Fast Tracker for Hadron Collider Experiments

FP7-PEOPLE-2012-IAPP: Grant Agreement Number 324318

FTK-IAPP Consortium http://ftk-iapp.physics.auth.gr/

***Abstract*— We present a project which aims to develop an extremely fast but compact processor, with supercomputer performances, for pattern recognition, data reduction, and information extraction in high quality image processing. The proposed hardware prototype features flexibility for potential applications in a wide range of fields, from triggering in high energy physics to simulating human brain functions in experimental psychology or to automating diagnosis by imaging in medical physics. In general, any artificial intelligence process based on massive pattern recognition could largely profit from our device, provided data are suitably prepared and formatted.**

The first goal is demonstrating the system can perform online track reconstruction of full events at the highest luminosities of the Large Hadron Collider at CERN beyond the limits of any existent or planned device. This task has to be performed inside a fixed latency of few tens of microseconds despite the overwhelming confusion due to the high track multiplicity produced by the exceedingly large number of proton-proton collisions overlapping in the same image recorded by the detector. With this goal we participate to the construction and the test for a high precision real-time tracker built for the Large Hadron Collider experiments, recently approved for the ATLAS experiment upgrade: the Fast Track (FTK) processor. FTK can improve the capability of the detectors to select the events with the greatest scientific potential. It uses FPGA and ASIC chips to implement real-time, complex track reconstruction algorithms. The track's trajectories are reconstructed in 3D, in few dozens of microseconds and the quality of the parameters is similar to the one of algorithms running minutes in the CPU farms.

In parallel we pursue challenging R&D & new real-time computing ideas for more complex applications.

We are funded by the FP7 Industry-Academia Partnerships and Pathways (IAPP) action. (*Abstract*)

Keywords— Pattern matching, Image processing, Parallel processing, Trigger circuits, Field programmable gate arrays, Application specific integrated circuits

I. INTRODUCTION (*HEADING 1*)

The physics program at CERN's Large Hadron Collider (LHC) [1] has been extremely successful since the early phase of data taking in the year 2010. The LHC detectors were designed to search for new discoveries in the head-on collisions of protons of extraordinarily high energy (14 TeV). Among the most interesting searches are the origin of mass, with the extremely important observation of the Higgs boson [2], extra dimensions of space, unification of fundamental forces, and evidence for dark matter candidates in the Universe.

Moreover, the LHC upgrade will widen our capability to search new phenomena that are beyond the scope of our current theory of matter and energy. In the next few years an impressive harvest of data will be collected and at the same time R&D at the technological frontier will be pursued for the upgrades. The partners of this project are active and essential participant in both aspects. These experiments will have a fundamental impact on physics and technology for the next 20 years.

In this scenario the electronics required to process the signals from the LHC complex detectors are taking a very important role, and they must be state of the art. The most interesting processes are very rare and hidden in an extremely high level of background. Implementing the most powerful selections in real-time (trigger) is therefore essential to fully exploit the physics potential of experiments where only a very limited fraction of the produced data can be recorded. The data flow is so massive (1.7MB/25ns ~ 80TB/s [3]) that a drastic real-time data reduction must be obtained. This makes on-line event reconstruction a critical component at any hadron collider experiment, in particular at LHC. A multi-level trigger [4] is an effective solution for an otherwise impossible problem. The level-1 (L1) has strong latency constraint (few microseconds). It is usually based on custom processors. It reduces the rate of events from the machine event production down to tens of kHz. In particular for LHC, the level-1 trigger reduces the event rate from 40MHz to 100 kHz. Subsequent levels, like level-2 and level-3 triggers, used to be well separated in previous experiments, with dedicated hardware used at L2. Current experiments, like CMS [5] and ATLAS [6], have them unified in one single high-level trigger (HLT) running on dedicated CPU farms.

This project directly addresses the main technological challenges of hardware, software and data analysis necessary to face the real-time reconstruction of particle trajectories at trigger level. This is the trigger task that requires the larger online computing power.

