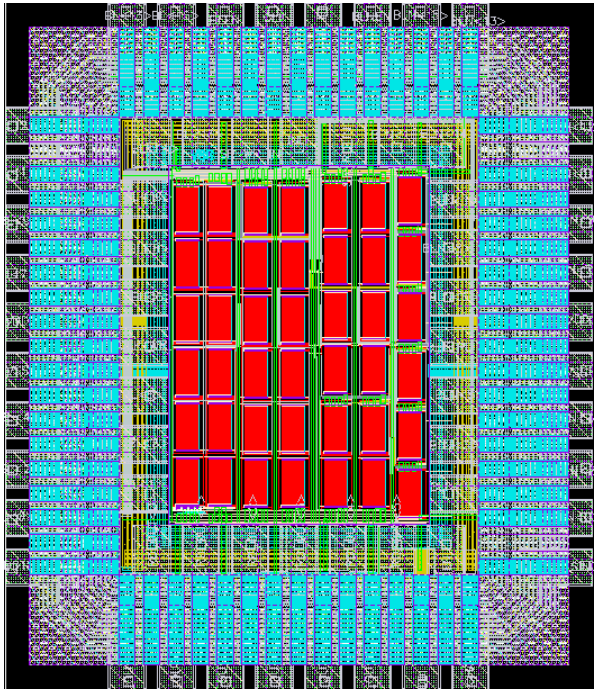
AM07A AND AM07B UPDATE

Alberto Stabile
INFN - Milano

IMPART Innovative Multi-chip system for multi-purpose PAttern Recognition Task

AM07A DEMONSTRATOR



A new little **AM07a already** designed in June 2015

- **Silicon area:** 0.6 mm^2
- **Memory depth:** ~50 kbit (336 patterns)
- **Main aim:** characterize performance technology at 28 nm

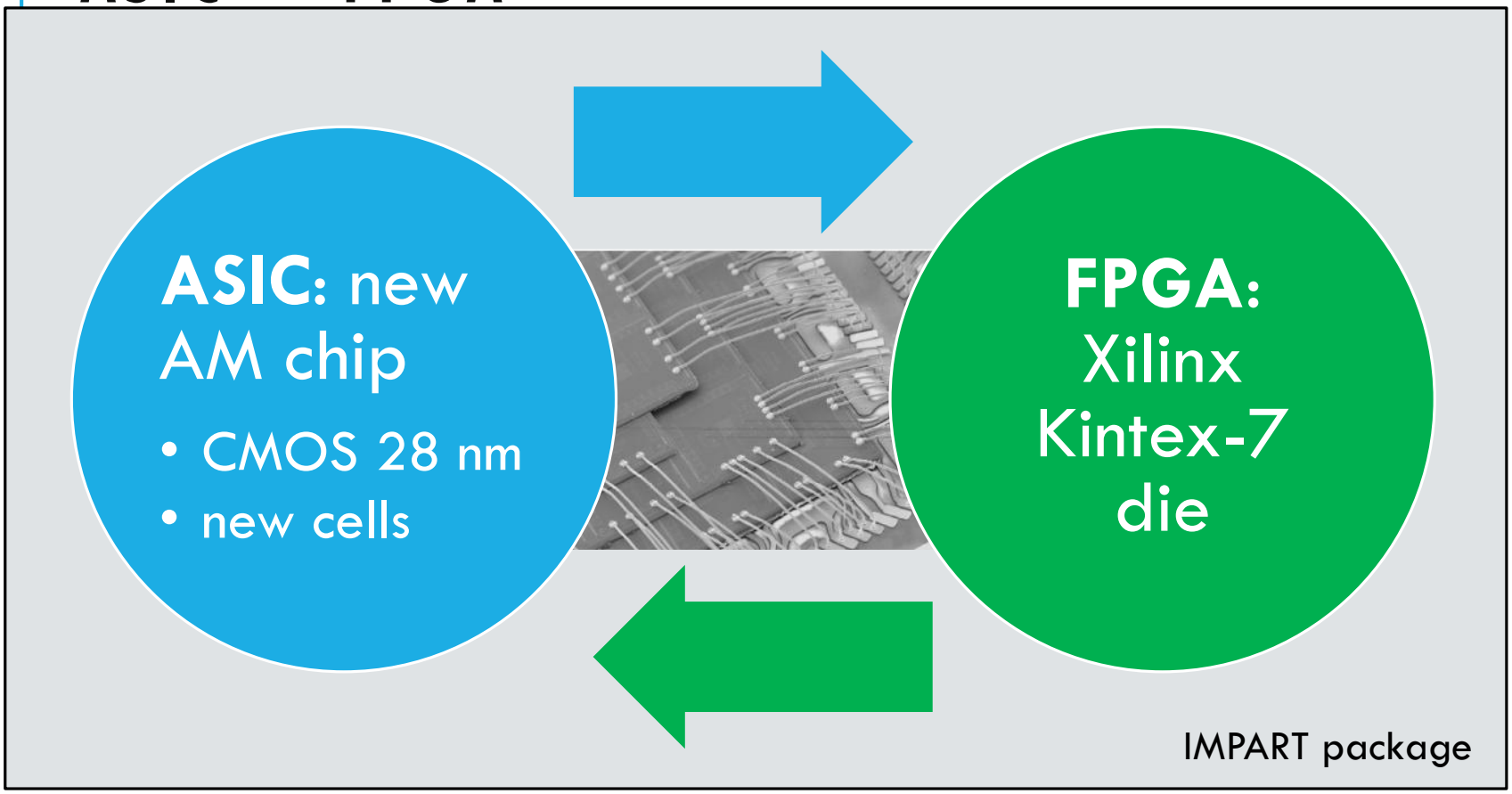
Next chip: AM07b will be submitted by 20 Apr 2016

- **Memory depth:** ~3.4 Mbit (~20 kpatterns)
- **Mixed approach:** full-custom + standard cells
- **Improve power save:** New cell tech (DOXORAM – KOXORAM – PIPECAM)

After 2017-2018: new larger chip AM08

(i.e., 0.5 Mpatterns of 8×16 bits)

ASIC + FPGA


ASIC: new AM chip
• CMOS 28 nm
• new cells
FPGA: Xilinx Kintex-7 die
IMPART package

FPGA TASKS

Transfer/receive external data through serial links at high frequency

Implement the **variable resolution**

Implement ***ad-hoc* algorithms**

- Computer vision **shape recognition**
- **DNA alignement** applications
- Refine HEP tracks

Complex algorithms

- Simple machine (microblaze CPU) inside the FPGA
- Linux module drivers

More demanding applications (low power and high speed)

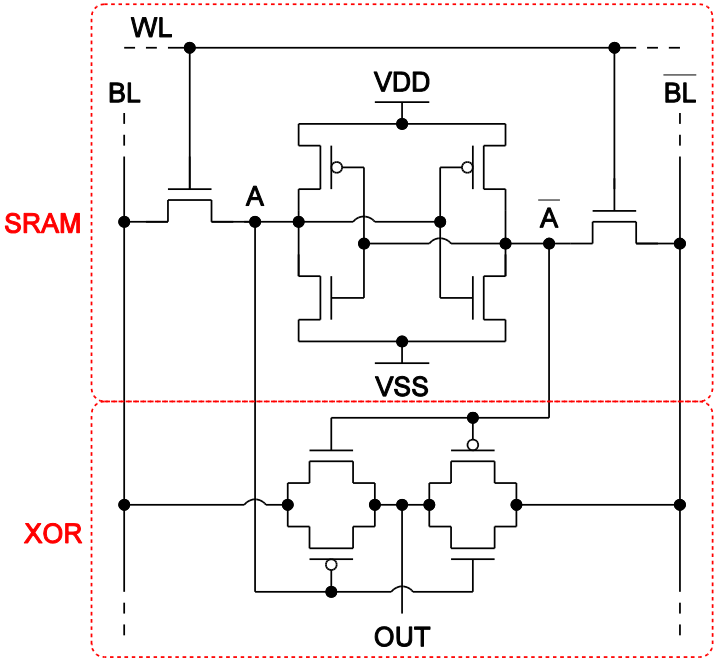
- Dedicated FPGA firmware

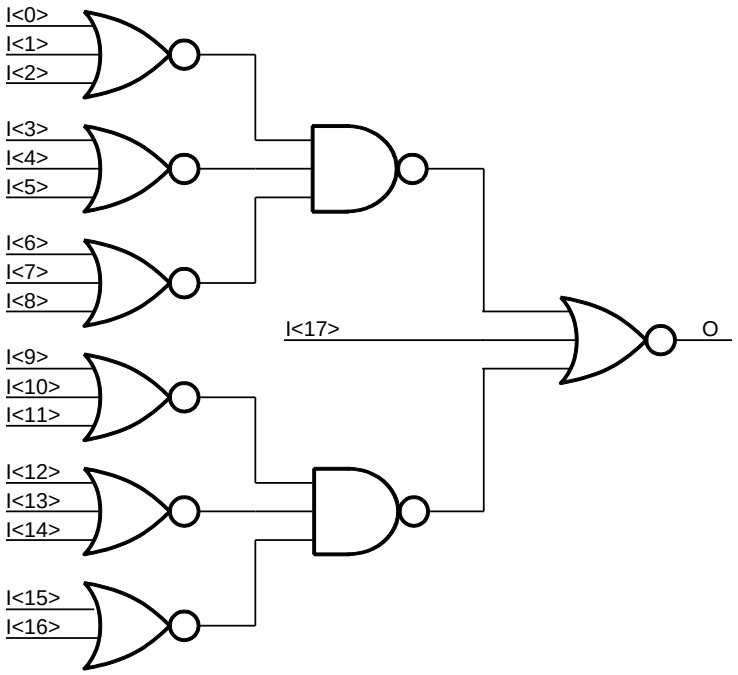
THE XORAM CELL

Based on the previous XORAM cell

Based on the XOR boolean function, instead of the NAND and NOR functions

Is made of a 6T SRAM cell connected to a 6T-XOR gate



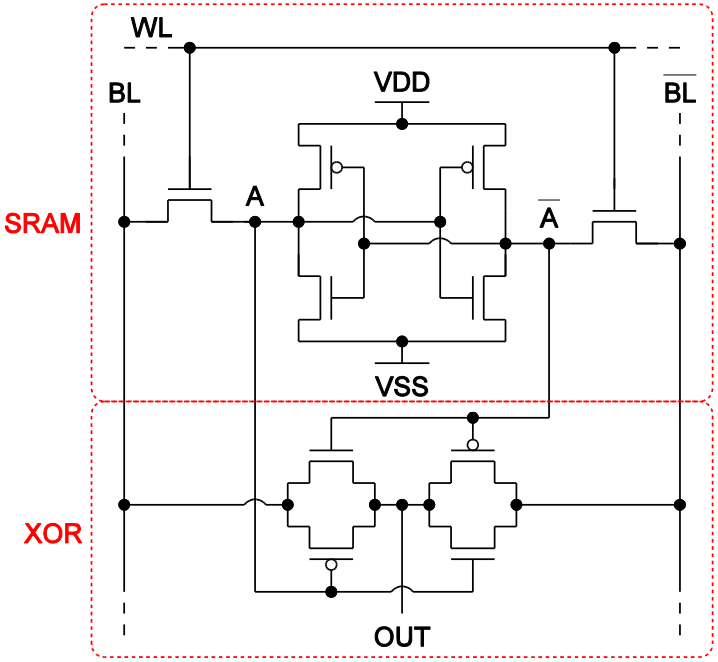


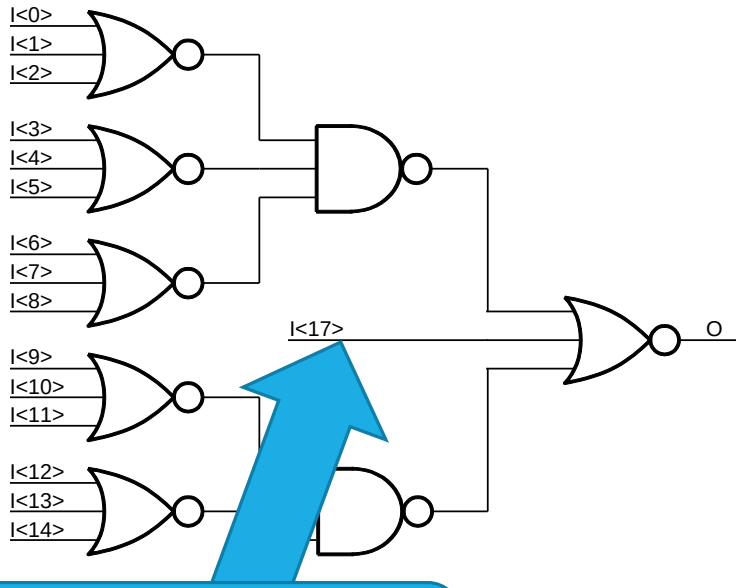
THE XORAM CELL

Based on the previous XORAM cell

Based on the XOR boolean function, instead of the NAND and NOR functions

Is made of a 6T SRAM cell connected to a 6T-XOR gate.





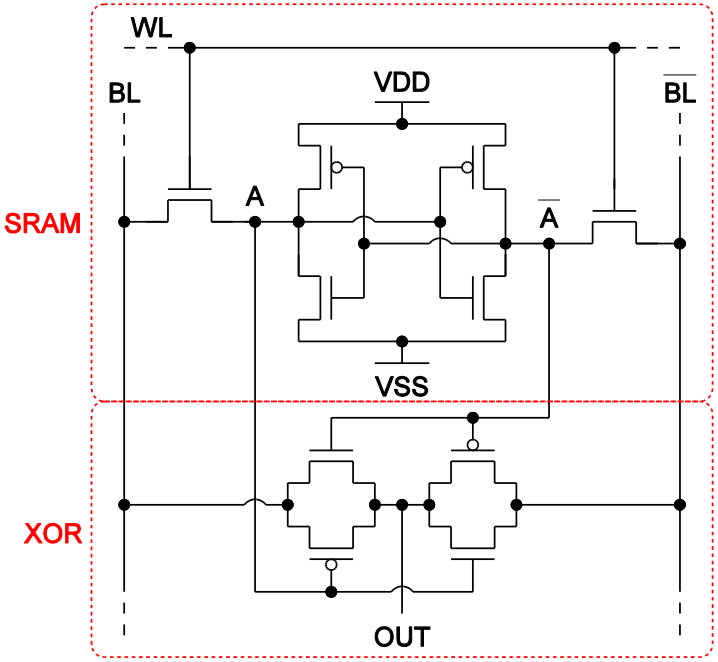
the 18th bitline is used as enable signal for the comparison operation

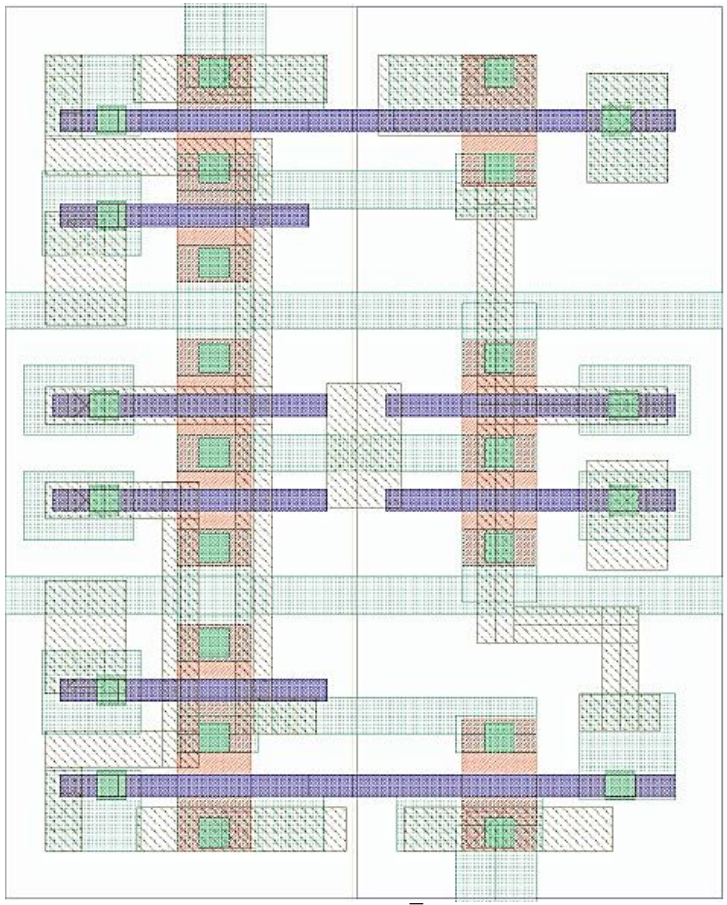
THE XORAM CELL

Based on the previous XORAM cell

Based on the XOR boolean function, instead of the NAND and NOR functions

Is made of a 6T SRAM cell connected to a 6T-XOR gate.





