

# Innovative Multi-chip system for multi-purpose PAttern Recognition Tasks (IMPART)

*Principal investigator: Alberto Stabile*

## *Short abstract:*

The goal of the “Innovative Multi-chip system for multi-purpose PAttern Recognition Tasks: IMPART” project consists in developing a cutting edge pattern recognition device for fast image analysis and future trigger processors for High Energy Physics (HEP). This project will achieve an innovative **multi-chip package** (System in Package, SiP) with the final aim of enhancing performance and power saving in electronics devices devoted to **pattern recognition** tasks for several **interdisciplinary applications**. The IMPART project consists in design, fabrication and characterization of a device made of a **Field Programmable Gate Array (FPGA) die and an *ad-hoc* Application-Specific Integrated Circuit (ASIC) assembled in the same package**. Internal wire and bump bonds connecting the FPGA and the ASIC will be designed for high-frequency communication (up to 200 MHz for single-ended wires and 10 Gbit/s for serial links). The **key role** in this **novel technology** is played by the custom multi-core ASIC which will intensively be used **to filter out the relevant information** of the data to be further processed by the FPGA executing high level algorithms. The ASIC implements a **maximally parallel architecture** and offers **best timing performance**, as it solves its task while data are loaded in it. Additionally, the IMPART system benefits from the FPGA computing power to eliminate the main drawbacks of the ASICs. In fact the ASIC performs a first low resolution filtering (pattern matching) function. The FPGA complements the AM task with **flexibility** and **configurability**, adapting its logic to perform any necessary “refining processing” of the low resolution ASIC result. In addition the IMPART ASIC has a specific architecture that helps to identify not only complete features, but also partial or noisy features.

The two key points of the project are: (1) to design an **innovative** version of the state-of-the-art Content Addressable Memory (CAM) with the capability of searching for correlation in the input data, and (2) to connect the CAM die with an FPGA die inside a single package. Repetitive tasks (e.g., bit-wise comparisons of billions of bits) will be performed by the CAM ASIC chip, while more complex algorithms will be performed by the FPGA. In this way, **flexibility and power saving will be strongly maximized**. Up to 200 bondwires will connect the FPGA with the ASIC. The high number of FPGA-ASIC connections and the high bandwidth allows **optimal integration of the two technologies**. This technique helps to speed up the ASIC development time since complex control logic, not easy to be simulated, but not-critical for the occupied area, can be implemented in the FPGA. The goals of IMPART multi-chip system are: (a) **maximum parallelism** exploitation; (b) **low-power** consumption; (c) execution times at least **1000 times shorter** than best commercial Central Processing Units (CPUs) performing the same task; (d) distributed debugging and monitoring tools suited for a **pipelined, highly parallelized structure**; (e) high degree of **configurability** to efficiently face **different applications**, some of which are listed below.

**Image processing for medical applications** (e.g., diagnosis based on 3D high resolution images which require long time for processing) will benefit from the increase in performance of this device. Industries and others research institutes will collaborate with this project with the aim of improving performance of **smart cameras** for smart cities and smart transports, and of speeding up **DNA alignment sequencing procedures**. The IMPART device develops from work in the HEP track trigger field. IMPART is a huge leap forward that will support a wide variety of applications including motion analysis, complex traffic monitoring, human tracking (e.g., to increase health and safety in work places).

## State-of-the-art and scientific background

A common problem in HEP applications, in particular at hadron colliders, is the identification of particle tracks inside the detector. Experiments at LHC, such as ATLAS<sup>1</sup>, produce a huge amount of data. Since a limited amount of events can be transferred to a storage system for subsequent off-line processing, an enormous data reduction must be performed. To this aim, a trigger system is used to recognize particle tracks in real time<sup>2</sup>.

Track detection can be performed by comparing data from detectors to a set of pre-computed patterns stored in a memory. The pattern recognition problem can be solved by an Associative Memory (AM) ASIC chip<sup>3</sup> which exploits parallelism to the maximum level.

The capability of performing real-time identification of tracks at a hadron collider leads to a very high background rejection. For example, many overlapping proton-proton collision can be separated using tracks to reconstruct the primary vertexes. Due to the complex structure of events, the trigger system must select few tracks among several tracks. Depending on available latency and output bandwidth, it is possible to use commercial CPUs or dedicated hardware. Theoretically, CPUs could provide the expected results (in terms of efficiency and resolution) but they require very large computing power to keep up with the event rate: ATLAS and CMS online tracking is performed only inside Regions of Interest (RoI), which are found by calorimetric or muonic selections executed before tracking. At high luminosity, it will not even possible to complete the tracking of all found RoIs<sup>4</sup>. Due to the event complexity increase and to the huge amount of data, CPU's tracking capability (with a reasonable number of CPUs per system) will not be sufficient to perform these tasks in future very high luminosity runs. Dedicated hardware allows to reconstruct tracks much faster than before. At the same time, algorithms have to maintain a good final quality. Speed and high quality requirements have led to the insertion of AM chips into the hadronic collider trigger systems, aiming at implementing fast and high quality tracking. The first dedicated AM chip for high-energy physics was designed in CMOS 180 nm<sup>5</sup> for the CDF experiment at the Tevatron, near Chicago. The AM concept was awarded the Panofsky prize<sup>6</sup> in 2009 since this technology was fundamental for the success of the CDF experiment in flavour physics. A new version of AM chip is under development in 65 nm CMOS for the ATLAS experiment at CERN<sup>7</sup> with the final submission expected in a few months. Several working prototypes of AM chips have been submitted under my supervision.

Currently, I am leading (with Francesco Crescioli, LPNHE, Paris) a group of people devoted to the design and to the characterization of Associative Memory chips within the FTK (FastTraKer) project. The FTK AMchip is designed in participation with LPNHE and funded by several institutions (INFN, Laboratoire de Physique Nucléaire et de Hautes Energies: LPNHE - Paris, University of Chicago, Desy, University of Heidelberg and University of Geneva). We expect to conclude the design of last version of the chip for the ATLAS application within few months (Dec. 2014).

This FTK AM chip, however, is very specific and lacks some features: (a) **flexibility**: the ASIC is tailored for pattern matching in the ATLAS experiment; (b) **fixed/low latency**: the chip is designed to meet the ATLAS level-2 time constraints. The IMPART project will expand our research toward a **more flexible ASIC, which will be assembled in a multi-chip package** together with an FPGA. The aim is to reduce

---

<sup>1</sup> ATLAS Collaboration, The ATLAS experiment at the CERN Large Hadron Collider, *IOP J. Instr.* 3 (2008) S08003.

<sup>2</sup> W. H. Smith, Triggering at LHC experiments, *Nucl. Instr. and Meth. in Phys. Res. – Sect. A* 478 (2002) 62–67.

<sup>3</sup> M. Dell'Orso and L. Ristori, VLSI structures for track finding, *Nucl. Instr. and Meth. in Phys. Res. – Sect. A* 278 (1989) 436–440.

<sup>4</sup> ATLAS Coll. Eur. Phys. J. C. 72, 2011.

<sup>5</sup> A. Bardi et al., A large Associative Memory system for the CDF Level 2 Trigger, *Proc. IEEE Nuclear Science Symposium* (1998): 312-313.

<sup>6</sup> <http://www.aps.org/programs/honors/prizes/panofsky.cfm>

<sup>7</sup> A. Annovi et al. Associative memory design for the fast track processor (FTK) at ATLAS. *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference - NSS/MIC* (2011): 141-146.

**latency and power consumption**, and to increase **hardware flexibility** (to cope with requirements of interdisciplinary applications).