The first important goal of this project is the realization and optimization of the Associative Memory (AM) system to make it work in the Fast Tracker (FTK) processor [7] recently approved by the Atlas experiment at LHC and its future evolution for new applications. FTK is a high-performance "super-processor" based on the combination of two innovative technologies: powerful FPGAs (Field Programmable Gate Arrays) working with standard-cell ASICs (Application-Specific Integrated Circuits), the Associative Memory (AM) chips [8], for utmost gate integration density. The target is to get the best results by combining the high performance of VLSI dedicated hardware with the distinctive flexibility of modern programmable logic. Optimal partitioning of complex algorithms on a variety of computing technologies has been

already proved to be a powerful strategy, which turned the past hadron collider experiment CDF [9] at the Tevatron accelerator in the Fermilab Laboratory, near Chicago (USA) into a major player in the field of B-physics, on par with dedicated experiments operating at e+e- colliders.

A complementary and not secondary goal is the dissemination and application of this technology outside of the high energy physics (HEP) research. We believe that HEP developments in this area are important to show the potential of these devices and to spread the skills needed to use them with top efficiency. Our system is an example of solution for a specific case of the "Big Data" problem. This solution is based on the organization of the trigger in different levels of selections, exploiting at low level parallelized, dedicated hardware for an extremely efficient preprocessing step.

This organization is similar to models of the vision processing task performed by the brain. We plan to study the possible impact of our devices for neurophysiologic studies of the brain. Understanding how the brain processes information or how it communicates with the peripheral nervous system could provide new potential applications, new computational systems that emulate human skills or exploit underlying principles for new forms of general purpose computing. Significant improvements could be gained in terms of performance, fault tolerance, resilience or energy consumption over traditional ICT approaches. The use of the associative memory processor for brain studies is particularly fascinating. The most convincing models that try to validate brain functioning hypotheses are extremely similar to the real-time architectures developed for HEP. A multilevel model seems appropriate to describe the brain organization for image processing [10]: "the brain works by dramatically reducing input information by selecting for higher-level processing and long-term storage only those input data that match a particular set of memorized patterns. The double constraint of finite computing power and finite output bandwidth determines to a large extent what type of information is found to be *meaningful* or *relevant* and becomes part of higher level processing and longer-term memory". The AM pattern matching has demonstrated to be able to play a key role in high rate filtering/reduction tasks. Simulations [10] have shown the potential of the pattern matching algorithm on static 2-D images. We are implementing the algorithm on our technology to extend its application to 3-D images and movies. These studies could have an impact in the area of medical imaging for real-time diagnosis and the study of this possible application is part of the project program. The computing power is still a limiting factor for some high quality medical applications. High-resolution medical image processing, for example, demands enormous memory and computing power to allow 3D processing in a limited time. One example is lung cancer Computed Tomography (CT) screenings that profits of Computer-aided detection (CAD) of pulmonary lesions to reduce the diagnosis times and the risks of errors. Our technology could be an interesting accelerator for such computations.

In brief this paper focus on the two main goals of the FP7 project: in section II we describe the AM system developed for FTK, its performances and its future evolution, in section III we describe the ongoing R&D for the brain study and its possible application to medical imaging.

II. PATTERN MATCHING AT HADRON COLLIDERS

A. The tracking problem at LHC and our proposed solution

Tracking devices, and in particular silicon detectors that are becoming the predominant tracking technology, play an essential role in the identification of interesting events. In fact, they provide very detailed information for charged particles and they can separate most of the different particle trajectories in the overlapping collisions recorded in the same image (see Fig. 1).

However, these detectors contain hundreds of millions of channels, so they require huge computing power for full track reconstruction. They make the problem of complete tracking a formidable challenge even for large computing farms [11]. Therefore, complete high-quality tracking for real-time event selection at very high rates has been considered impossible in LHC experiments at the time they were built. Real-time tracking was planned for limited detector regions or on a small subset of events, previously selected using other detectors [6].