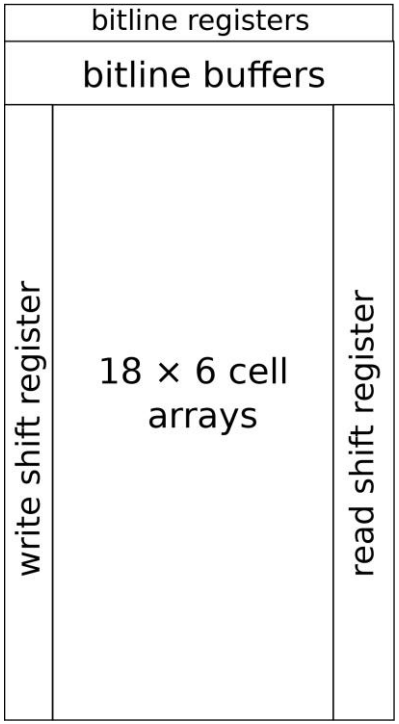
AM07A BLOCK ARCHITECTURE

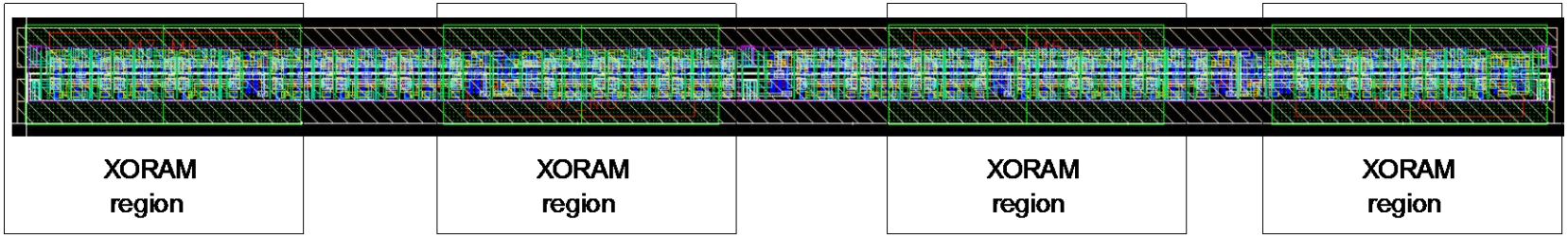
The 18-bit word line blocks have been piled up to compose and array of **18 × 64 bits**

Buffers placed at the top of the array

The **shift registers** is used to write and read the memory bank just for prototype characterization

Single cell area is 1.17 µm × 0.98 µm, and the block area is 6925 µm^2 : **factor 3** w.r.t. old cell designed at 65 nm

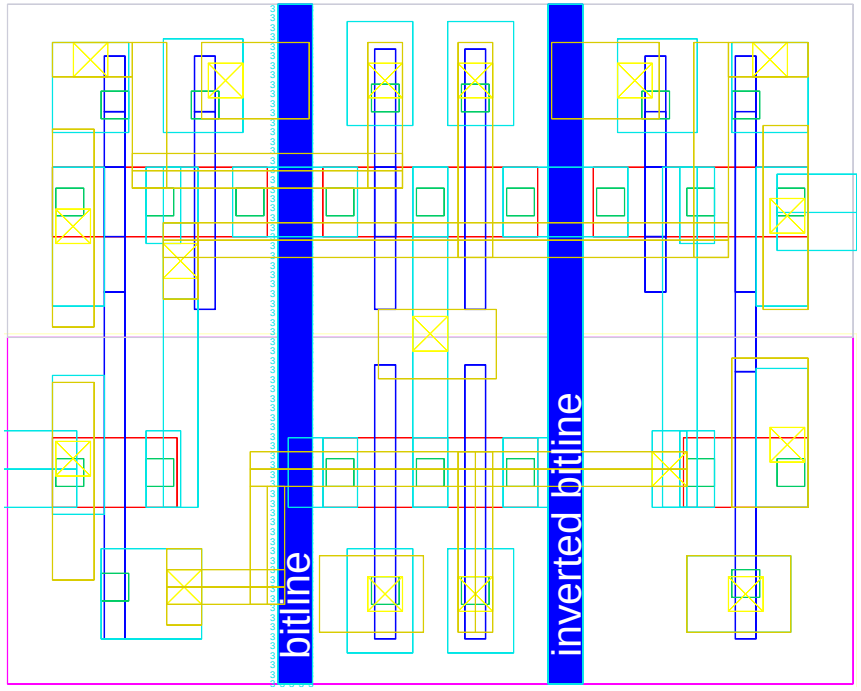




POWER SAVE METHODS

Minimization of parasitics

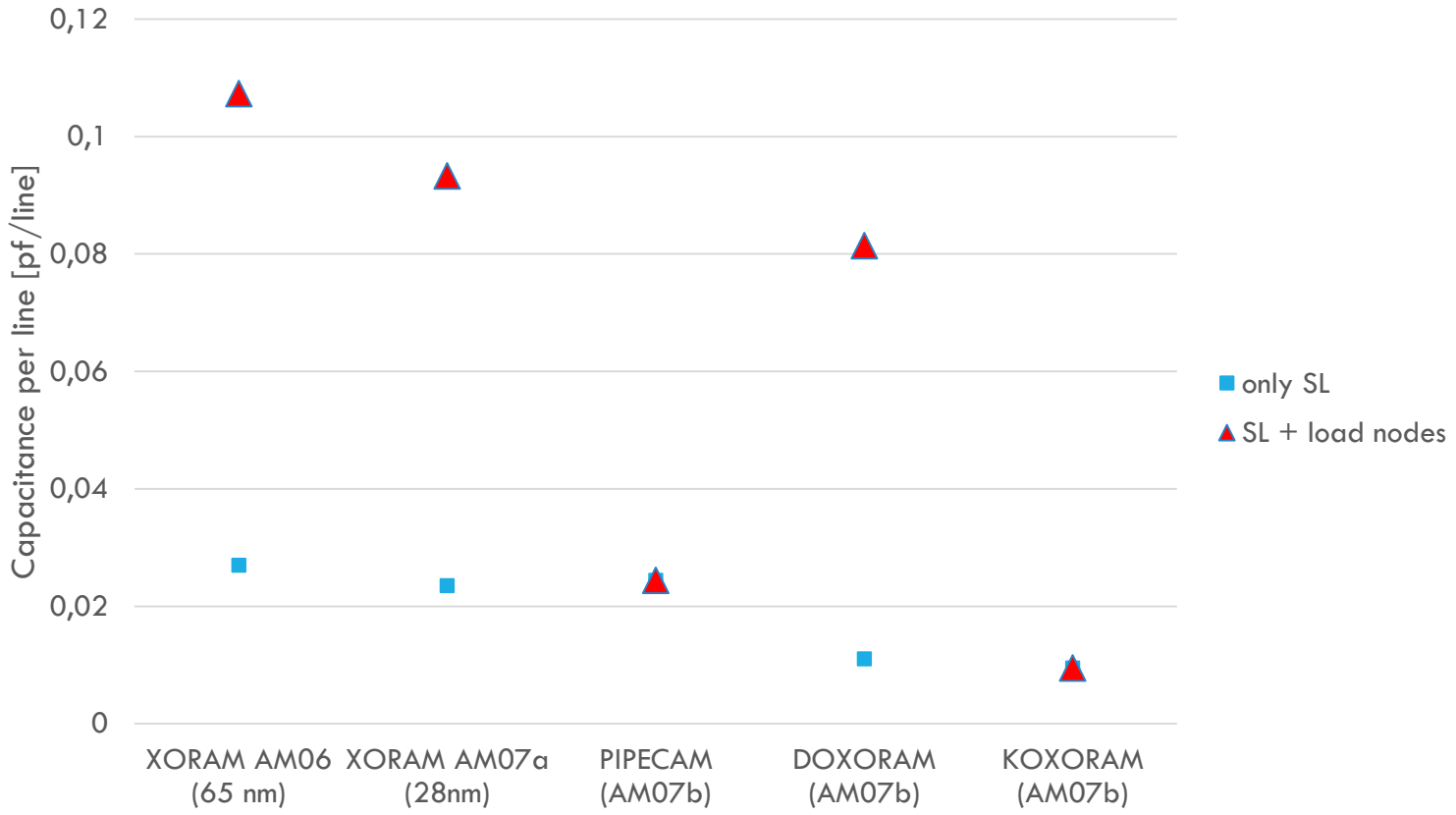
- 70% of dynamic power consumption is due to **propagation of bitline data**
- Block layout has been designed aiming at **minimizing bitline RC parasitic**
- Bitline wires have been designed in the **first metal layer (M1)**
- Bitline capacitance: 23 fF at CMOS 28 nm w.r.t 27 fF at CMOS 65 nm



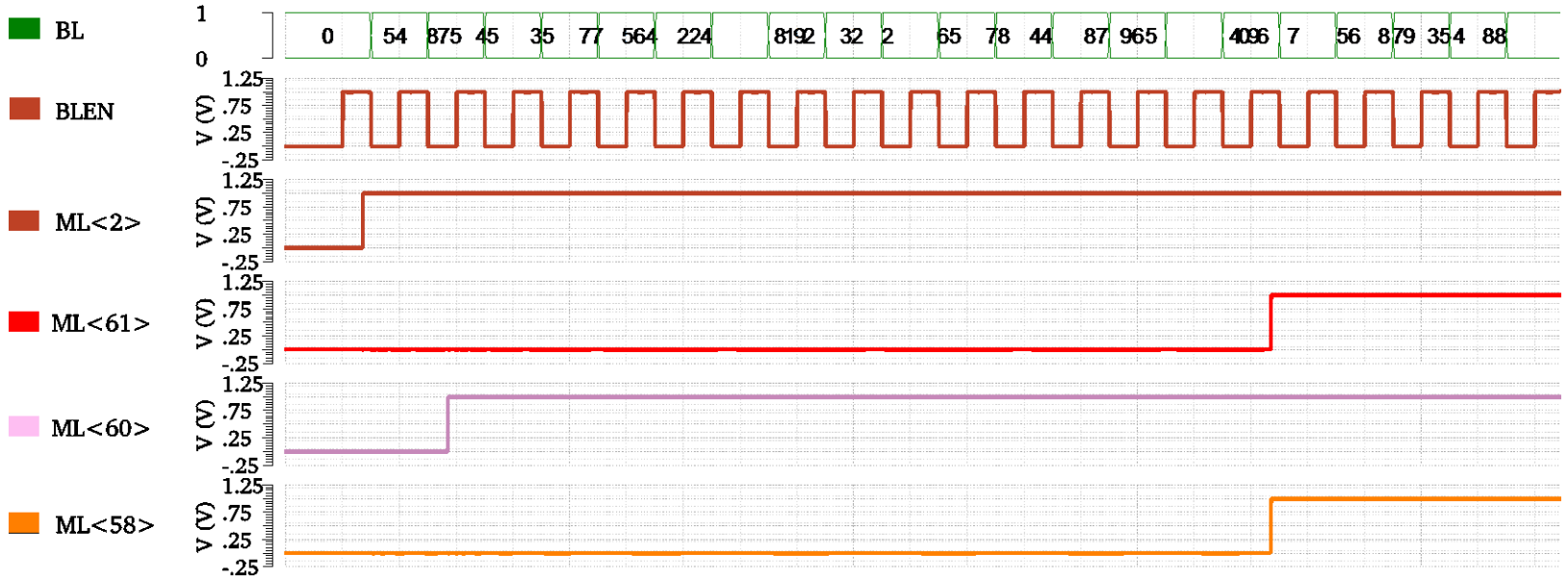
Clockless circuit

- enable is given on the 18th bitline

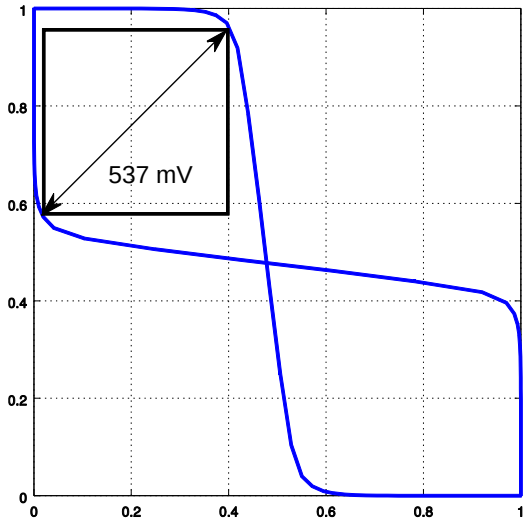
CAPACITANCE COMPARITION ON THE BLS



AM07A SIMULATION RESULTS



Case	Temperature	Supply voltage	MOS models	Active power	Leakage power
typical	27 degrees	1.0 V	TT	22.9 µW	144 nW
Worst power	0 degrees	1.1 V	FF	23.8 µW	731nW
Worst speed	0 degrees	0.9 V	SS	22.1µW	22 nW
Worst zero	0 degrees	0.9 V	FS	17.1µW	119 nW
Worst one	0 degrees	0.9 V	SF	26.7 µW	93 nW



THE CHIP PROTOTYPE

A new small **AM07a already** designed in June 2015

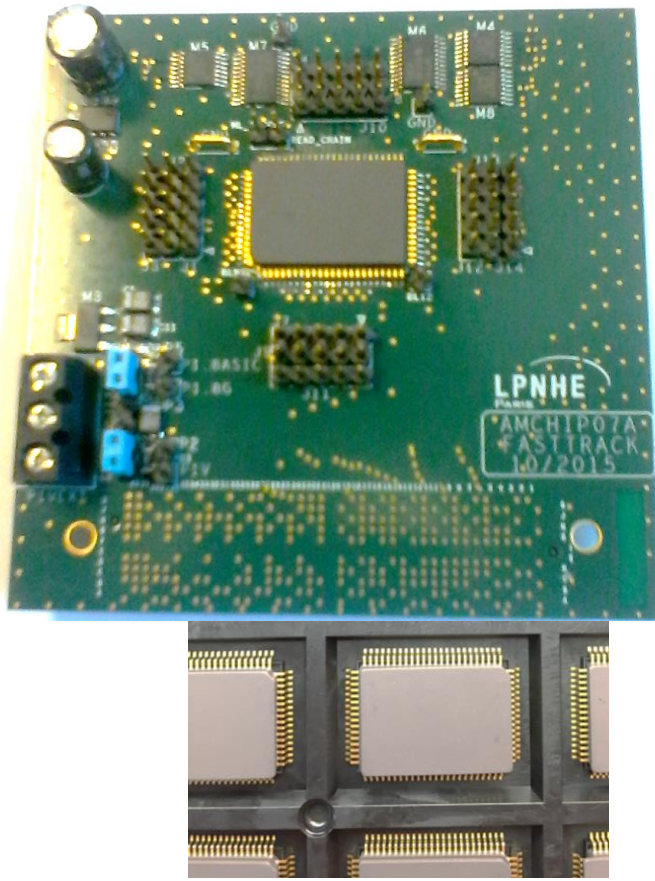
- **Silicon area:** 0.6 mm^2
- **Memory depth:** ~50 kbit (336 patterns)
- **Target:** characterize performance technology at 28 nm

Four different clock signals have been routed across the whole chip

- clock for the bitline registers
- clock for the BLEN register (18th bitline)
- read clock
- write clock

