Reconfigurable and Rad-Hard Accelerated Computing in Space

Andrea Portaluri
Eleonora Vacca
Luca Sterpone

Politecnico di Torino

Outline

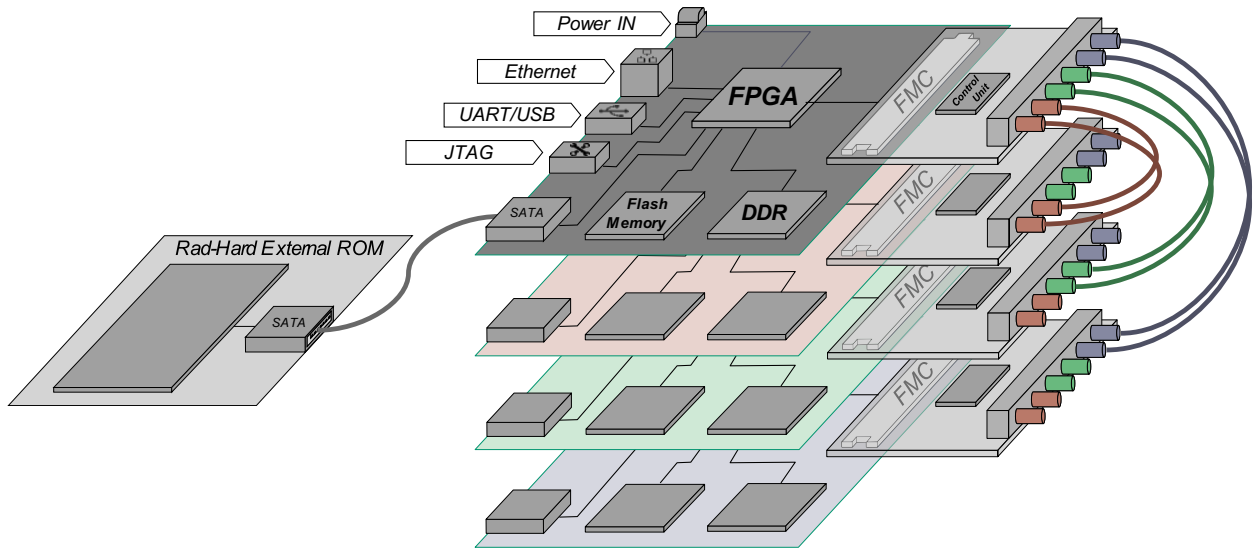
- Exascale Computing in Space
- High-Performance Computing Node
- NG-Medium Overall Placement
- Implementation and analysis of VLIW processor on NG-Medium
- Conclusions
- Future works

Exascale Computing in Space

- **The next generation of flight computing electronic systems will address high computational performance**
  - High computational demanding algorithms
  - Telecommunication payload
  - Earth or Planets observation
  - Quantum communication
  - Autonomous operativity (e.g., Artificial Intelligence algorithm)
- **The computing architecture will be oriented to multicore computing chips with multiple processing cores**