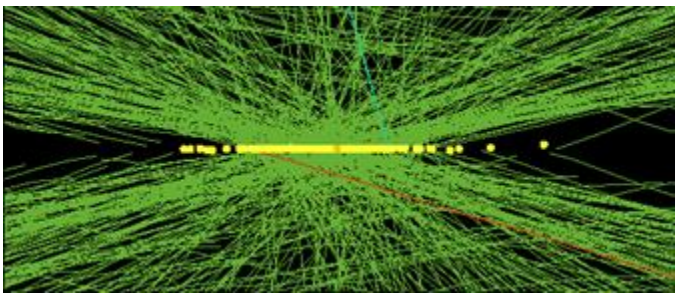

Fig. 1. An event produced at LHC. Image credit: Andre Holzner (http://cms.web.cern.ch/news/reconstructing-multitude-particle-tracks-within-cms).

We overcome the problem by providing real-time tracking using a massively parallel high performance system [7].
Our goal today is the real-time reconstruction of all the tracks above a minimum energy for all the events selected by the L1 trigger, which correspond to an event rate of 100 kHz. Given the complexity of the events, the input data rate to the system is expected to be ~200 GB/s at the maximum LHC luminosity. The maximum supported data rate is 400 GB/s

A key role in the architecture is played by high-performance field programmable gate arrays (FPGAs), while most of the computing power is provided by full-custom ASICs, the Associative Memory (AM) chips [8]. AMs exploit massive parallelism in data correlation searches by comparing the input data simultaneously to a list of pre-calculated "expectations" (pattern matching). These correlations (i.e. the matched patterns) are found by the time input data are loaded into the devices.

The AM chip has characteristics similar to a Content-Addressable Memory (CAM) [12]. However, the design of the AM is conceptually different to that of a CAM. In the AM each pattern is not stored in a single memory location, like in the commercial CAM, but it consists of 8 independent 16-bit memory locations, in which the coordinate locations of the

position where the particle hits the silicon detector (hit) [13] can be stored. The innovative characteristic of the AM is that each one of these 8 words has a comparator and a match flip-flop to compare continuously the stored data with its own input data stream. Data are sent on 8 parallel buses, one for each word of the pattern. All words in the AM make independent and simultaneous comparisons with the data serially presented on its own bus. Every time a match is found, the match flip-flop is set and remains set until the end of the event processing, when a reset signal is propagated. A pattern matches when a predefined number of the flip-flops is set (user defined threshold). All the matched patterns are read out. An extensive description of the AM and its operation can be found in [8].

FPGAs configure and control the AMs and their I/O, providing the flexible computing power to process the selected patterns. Distributed debugging and monitoring tools suited for a pipelined, highly parallelized structure and a high degree of configurability can cope with a variety of applications.

B. System Segmentation and Scalability

The input bandwidth sets an upper limit either on the event rate or on the size of the detector connected to the processor. In order to sustain very high event rates, it is necessary to organize the system as a set of independent engines (typical input bandwidth of each one 1.6 GB/s), each one working on a different sector of the silicon tracker. Let us imagine dividing the detector into azimuthal sectors. This segmentation generates some inefficiency at sector boundaries that can be removed by allowing a small overlap region between adjacent sectors. Thus, the system is scalable and can grow to provide higher computing power to cope with the detector occupancy increases due to the increment of the LHC luminosity. The AM system that is approved for the current ATLAS detector upgrade stores 1 billion (109) AM patterns [7] into 128 AMBSLP boards.

Fig. 2 shows the AMBSLP motherboard, equipped with 4 mezzanines. Each mezzanine has 16 AM chips (see fig. 2 on the right) for a total of 64 chips per AMBSLP. We foresee an enlargement of more than a factor 10 for the LHC future upgrades without a significant increase of boards thanks to the miniaturization process due to the technology advancement.

