resolution hits and sends them and the road ID to the Track Fitter (TF). Because each road is quite narrow, the TF can provide high resolution helix parameters using the average parameters across the relevant tracking modules, adjusted by corrections that are linear in the actual hit position in each layer. Fitting a track is thus extremely fast since it consists of a series of multiply-and-accumulate steps. A modern FPGA can fit approximately 10^9 track candidates per second. Finally, duplicate-track cleanup is performed by the Hit Warrior (HW).

After processing the hits, the PU sends all found tracks with transverse momentum P_T above a minimum value, typically 1 GeV/c, over S-LINK to the second FTK stage.

2.1 The First PU Prototype

The associative memory board has a long development history [6, 7] while the use of an AUX card containing multiple high-level functions constitutes a totally new development. For this reason, this first PU prototype includes a fully functional AMBFTK, with a much simpler AUXFTK, the proto-AUX, where the DO and TF functions are missing. The goal of the proto-AUX is to test the high frequency serial links and provide data to the AMBFTK board at full speed in order to verify the performance of the Associative Memory system. The AMBFTK will have to face large input (hit) and output (roads) data traffic and a challenging fan-out. problem.

The proto-AUX card provides hits on 12 buses, covering the 8 logical layers used by the AM for pattern recognition. The connection allows a maximum of 12 Gbits/sec input to the AMBFTK through 12 high frequency serial links and a maximum of 24 Gbits/sec output (found roads) through additional 16 high frequency serial links. These buses are provided by the AUX card through a high frequency ERNI P3 connector (see figure 1b). A custom board profile has been studied and simulated at the CAD to guarantee a perfect board-to-board closure of the P3 connector without backplane support in that region. A network of high speed serial links characterizes the bus distribution on the AMBFTK.

2.2 The AMBFTK

The motherboard has flexible control logic placed inside a group of powerful FPGA chips visible in figure 2a. All the FPGAs are Xilinx Spartan6 chips [8] which have Low-Power Gigabit Transceivers (GTP). Ultra-fast data transmission between chips, over backplanes, or over longer distances is becoming increasingly popular and important. However, it requires specialized, dedicated on-chip circuitry and differential I/O capable of coping with the signal integrity issues present at these high data rates. All Spartan-6 LXT devices have two 8 gigabit transceiver circuits. Each GTP transceiver is a combined transmitter and receiver capable of operating at data rates up to 3.2 Gb/s. The transmitter and receiver are independent circuits that use separate PLLs to multiply the reference input frequency by programmable numbers between 2 and 25, generating the bit-serial data clock. Additionally, each GTP transceiver has a large

number of user-definable features and parameters. All of these can be defined during device configuration, and many can also be modified during operation.

The incoming hits (12 serial links, the red arrows in figure 2a) are received by the GTPs in the two input FPGAs and saved in large derandomizing FIFOs. Outgoing road IDs from the LAMBs (4 links/LAMB) are sent to the FPGAs (in the blue boxes) near the P3 on serial links (blue arrows). The FPGA inside the grey box is the Control chip. It is connected to all of the FPGAs in the AMBFTK to control the event processing.
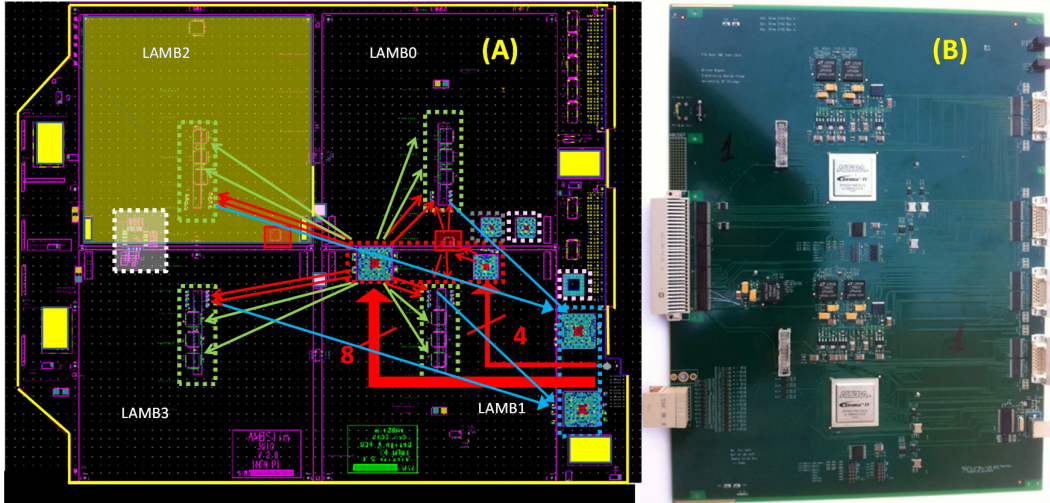

Figure 2: (a) AMBFTK layout. The red arrows show the hit paths from the P3 connector to the LAMBs and the blue arrows show the road paths from the LAMBs to the P3 connector. The yellow square in the top-left corner shows the size of the LAMB mezzanine. (b) the Proto-AUX