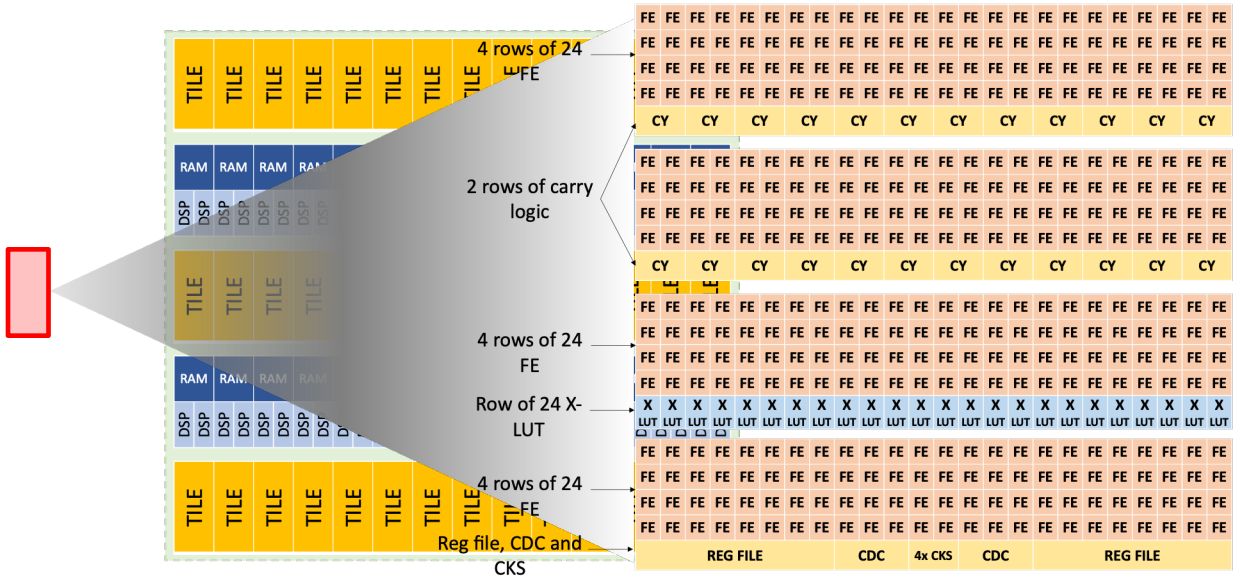
High-performance Computing Node

- **A cluster of FPGAs as a unique entity taking advantage of the resource abundance to support**
  - High performance and efficient computation
  - Flexibility to introduce radiation mitigation techniques
- **Increasing system robustness while efficiently balancing computation workload and envisioning increased scalability**



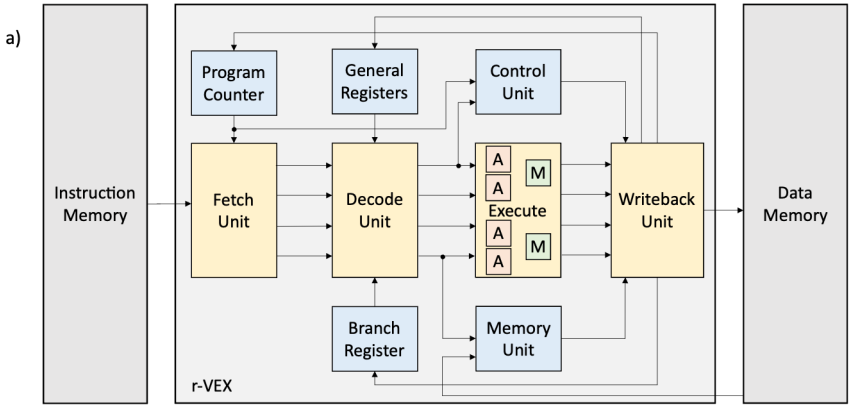
NG-Medium Overall Placement

- Space-grade FPGA technologies from STM 65nm to 28nm FD-SOI
- Cluster-oriented programmable logic with tight placement constraints



VLIW processor on NG-Medium

- A Reconfigurable VLIW Example (r-VEX) is a Very-Long Instruction Word reconfigurable processor
- Number of registers, ALUs, memory sizes, and multipliers is parametrizable
- Its architecture can loosely mimic the latest AMD Xilinx AI Engine control unit based on VLIW processors



b)