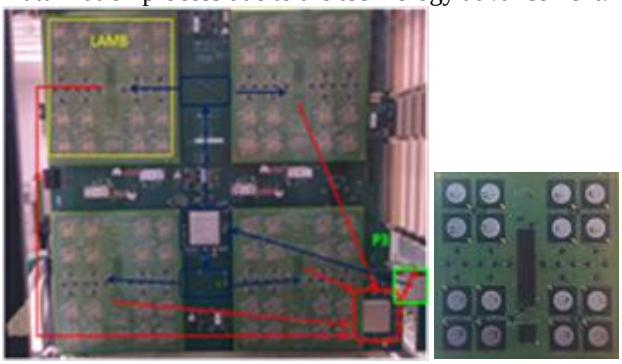

Fig. 2: The AMBSLP 9U VME board and its mezzanine

B. Implementation

The design of the AMBSLP has been a challenging task, due to the following factors: (1) the high pattern density (8 million patterns per board today), which requires a large fraction of the board to be filled by AM chips; (2) the I/O signal congestion at the board level, which requires the use of a huge network of serial links; (3) the power limitation due to the available cooling system: as we can fit up to 20 AMBSLPs in a VME crate, the power should not exceed 250 W per AMBSLP. We have on purpose chosen the 9U VME standard to spread around computing elements and I/O resources on the large area provided by the 9U mechanics. We implement a large number (~850) of medium frequency (2 Gbit/s) serial links to keep low the consumption per unit area on the PCB. However, the total traffic on the board is above the Tbit/s, for a total of more than 100 Tbit/s in the whole AM system used by the ATLAS experiment. Each AMBSLP executes more than 67 million 16-bit word comparisons every 10 ns. At each clock cycle data are distributed in parallel to the large number of patterns with fan-outs of 1:8 million.

C. Performances

Candidate tracks are found exploiting the detector readout time, few clock cycles after the arrival of the corresponding detector channels belonging to the track.

This powerful highly parallel dedicated hardware has been demonstrated using the experiment simulation [7] to provide excellent performance, reaching resolutions, efficiencies and fake track rejection typical of the best tracking algorithms. For this reason, the use of the system in offline simulation has also been proposed [14], with the advantage of a low power usage (250 W/board). The system in fact is very compact and requires simplified infrastructures [15] compared to the ones necessary for the huge CPU farms executing an equivalent task. Four racks of electronics, for a total power of ~40 kW, are able to reconstruct events with an average latency of ~100 μs [7], while offline tracking requires several seconds when performed on events containing 60 p-p collisions [11].

One interesting technology which recently has attracted the attention of the high energy physics community for real-time applications is graphic processing. Both ATLAS ([16], [17]) and CMS [18] are studying the performance of real-time tracking at LHC executed on modern Graphic Processing Units (GPUs). Even if the comparison with the CPU performances is promising, the latency to execute tracking is at least tens of milliseconds for simplified algorithms and reduced detector occupancies, with a fast grow above hundreds of milliseconds when the occupancy increases. In conclusion, our hardware dedicated approach is today thousands of times faster than any available commercial computing device.

D. R&D for the Future Evolution

The short latencies, reachable by the parallelized AM system, push both CMS and ATLAS to study its possible application at L1 [19] for the future accelerator upgrades, when the LHC luminosity will cause the superimposition of hundreds of collisions in the same image and will require much faster and more efficient trigger selections. This new possible area of

application requires to achieve further technology performance, miniaturization and integration of the current state of the art prototypes. The L1 trigger is much more demanding in terms of a higher rate of events to be processed and a shorter processing latency time.

To face the increase of complexity we plan to increase the FPGA parallelism by associating one single FPGA to each AM chip. The FPGA configures and handles the AM and provides a flexible computing power to process the shapes selected by the AM [20]. An innovative multi-chip package (System in Package, SiP) including both the AM chip and the FPGA in the same space, should provide the final necessary miniaturization with the aim of enhancing performance and power saving without increasing the volume of the hardware.