Each clock signal has been routed with the H-tree method

- The test chip available since December 2015 for characterisation

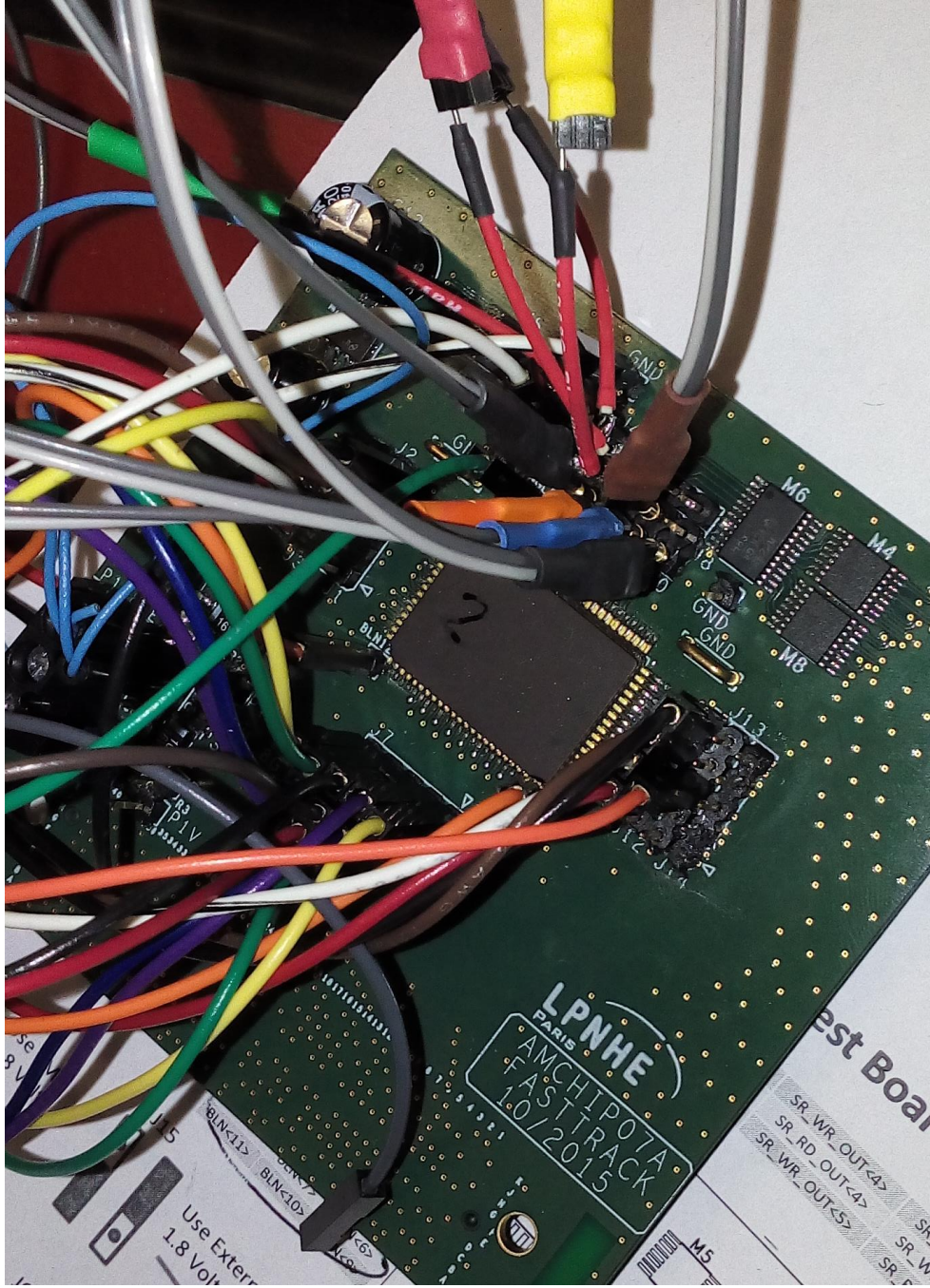


AM07A CHARACTERIZATION

Preliminary characterization tests have been performed in Paris and Milano and demonstrated that the **shift register for writing and reading are fully functional.**

Now, the tests are moved to INFN Milano for an exhaustive characterization.

Characterization will be performed by means of a pattern generator and a logic analyzer.



AM07A CHARACTERIZATION

Shift register for writing and reading are fully functional up to 97 MHz

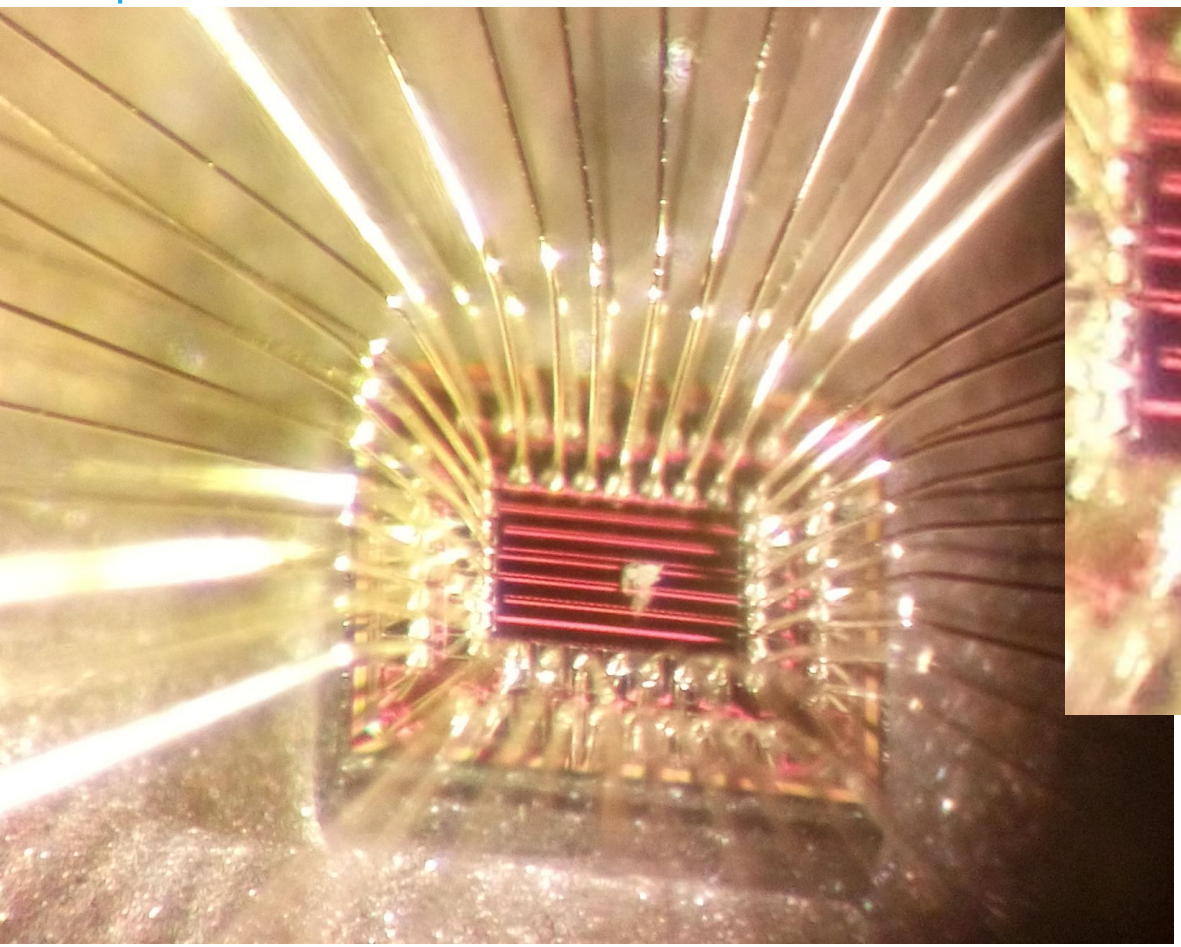
The MUX to select data from the output bus of CAM or from the shift register is fully functional

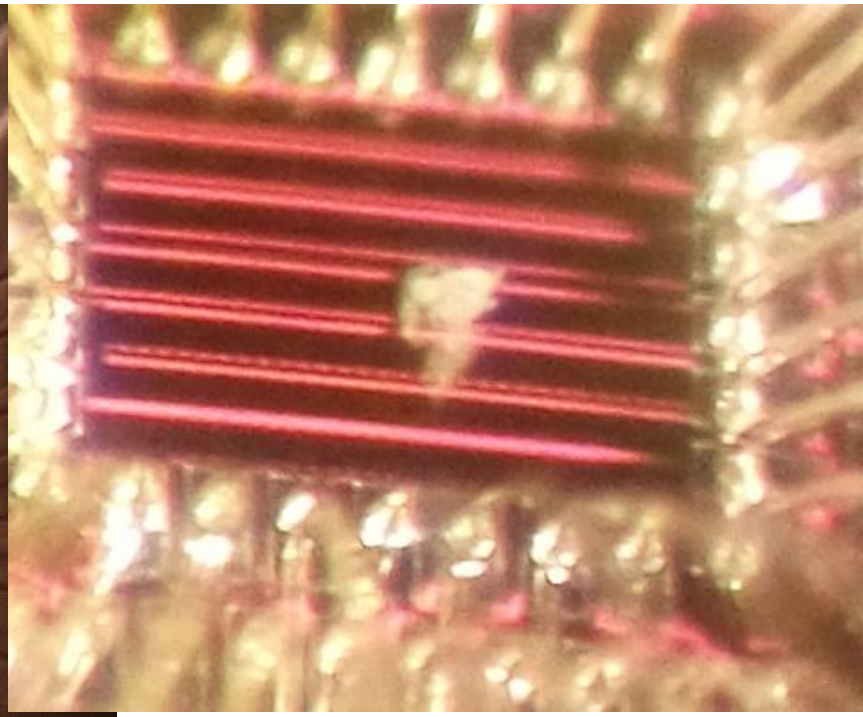
All bits of output bus (load of Set Reset FF) are still 0 → **At the microscope: There is a large spot of dust with irregular perimeter in the middle of the die!!!** 😐

- We don't have a ZIF socket
- We try to remove the dust BUT still does not work ☹
- Now I will try a new chip + new board

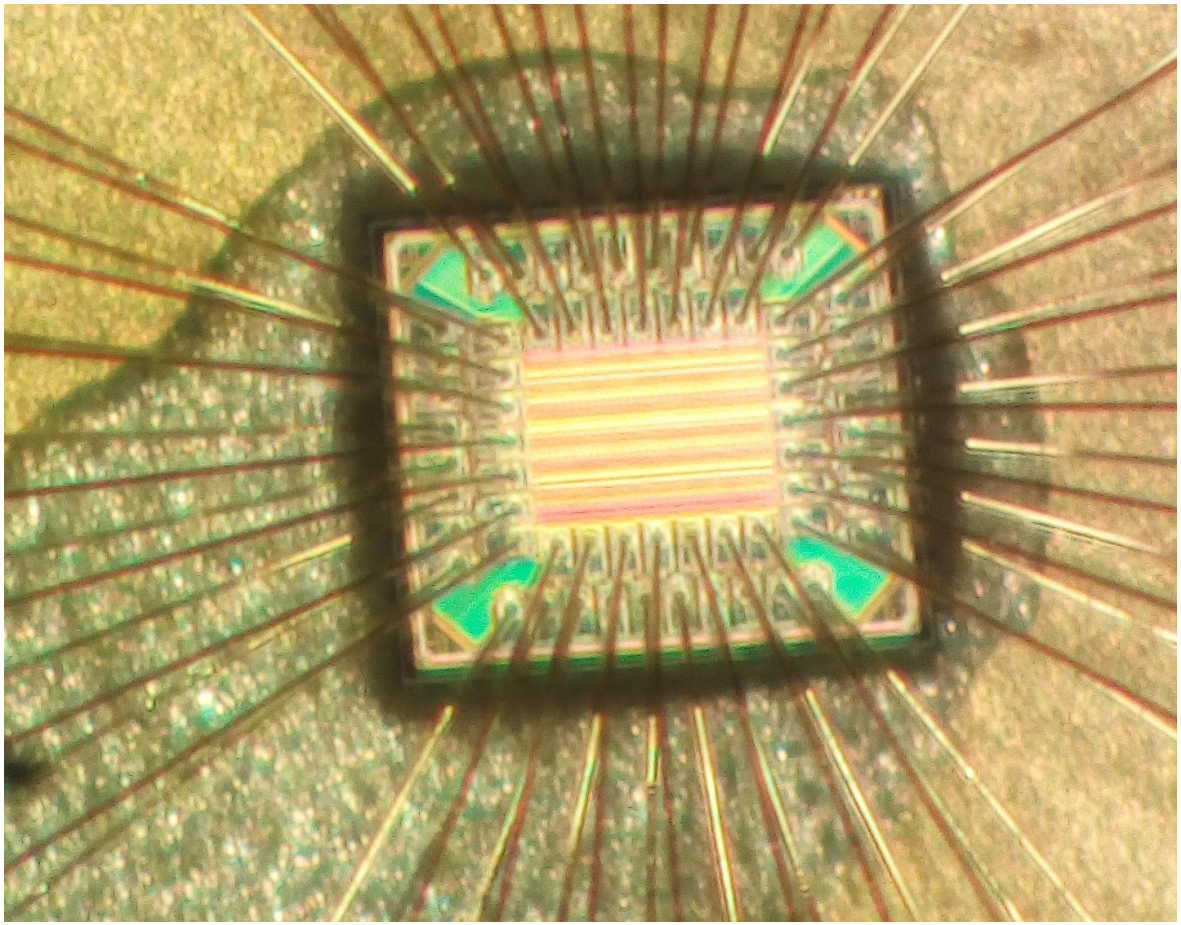
The "pull-medium" resistances toward VTT = 0.56 V create paths with a lot of inter-noise → We have removed and now is ok!