The IMPART idea was born inside the INFN institutions leading the FTK collaboration. Hence, three main groups will participate under my leadership: my division INFN-Milan, LNF (under supervision of Matteo Beretta), and Pisa (under supervision of Calliope-Louisa Sotiropoulou). The 3 institutes will provide infrastructures for new developments, as an important FTK spin-off. Our goal is to exploit the knowledge acquired in the HEP/FTK field to implement a new multi-chip system. A fruitful collaboration with LPNHE group in Paris has been historically established on FTK AM chip design, and will continue for the ASIC development in this project. The LPNHE group will also be a user of the IMPART device for DNA sequencing. The IMPART system will be useful for several interdisciplinary fields (e.g., image recognition for medical and transport applications, and DNA alignment sequencing). The EU Horizon 2020 program offers large opportunity for research on smart embedded systems to prevent/mitigate transport accidents and environmental disasters. For this reason, in the IMPART project we propose a work package that aims at using the IMPART hardware for smart cameras.

Smart cameras capture high-level description of a scene and perform real-time analysis of what they see. Moving well beyond pixel processing and compression, these systems run a wide range of algorithms to extract meaningful information from streaming video. These devices can support a wide variety of applications including surveillance, motion analysis, traffic monitoring, and so on. However, a system with a single camera has a limited field of view. Furthermore, some objects inside the scene may be blocked from the point of view of a single camera. Distributed cameras are often used to aid the analysis process, for example, to avoid blockages. When analysing video streams from multiple cameras, data fusion is a design challenge. Traditionally, multiple camera systems for video processing have relied on central servers: the captured data is sent to a central server (or perhaps a cluster of servers) for processing. Server-based video analysis systems simplify synchronization and data-sharing problems, but the centralization of images can lead to bottlenecks for very large data rate. For this reason, data compression is performed before transferring images.

Currently, two basic types of compression exist: temporal and frame-by-frame. Temporal compression, such as H.264, periodically saves a full picture and uses complex algorithms to interpret what is happening between two full picture frames. Frame-by-frame compression, for example MJPEG, takes a full picture for each frame. Most common compression algorithms are MJPEG, H.264, and MPEG4. MPEG4 is used for low-quality, low-bandwidth video, while MJPEG and H.264 work with higher quality HD/Megapixel video. It is important to notice that a good resolution is required to monitor events and to analyse detailed region of interests. For this reason, data reduction would require MJPEG or H.264 implementation, as both algorithms are able to reduce data by a factor of 5. H.264 has to collect a bunch of pictures (some full picture frames, some differential prediction frames) before it begins the compression process: **few seconds** are required before the event is displayed on the monitor.

Latency should have no impact on a system that needs to record videos from cameras only. **However, for safety-critical applications where real-time monitoring is essential (e.g., transports, or personnel tracking in a dangerous environment), latency could lead to serious repercussions.**

A solution could consist in sending raw data to the central server(s). However this solution has severe penalties, as a high-performance network is required to connect the camera nodes to the server, and the system may consume a large amount of energy which may exceed supply capabilities. For this reason, new implementation algorithms can provide new solutions and improve the smart camera systems performance in critical situations. We plan to use the IMPART system to implement a huge, very fast and general

compression of data, exploiting the algorithm proposed by Del Viva et al.<sup>8</sup> They have demonstrated through a Monte Carlo simulation that pattern matching can extract the relevant features of an image, reducing the amount of the to 10% or even further, if needed. Further details will be given in the next section. The AM chip will select the relevant Regions of Interest and the FPGA will work on them.

Bioinformatics and genomics are other important scientific fields where the IMPART chip can be used. These fields are a perfect example of application, since high bandwidth data must be processed in real time using pattern recognition and other data reduction algorithms. The analysed data is typically described as a sequence of nucleotides (DNA, RNA) or proteins. These sequences can be converted into character strings. Characters are taken from a limited alphabet: 4 symbols for DNA and RNA and about 30 for proteins. Another important consideration is the number of genome characters: a simple organism such as the sea squirt *Ciona savignyi* has a genome represented by a string of about 170 million characters, while the human genome is made of about three billion characters.

Recently developed techniques have strongly reduced the cost of DNA sequencing. In addition, the amount of genetic data available for studies has increased in the last few years. Normally, produced data is made of several millions of short sequences. For example, the SOLiD<sup>9</sup> sequencing method generates about 1.2 billion sequences of 35-50 characters per run. Such short sequences must be mapped over the reference genome in order to be analysed. Mapping algorithms may take care of two main aspects: sequencing errors and natural variability of the sequence among different individuals. It is worth saying that this is a very hard computational problem. In addition, the current approach, consisting in running sophisticated software over big and powerful clusters of PCs, is rapidly becoming inadequate to deal with the ever increasing amount of data. There are several examples of optimization of the sequence data analysis by using more specialized hardware such as GPGPU (General-purpose graphics processing unit)<sup>10</sup>, Cell Broadband Engine and FPGA<sup>11</sup> showing that is possible to increase the speed of the analysis algorithms by several orders of magnitude. Our proposed hardware composed by AM chip and FPGA is an ideal platform to develop a novel and competitive sequence processor for bioinformatics applications.

The IMPART project will continue the R&D in the HEP experiments aiming at improving tracking performance. A severe problem is the increase in complexity of online tracking for experiments at hadron colliders with high instantaneous luminosity such as the LHC after Phase-II upgrade<sup>12</sup>. The last version of the AM chip (that we are using now for ATLAS level-2 trigger) is not powerful enough for future level-1 triggers. For this reason, the IMPART system is an interesting evolution for future projects, for example track triggers in the harsh environment of future LHC runs when up to 200 proton-proton collision will overlap in a single bunch-crossing.

## Objectives and detailed description of the proposal

The aim of the IMPART project consists in delivering a System in Package (SiP), made of an FPGA die (Xilinx Kintex-7) and an Associative Memory (AM) ASIC which will be designed in a 28 nm CMOS technology. The IMPART system will be useful to compute pattern recognition tasks for several interdisciplinary applications.

---

<sup>8</sup> M. Del Viva, G. Punzi, and D. Benedetti. Information and perception of meaningful patterns. *PloS one* 8.7 (2013): e69154.

<sup>9</sup> The SOLiD™ System: Next-Generation Sequencing. <http://solid.appliedbiosystems.com/>

<sup>10</sup> P. Klus et al. BarraCUDA - A fast short read sequence aligner using graphics processing units. *BMC Research Notes* (2012): 5-27

<sup>11</sup> J. Xianyang et al. A Reconfigurable Accelerator for Smith–Waterman Algorithm. *IEEE Trans. Circuits and Systems II: Express Briefs* 54:12 (2007): 1077-1081

<sup>12</sup> B T Huffman. Plans for the Phase II upgrade to the ATLAS detector. *Journal of Instrumentation* 9 (2014): C02033

The SiP assembly technique is essential to reduce latency and power consumption. Moreover, the bandwidth for communication between AM and FPGA will be increased. SiP assembly solution is feasible and IMEC (Belgium) has the facilities for SiP fabrication. Up to 200 wire bonds will connect the FPGA with the ASIC. The high number of FPGA-ASIC connections and the high bandwidth will lead to better integration. The interconnections will be designed for high-frequency communication: up to 200 MHz for single-ended wires and 10 Gbit/s for serial links.

The new AM chip will store more patterns per chip, will reduce latency by running faster and will consume less power per pattern. The IMPART project will study the implementation of a new chip based on smaller deep submicron technology (Taiwan Semiconductor Manufacturing Company 28 nm CMOS or similar). In fact the pattern density increases roughly by a factor of 5 for a 28 nm technology, compared to the last version of the chip which was designed in 65 nm<sup>13</sup>. Even more important, the scaled technology will reduce power consumption better than linearly with feature size. The increase of speed from scaled technology will also be important for level-1 trigger applications. A clock frequency of 200 MHz or so will be used in the new design, to reduce latency and to increase bandwidth both at the chip level and at the system level. Finally, serializer/deserializer blocks and all the logic that is easily provided inside FPGAs and hard to implement inside ASICs will be removed from the AM chip, since we plan to use the FPGA internal resources to communicate with the external world. In this way, the total number of serializer/deserializer links and logic used inside the system will decrease, with a subsequent reduction in power consumption, silicon area, latency and cost.

Multiple IMPART systems could fit on a compact PC board that could be used as coprocessor for several interdisciplinary applications, thus providing a spin-off of the research activity.