Our long-term future goal is to produce a powerful and miniaturized hardware that can be used as coprocessor also for offline event simulation [14].

III. Emulating the Brain for Image Processing

In [10] a study of a brain model for image preprocessing is presented and successfully applied to static 2-D images. Since the needed computational time cause serious limits to the capability to extend these studies to 3-D images and movies, we plan to use the AM-based processor for a real-time hardware implementation of fast pattern selection/filtering of the type studied in these models of human vision.

A. The filtering Algorithm

Fig. 3 shows the results of the simulations of the model described in [10] where pattern matching with relevant patterns is used to filter the main features of the image.

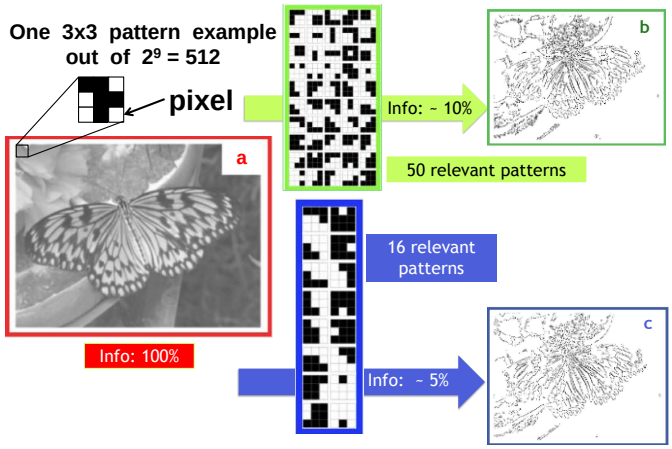

Fig. 3: natural image (a) and corresponding filtered images(b,c)

The pictures on the right (b,c) show the quality of the filtered images. The butterfly can be clearly recognized even if the image information is reduced at the level of 10% or less of the original content. The associative memory works as an edge detector implementation able to extract the salient features.

The pattern is defined as the collection of pixels contained in a 3×3 pixel square, as shown above the butterfly image (a) in fig. 3. Each square is converted in a 9 bit sequence (each bit is 1 for a black pixel and zero for a white one in the case of B/W) or an 18 bit sequence in case of 4 level greys (2 bits/pixel). The bit sequence is used to identify the pattern.

Starting from the left top corner the image is scanned by the 3×3 square that is moved in step of one pixel toward the right. When the row is finished, the square is moved one pixel down to scan again the raw from the left to the right. Each pattern detected in the figure during the scan is compared to the set of "relevant patterns" predefined by a training phase. It is rejected if it does not match any of them; it goes back in its position in the picture if it is accepted. Fig. 3 shows two collections of relevant patterns for two different selections. The 16 patterns in the blue box produce a larger reduction of information in the final image than the 50 patterns in the green box. Smaller is the set of chosen patterns stronger is the information reduction.

Analyzing 3-D images or movies increases enormously the number of possible and relevant patterns. The pattern in this case is not a square, but a cube of pixels: a set of three 3×3 squares taken from 3 subsequent frames. Each pattern for B/W is made of 27 bits corresponding to 2^{27} possible patterns. If 4 levels of grey are necessary the total number of patterns becomes 2^{54}. One goal of the study is to understand which is the minimum set of "relevant patterns" in these complex cases and how much large has to be the memory to contain them.

B. Implementation

Our initial plan was the use of exactly the same hardware (see Fig. 2) developed for HEP, adapted for generic imaging. The VME solution is very powerful and offers a lot of computational power but it is large (not portable) and requires a specific interface (VME standard). It is not easy to use for every day applications. For these reasons we decided to try a more modern, compact solution. One of our key goals is the miniaturization of the system in new modern standards with the objective to make the system suitable for an open range of applications in which massive and parallel data processing makes the difference.