AM07A CHARACTERIZATION





AM07A CHARACTERIZATION

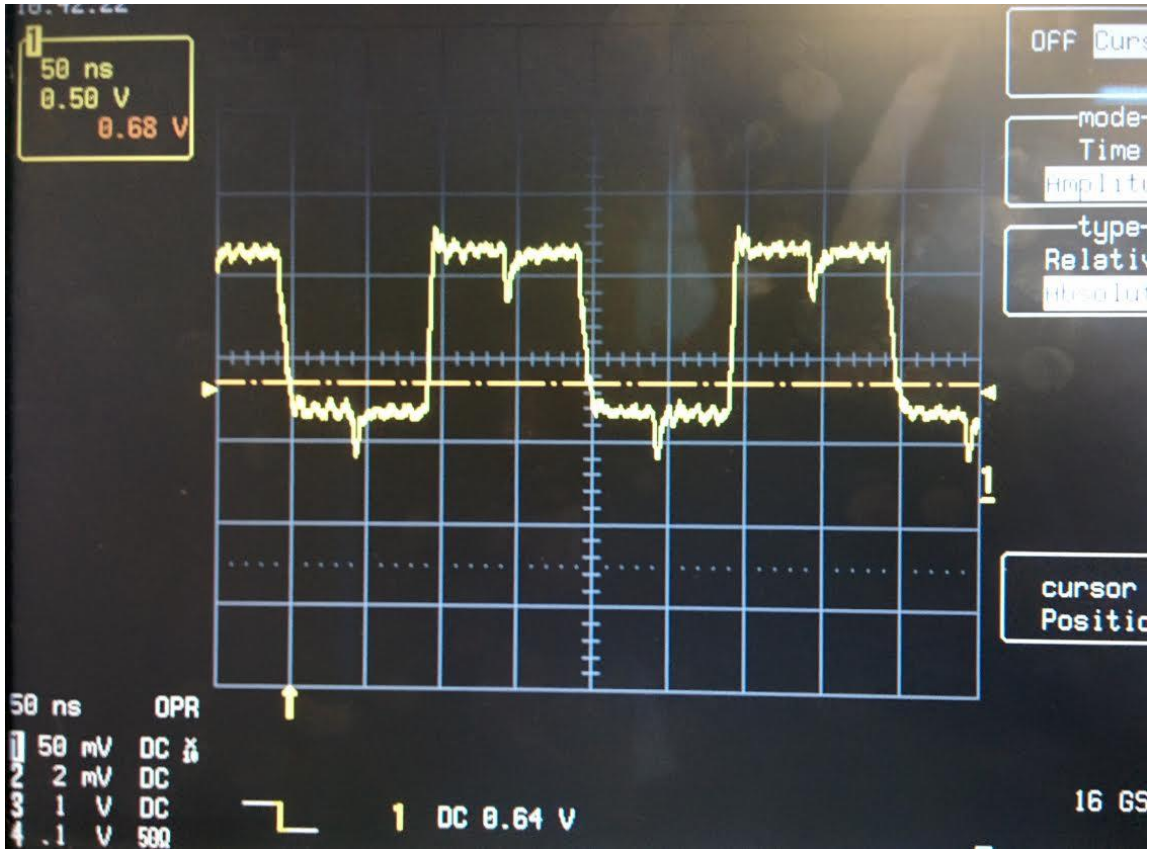


AM07A CHARACTERIZATION PLAN

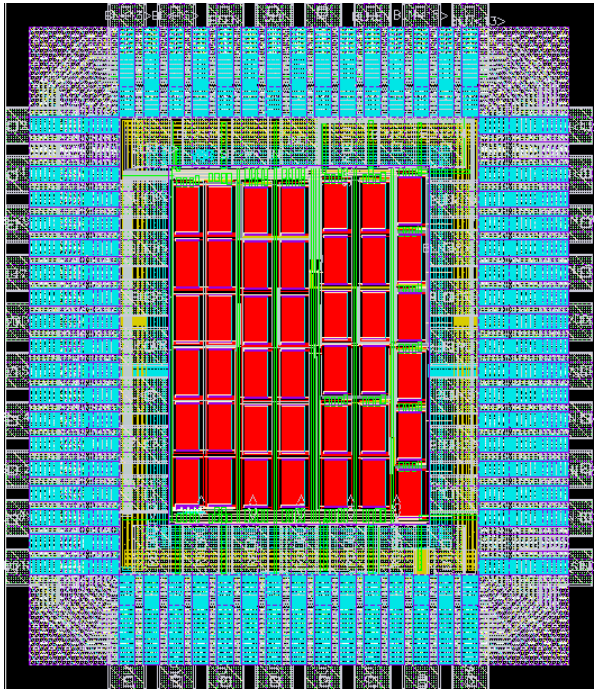
We are looking for a group to characterize the AM07a:

- Design a new board with a ZIF.
- Better understand how to transfer signal to/from Logic Analyser at 200-400 MHz
- Meas the current in multi modes (write, compare) multi corner (low and high temperature – low and high Vdd – different frequencies)
- Have sufficient data a month before next chip submission (20 March 2016)
  - In this way we can tailor the new chip.

PULL-MEDIUM ISSUE



AM07A DEMONSTRATOR



A new little **AM07a already** designe[illegible]5

DONE

- **Silicon area:** 0.6 mm^2
- **Memory depth:** ~50 kbit (336 patterns)
- **Main aim:** characterize performance technology at 28 nm

Next chip: AM07b will be submitted in Apr 2016

- **Memory depth:** 20 kpatterns
- **Mixed approach:** full-custom + standard cells
- **Improve power save:** New cell tech

We are afraid on the AM07a tests. It is important to understand the technology before the next submission. It is necessary to reach results as soon as possible to met the submission of April 2016!

OTHER NEW CELLS

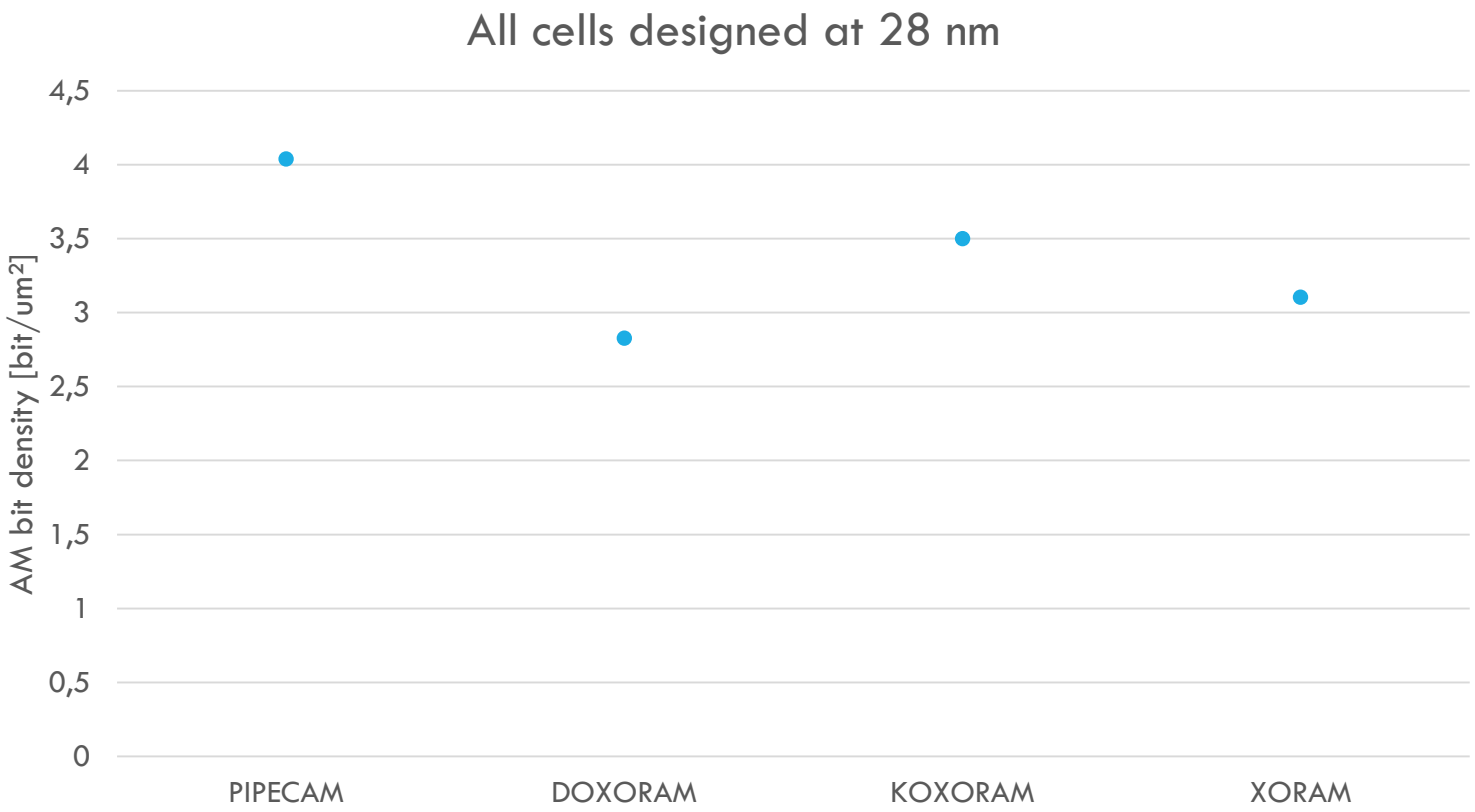
With similar power save methods two other cells are under investigation:

- DOXORAM
  - Design in Milano Celoria
- KOXORAM
  - Design in Milano Celoria

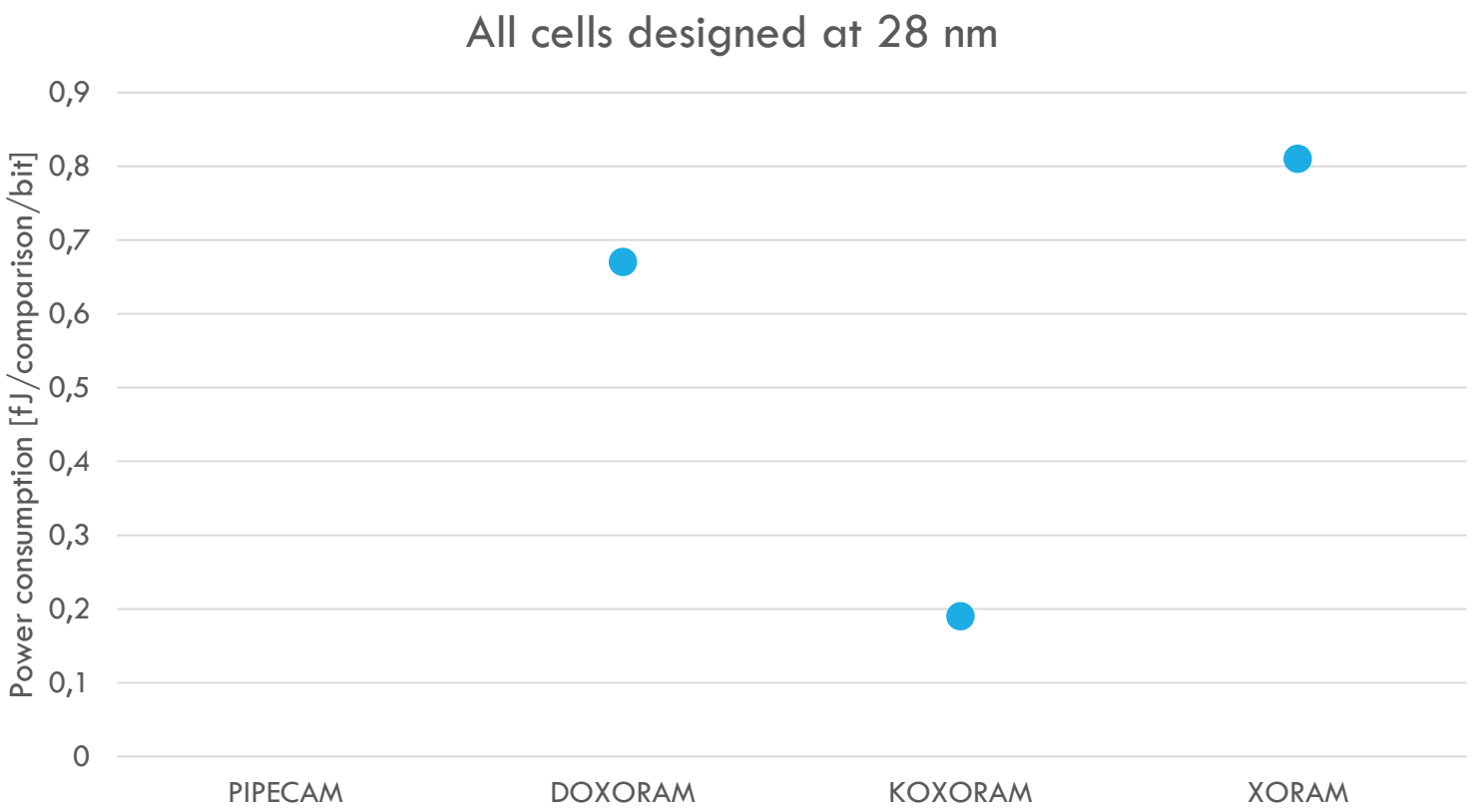
Based on the pipeline NOR cells methods

- PIPECAM
  - Design in Milano Bicocca with the help of Milano Celoria

AREA COMPARITION



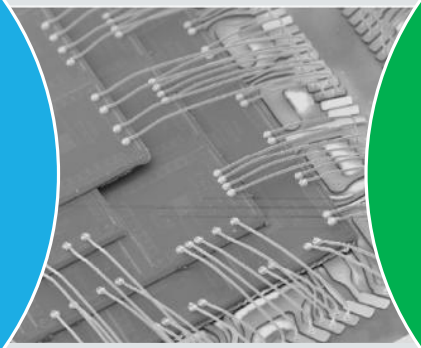
POWER CONSUMPTION COMPARITION



ASIC + FPGA

ASIC: new AM chip

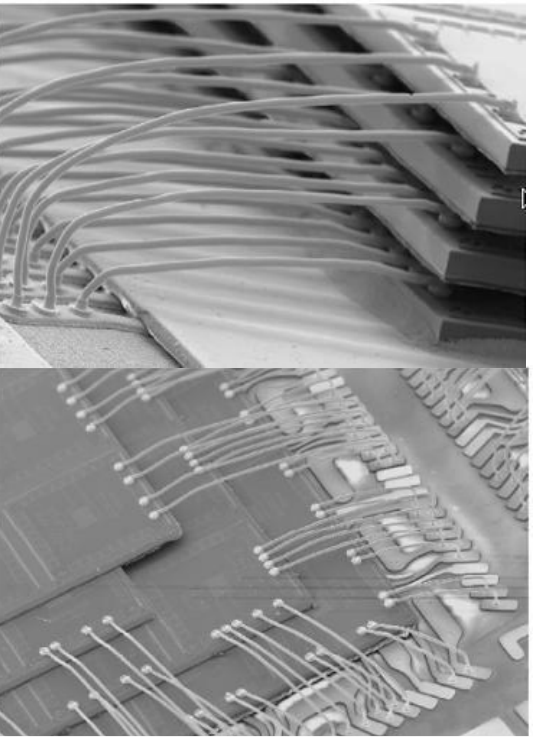
- CMOS 28 nm
- new cells

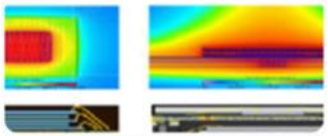


FPGA: Xilinx Kintex-7 die

IMPART package

MULTI PACKAGE AND POWER CONSIDERATIONS





3D assembly technology **studies:**

- **Choose technology** which optimize power consumption
- Electrical and thermal **3D simulations** needed
- **Design and test** of the package
- **We need a LVDS and cross-talk simulation on the substrate!!!**
- **Analyse the SSN on the FPGA!**

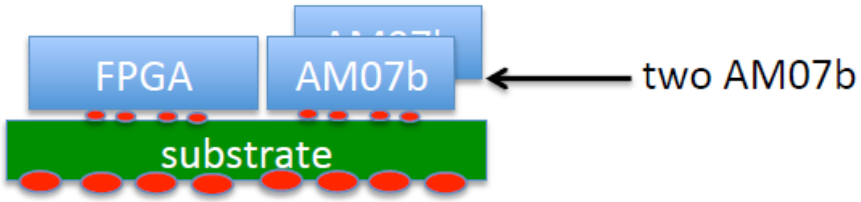
We need help to have result before the chip submission.

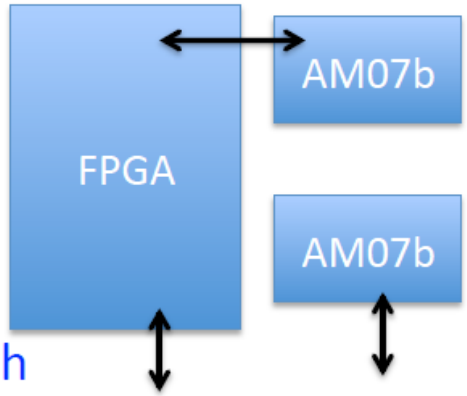
MULTI PACKAGE STUDIES

Explore multipackaging

- 1 or 2 AM chip + 1 FPGA in the same package

Flip Chip





- Dramatically increases AM to Track Fitter bandwidth
  - "Through Silicon Vias not required!"