### **Pattern recognition approach for low power systems:**

Within the IMPART project, the Associative Memory (AM) device will be remarkably enhanced. The AM is a VLSI processor for pattern recognition based on Content Addressable Memory (CAM) architecture. However the AM and commercially available CAMs differs substantially. **The AM provides the unique capability to store partial matches as they are found and to use these partial matches to find correlation among data received at different times.** The AM stores each pattern in a single memory location like in the commercial CAMs, but the total number of available bits can be organized as an array of  $N$  independent words of  $M$  bits each. Each stored word contains a particular item to be identified in a flux of data that is delivered at the AM input. In fact, data are sent to 8 parallel buses, one for each word or group of words of the pattern. Each word is provided with dedicated hardware comparators and a match flip-flop.

A large AM cell bank stores all interesting patterns, for a given input resolution. The AM extracts pattern addresses, when a sufficiently high number of words have matched the incoming data (Fig. 1). For most practical problems the complete set of patterns with full resolution is extremely large. A smart approach consists in performing pattern matching at reduced resolution first, with a resolution adequate to simplify and reduce the amount of data, and then to refine matching using the FPGA. Essentially, this constraint arises from memory and bandwidth limitations of electronic devices. Mobile devices that work with battery or solar panels exhibit additional limitations because they cannot compute a large number of operations due to the reduced available power. For this reason, an elegant solution is proposed in IMPART project to solve this problem: the “variable resolution patterns”. A “don’t care” feature/bit is used to increase the pattern recognition efficiency at different resolutions. In other words, we can use patterns of variable shape. As a result of variable resolution, both the number of fakes and the bank size decrease, while efficiency is high. Hence, for a given efficiency, the number of required patterns will decrease, leading to a lower power

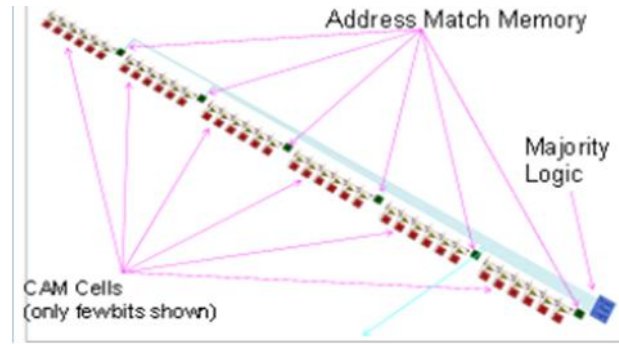
---

<sup>13</sup> L Frontini, S Shojaii, A Stabile, and V Liberali. A new XOR-based Content Addressable Memory architecture. *Proc. IEEE Int Conf. on Electronics, Circuits and Systems - ICECS* (2012): 701-704

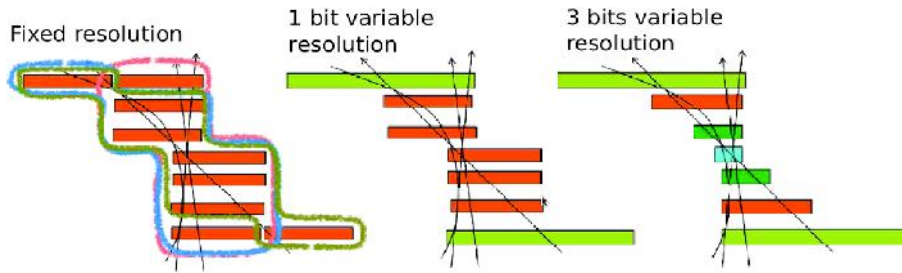


consumption.

Fig. 2 shows a qualitative example of the detection of 4 different particle trajectories. To achieve good efficiency, the fixed resolution approach requires to store 3 patterns, while with 1-bit or 3-bit variable resolution it is necessary to store a single pattern. In addition, for the 3-bit variable resolution case the detector volume for track identification is much smaller, since very thin bins are used when possible, leading to a good noise filtering and thus a fake decrease. Simulations demonstrated that an AM system with variable resolution can be as effective as a 5 times larger AM system with fixed resolution<sup>14</sup>.



**Fig. 1:** Structure of one pattern AM for the IMPART project.



**Fig. 2:** Diagram illustrating the multi-bit implementation of variable resolution patterns: **(left)** fixed resolution patterns - red rectangles are patterns without “don’t care” bits; **(center) 1-bit variable resolution pattern** - green rectangles are pattern subset with one “don’t care” bit, red rectangle are pattern without “don’t care” bits; **(right) 3-bit variable resolution pattern** - green rectangles are pattern subset with three “don’t care” bits, red rectangles are pattern subset with two “don’t care” bits, dark green rectangles are pattern subset with one “don’t care” bits, and cyan rectangle is a pattern subset without “don’t care” bits .

## The architecture of AM chip:

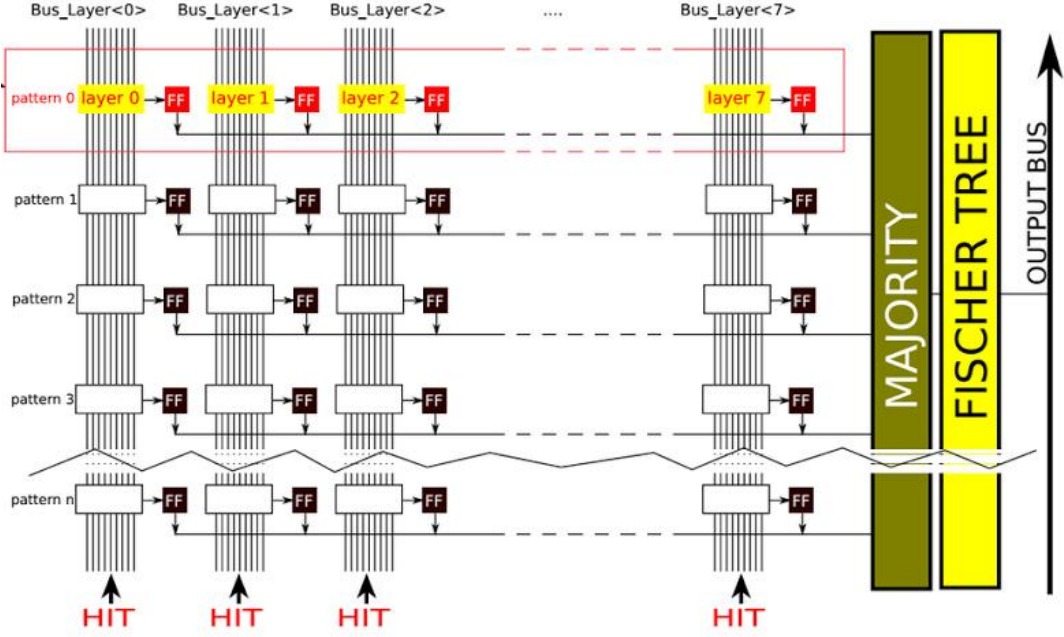
As described above, the AM chip is made of an array of CAM cells. The total number of bits assigned to a pattern is divided between a group of cells that can be configured of different size (16, 32, or 64 bits per cell) depending on the application. The number of the bit per word can be parameterized to obtain a good level of flexibility. The AM cells perform a bit-wise comparison between the input data (the input bit buses) and the pre-stored bits in the cells. If all the pre-stored bits match the input bits exactly, the cell flip-flop register is set to a high logic value. After that, a majority block counts how many flip-flops have been set to a high logic value, and the sum of ‘1’ is compared to a threshold parametric value. If the sum is greater than the threshold, the **corresponding** pattern address is transferred to the output bus. Obviously, many patterns addresses could pass the majority threshold. For this reason, a Fischer tree<sup>15</sup> readout method is used to organize the output data with a priority approach. Since the fraction of selected patterns found at the output is significantly smaller than the total number of patterns registered inside the bank, the AM chip will be

<sup>14</sup> A Annovi, et al. A new variable-resolution associative memory for high energy physics. *Proc. IEEE Int. Conf. on. Advancements in Nuclear Instrumentation Measurement Methods and their Applications - ANIMMA* (2011).