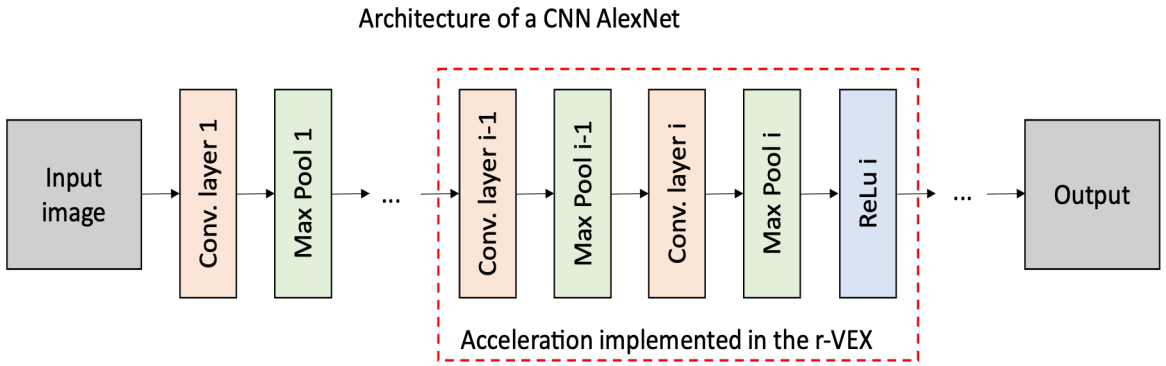
31					
opcode [31:25]	0 0	dest GR [22:17]	src1 GR [16:11]	src2 GR [10:5]	dst BR [4:2]
opcode [31:25]	0 1	dest GR [22:17]	src1 GR [16:11]	short immediate[10:2]	
opcode [31:25]	1 0	link reg [22:17]	branch offset immediate [16:5]		dst BR [4:2]
opcode [31:25]	1 1	dest GR [22:17]	src1 GR [16:11]	long immediate[10:1]	
opcode [31:25]	dest GR [24:3]				

Instructions are directly coded into the VHDL through Assembly

5 types of instruction packed in syllables of 4 (parametrizable)

Convolution in VEX Assembly

- **A portion of the AlexNet CNN architecture has been chosen to be accelerated by the r-VEX as a benchmark application.**
- **Assembly equivalent of the convolutional steps has been written and coded into the instruction memory of the r-VEX**



Layer	Input	Stride	Padding	Filter	Output
Conv. Layer 1	12 x 12	1	1	3 x 3	12 x 12
Max Pool 1	12 x 12	0	1	2 x 2	11 x 11
Conv. Layer 2	11 x 11	0	1	3 x 3	9 x 9
Max Pool 2	9 x 9	0	1	2 x 2	8 x 8