AM07B FLIP CHIP

Area: 10 mm²

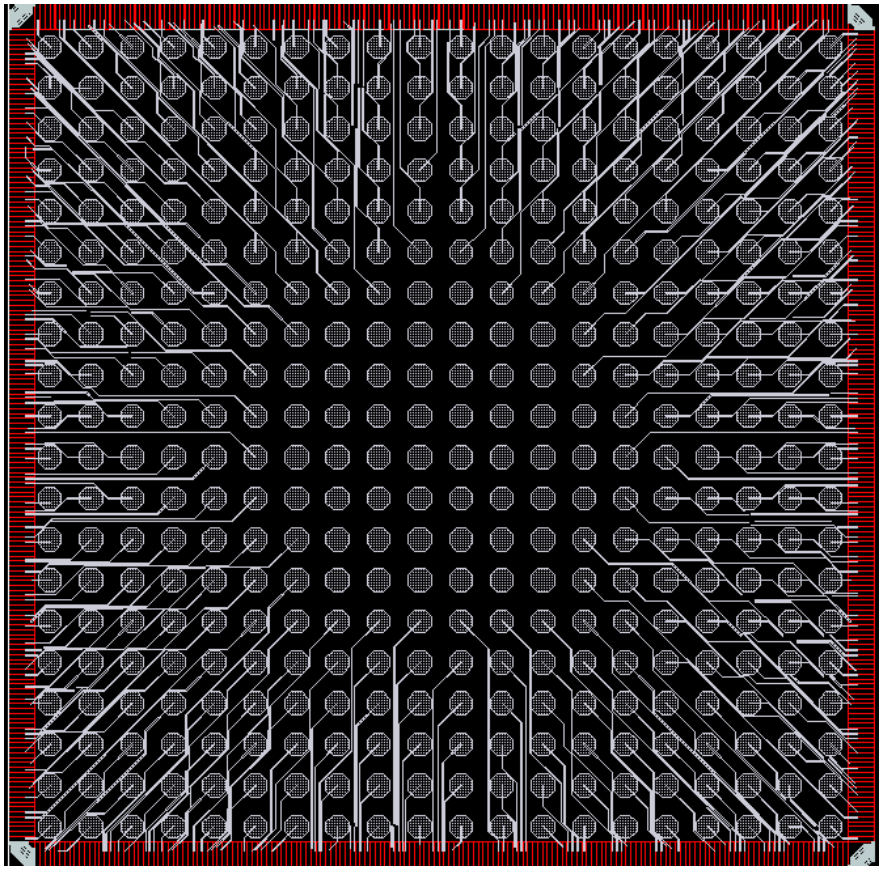
Memory depth: 20 kpatterns

400 bumps

- 341 signals bumps – 59 power bumps

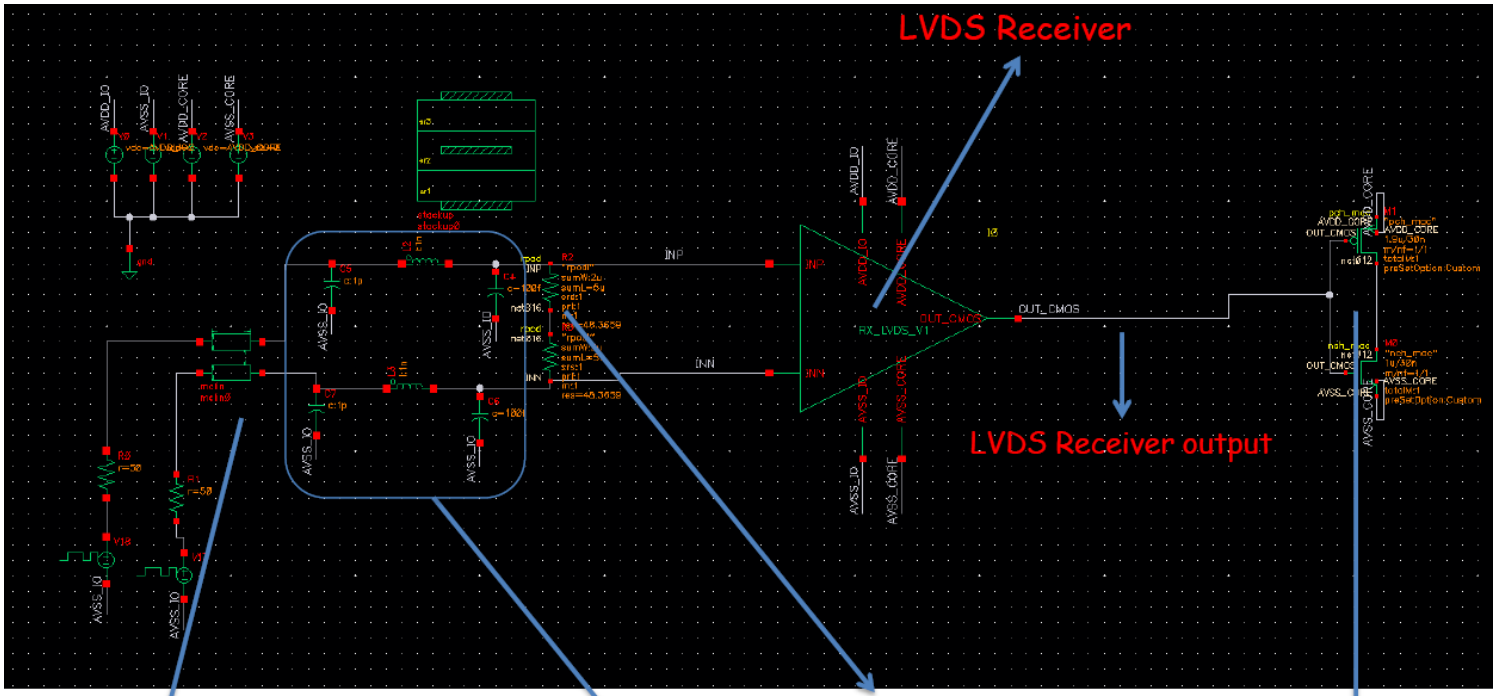
Ext. cost 97 k€ + 10 k€ for the bumps

Max current density at 125 degrees: 4.19 A



LVDS INTERFACE

Update on the activity of BG/PV group - G. Traversi[1], F. De Canio[2]
[1]University of Bergamo - INFN Pavia gianluca.traversi@unibg.it [2]University of Pavia – INFN Pavia francesco.decanio01@ateneopv.it



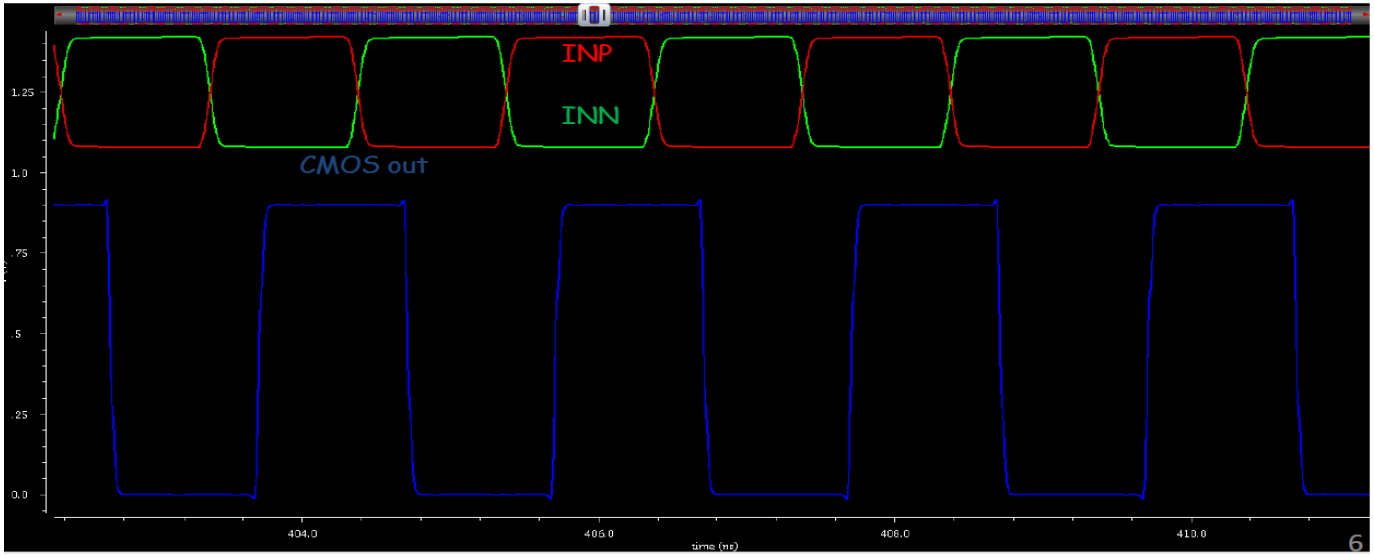
LVDS INTERFACE

Update on the activity of BG/PV group - G. Traversi[1], F. De Canio[2]
[1]University of Bergamo - INFN Pavia gianluca.traversi@unibg.it [2]University of Pavia – INFN Pavia francesco.decanio01@ateneopv.it

LVDS receiver – simulation results - Clock (V_{ID} = ±350 mV)

- VDD_{IO} = 1.8 V, VDD_{CORE} =0.9V
- Temp: 27°C – TT Corner
- Input: Clock @ 500 MHz (t_r = t_f 100ps) V_{CM}=1.25 V, **V_{ID}=±350 mV**
- Output Load = 2fF

Duty Cycle	50.2%
Rise Time	55 ps
Fall Time	51 ps
rms Jitter	1.8 ps
Propagation delay H2L	336 ps
Propagation delay L2H	330 ps
Total Power Dissipation	**0.2 mW**



QUASI DEFINITIVE INTERFACE OF AM07B

```
inout [17:0] \bus_n[0] ;// LVCMOS18
inout [17:0] \bus_n[1] ;// LVCMOS18
inout [17:0] \bus_n[2] ;// LVCMOS18
inout [17:0] \bus_n[3] ;// LVCMOS18
inout [17:0] \bus_n[4] ;// LVCMOS18
inout [17:0] \bus_n[5] ;// LVCMOS18
inout [17:0] \bus_n[6] ;// LVCMOS18
inout [17:0] \bus_n[7] ;// LVCMOS18
inout [17:0] \bus_p[0] ;// LVCMOS18
inout [17:0] \bus_p[1] ;// LVCMOS18
inout [17:0] \bus_p[2] ;// LVCMOS18
inout [17:0] \bus_p[3] ;// LVCMOS18
inout [17:0] \bus_p[4] ;// LVCMOS18
inout [17:0] \bus_p[5] ;// LVCMOS18
inout [17:0] \bus_p[6] ;// LVCMOS18
inout [17:0] \bus_p[7] ;// LVCMOS18
inout        init;      // LVCMOS18
inout        clk;       // LVCMOS18
inout [22:0] addr_out;  // LVCMOS18
inout [14:0] controls;  // LVCMOS18
input [2:0]  RX_P;      // LVDS18
input [2:0]  RX_N;      // LVDS18
output [2:0] RX_OUT;    // LVCMOS10
output       TX_P;      // LVDS18
output       TX_N;      // LVDS18
input [1:0]  biccoca_test //analog
             //tot signal wire 341
             // Power will be decided as soon as possible.
```

AM07B OPEN QUESTIONS

LVDS pads will be added just only for test the Bergamo Receiver and Drivers: 3 receivers and 1 driver!

How many cells? 2, 3 o 4? List of priority…

- AM07a XORAM design by Alberto & Federico (Mi-Celoria + Mi-Biccoca)
  - Silicon proof??
- PIPECAM designed by Marcello & Federico (MI-Bicocca)
- DOXORAM® (Mi-Celoria)
- KOXORAM® (Mi-Celoria)

How to build the characterization board (really complex like an Evalution board of Xilinx)?

- Lyon University (Sebastien + Williams) will responsible for the PCB test design + testing procedure

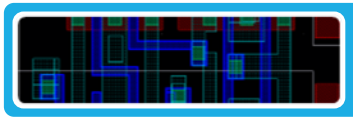
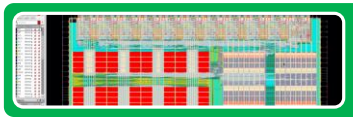
WP1A. AM CHIP DESIGN AND SIMULATION

Alberto Stabile (INFN – Milano)

This work package consists in the **design of the Associative Memory 07a** in a standard **28 nm CMOS** process.

Starting time: Mar. 2015

deadline

Deliverable		**WP1a.1:** AM cell design (Bicocca) • Minimize the silicon area and the power consumption	DONE May 2015
Deliverable		**WP1a.2:** Front-end design (A. Stabile) • Writing a behavioral netlist	DONE May 2015
Deliverable		**WP1a.3:** Floorplan and back-end design (A. S • Automatic place & route	DONE Jun. 2015
Deliverable		**WP1a.4:** Functional simulation (Bicocca) • Validate through digital simulations to confirm the functionality of the la	DONE Jun. 2015
Milestone		**WP1.5:** Chip integration and submission (A. St • Final checks: Design Rule Check (DRC) and Layout Versus Schem	DONE Jun. 2015
Fabrication		**WP1.Fab:** AM chip will be fabricated by TSM	In work Feb. 2016
Milestone		**WP1.6:** Characterization of the AM chip (A. Stabile) • AM chip will be characterized in a single-chip package.	Feb. 2016

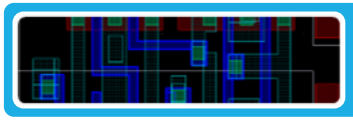
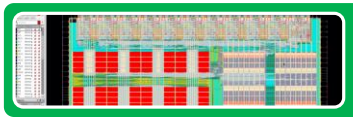
WP1B. AM CHIP DESIGN AND SIMULATION

Alberto Stabile (INFN – Milano)

This work package consists in the **design of the Associative Memory 07b** in a standard **28 nm CMOS** process.

Starting time: Jun. 2016

deadline

Deliverable		**WP1.1:** AM cell design (M. Dematteis + L. Frontini) • Minimize the silicon area and the power consumption	Jan. 2016
Deliverable		**WP1.2:** Front-end design (A. Stabile + T. Kubota) • Writing a behavioral netlist	Feb. 2016
Deliverable		**WP1.3:** Back-end design (A. Stabile + L. Frontini) • Automatic place & route	Mar. 2016
Deliverable		**WP1.4:** Functional simulation (A. Stabile + T. Kubota) • Validate through digital simulations to confirm the functionality of the layout	Mar. 2016
Milestone		**WP1.5:** Chip integration and submission (A. Stabile) • Final checks: Design Rule Check (DRC) and Layout Versus Schematic (LVS)	Apr. 2016
Fabrication		**WP1.Fab:** AM chip will be fabricated by TSMC	May. 2016
Milestone		**WP1.6:** Characterization of the AM(Paris + Milano + Lyon) • AM chip will be characterized in a single-chip package.	Sept. 2016

A CHIP FOR INTERDISCIPLINARY APPLICATIONS

Alberto Stabile
INFN Milano

MOTIVATIONS FOR PHASE II

Big challenges from phase-II conditions

- Goal 7e34 Hz/cm2
- Pileup 140 (max 200)
- Move part of HLT including Traking to L1

Track Triggers are a crucial piece of the phase-II upgrade plan

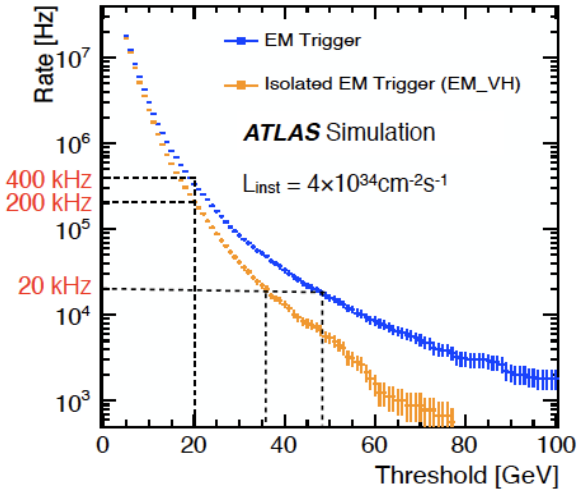
- ATLAS: L1Track and FTK upgrade -> Architecture still evolving
- CMS: L1Track Trigger is the baseline