2.3 The LAMB Mezzanine

The LAMB and the AMBFTK communicate through an SMD connector placed in the center of the LAMB (inside the yellow central rectangle in figure 3 corresponding to the 4 green boxes in figure 2a). One of the LAMBs is shown as a transparent yellow square in the top-left quarter of figure 2a. Each LAMB contains 32 Associative Memory (AM) chips, 16 per face, as in past designs. However, the AM chips are new [9]. They are 65 nm standard cell devices except for the individual patterns each of which is a single custom cell designed to maximize the pattern density and minimize power consumption. They are in LQ208 packages (1.4 mm thick) and contain the stored patterns along with the readout logic. They have a core voltage of 1.2 V and an I/O voltage of 3.3 V. A large current (up to 25 A per LAMB) is provided through the two connectors in the yellow boxes at the bottom of figure 3 to support the 1.2 Volts. The 3.3 V current is provided by the LAMB central SMD connectors, together with the signals.

The 32 AM chips in each LAMB are connected as eight 4-chip pipelines (4 on the top side and 4 on the bottom side). Each AM chip receives an input road bus from its upstream chip (in figure 3 the green arrows show the direction of the road flux) which it multiplexes with its internally matched roads onto a single output bus to its downstream chip. The roads flow down the 8 pipelines and are collected and merged into 4 streams by the two GLUE chips shown on the bottom of the LAMB inside green boxes. The Glues are Spartan6 FPGAs, containing high-speed serial links, the GTPs, used to transfer roads to the motherboard through the SMD connector.
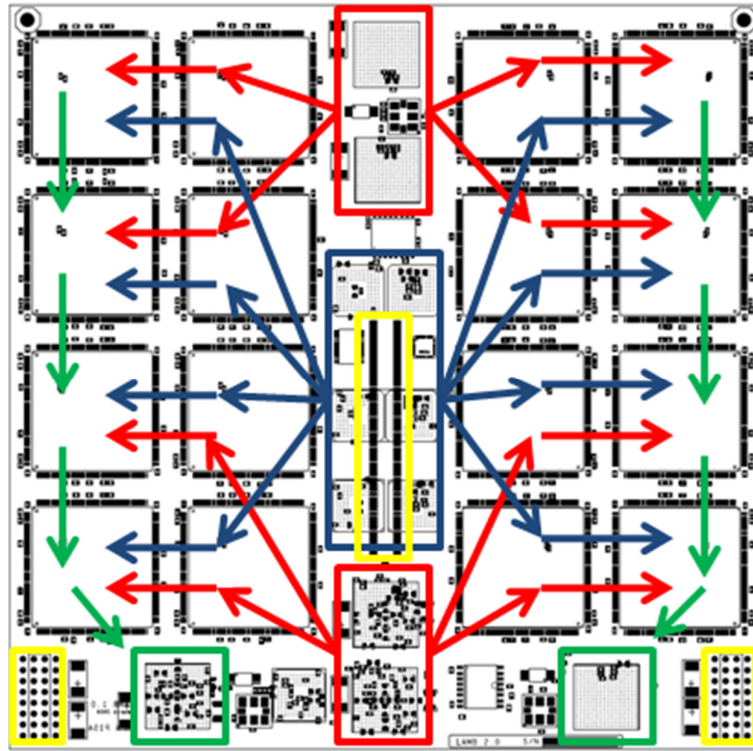
Figure 3: The LAMB mezzanine layout

The hits for each event, organized into 8 buses for 8 detector layers, arrive at the LAMBs from the AMBFTK through the SMD connector and are fed partially in parallel, partially serially. They are distributed to the 32 AMchips with a four-fold fan-out through the Input Distributor (INDI) chips shown in the red (Spartan6 FPGA, connected through GTPs) and blue (Xilinx CPLDs) boxes. The 4 buses that are distributed serially by the AMBFTK are received by the GTPs of the Spartans, while the CPLDs receive the parallel buses. The CPLDs are located in the center of the board; the left column distributes the 4 buses to the left half mezzanine (see blue arrows), while the right column distribute the same buses to the right. There are also multiple Spartan chips, one placed at the top and one at the bottom of the LAMB, and they distribute their outputs to the bottom or top half of the board. The red arrows show how each bus is multiplied by 4 and distributed by the Spartans while the blue arrows show the bus distribution by a CPLD couple.

2.4 Event Processing

When the Control chip starts to process an event, the hits are popped in parallel from all the hit input FIFOs and are simultaneously sent to the four LAMBs.

An End Event (EE) word separates hits belonging to different events. The EE contains the event tag. The data on different streams have to be synchronized. The input FPGAs are provided with deep FIFOs for this purpose, one for each serial link. If, occasionally, a FIFO becomes

"Almost Full", a HOLD signal is sent to the upstream board, which suspends the data flow until more FIFO locations become available. The Almost Full threshold is set to give the upstream board enough time to react. The 12 Holds for the input serial links are sent back to the AUXFTK through the P2 connector. Holds are also sent by the AUXFTK to the AMBFTK to stop sending roads if an AUX FIFO receiving one of the road links becomes half full.