<sup>15</sup> P. Fischer. First implementation of the MEPHISTO binary readout architecture for strip detectors. *Nucl. Instrum. Meth. - A* 461 (2001): 499.

designed to exploit the full bandwidth in input (working frequency up to 200 MHz), with 8 parallel input buses and a single output bus.

Overall, the AM chip will contain three blocks: the CAM cell array, the majority blocks, and the readout tree. Fig. 3 shows the architecture of the proposed AM chip. The IMPART chip will be provided with the boundary scan standard interface to the other devices in the board. JTAG control logic will be placed into the ASIC to configure the chip, for example to write patterns in the CAM cell array and to set the majority logic threshold mentioned before. Both the ASIC and the FPGA JTAG interfaces will be accessible either separately for test purpose, or in the same JTAG chain to perform all configuration operations through (and with the help) of the JTAG system.

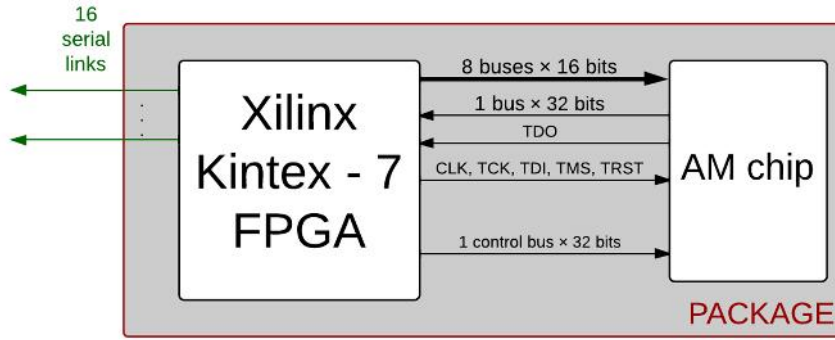


**Fig. 3:** Architecture of the proposed AM chip. The HIT buses are the input data: the number of bit per HIT is a parametric value. The number of layers per pattern is a parametric value also.

To demonstrate feasibility, we plan to design a small test chip prototype with a standard size ( $1416 \mu\text{m} \times 1416 \mu\text{m}$ ), containing 262.14 kbits (e.g., 16388 words if each word has a 16-bit width). This means that we expect to perform  $32.8 \times 10^{12}$  word comparisons in parallel per second. In future (after year 2017) we will plan to design a larger chip for several applications (e.g., for level-1 triggers we plan to design a chip containing 0.5 Mpatters of  $8 \times 16$  bits).

### System-in-Package interconnects

Communication between the FPGA and the AM chip will occur mainly through single-ended bondwires, at a frequency up to 200 MHz. The overall bandwidth will be 40 Gbit/s. Data from the FPGA to the ASIC will occupy the major part of wires in the IMPART system. We expect to use 8 input buses, each of them made of 16 bits, transmitted in parallel from the FPGA to the ASIC. In addition, the conventional 5 JTAG bits (TDI, TMS, TCK, TDO, TRTS) will be used to control the AM chip via JTAG and 32 control lines will be used to change some AM chip parameters at high speed from the FPGA (e.g., the pattern matching threshold). Matching results will be transferred from the AM chip to the FPGA with a single bus of 32 bits. A single-ended clock (working at 200 MHz) will be used to enable or disable the real-time pattern matching (performed by the AM chip). Writing operations on the AM chip will be performed via JTAG. Fig. 4 shows the main interconnections inside the IMPART multi-chip package.



**Fig. 4:** Data and control interconnections.

### The FPGA tasks

The FPGA will run all the tasks required by the specific application complementary to the pattern matching function that will be performed very efficiently by the AM chip. Depending on the application, tasks will differ: (1) for recognition of natural images, the FPGA will process the image contours filtered out by the AM; (2) for DNA alignment detection, the FPGA will serve as input-output manager for the AM; and (3) for track reconstruction at CMS and ATLAS experiments, the FPGA will perform high resolution track fitting among track candidates found at low resolution by the AM and all interfacing functions between pattern matching and track fitting steps.

For some applications, we plan to implement a very simple machine (Microblaze CPU + AXI system + registers) inside the FPGA. In addition, a firmware provided with a Linux module driver will be written to link the FPGA to the AM chip. In this way the FPGA could be used as a computing machine linked with a boost associative memory device. This allows to execute different algorithms on a standard Linux system, thus providing easy access to non-expert users (not only to the designers of the AM chip). However, the microblaze CPU requires FPGA resources, thus decreasing performance of the parallel pattern recognition task. For more demanding applications requiring low power and high pattern recognition performance, we will develop a dedicated FPGA firmware. Essentially, the FPGA will: (1) transfer/receive data outside the chip through serial links at high frequency; (2) implement the variable resolution control logic, so that the AM chip will be used in different ways with different resolutions, and configure all other AM chip functions; (3) implement *ad-hoc* algorithms (such as principal component analysis, clustering, etc.) on the microblaze CPU or with dedicated firmware; (4) allow to write patterns on-the-fly into the AM, thus combining learning and processing capability at the same time.

The FPGA that we plan to use is a device of the Xilinx Kintex-7 family: namely, the XC7K325T with 326080 logic cells, 50950 slices, 4000 kB distributed RAM, 840 DSP Slices, 16 serial GTX links, 1 ADC, and 500 user I/O. From these numbers, it is clear that the number of wire bonds between the FPGA and the external pins will be quite larger than the number of wire bonds between the FPGA and the AMchip.

### Tracking system for the high luminosity LHC

The development of the FPGA firmware is already ongoing, and it has been presented at the Technology and Instrumentation in Particle Physics conference (TIPP 2014) in Amsterdam<sup>16</sup>. The firmware will be first tested on a mini board and then implemented in the IMPART system, when ready.

<sup>16</sup> C. Gentsos et al. Future evolution of the Fast TracKer (FTK) processing unit. *PoS Int. Conf. on Technology and Instrumentation in Particle Physics* (2014).



## Computer vision for Smart Cameras and Medical Imaging Applications

One of the goals of the project is to use the IMPART multi-chip system to reproduce the initial stages of the brain visual processing. In particular, we plan to use the ASIC to extract **object contours**. Del Viva's paper<sup>17</sup> describes convincing models about the brain functioning at this initial stages. The paper points out a strong correlation between data reduction performed by pattern matching in HEP experiments and first level data reduction performed by the human brain. In particular, strong data reduction of external stimuli before higher level processing is essential in applications characterized by a high-rate data flux, typical of the brain vision processing. Our system will be able to process complex coloured and highly textured natural images with multiple luminosity levels varying over time due to changes in illumination or motion of objects. Simulations of such visual analysis are an onerous task to implement, if possible at all, in usual computers. With the system in package (FPGA + AM chip) we can build a small coprocessor which will make a huge improvement in this kind of computation, by achieving both good efficiency and low power consumption.

Different computing phases are necessary.

1. **Training phase:** The embedded system receives the image bit-streams (e.g., data from a smart camera). The FPGA partitions/reorganizes the input data into small patterns (3×3 square pixel matrices in black-and-white (B/W, 1-bit depth) for static images, 3×3×2 to use four levels of grey in static images, and 3×3×3 cubic pixel matrices, 3 black-and-white frames taken at 3 subsequent instants, for movies). Then, for each pattern, the FPGA calculates the occurrence of the analysed pattern in the processed images/frames. This calculation is iterated for all possible patterns<sup>18</sup> in a large set of training images. In this way, different Probability Density Histograms (PDHs) are computed for different training image sets. Of course, medical images have different PDHs than natural images.
2. **Pattern selection:** In this step, the system must decide which set of patterns has to be selected for memory storage (in the AM bank). To maximize the capability to recognize shapes (both human-brain recognition and artificial recognition), we adopt the hypothesis described in Del Viva's paper, i.e., that the principle of maximum entropy is a measure of optimization. The set of patterns that produces the largest amount of entropy allowed by system limitations<sup>19</sup> is the best set of patterns that we can select to filter our images or videos. From Del Viva's paper, the entropy yield per unit cost and for each pattern is given by:

$$f(p) = \frac{-p \log p}{\max(1/N, p/W)}$$

where  $p$  is the probability that a given portion of the input data matches a specific pattern,  $N$  is the maximum number of storable patterns and  $W$  is the maximum allowed total rate of pattern acceptance. Fig. 5 shows three entropy yield functions for different values of  $W$  and  $N$ .