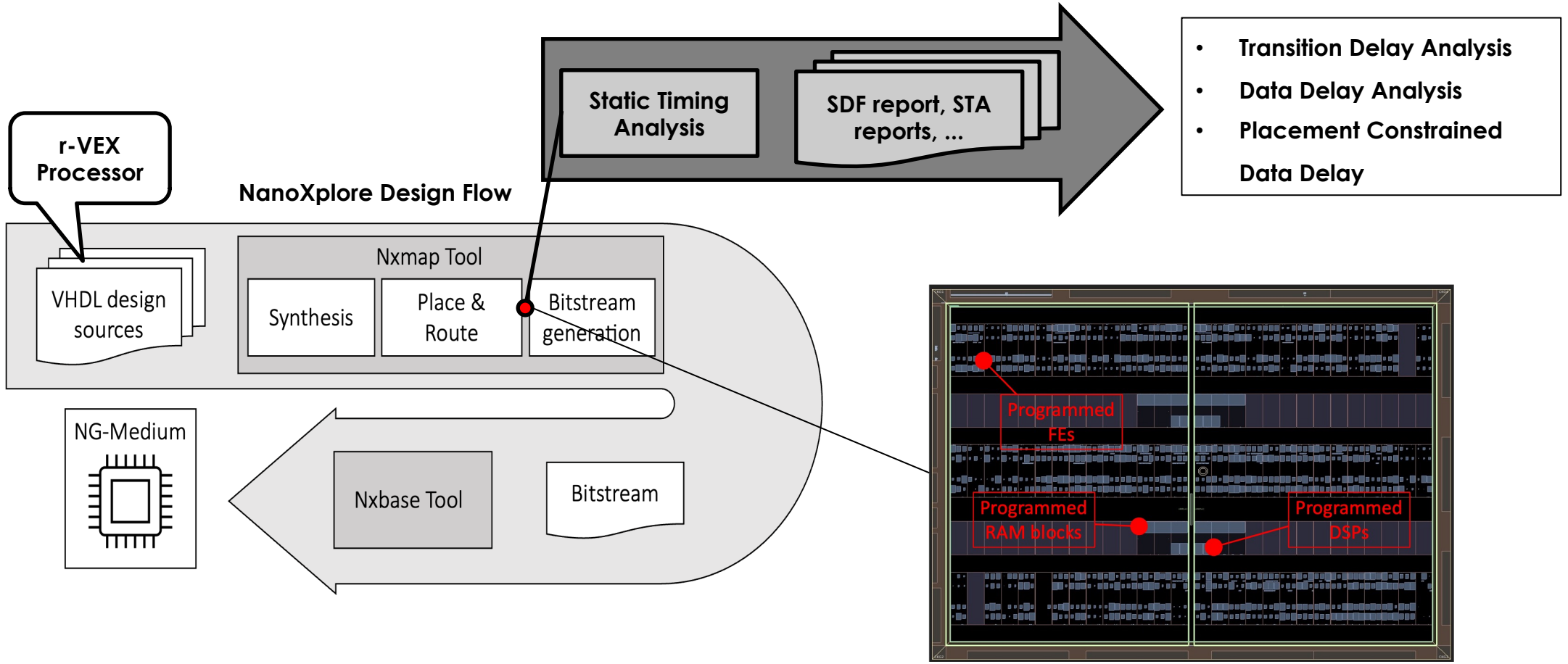
Convolution in VEX Assembly

- NanoXplore NG-MEDIUM chip has been the target of the analysis
- Three different designs have been implemented varying the number of r-VEX cores (1-core, 2-core, and 3-core, respectively)
- A USB-to-UART bridge has been used to read serial data from a UART VHDL module implemented along the design

General Registers [#]	64
Branch Registers [#]	8
ALUs [#]	4
Multipliers [#]	2
Syllable-Issues [#]	4
Data Memory [kB]	32

Design	4-LUT	DFF	Carry Logic	DSP	BRAM
1-core	6,098 (19%)	1,901 (6%)	298 (4%)	3 (3%)	4 (8%)
2-core	11,714 (37%)	3,801 (12%)	596 (8%)	6 (6%)	8 (15%)
3-core	17,946 (56%)	5,703 (18%)	894 (12%)	9 (9%)	12 (22%)

The Implementation Flow



- Transition Delay Analysis
- Data Delay Analysis
- Placement Constrained Data Delay

Experimental Analysis

- **The Standard Delay Format (SDF) analysis allows to estimate transition timing delays**
  - Internal Routing: routing resources into the same tile
  - General Routing: routing resources for inter-tile communication
  - Low-Skew Routing: routing resources for the distribution of global signals

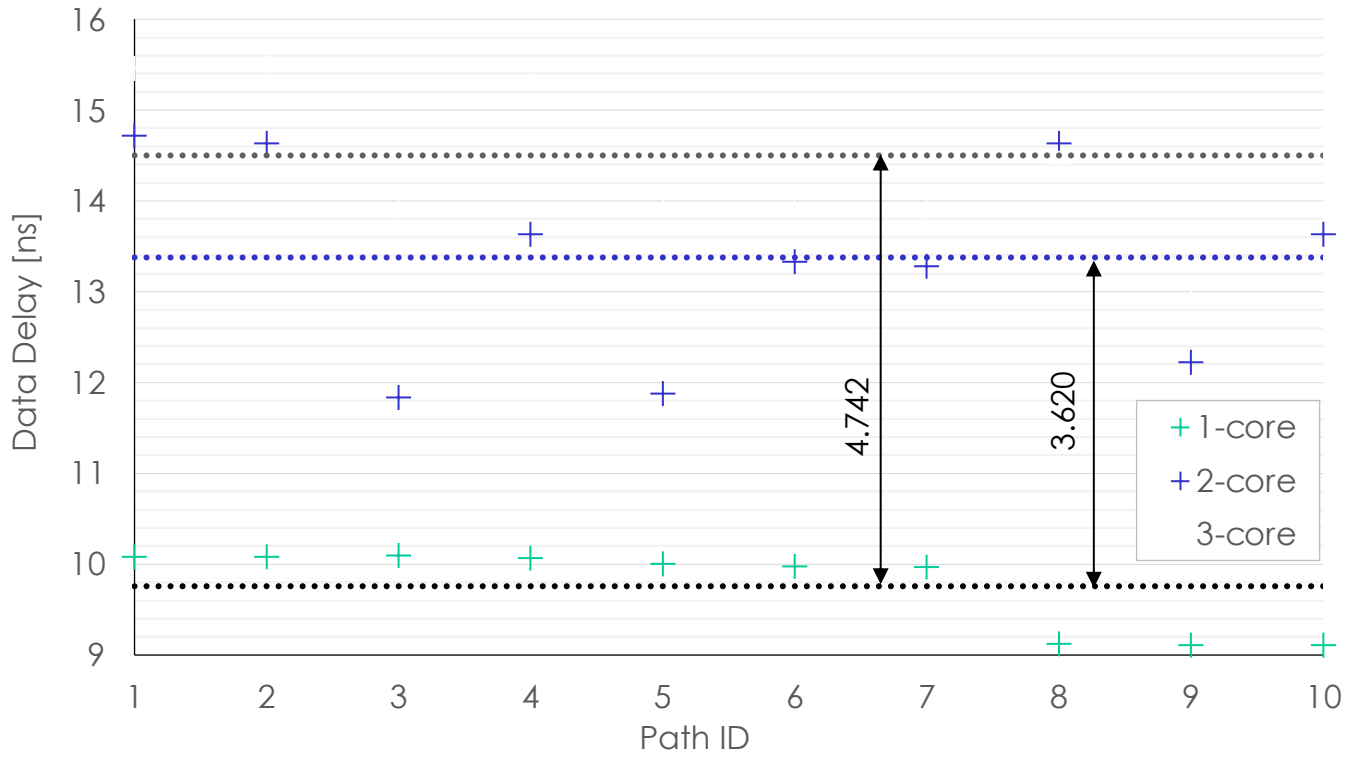
Design	$0 \rightarrow 1_{min}$ [ns]	$0 \rightarrow 1_{max}$ [ns]	$1 \rightarrow 0_{min}$ [ns]	$1 \rightarrow 0_{max}$ [ns]
1-core	2.278	2.378	2.278	2.378
2-core	2.497	2.601	2.497	2.601
3-core	2.537	2.643	2.537	2.643