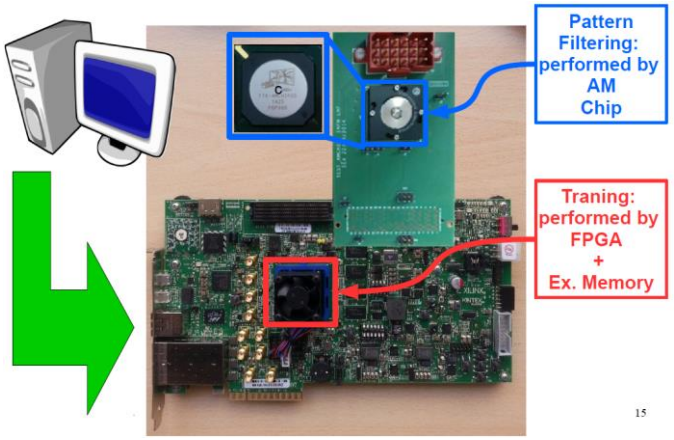

Fig. 4: The hardware setup for image processing: this is an evaluation board with a Kintex Ultrascale, which is the state of the art in FPGA devices, with a single AM chip mezzanine placed on the FMC connector.

The new Control Board needs the following characteristics:

1) powerful FPGA (Field Programmable Gate Arrays) with large on-board memory,

2) Ethernet & PCI Express I/O,

3) handling (distribution and collection) of all AM chip serial links,

4) configuration and control of the AM pattern bank,

5) provision of extra functionality to complete the AM functions in real-time.

While the AM chip needs challenging developments, one of the advantages of the FPGA imaging task is that boards already available on the market are powerful enough to cover the above listed specifications. New generation commercial FPGAs are already available (e.g. Xilinx Ultrascale FPGAs) and will allow us to develop the high performance embedded system required in parallel with the development of the new generation AMchip. Fig. 4 shows the computing unit used today, based on a Xilinx Ultrascale evaluation board. A new mezzanine with multiple AM chips can be connected to the large connector on the top of the board, but for the moment we start with a single chip mezzanine as shown in the figure.

The algorithm is divided in two main parts: "Training Phase" and the "Real-Time patterns recognition phase", what we call the "data taking phase". Most of the functions are executed by the FPGA with the only exception of the "Real-time patterns matching phase", that is executed by the associative memory.

A challenging task of the implementation is the "Training Phase". It is subdivided in the following steps:

1. *Calculation of the pattern appearance frequencies*: The embedded system receives the image bit-streams (e.g., data from a PC or a video camera). The FPGA partitions/reorganizes the input data into the small 3×3 pixel patterns. Then, for each pattern, the FPGA calculates the occurrence frequency in the processed images/frames. This calculation is iterated for all possible patterns in a large set of training images. In this way, different Probability Density Histograms (PDHs) are computed for different training image sets. PDHs are different for different types of images, from different applications and sources. Medical images have different PDHs than natural images, security images etc. The training is required for the choice of relevant patterns. When the environment and the lighting conditions change, especially for security and machine vision applications for streaming video the training has to be executed continously in real-time. In this way the device adapts itself autonomously to the different conditions of the images that it observes.
2. *Pattern selection*: the system must decide which set of patterns must be selected for memory storage (the relevant patterns). To maximize the capability to recognize shapes (both human-brain recognition and artificial recognition), we adopt the hypothesis described in [10], i.e., the principle that maximum entropy is a measure of optimization. The set of patterns that produces the largest amount of entropy allowed by system limitations is the best set of patterns that we can select to filter our images or videos. The system limitations can be summarized in two main parameters: N maximum number of storable patterns, and W, maximum bandwidth. In [10] are described the details of the selection.
3. *Pattern Writing operation:* the relevant patterns (selected in step 2) are written in the AMchip bank. The writing operation is made via JTAG by means of a system controller. This is the last step of the Training Phase.