R&D for a feasible design

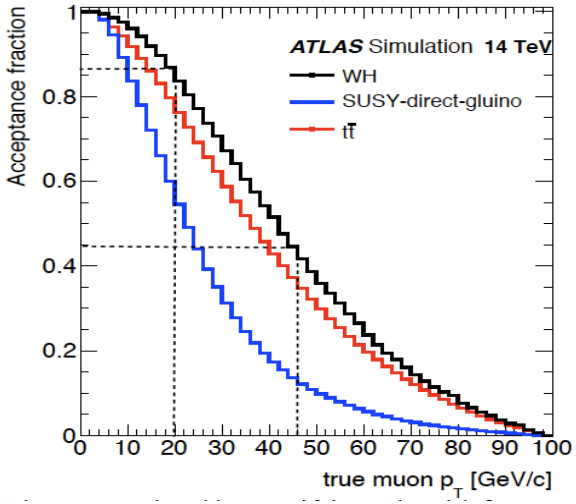
- ATLAS TDAQ IDR (begin 2016)
- ATLAS TDAQ TDR (2017)
- CMS Tracker TDR (2017)

Goal: evolve FTK design to phase-II environment

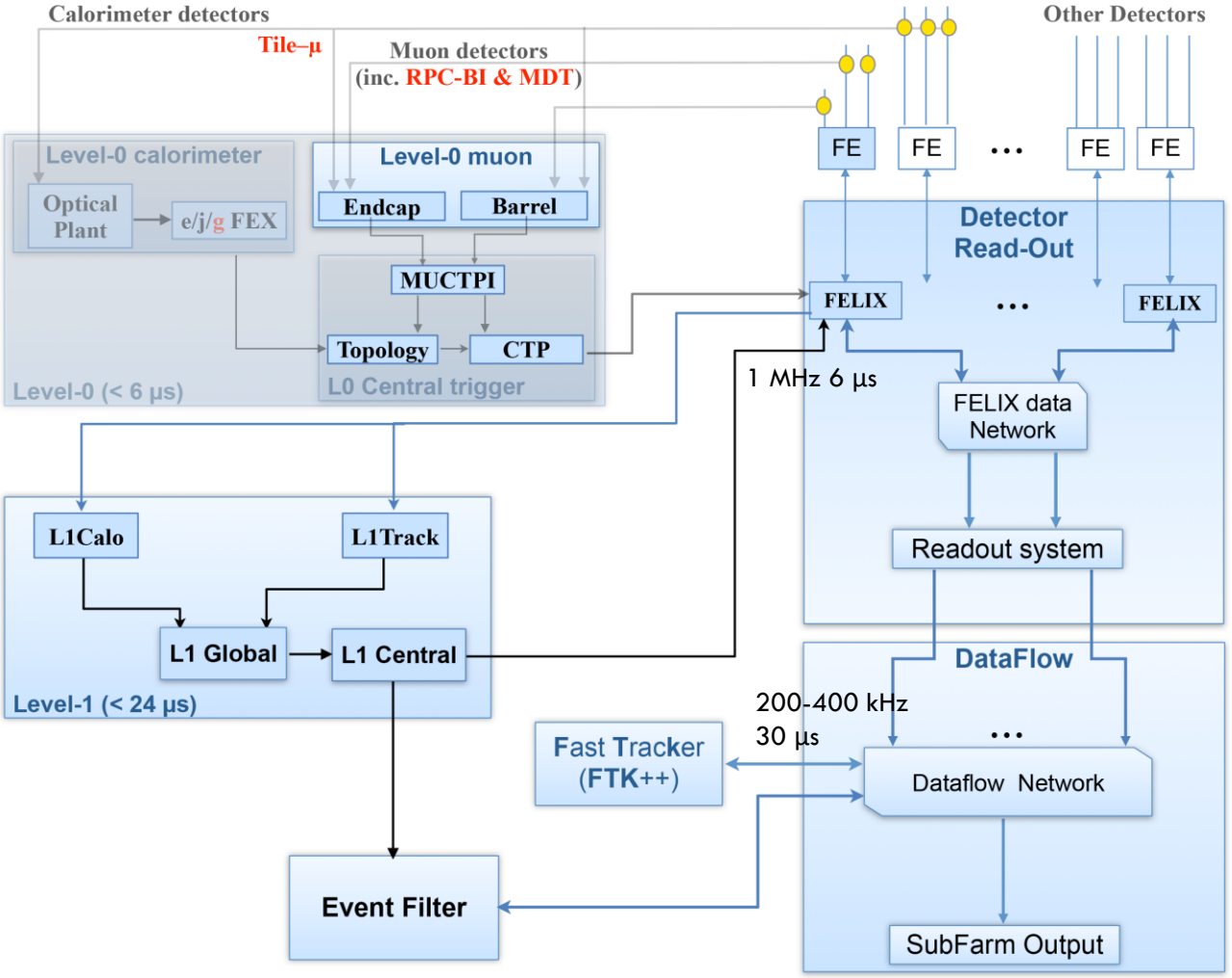
- AM chip R&D
- Fully exploit ATCA potentialities



Need a L1Track to save EM pt>20 GeV

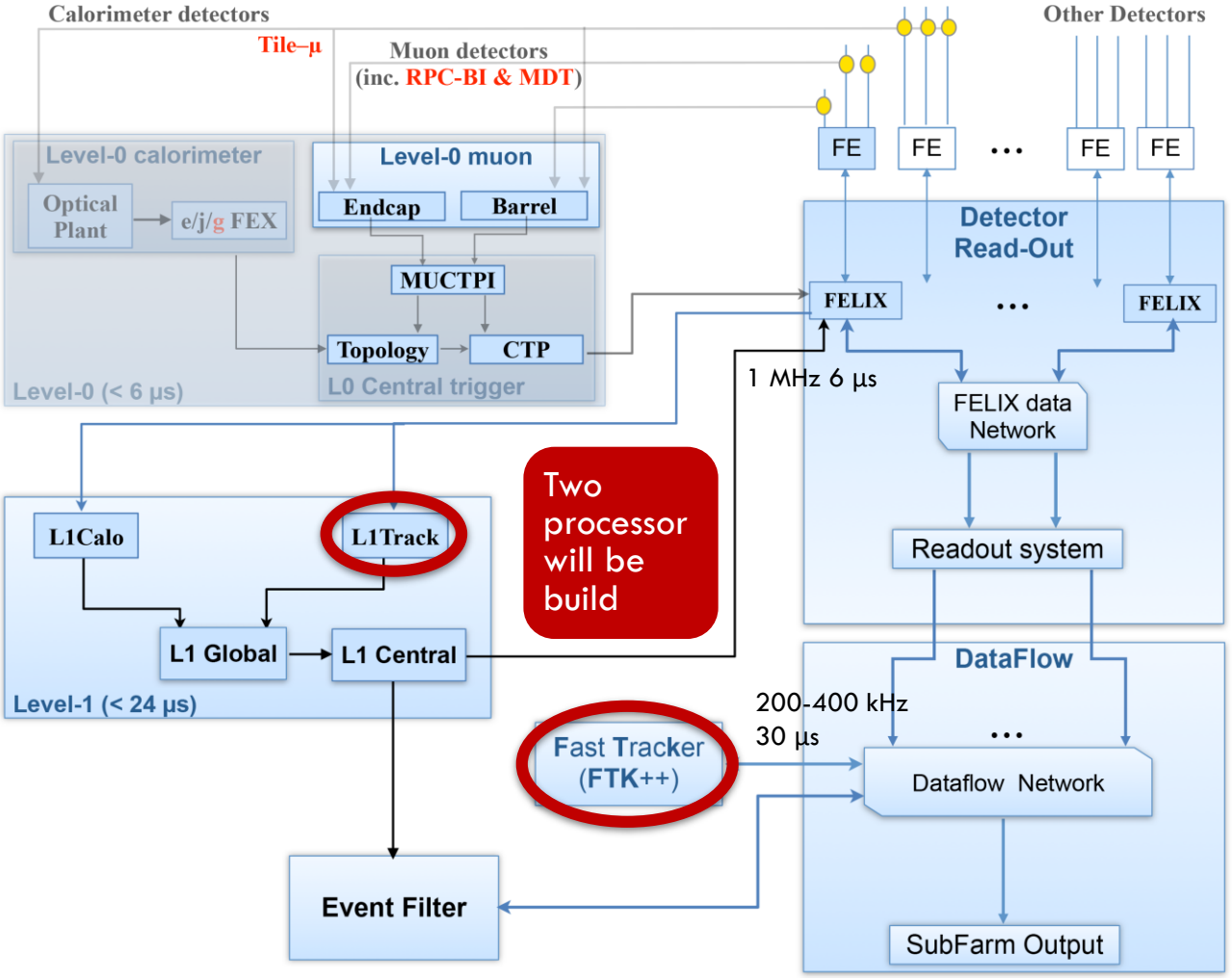


ATLAS APPROACH



New electronics being designed for 6 µs and 30 µs respectively

ATLAS APPROACH



LAST SIMULATION RESULTS

Need to match AM chip performance with …

- # pattern requirement from simulation studies and # of AM chips in the system (negotiable)

CMS:

- # pattern/trigger tower 1.5 - 2.5 M
- ~ N*120 Mpatterns total
- N (~10-20) number of parallel processors to reach 40MHz

ATLAS LV1 Track:

- A few M patterns / "trigger tower"
- ~ 800 "trigger towers"

ATLAS FTK++ order of 4 billion patterns

All 3 projects need a few billion patterns

	Year	Density (No. Mbits)	Working Frequency (MHz)	Power (W)	Voltage (V)	Technology	Area (cm^2)
AM03	2004	0.5	40	1.26	1.8	180 nm	1
AM04	2012	1.2	100	3.70	1.2	65 nm	0.12
AM06	2014	19	100	2-3	1.0/0.8	65 nm	1.6
AM2020	**2020?**	**76?**	**Up to 200**	**4-8 @200 MHz**	**0.9**	**28 nm**	**?**

A NEW MULTI-PURPOSE IDEA

AM chip v.6: Tailored for pattern matching in the **ATLAS level-2** experiment

The IMPART project will **expand** our research toward a **more powerful ASIC** (CMOS 28 nm) and a **more flexible device**

More patterns per chip and reduced latency and power consumption

Repetitive task (i.e., bit-wise comparison) performed by new AM chip

Complex task (i.e., image analiysis algorithms, serialization procedures) perfomed by FPGAs

IMPART device on a compact PC board as **coprocessor** fro several **interdisciplinary** application

Computer vision for **smart cameras** and **medical** imaging

DNA sequencing applications

High Energy Physics (HEP) **Trigger DAQ** R&D

IMPART PROJECT: AMCHIP TREND

Since 2010:

- **AM chip v. 4** designed and characterized 2012 – area: 14 mm² – cost: 50 k€ – 8 kpattens

65 nm

65 nm

28 nm + FPGA

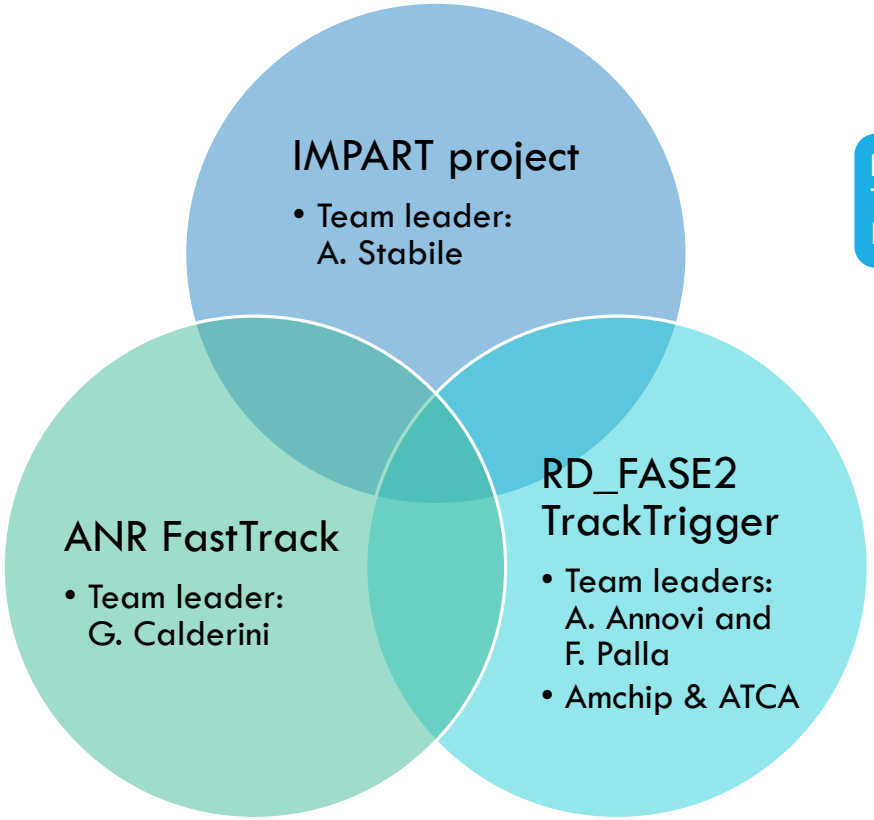
2015-2016 IMPART

- **IMPART chips will be** designed and characterized within the IMPART project for several disciplines:
  - **Image analysis**
  - **DNA sequencing**
  - **Trigger DAQ**

Since 2011:

- **AM chip v. 5-mini@sic** designed and characterized 2012 – area: 4 mm² – cost: 20 k€ – 256 pattens for cell test
- **AM chip v. 5-mpw** designed and characterized 2013 – area: 12 mm² – cost: 75 k€ – 5 kpattens
- **AM chip v. 6** under design; to be submitted in Dec. 2014 – area: 160 mm² – cost: 550 k€ – 128 kpattens

R&D NEW AMCHIP 28NM



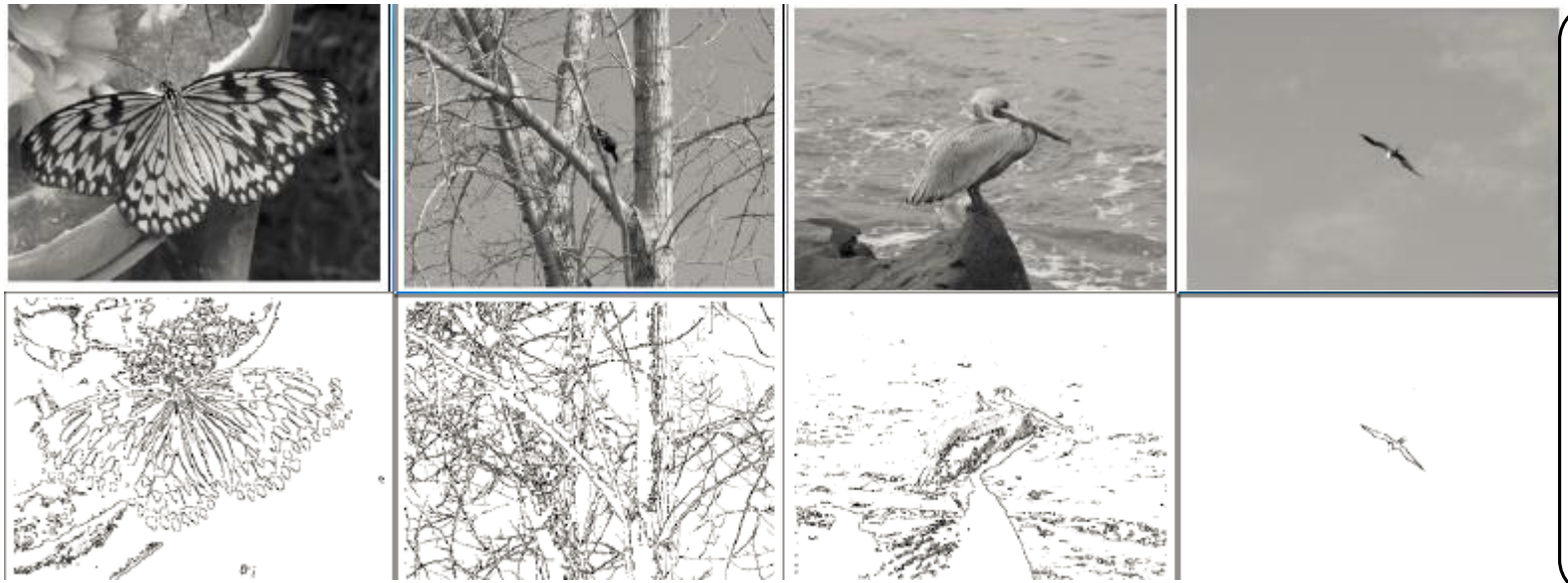
Firmware: C. GENTSOS
TIPP2014:The Fast TracKer Processing Unit future evolution