Design	$0 \rightarrow 1_{min}$ [ns]	$0 \rightarrow 1_{max}$ [ns]	$1 \rightarrow 0_{min}$ [ns]	$1 \rightarrow 0_{max}$ [ns]
1-core	2.742	2.827	2.742	2.827
2-core	2.801	2.922	2.801	2.922
3-core	2.905	2.985	2.905	2.985

Design	$0 \rightarrow 1_{min}$ [ns]	$0 \rightarrow 1_{max}$ [ns]	$1 \rightarrow 0_{min}$ [ns]	$1 \rightarrow 0_{max}$ [ns]
1-core	6.605	6.667	6.605	6.667
2-core	6.587	6.646	6.587	6.646
3-core	6.683	6.743	6.683	6.743

Experimental Analysis

- **Through the data delay analysis, the 10 worst routing paths in terms of timing have been evaluated**



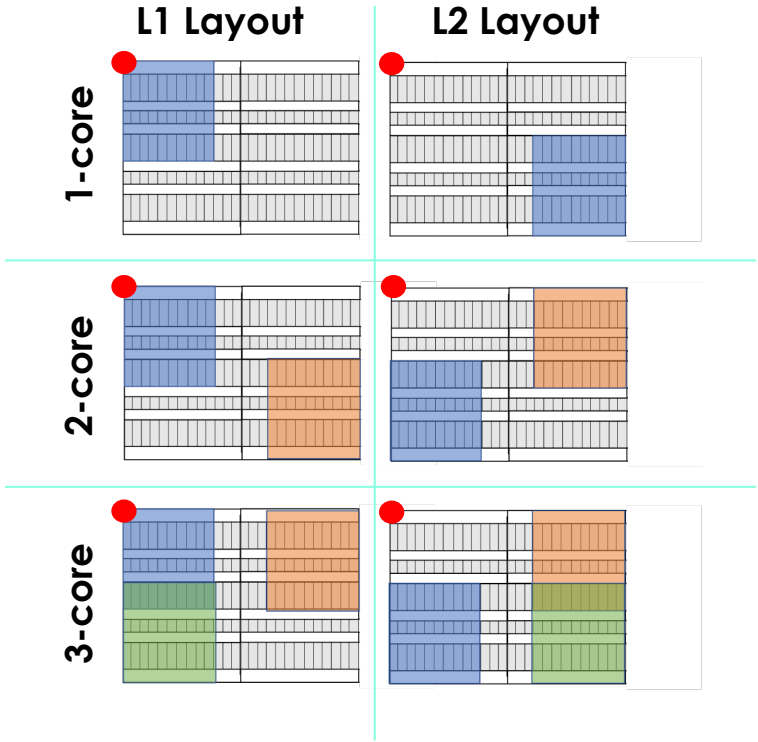
Resource utilization (and number of cores) affects negatively delays of nets