The optimal solution in the selection of pattern is made by choosing the set of pattern such that  $f(p) > c$ , where  $c$  is determined by the computational constraints:

$$\int_{f(p)>c} f_n(p) dp < N \quad \frac{1}{N_{tot}} \int_{f(p)>c} p f_n(p) dp < W$$

where  $f_n(p)$  is the density of patterns having probability of occurrence  $p$  normalized to the total number  $N_{tot}$  of patterns in a set of mutually exclusive patterns  $Q$ . Fig. 6 shows an example of pattern selection performed on natural digitized images (Fig. 7-b). This task will be performed by the FPGA inside the

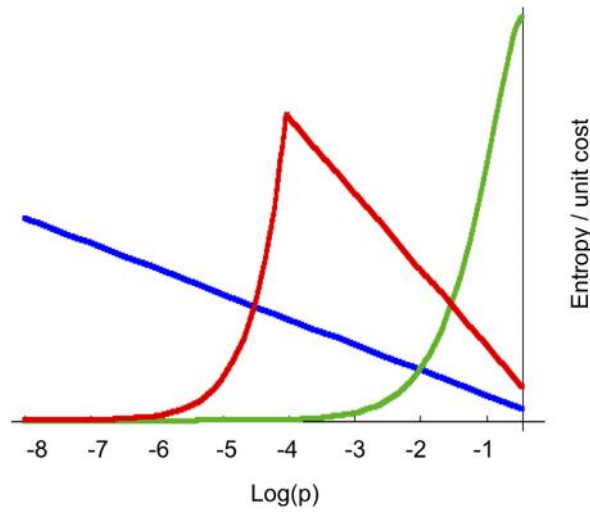
<sup>17</sup> M. Del Viva, G. Punzi, and D. Benedetti. Information and Perception of Meaningful Patterns. *PloS one* 8.7 (2013): e69154.

<sup>18</sup> The total number of possible patterns is  $2^9 = 512$  for static image and  $2^{27} = 134$  Mpatters, if B/W is used.  $2^{18}$  is the total number of patterns for static images using 4 levels of grey.

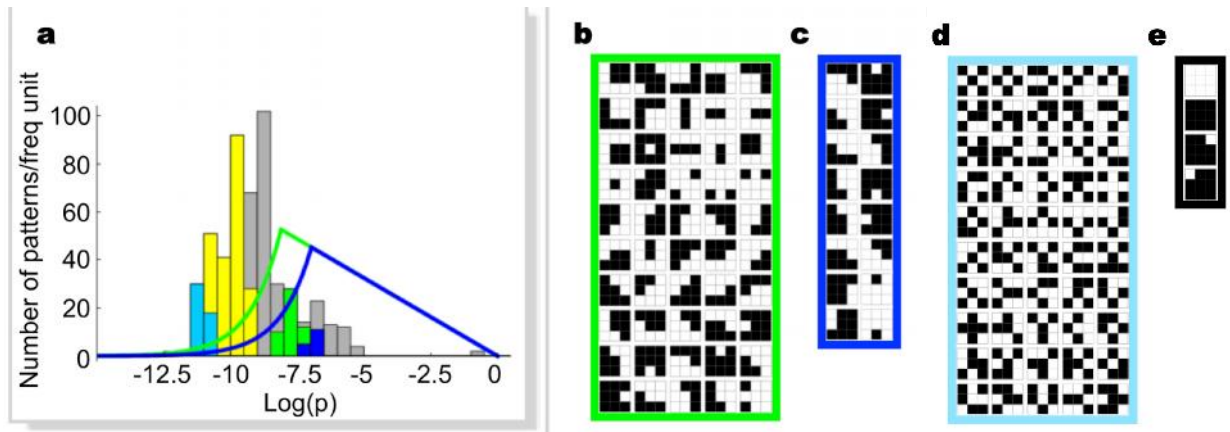
<sup>19</sup> System limitation can be summarized in two main parameters:  $N$  maximum number of storable patterns, and  $W$  maximum bandwidth.

package.

3. **Writing operation:** The relevant pattern (selected in the second step) is **written** in the AMchip bank. The writing operation is made via JTAG by means of a system controller inside the FPGA.
4. **Real-time pattern recognition:** After training, selection and writing operations, the system is able **to work in real-time at the maximum working frequency** and to perform **parallel recognition** of patterns in the data stream. **The filtered pattern is recognized** by the AM chip (through a bit-wise comparison) and the pattern address is transferred at the output from the AM chip to the FPGA.
5. **Output formatting operation:** The resulting patterns are reorganized into **a new compressed image** by the FPGA, to produce the filtered images/videos, called “sketches”, where only the boundaries of the relevant objects are kept.
6. **Clustering operation:** A clusterization of contours is performed to convert the image from a raster format into a vector image. This task is performed by the FPGA inside the package.
7. **A shape recognition algorithm** will be implemented in the FPGA, aiming at recognizing the salient shapes which will be used for the final high-level elaboration.



**Fig. 5: Entropy yield per unit cost, as a function of the pattern probability.** Blue curve: limited bandwidth and unlimited pattern storage capacity ( $W=0.001$ ,  $N=$  ); green curve: limited storage and unlimited bandwidth ( $N = 100$ ,  $W=$  ); red curve: limited bandwidth and storage ( $N=100$ ,  $W=0.001$ )



**Fig. 6: Probability distribution of the  $N_{tot}=2^9$  possible  $3 \times 3$  square pixel matrices in black-and-white (1-bit depth) for natural images.** a, Blue curve and histograms is obtained setting  $W=0.05$ ,  $N=15$ ; green curve and histograms with  $W = 0.05$ ,  $N=50$ ; b,c Visualization of the pattern sets shown in (a), in green and blue respectively (resembling to edges, bars, or end-stops in several orientation); d, Visualization of the lowest-probability patterns (resembling visual noise); e, Visualization of the highest-probability patterns (resembling uniform luminance)



**Fig. 7: Examples of images from the database and sketches used.** a, Example of full color natural images extracted from the database<sup>21</sup>; b Digitized versions of images in (a); c, Sketches obtained from the images in (b) by using the optimal pattern set in Fig. 6-b

During the real-time image processing phase, the FPGA continuously monitors the probability distributions of salient patterns and, if the downloaded patterns are outside the correct frequency range (i.e., predictive of their importance), image processing will stop and a new training phase is performed to update the pattern list. Aiming at avoiding “dead time” due to non-fully parallel training steps, the bank can be divided into two equal sub-banks. In this way, the system could use the sub-bank 1 in real-time, while writing and training operation could be performed on the sub-bank 2 with updated values. When training and writing steps are over, the system could switch real-time operation from bank 1 to bank 2 and so on, iteratively.

Dedicated hardware based on Del Viva’s algorithm could be used for medical imaging as well. It is clear that medical imaging techniques produce images which are somehow very similar to natural images, but they could be more complex. **The IMPART system filtering function, combined with the computing power of parallel arrays of FPGA, and applied to image reconstruction, could heavily reduce execution time and power consumption of image processing algorithms. This could be useful to filter complex/crowded images for critical applications.**

One interesting application area is in the realm of automated medical diagnosis by imaging. Possible application targets are the analyses of time-varying images and of the output of very accurate instruments producing a huge amount of data (e.g., optical coherence tomography). A future R&D activity could be focussed on real-time applications for adaptive radiotherapy in collaboration with the INFN-Torino research group working on medical physics (Simona Giordanengo's CSN5 funded project in 2013).