The new AM chip is able to process two events in parallel. While hits of one event (the "N+1st" event) are downloaded into the LAMBs, locally matched roads of the previous event (the "Nth" event) are collected from the LAMBs and sent to the AUXFTK.

When the EE word is received on a hit stream, no more words are popped from the relative FIFO until the EE word containing the same tag is received on all 12 hit streams (event tag "N+1"). At that time, the "N+1st" event is fully loaded into the AM chips. Once the LAMBs have made the last of the matched roads of the "Nth" event available to the AUXFTK, including the 16 EE words on all the 16 road serial links, the "Nth" event is considered fully processed. At that point, the Control chip sends INIT to push events ahead in the pipeline: the "Nth event" will be deleted by the AMBFTK since it is fully processed, the matched pattern information of the "N+1st" event will be copied to reset the pattern bank for a new event and start the read-out of the matched roads of the "N+1st" event (complete the "N+1" event processing). Finally, a new event, the "N+2nd" event, will be popped out of the input FIFOs and sent to the AM chips to register its matches with patterns.

In conclusion, the INIT signal will be provided with the logical AND of these 2 conditions:
(1) All EE words for the "Nth" event have been delivered to the output on the 16 serial road links going to the AUXFTK.
(2) All EE words have been received for the "N+1st" event from the 12 serial links coming from the AUX board.

2.5 Proto-AUX

The Proto-AUX, pictured in figure 2b, was designed to test the high-speed links to the AMBFTK and VME access to the AUXFTK. It contains two Altera Stratix IV [10] FPGAs, each of which has sixteen 8.5 Gbit/s transceivers. A small test card was also produced to test the VME access to the Proto-AUX. There are roughly 300 MB of ROM memory on each AUXFTK which will be written with VME.

Although each transceiver can be configured as independent receiver and transmitter, the StratixIVs are configured such that one has 16 receivers and the other has 16 transmitters. The transmitter FPGA will send fake data, which it will have read from a buffer on the board, on 12 of its transmitters to the AMBFTK over the P3 connector. The receiver FPGA will receive data from the AMBFTK on all 16 of its receivers. Data integrity will be checked by comparing the received road IDs to the expected IDs corresponding to the sent data.

3. PU's diagnostic: the Spy Buffers

Each main FTK functional block is provided with circular memory buffers called Spy Buffers. They are deep enough to record, as a logic state analyzer, a few events sent to the block output or received on the block input. Our PU prototypes implement them. Comparing a sender's output buffer with a receiver's input buffer checks data transmission. Comparing a block's input and output with emulation software checks data processing. The memories also serve as sources and sinks of test patterns for testing single functions, single boards, a small

chain of boards, a slice of FTK, FTK as a standalone system, or the data paths to FTK's external sources and sinks. The buffers can be frozen and read by monitoring software parasitically during data-taking, and buffers inside contiguous boards or block functions can be frozen together via dedicated signals when any piece of logic detects an error condition, such as 'invalid data' or 'lost synchronization'. By polling the PU's circular memories during beam running, large samples of track and hit data, pattern IDs, etc., unbiased by L2 or L3 trigger decisions, are sampled and statistically analyzed to monitor data quality.

4.   Conclusions

We report on the design of the first Processing Unit prototypes for the FTK processor. This design had to face the most challenging aspects of this technology: the huge volume of detector clusters ("hits") distributed at high rate with large fan-out to all patterns (10 million patterns will be located on 128 chips placed on a single board) and the large number of roads collected and sent back to the FTK post-pattern-recognition functions. A network of high-speed serial links has been used to solve the data distribution problem.

Acknowledgments

References

[1] W. Smith, *Triggering at LHC Experiments*, 2002 *Nucl. Instr. and Meth. A*, vol. 478, pp. 62–67

[2] G. Aad et al [ATLAS Collaboration], *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512 [hep-ex], pp. 549, 2008.

[3]   Andreani et al., *The FastTracker Real Time Processor and Its Impact on Muon Isolation, Tau and b-Jet Online Selections at ATLAS*, 2012 TNS Vol.: 59 , Issue:2, pp, 348 – 357

[4] *http://hsi.web.cern.ch/HSI/s-link/*

[5] Aad, G, et al., *Performance of the ATLAS Detector using First Collision Data,* JHEP 1009 (2010) 056

[6] Annovi, A, et al. *The AM++ board for the silicon vertex tracker upgrade at CDF,* Nuclear Science Symposium Conference Record, 2005, NSS '05, IEEE Volume: 1, Page(s): 598 - 602

[7] Batignani, G et al. *The associative memory for the self-triggered SLIM5 silicon telescope*, Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE, Page(s): 2765 - 2769

[8] http://www.xilinx.com/support/index.htm#nav=sd-nav-link-114723&tab=tab-sd

[9] Annovi, A et al., *Associative memory design for the fast track processor (FTK) at ATLAS,* Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE Page(s): 141 – 146

[10]   http://www.altera.com/devices/fpga/stratix-fpgas/stratix-iv/stxiv-index.jsp