2-core and 3-core designs delays are comparable with clock period

Design	Max Frequency [MHz]
1-core	32.753
2-core	27.586
3-core	23.384

Experimental Analysis

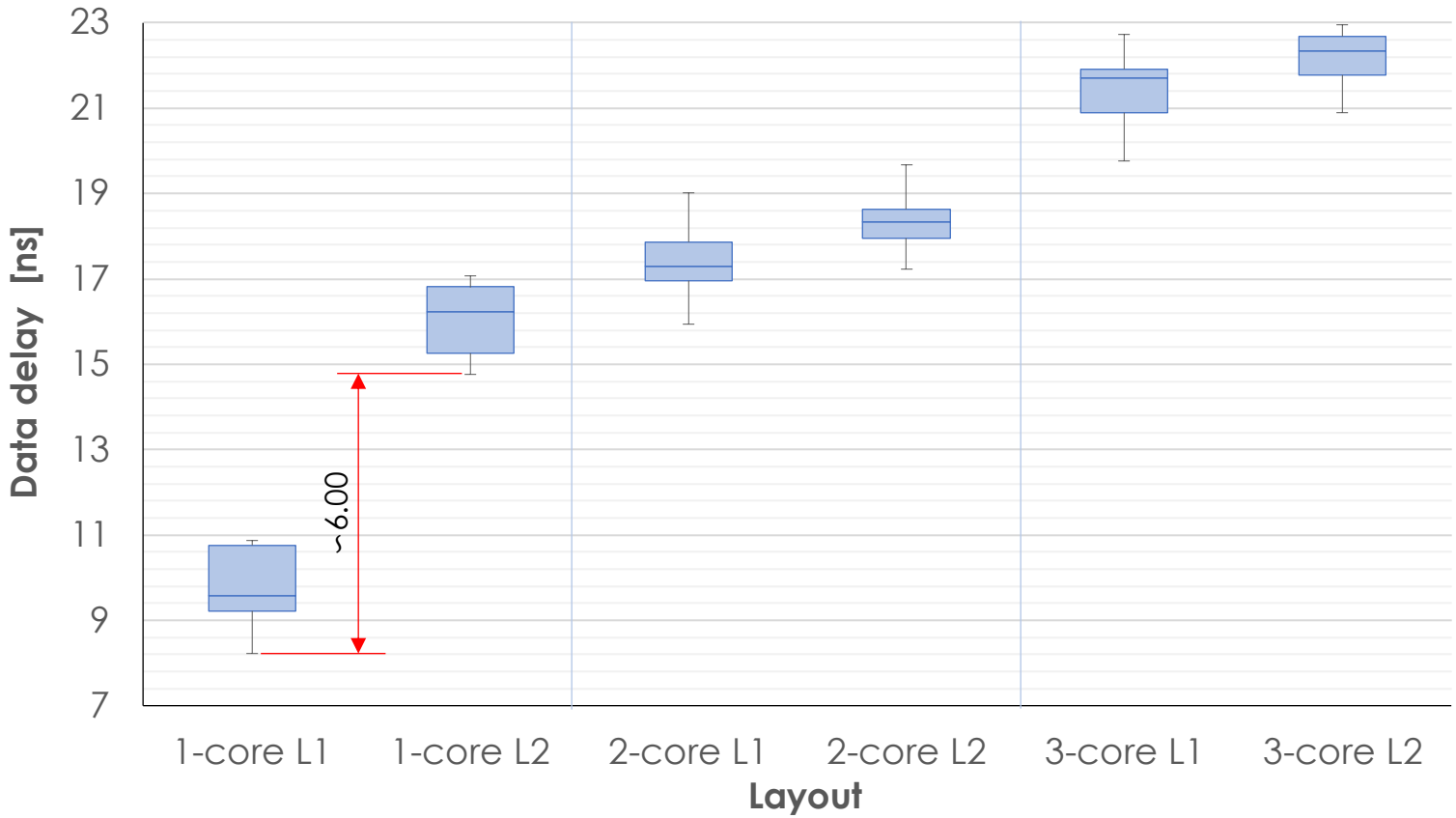
- **Finally, dependencies from core locations have been analyzed in terms of timing:**
  - 2 different layouts have been implemented for each design
  - L1 minimizes the distance from the I/O buffers while L2 maximizes it
  - Each color represents a r-VEX core
  - The red circle represents the I/O buffer