When the training is complete, the "Real-time pattern matching phase" or "Data taking" can start: the system is able to work in real-time at the maximum frequency and is able to perform:

1. Parallel recognition of patterns in the data stream. Input Patterns are sent to the AM bank and addresses of matched patterns are transferred at the output of the AM chips.
2. Output formatting operation: The matched patterns are reorganized into a new image, to produce the filtered images/videos, called "sketches", where only the boundaries of the relevant objects appears, while uniform areas are suppressed.

C. Logic Description

The system needs to be able to perform both training and pattern identification in real-time for demanding streaming video applications. Several optimization techniques are used to achieve the best performance possible in the hardware implementation. The videoframes are stored in the external memory before being transferred in an internal frame buffer. As soon as enough data has been tranfered for the 3x3 patterns to be formed, a pattern identification matrix begins to be loaded that identifies and propagates two patterns per clock cycle to the pattern accumulators. The accumulators are designed to facilitate successive accumulation in the same memory location ("fall through" data logic). As soon as the whole image sample has been read, the pattern frequency is calculated by taking advantage the FPGA DSP slices. The pattern selection process is done by using logic with principle similar to the one used for pattern identification in the HEP FTK implementation, but appropriately optimized for image processing applications. The selected patterns are then loaded to the AM chips for the execution of the pattern matching process. The prototype of the system is being developed on a last generation FPGA device, a Xilinx Kintex Ultrascale *XCKU040* using the KCU105 evaluation board.

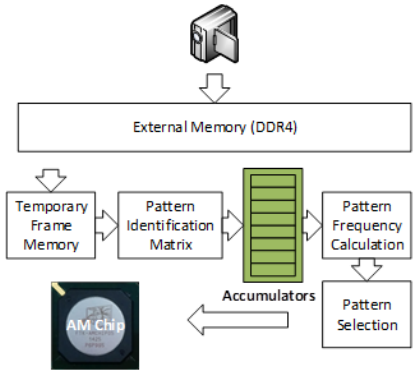

Fig. 5: Training Phase Block Diagram

D. Reconstruction of contours

The extracted features can be processed with fast but complex reconstruction algorithms implemented on FPGA devices as we do in the FTK project to find clusters of contiguous pixels above a certain programmable threshold [13]. As we process them producing measurements that characterize their shape, we can measure quantities of interest in medical applications like the size of the found spots, how circular or irregular the spot is. The algorithm can be extended to 3-D images.

E. Lung cancer diagnosis: an interesting example application

A lung cancer Computed Tomography (CT) screening produces 300-400 noisy slices per subject to be reviewed (fig. 6). This is a huge amount of difficult work for radiologists. Computer-aided detection (CAD) of pulmonary lesions used as second reader can improve the radiologists' detection ability. Nodules are identified because of their sphericity, so a 3-D reconstruction could be particularly interesting to distinguish them from vessels and bronchial tubes that have long shapes in the lung images. We plan to try this 3-D reconstruction on our pattern matching machine.

Fig. 6 CT produces 300-400 noisy slices per subject to be reviewed

IV. Conclusions

Our project has been developed to improve the real-time tracking at hadron colliders, making it thousands of times faster than any other solution available today. It is a succesful example of Big Data processing. Our experience in this field shows the importance of highly parallelized dedicated hardware to reach extraordinary high computing performances with limited infrastructures and consumptions. Our long-term goal is the dissemination of our embedded systems outside high energy physics. They can be powerful accelerators for high performance computing in scientific areas which have to solve a specific task, repeated an enormous amount of times, always the same.