Another important application (also taking Horizon 2020 expectations into account) are smart cameras for smart cities and smart transportation systems. As described above, smart camera systems will be widely used in the future to avoid limitations due to server-based video. Each smart camera is essentially made of an image acquisition block followed by a processing module built with a general-purpose processor and/or a digital signal processor. Our idea is that these systems will run the aforementioned algorithms in real-time with the goal to extract high-level descriptors from the streaming video. The IMPART system could extract the relevant features and substantially reduce the amount of data to be shared between many devices or to be analysed centrally. The “filtered image” will be shared among the smart camera grid, aiming at detecting salient events.

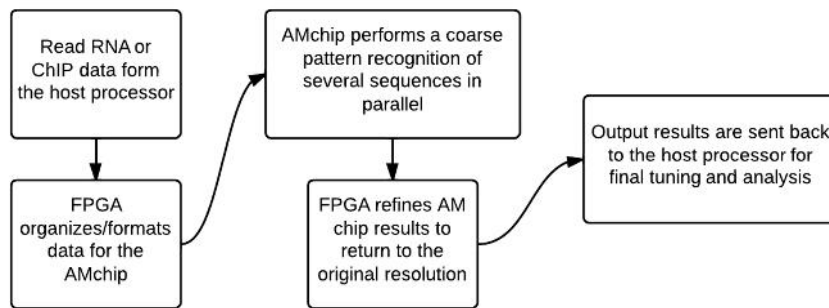
<sup>21</sup> A Olmos. A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception* 33.12 (2003): 1463-1473.

Our idea is to collaborate with a SME, namely, EMC s.r.l.<sup>22</sup> (headquartered in Poggibonsi, Siena, Italy), with the scope to employ our multi-chip system in the EMC smart camera devices, thus achieving a new prototype of smart camera that demonstrates and exploits the validity of the Del Viva's algorithm. Smart cameras will operate in different environments, including outdoor environments requiring solar panels or batteries. For this reason, a system that minimizes the power consumption is essential.

## DNA alignment sequencing

As presented in state-of-the-art section, bioinformatics and genomics are rapidly reaching a point where the amount of data that can be processed is limited by the available computing power. Currently, a lot of services (e.g., dedicated cloud services) compute every day a huge quantity of bioinformatic alignment data. For this reason, hardware accelerated solutions are expected to be the next research step.

The HEP experience might be very valuable and the IMPART system solution could provide a novel approach, which is somehow complementary to what is currently in use in the bioinformatics field. The IMPART system can be a DNA co-processor to speed up critical parts of the alignment algorithm and it will allow mapping algorithms to run on a small computer cluster.



**Fig. 8:** Possible flow diagram for DNA sequencing.

As for the HEP application, we expect to use the AM chip for a coarse alignment on whole reference genome, while optimal alignment of short segments is performed by the FPGA. Fig. 8 shows a possible flow diagram for DNA sequencing.

A rough estimate of the performance of an IMPART-based system for DNA sequencing can be extrapolated with some back-of-the-envelope calculations. The human exome (a 1.5 % subset of the human genome) is very important to study mutation with an impact on human health. Human exome data ranges from 25 to 40 million nucleotide pairs. To encode a nucleotide in the FASTA format<sup>23</sup> at least 4 bits are needed. With the 8 input bus  $\times$  16 bits of the IMPART system is possible to compare nucleotide sequences (reads) of 64 bases at 200 MHz rate, thus is possible to process the whole exome data in about 0.3 s comparing it against 2 million bases stored in the IMPART chip associative memory. Hence, about 13 interactions are needed to align a entire human exome data (of 25 million of reads) for a total computation time of about 4 s. The IMPART system is a great improvement with respect to the time required by commercial machines, such as Bowtie-based machines, which can align more than 25 million reads in 1 CPU hour (3600 s)<sup>24</sup>: the improvement factor is about 900x. The IMPART system fully explores the possibility to perform sequencing of RNA, ChIP data in different format (e.g., FASTQ with the phred-33 quality scores that could be implemented with

<sup>22</sup> <http://www.emcr.it>

<sup>23</sup> <http://www.ncbi.nlm.nih.gov/BLAST/blastcgihelp.shtml>

<sup>24</sup> B Langmead et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10.3 (2009): R25.

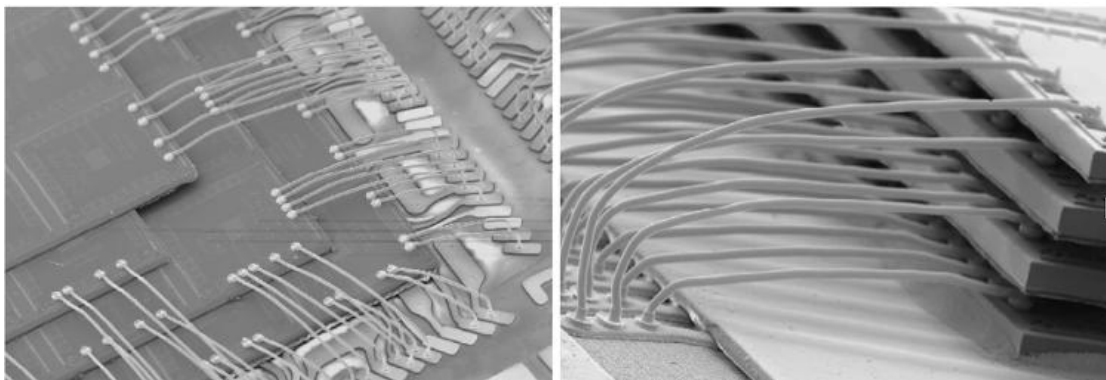


a sort of “don’t care” bit map). We will start working on non-splicing sequences (e.g., ChIP-seq<sup>25</sup>). If results are promising we will move our research on splicing sequences (e.g., RNA-seq<sup>26</sup>). In addition, we expect to use the IMPART system for *de novo* transcriptome assembly<sup>27</sup> which is the method of creating a transcriptome without the aid of a reference genome. The latter case requires more computational effort than conventional techniques.

### Multi-chip package technology and power considerations

In the IMPART project we will study/evaluate which kind of 3D assembly technology is best suited to optimize power consumption and latency. Fig. 9 shows two possible technologies suitable for the IMPART multi-chip system. The evaluation procedure will consist in performing electrical and thermal 3D simulations. Studies will be made in collaboration with IMEC, a micro- and nanoelectronics research center with headquarters in Leuven, Belgium. We previously collaborated with IMEC to integrate the 65nm AM chip with BGA package. IMEC has a broad expertise in multi-chip package technologies and will help us to carry out this task. A significant amount of IMPART funds will be used for bonding and packaging services from IMEC. Details are provided in the cost document.

Finally, a power consumption study will be carefully performed at system level (SiP level). Our expectation is to have an AM chip containing about 262 kwords and consuming less than 0.1 W. In this way, the main contribution to power consumption will occur inside the FPGA (about 12 W at maximum performance). The power figure of 0.1 W is an estimate for the small AM die (2 mm<sup>2</sup>) in the first IMPART device. In the future, larger area AM dice will have a still manageable consumption of a few watts, as power consumption scales near linearly with die size.



**Fig. 9:** Example of Wire Bonding of Stacked Die and Reverse Wire Bonding with Au Studs on the Die<sup>28</sup>.

## Description of the proposed research activity underlying Research units description and duties

The IMPART project will be divided in work packages. Essentially, for each work package a person has been assigned as WP leader.

The resources that we expect to use to perform work packages are:

1. INFN (Sezione di Milano) - **Principal Investigator: Alberto Stabile**. The microelectronic lab in

<sup>25</sup> PJ Park. ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics* 10.10 (2009): 669-680.

<sup>26</sup> Z Wang, M. Gerstein, and M. Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10.1 (2009): 57-63.

<sup>27</sup> I Birol et al. De novo transcriptome assembly with ABySS. *Bioinformatics* 25.21 (2009): 2872-2877.