The 3 cores overlap due to unavailability of resources

Experimental Analysis

- **Data Delay analysis of the 10 worst paths with placement influence**



Overall deterioration of the performances with higher distances to the I/O buffers

Highest difference between 1-core layouts due to the highest average distance from I/O

Degradation of other L2s appears to be less since routing paths are shorter

Conclusions

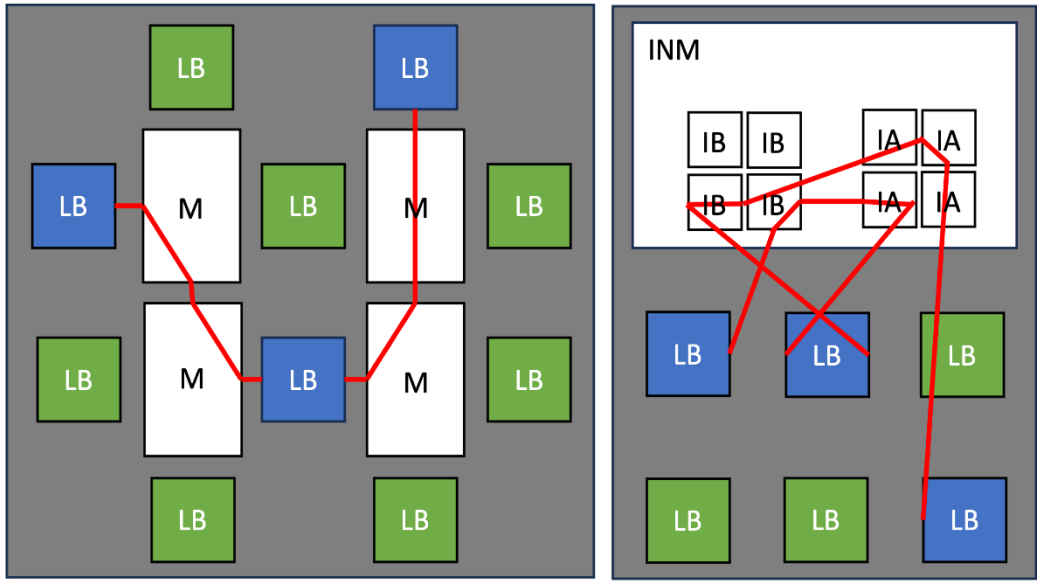
- **We implemented single and multi-core solutions of a convolutional accelerator in the r-VEX soft processor on a NanoXplore NG-Medium device**
- **The timing performances have been evaluated**
- **Experimental results show that the placement location of the cores plays a key role in the timing performances, which can be seriously affected,**
  - Timing constraints are often not enough to fill this gap

Future Works (i)

- To develop a routing model reflecting the internal routing details, managing internal routes and auxiliary logic

(a) Routing using switch matrices, where the delay is dependent on the physical distances among the routed logic blocks