COMPUTER VISION FOR SMART CAMERAS AND MEDICAL IMAGING APPLICATIONS

Smart cameras capture high-level description of a scene and perform **real-time extraction of meaningful information**

- Current compression algorithms: **few seconds** are required
- **For safety-critical applications (e.g., transports, or personnel tracking in a dangerous environment), latency could lead to serious problems.**

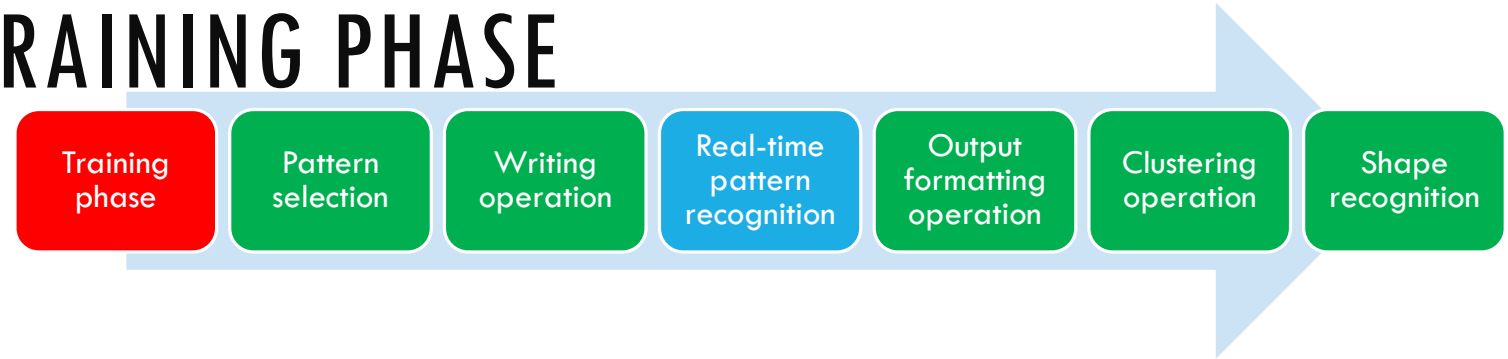
Del Viva et al algorithm[1] studied how to reproduce initial stage of the brain visual processing: find contourns



off-line simulation results

[1]*M. Del Viva, G. Punzi, and D. Benedetti. Information and Perception of Meaningful Patterns. PloS one 8.7 (2013): e69154.*

TRAINING PHASE



FPGA reorganizes the input data in small patterns: 3×3×M square pixel matrices

- M ranges from 1 to 3
  - 1 static b/n image
  - 2 four gray level image
  - 3 mov

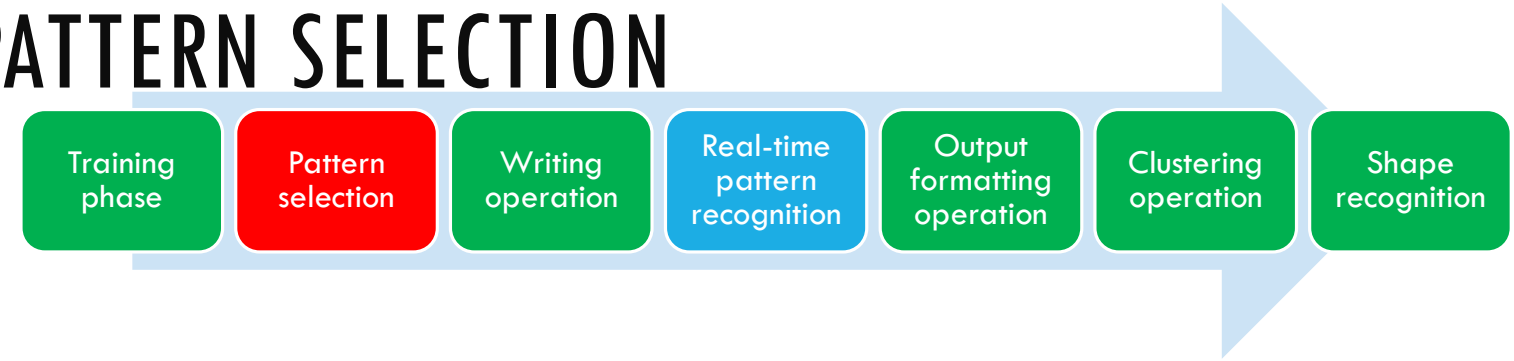
For each pattern, the FPGA **calculates the occurrence** of the analyzed pattern in the processed images/frames.

Calculation is **iterated** for all possible patterns in a large set of training images.

Different Probability Density Histograms (PDHs) computed for different training image sets.

- **Medical** images have different PDHs than **natural** images

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 shape (both human and artificial) we adopt the hypothesis described in Del Viva's paper:

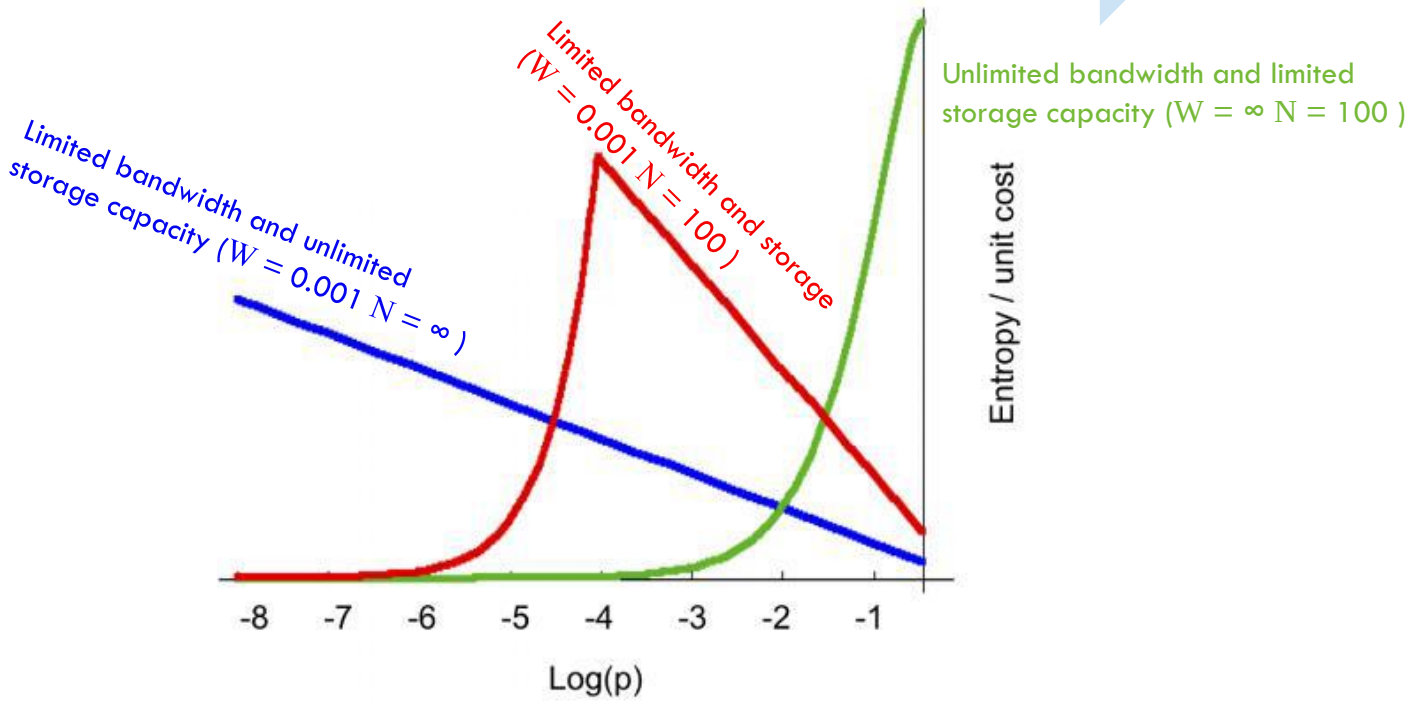
- Maximum entropy is a measure of optimization: **the set of patterns that produces the largest amount of entropy is the best set** of patterns that we can select to filter our images or videos.
- Entropy yield per unit cost and for each pattern is given by:

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

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.

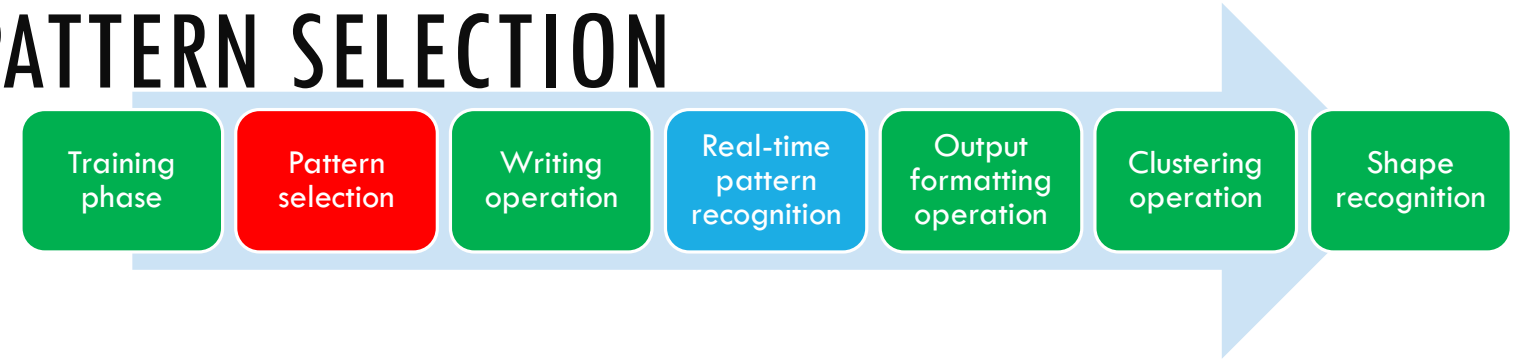
PATTERN SELECTION

Training phase | **Pattern selection** | Writing operation | Real-time pattern recognition | Output formatting operation | Clustering operation | Shape recognition



Entropy yield functions *for different values of* W *and* N.

PATTERN SELECTION



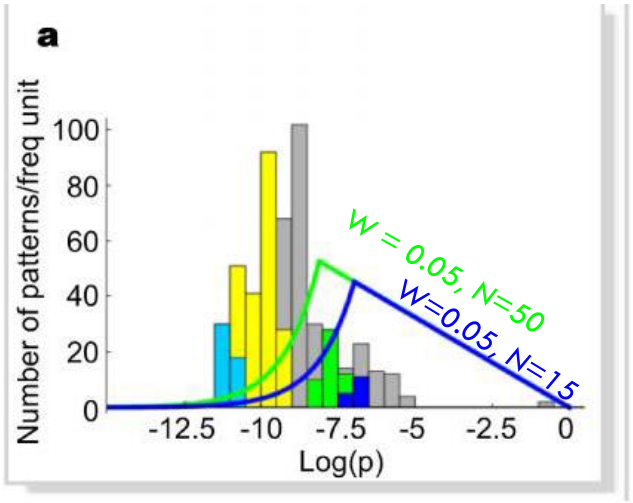
The **optimal solution** in the selection of pattern is made by choosing the set of pattern such that $f(p)>c$, where c is determined by the computational constraints:

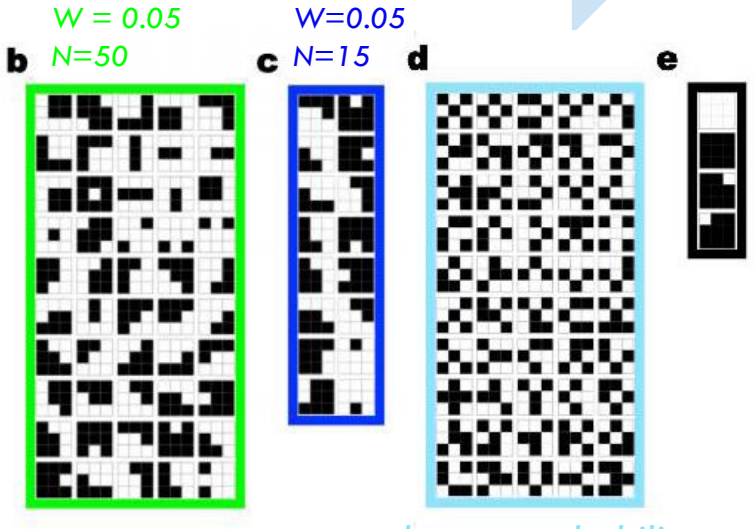
$$\int_{f(p)>c} f_n(p)dp < N \qquad \frac{1}{N_{tot}} \int_{f(p)>c} p f_n(p)dp < W$$

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.

PATTERN SELECTION

Training phase | **Pattern selection** | Writing operation | Real-time pattern recognition | Output formatting operation | Clustering operation | Shape recognition



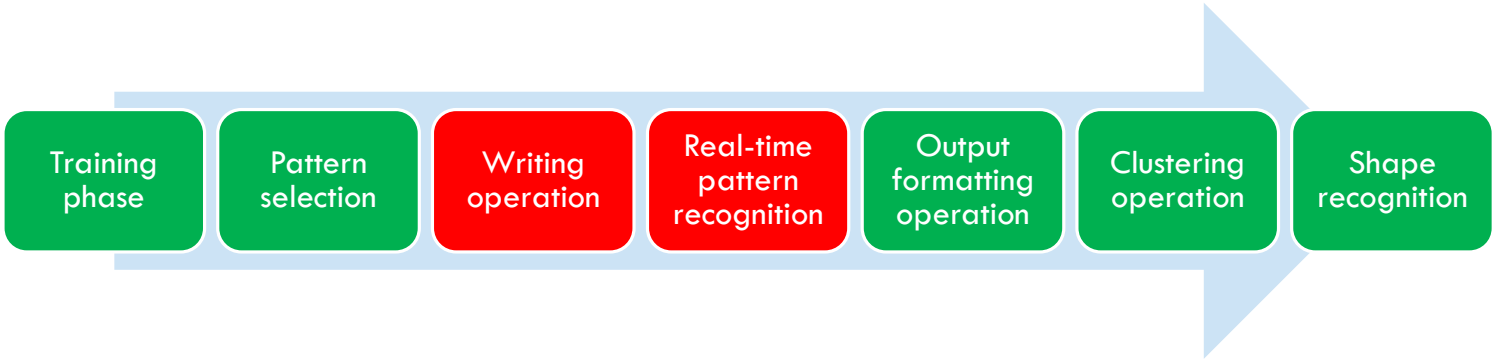


highest probability patterns (resembling uniform luminance)

lowest-probability Patterns resembling visual noise

Probability distribution of the $N_{tot}=2^9$ possible 3×3 square pixel matrices in black-and-white (1-bit depth) for natural images.

WRITING OPERATION & REAL-TIME PATTERN RECOGNITION



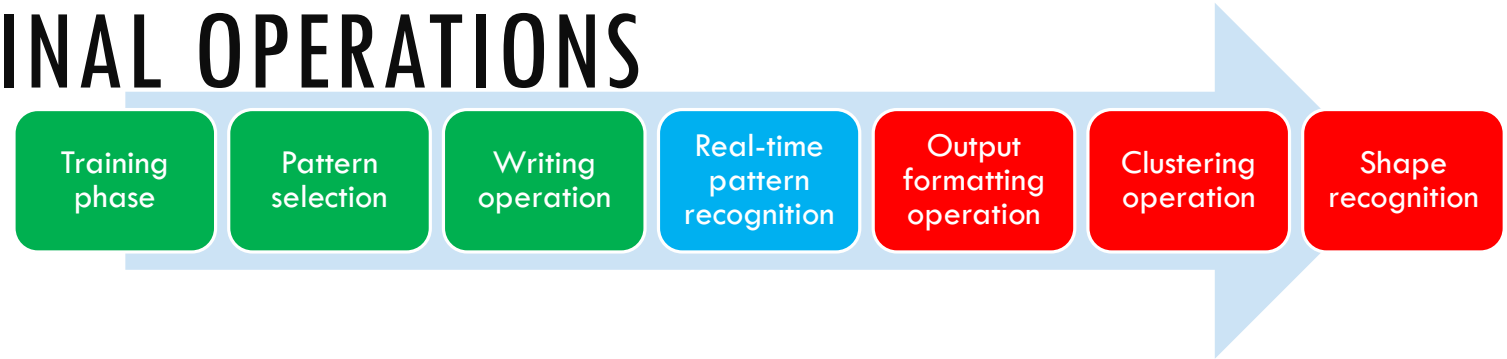
The relevant pattern (selected in the second step) is **written** in the AMchip bank

The IMPART system works 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)

The **found pattern address** is transferred at the output from the AM chip to the FPGA.