Acknowledgment

The Fast Tracker project receives support from Istituto Nazionale di Fisica Nucleare; the US National Science Foundation and Department of Energy; Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science and MEXT, Japan; the Bundesministerium für Bildung und Forschung, FRG; the Swiss National Science Foundation; the European community FP7 People grants ARTLHCFE 254410 FP7-PEOPLE-2009-IOF and FTK 324318 FP7-PEOPLE-2012-IAPP; the Agence Nationale de la Recherche, project "FastTrack" n. ANR-13-BS05-0011, France; ARC, Australia; and GSRT, NSRF and IKY, Greece.

References

[1] L. Evans and P. Bryant, "LHC machine," *Journal of Instrumentation* 3 S08001, 2008.

[2] ATLAS Collaboration Phys.Lett. B716, 1 (2012); CMS Collaboration, Phys.Lett. B716, 30 (2012)

[3] The ATLAS Collaboration, "ATLAS high-level trigger, data-acquisition and controls: Technical Design Report", ATLAS-TDR-16, CERN-LHCC-2003-022, available online: http://cds.cern.ch/record/616089?ln=en.

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

[5] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *Journal of Instrumentation* 3 S08003, 2008.

[6] The CMS Collaboration, "The CMS experiment at the CERN Large Hadron Collider," *Journal of Instrumentation* 3 S08004, 2008.

[7] A. Andreani, et al., "The FastTracker Real-Time Processor and Its Impact on Muon Isolation, Tau and b-Jet Online Selections at ATLAS," IEEE Trans. on *Nuclear Science*, vol.59, no.2, pp. 348-357, April 2012.

[8] M. Dell'Orso and L. Ristori, "VLSI Structures Track Finding", *Nucl. Instr. and Meth. A*, vol. 278, pp. 436-440, 1989.

[9] W. Ashmanskas et al., " The CDF online Silicon Vertex Tracker", *Nucl. Instr. and Meth. A*, vol. 485, pp. 178-182, 2002.

[10] M. Del Viva, G. Punzi, and D. Benedetti. Information and perception of meaningful patterns. PloS one 8.7 (2013): e69154.

[11] The CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *Journal of Instrumentation* 9 P10009, 2014.

[12] K. Pagiamtzis and A. Sheikholeslami, "Content-addressable memory (CAM) circuits and architectures: A tutorial and survey," in IEEE Journal of Solid-State Circuits, vol.41, no.3, pp. 712-727, March 2006.

[13] C.-L. Sotiropoulou, S. Gkaitatzis, A. Annovi, M. Beretta, P. Giannetti, K. Kordas, P. Luciano, S. Nikolaidis, C. Petridou and G. Volpi "A Multi-Core FPGA-based 2D-Clustering Implementation for Real-Time Image Processing", in IEEE Trans. on *Nuclear Science*, vol. 61, no. 6, pp. 3599 - 3606, December 2014.

[14] A. Annovi, et al., "Associative memory design for the Fast TracK processor (FTK) at Atlas," in IEEE NSS/MIC, 2009, Orlando, pp. 1866 – 1867.

[15] D. Calabro, et al., "The Associative Memory Boards for the FTK processor at ATLAS," in IEEE *NSS/MIC*, 2013, Seoul, pp. 1 – 5.

[16] D. Emeliyanov, et al., "GPU-based tracking algorithms for the ATLAS high-level trigger," in Journal of Phys. Conf., Ser. 396, 012018, 2012.

[17] J. Mattmann, et al., "Track finding in ATLAS using GPUs," in Journal of Phys. Conf., Ser. 396, 022035, 2012.

[18] V. Halyo, et. al., "GPU Enhancement of the Trigger to Extend Physics Reach at the LHC," Journal of Instrumentation 8 P10005, 2013.

[19] A. Annovi, et al., "Associative Memory for L1 Track Triggering in LHC Environment," in IEEE Trans. on Nuclear Science, Vol. 60, No. 5, pp. 3627 – 3632, 2013.

[20] C. Gentsos, F. Crescioli, P. Giannetti, D. Magalotti, S. Nikolaidis, "Future evolution of the Fast TracKer (FTK) processing unit", PoS (TIPP2014) 209