<sup>28</sup> Amkor, Stacked CSP Technical Data Sheet. <http://www.amkor.com/go/StackedCSP>



Milan consists of 2 young postdocs and some undergraduate/master thesis students. The lab is equipped with all needed services for CAD design and ASIC/system characterization;

2. INFN (Laboratori Nazionali di Frascati) - **Group team leader: Matteo Beretta**. The microelectronic lab in Frascati consists of some master thesis students. The lab is equipped with all needed services for CAD design and ASIC/system characterization;
3. INFN (Sezione di Pisa) - **Group team leader: Calliope-Louisa Sotiropoulou**. The microelectronic lab in Pisa consists of some master thesis students. The lab is equipped by with all needed services for CAD design and ASIC/system characterization.

**The 3 laboratories are funded by the FTK project and will provide their facilities for these new developments. The costs for the IMPART project will be used to fabricate the new IMPART system.**

## **WP1: AM chip design and simulation (A. Stabile)**

This work package consists in the design of the Associative Memory ASIC in a standard 28 nm CMOS process. In addition, this task includes the pre-fabrication simulations (digital and analog). We plan to use a mixed design approach: more repetitive blocks (e.g., cell arrays) will be designed by hand with a full-custom approach, while the more complex design module (from the logical point of view) will be placed and routed with and automatic software (e.g., using Cadence Encounter). The work will be essentially divided in six tasks:

### **WP1.1: AM cell design (M. Beretta)**

In this task an associative memory cell will be designed with the aim to minimize the silicon area and the power consumption.

### **WP1.2: Front-end design (A. Stabile)**

The front-end design consists in writing a behavioral netlist (e.g. written in VHDL) of the whole chip. After that, the aforementioned netlist will be used as input of a synthesizer (e.g., Synopsis dc\_shell) to obtain an RTL netlist which is composed of standard cells, full-custom blocks and their interconnections. Finally, the RTL netlist functionality will be verified by means of digital simulations of the entire chip.

### **WP1.3: Floorplan and back-end design (A. Stabile)**

The floorplan (first step of an ASIC design) consists in defining the chip area, the I/O pad and block position, and in designing the power supply interconnections. After this step, the RTL generated in the WP1.2 will be used to place and route the chip automatically. We plan to use the Foundation Flow of Cadence which performs a lot of steps such as clock tree place and route, static analysis of time, signal integrity analysis, RC analysis in multi corner multi mode, etc.

### **WP1.4: Functional simulation (N. Biesuz)**

The back-annotated netlist (with RC and crosstalk delays) will be validated through digital simulations to confirm the functionality of the layout produced by the place & route step.

### **WP1.5: Chip integration and submission (A. Stabile)**

In this task, final checks will be performed: Design Rule Check (DRC) and Layout Versus Schematic (LVS). The chip design will be sent to the foundry for fabrication.

### **WP1.6: Characterization of the AM chip (M. Beretta)**

In this task, the AM chip will be characterized in a single-chip package.

## **WP2: System in Package (SiP) design, simulation, and tests (M. Beretta)**

### **WP2.1: FPGA to AM connections (M. Beretta)**

Electrical and thermal studies will be performed to verify functionality and signal integrity of wire bonds. In addition, power consumption will be simulated. **E. Rossi** will be responsible of interconnection design, while **M. Beretta** will be responsible of power consumption and thermal dissipation simulations. These tasks will be performed in collaboration with IMEC.

After simulations, IMEC will produce the multi-chip package in collaboration with ASE<sup>29</sup>.

### **WP2.2: Characterization of the System in Package (M. Beretta)**

The IMPART chip will be characterized by means of an external PC directly connected to the FPGA inside the package. We plan to use an I2C protocol for the ad-hoc firmware and a simple SSH protocol for the connection to the microblaze machine. Characterization results will confirm functionality of the IMPART system and will provide power consumption measurements.

*We have submitted 3 over 3 successfully working devices in the past 3 years. Based on our previous experience a major problem in the ASIC design is a very remote possibility.*

*However for a risk analysis consideration, we evaluate that this possibility would have a small impact on the project because the ASIC cost is a small fraction of the total cost (20%) and a resubmission would add about 3 month to the planned schedule.*

## **WP3: Smart Cameras and DNA sequencing (C. L. Sotiropoulou)**

LPNHE will collaborate to this work package. In particular, DNA applications will be led by Francesco Crescioli.

### **WP3.1: Study of the techniques (C. L. Sotiropoulou)**

This task is devoted to study techniques used to filter images for smart cameras applications.

### **WP3.2: Firmware and software implementation (C. L. Sotiropoulou)**

In this task, firmware and software will be implemented for the Kintex-7 FPGA, aiming at compatibility with smart cameras data and the IMPART system.

### **WP3.3: Characterization of the IMPART chip (A. Stabile)**

A characterisation of the firmware + software (written to be compatible for smart cameras and DNA alignment algorithms) will be performed by using the fabricated IMPART chip.

### **WP3.4: Field tests with image and video acquisition from smart cameras (C.L. Sotiropoulou)**

The IMPART system will be used to acquire real-time images and videos in collaboration with EMC s.r.l.

## **Milestones/deliverables and timeline for each research institutions**

The deliverables for each research institutions correspond to the Work Packages. More important deliverables and milestones are summarized in this list:

**Deliverable (end of May 2015, M. Beretta):** full custom layout design of the array of AM cells.

**Deliverable (end of May 2015, A. Stabile):** complete design of front end (synthesized netlist) of the AM chip.

**Milestone (July 2014, C.L. Sotiropoulou):** after the study on smart cameras techniques, we will choose the optimal technology for the IMPART prototype.

---

<sup>29</sup> <http://www.aseglobal.com/en/>

**Milestone (August 2014, M. Beretta):** after the study on 3D technologies, we will choose the optimal assembly technology for the IMPART purposes.

**Deliverable (September 2015, A. Stabile):** design of back end (automatic place & route) of the AM chip.

**Deliverable (end of November 2015, N. Biesuz):** functional simulation (back-annotated simulation with parasitic components) of the AM chip.

**Milestone (end of December 2015, A. Stabile):** signoff controls (DRC & LVS) of the AM chip and submission to the foundry.

**Deliverable (end of February 2016, M. Beretta):** design and simulation of the IMPART multi-chip package.

**Deliverable (end of February 2016, IMEC - TSMC):** the AM chip will be fabricated.

**Milestone (April 2016, M. Beretta):** the AM chip will be characterized in a single-chip package.

**Deliverable (April 2016, IMEC):** the IMPART multi-chip system will be ready for test.

**Milestone (August 2016, M. Beretta):** characterization of the IMPART multi-chip system and characterization protocol has to be written for further tests.

**Milestone (September 2016, A. Stabile and C.L. Sotiropoulou):** the IMPART multi-chip system will be evaluated for smart cameras and DNA sequencing applications.

**Milestone (December 2016, A. Stabile and C.L. Sotiropoulou):** the IMPART multi-chip system will be used to acquire data for smart cameras and DNA sequencing applications.

Fig. 10 shows the IMPART Gantt diagram with detailed description of milestones for each work package and task.