FINAL OPERATIONS



Output formatting operation: The resulting patterns are reorganized into a new compressed image by the FPGA

- Produced the filtered images/video (sketches)
  - only the boundaries of the relevant objects are kept

Clustering operation: A clusterization of contours is performed to convert the image from a raster format into a vector image

A shape recognition algorithm recognizes the salient shapes used for the final high-level elaboration

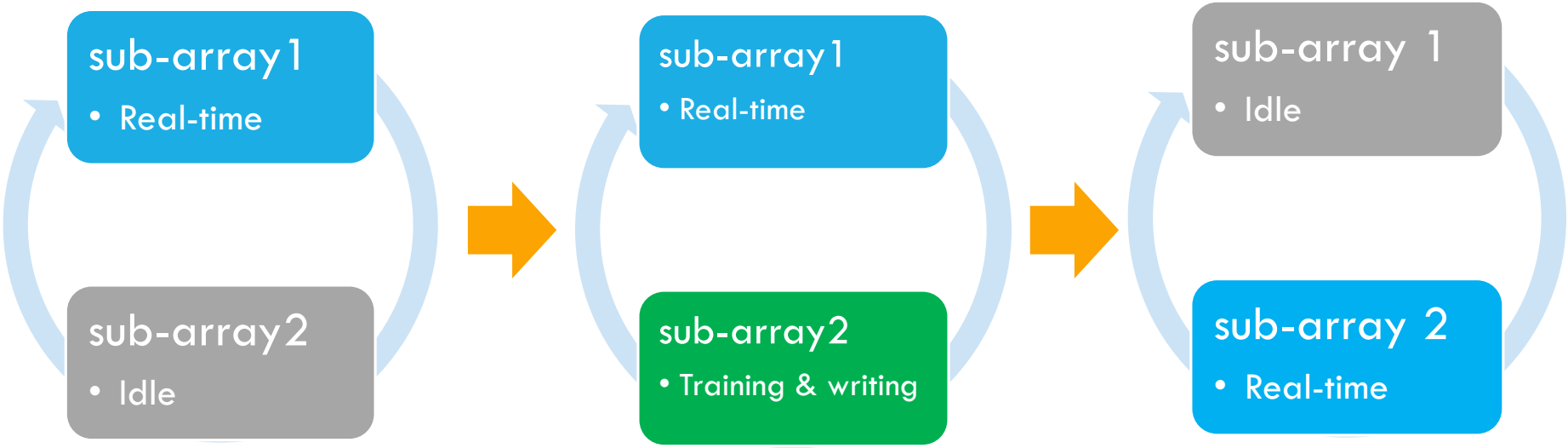
FPGA CONTINUOUS MONITORING

Commonly, **training** is performed **before pattern recognition** causing "dead time"

Solution

Training online (in the same time of real-time pattern recognition)

Split AM chip in two equal sub-array



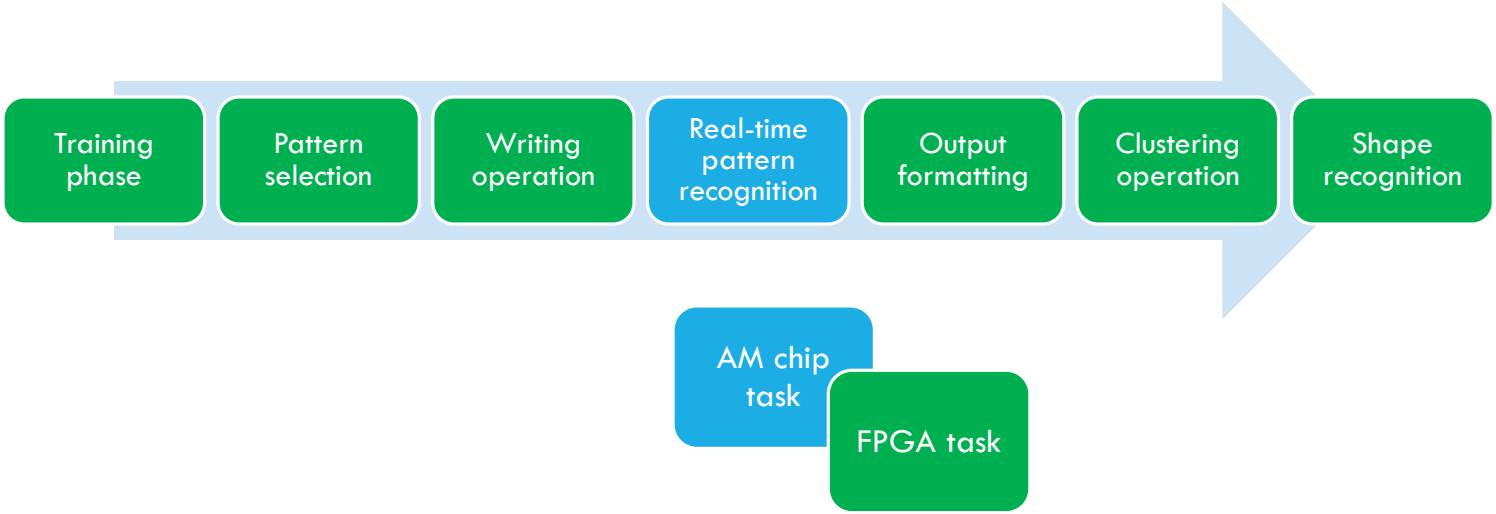
OFF-LINE PRELIMINARY RESULTS

IMPART system a **parallel hardware** for Del Viva[1] et al. algorithms:

- huge, very fast and general compression of data
- **suitable for the AM chip architecture**

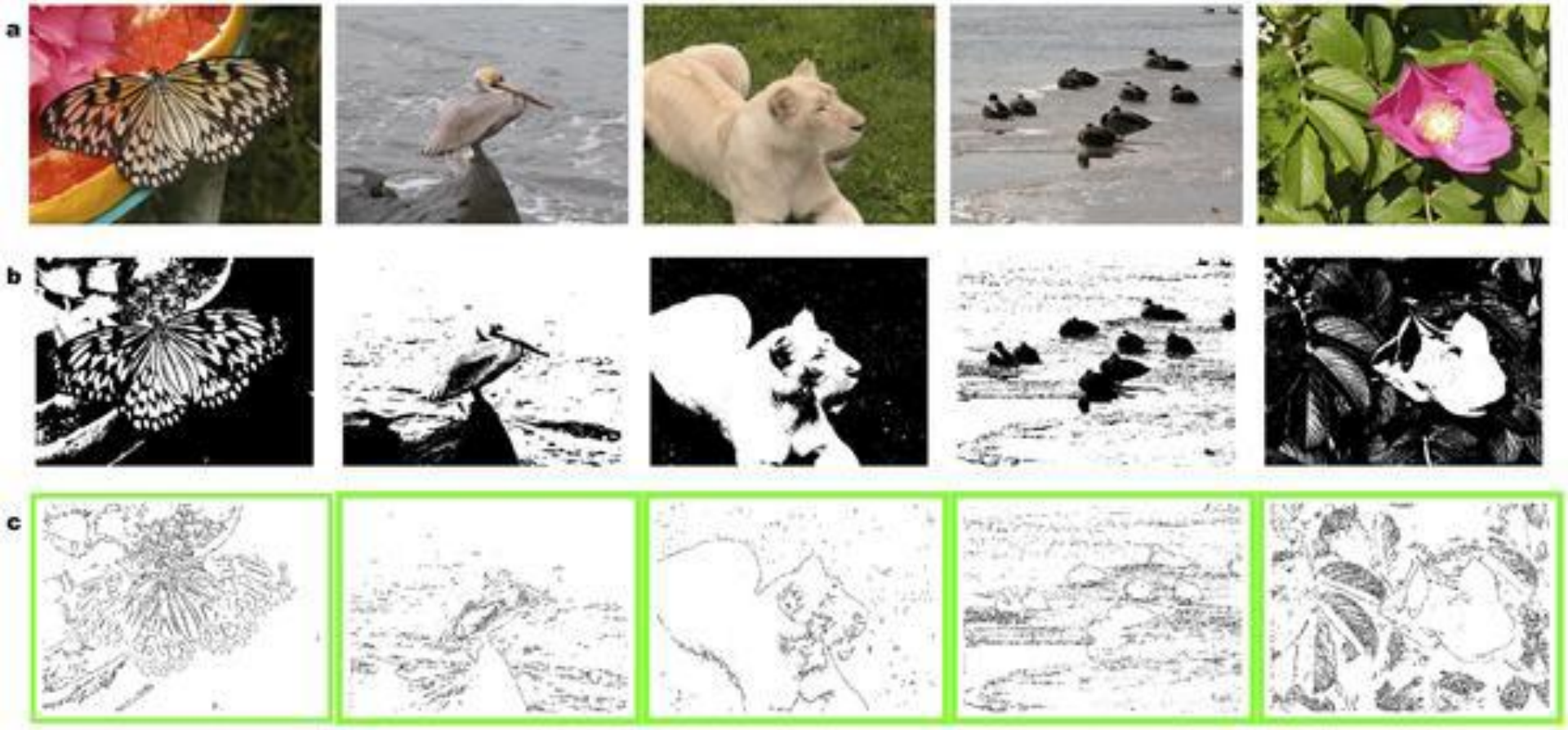
Outdoor environments requiring **solar panels and batteries.**

- For this reason, a system that **minimizes the power consumption** is essential.



[1]*M. Del Viva, G. Punzi, and D. Benedetti. Information and Perception of Meaningful Patterns. PloS one 8.7 (2013): e69154.*

OFF-LINE PRELIMINARY RESULTS



OFF-LINE PRELIMINARY RESULTS

four levels of grey

SMART CAMERAS IMPACT

IMPART hardware used for **smart cameras for smart cities** and **smart transportation systems** (Italian Government & Horizon 2020 expectations)

Collaboration with a SME, namely, EMC S.r.l. (Siena, Italy)

- IMPART system will be installed **in smart camera** devices to demonstrate and to exploit the validity of the Del Viva's algorithm

FUTURE DEVELOPMENTS

Medical application

Automated medical diagnosis:

- Huge amount of image data
  - time-varying images
  - very accurate resolution

Real-time applications for **adaptive radiotherapy** in collaboration with the INFN-Torino research group

- Simona Giordanengo's CSN5 funded project in 2013

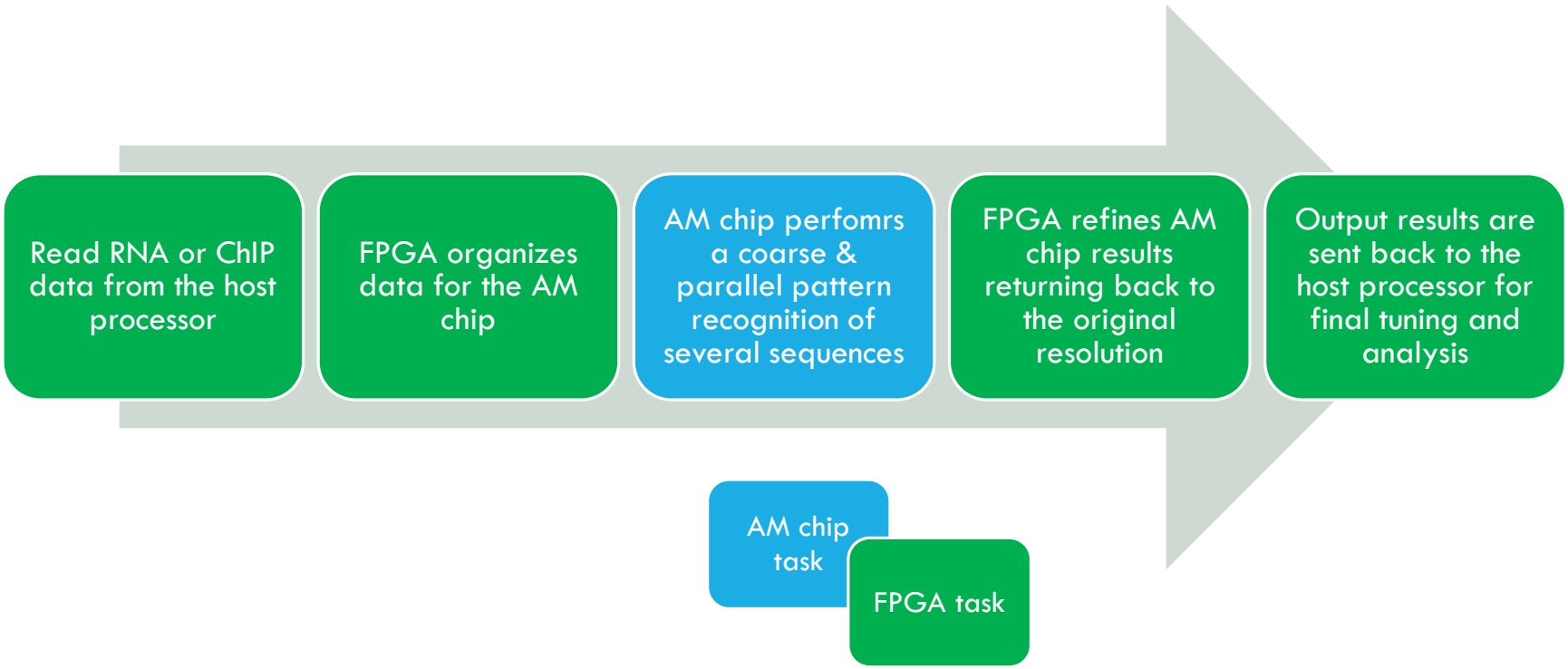
IMPART-based system **performance:**	Commercial machines **performance:**
• **Human exome:** 1.5 % subset of the human genome (25 million nucleotide pairs) • **Nucleotide encoding:** FASTA format (at least 4 bits are needed) • Whole exome alignment with **IMPART device: ~ 4 s**	• Bowtie based machines: **1 CPU hour**

Improvement factor is about 900x

DNA application

DNA SEQUENCING APPLICATION

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.



SOCIAL IMPACT OF IMPART PROJECT

These innovative smart systems could ameliorate the life quality and safety in congested and overcrowded environments

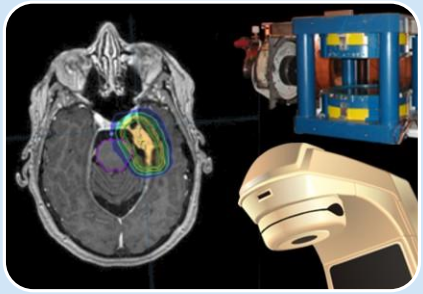
- cities, construction sites, stations, and airports

Smart cameras with the IMPART system could be installed in remote environments (forests or mountains)

- supplied by batteries and solar panels
- to **monitor environmental phenomena** such as avalanches, landslips, summer fires

Transports could benefit from the project (safety point of view)

- Monitor the speed, the distance between vehicles, and the temperature of the air.
- Alternative choice to the state-of-the-art systems



IMPART system could be also used to **monitor medical personnel and patients**

- nurse hand recognition to reduce **the possibility of biological contamination**
- to monitor body movements of a patient under **adaptive radiotherapy.**

Several applications could improve from the project outcome.

OTHER NEW CELLS