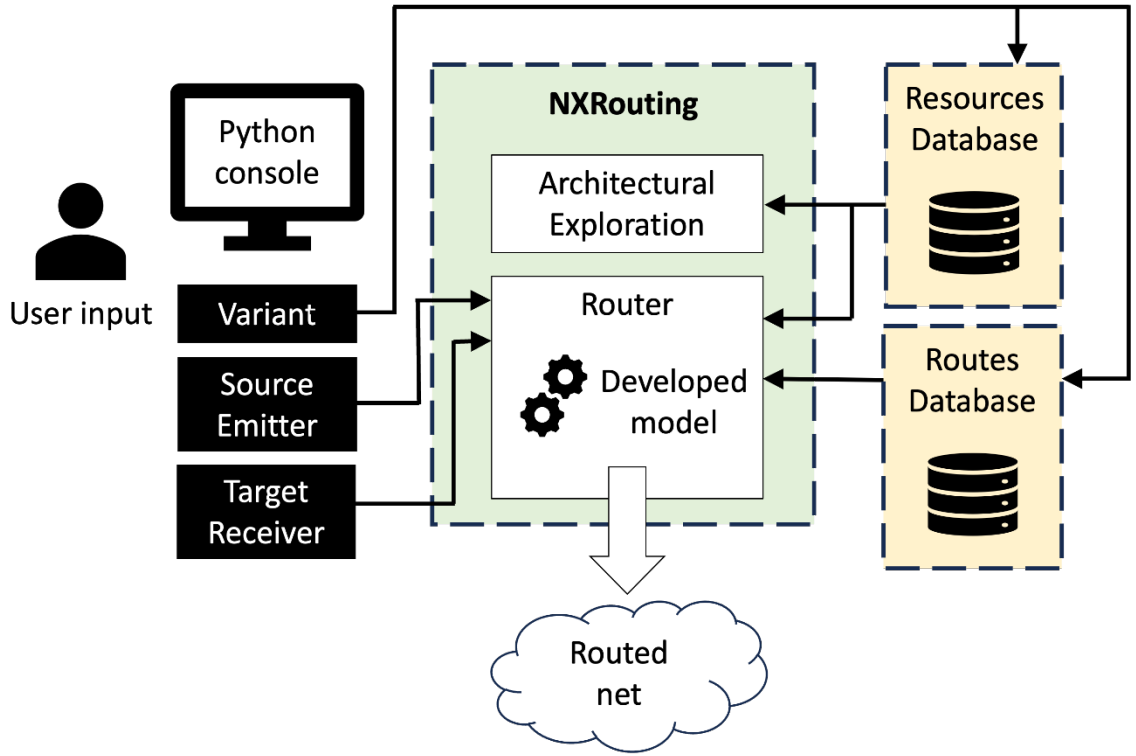
(b) Routing using RI, where the delay only roughly depends on how many times the route enters the RI



Future Works (ii)

- **A routing feature has been added to the tool to retrieve all the connection branches between two points in the programmable logic**

Benchmark	Routing Time [s]			
	B03	B06	B09	B12
Sequential Router	143.17	62.73	300.05	433.14
Fine-Grained Router	0.81	0.47	0.89	4.62
Fine-Coarse-Grained Router	0.60	0.47	0.58	2.64



Future Works (iii)

- **On-going ESA OSIP Project**

Towards Exascale Reconfigurable and Rad-hard Accelerated Computing in space

asaclab.polito.it

Home People

ASAC LAB

AEROSPACE SAFETY AND COMPUTING LAB

The **Aerospace, Safety, and Computing (ASaC) Lab** is a research team of the Electronic CAD & Reliability Group in the Department of Control and Computer Engineering (DAUIN) of Politecnico di Torino. Its mission is to support designers of electronic circuits and systems with expertise, techniques, methodologies, and tools. The ASaC Lab's research involves several multidisciplinary topics such as:

IEEE FPL 2024

REGISTRATION FOR FPL 2024 IS OPEN!

CALL FOR TUTORIAL AND WORKSHOP IS OPEN

The International Conference on Field-Programmable Logic and Applications (FPL) was the first and remains the largest conference covering the rapidly growing field-programmable logic and reconfigurable computing area. During the past 33 years, many of the advances in reconfigurable system architectures, applications, embedded processors, design automation methods, and tools were first published in the proceedings of the FPL conference series. The conference's objective is to bring together researchers and practitioners from academia, industry, and worldwide.

The 34th edition of the International Conference on Field-Programmable Logic and Applications (FPL 2024) will be held in Turin from 2nd to 6th September 2024. The 2024 edition of the conference is organized by ASAC (Aerospace, Safety, and Computing) Lab of Politecnico di Torino, Turin, Italy.