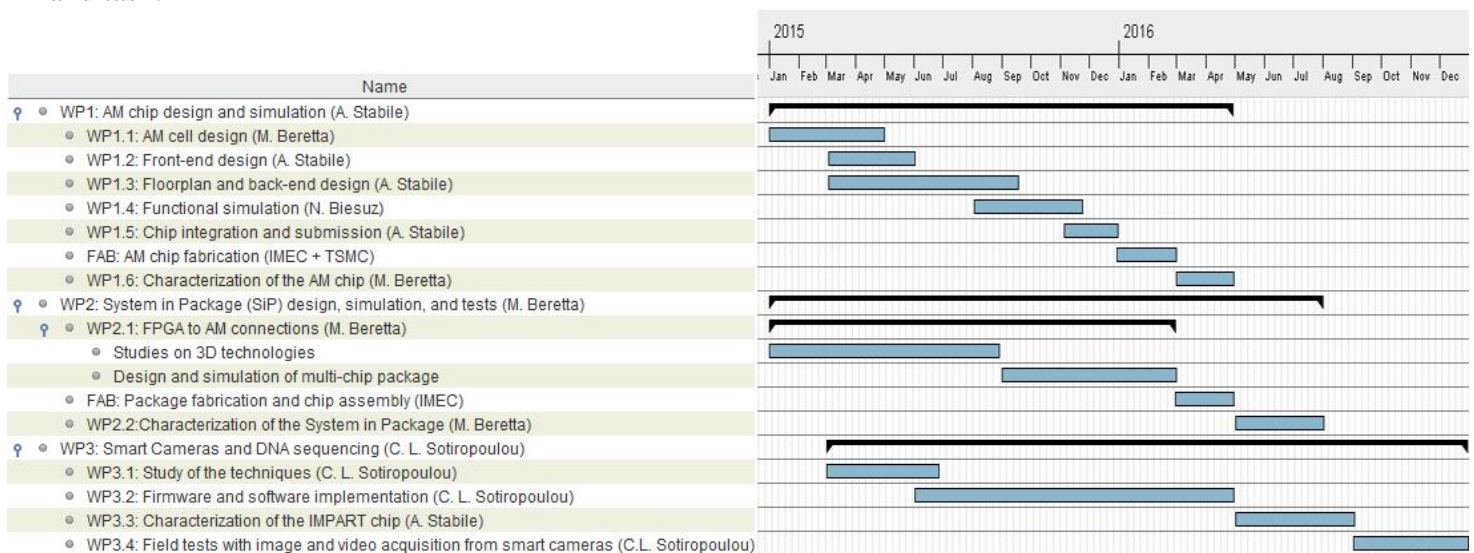


Fig. 10: IMPART Gantt diagram.

## Impact (scientific, technological, and socio-economic)

The IMPART project is expected to have an impact on different fields. Thanks to the implementation of a CPU + AXI system inside the FPGA, the system will be accessible (if needed) through a simple Ethernet interface from remote sites (either from research institutes or companies). The IMPART hardware will be useful for all applications requiring to run pattern recognition on a huge amount of input data with low power constraints. In addition, the mixed approach (FPGA + ASIC) follows the trend to use FPGAs in physics applications, while adding the benefit of reduced power consumption and increased bandwidth for more repetitive task, such as pattern recognition, which are performed by the low-power and high-density associative memory ASIC.

The scientific impact on HEP experiments will be extremely important. The recent discovery at the LHC of a new boson having a mass of  $(126.0 \pm 0.6)$  GeV opened a new era in particle physics. Several studies are underway to assess its properties and, to date, all measurements are consistent with the hypothesis of the Standard Model (SM) Higgs boson. Based on this working hypothesis, the discovery opened the era of precision measurements of its couplings. The observation of the tiniest deviation from SM predictions could bring evidence of new physics. The IMPART project is pursuing the development of advanced techniques of trigger, in order to improve the present precision levels (from 20% to 50%) on the Higgs couplings, hopefully by one order of magnitude, with the goal of probing possible effects of new physics. According to the current plans, LHC is going to operate at 13-14 TeV during the run starting in 2015. With the forthcoming LHC upgrade, the instantaneous luminosity will also increase with the goal to deliver nearly 100 times more statistics with respect to what has been achieved so far. The machine improvement alone would not be sufficient to reach our target of a few percent precision. To achieve this goal we need to improve significantly the event selection techniques. In particular, the online event selection should yield higher efficiency and better signal-to-noise ratio, in order to fully exploit the increased statistics offered by the accelerator upgrade. The IMPART project will play a key role in efficient selection of the very rare Higgs events hidden in an extremely large level of background, thus substantially improving the Higgs sample statistics.

Nevertheless, the IMPART project will have a high impact also on the society and it will be an important step ahead in technology, as low-cost, low-power and extremely parallelized pattern matching mimics the low level brain functions for vision. Understanding how the brain processes information or how it communicates with the peripheral nervous system (PNS) could provide new potential applications, new computational systems that emulate human skills (e.g., by using the directed fusion of information from different sensors) 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 IMPART project is a step in this direction, as it targets applications of low-cost, low-power and extremely parallelized pattern matching in many fields.

An important impact is expected in the field of the smart cameras, for smart cities and smart transports. The IMPART project could speed up state-of-the-art systems. We plan to use the IMPART system in collaboration with the EMC s.r.l., a small company interested to produce Smart Integrated Systems, including the development of smart cameras. Currently, smart cameras for smart cities and transportation are one of the key research topics proposed by the EU. National governments are also moving to install smart vision systems. The Italian government promotes the development of researches and spin-offs in this field<sup>30</sup>. Research interest in this field is very high and there is space for innovation and groundbreaking systems<sup>31</sup>. Innovative smart systems could ameliorate the life quality in congested/overcrowded cities, as well as safety in construction sites, stations, and airports. Smart cameras could be installed in remote environments, such as forests or mountains, and could be supplied by batteries or solar panels and used to monitor environmental phenomena such as avalanches, landslips, summer fires, and so on. Transports could benefit from the project, especially from the safety point of view. As an example, in road tunnels, the driver's reaction time is essential to avoid serious accidents. Smart cameras can monitor the speed, the distance between vehicles, and the temperature of the air. We already have a working example under Monte Bianco; however, a low cost, low power system could be spread around everywhere. Smart cameras could be also used to monitor medical personnel and patients, e.g., by performing nurse hand recognition to reduce the possibility of biological contamination<sup>32</sup> or to monitor body movements of a patient under adaptive radiotherapy<sup>33</sup>. It is worth pointing out that several applications could improve from the IMPART project outcome.

---

<sup>30</sup> [http://www.camera.it/leg17/537?shadow\\_mostra=23910](http://www.camera.it/leg17/537?shadow_mostra=23910)

<sup>31</sup> A. Zanella, et al. Internet of Things for smart cities. *IEEE Internet of Things Journal* (2014).

<sup>32</sup> <http://smartcity.csr.unibo.it/wp-content/uploads/2012/04/Hand-free.pdf>

<sup>33</sup> Simona Giordanengo's CSN5 funded project in 2013

Another important application to be investigate, in collaboration with LPNHE, is DNA sequencing. The DNA alignment machine architecture will benefit from the IMPART system, because each step in alignment and mapping processes will be performed by the most suitable hardware. High level data access, final tuning and interaction with the user will be performed by a standard PC, while coarse pattern matching of a large input database will be performed by the AM chip, and fast weight computation for optimal alignment will run on the Xilinx Kintex-7 FPGA.

The IMPART system could be also installed as a part of FrontEnd or BackEnd system inside detectors used in different applications. A study on the radiation hardness will be performed to quantify the level of radiation tolerance of the system. I will exploit the knowledge that I have acquired on radiation effects to perform this task. The Kintex-7 FPGA is resistant to radiation, as demonstrated by radiation tests made by INFN - Milano. Preliminary results confirm a good level of radiation hardness of Kintex-7 FPGAs against cumulative effects<sup>34</sup>.

The IMPART project will increase group synergies between HEP experiments (INFN group I) and medical and environment disciplines (INFN group V). It is worth noting that IMPART is an electronics/informatics project which starts from knowledge acquired in HEP field, and aims at exploring new disciplines which are in the field of interest of INFN group V. In addition, several aforementioned applications are also mentioned in the Horizon 2020 program. The IMPART R&D plan will give us a competitive advantage in Horizon 2020 calls since it will allow us to deal with several case studies in line with the Horizon 2020 goals. In the future, our idea is to apply for ICT calls, ITN (Ph.D. calls within the Marie Skłodowska-Curie program), and COFUND calls involving IMEC and Cadence for high quality trainings, SMEs (e.g., EMC s.r.l. and Microtest s.r.l., and the Sant'Anna School of Advanced Studies, which is an excellent research centre for most advanced technological topics, including robotics, real time processing and Artificial Intelligence (AI). Results of the IMPART project could be used by SMEs such as EMC s.r.l., CAEN, Microtest s.r.l., and Kaiser s.r.l., to improve their products in the future.

## Bibliography

The bibliography has been included in footnotes along the proposal text.

---

<sup>34</sup> MJ Wirthlin, H Takai, and A Harding. Soft error rate estimations of the Kintex-7 FPGA within the ATLAS Liquid Argon (LAr) Calorimeter. *Journal of Instrumentation* 9.01 (2014): C01025.