CEI UPM | Centro de Electrónica Industrial

Università degli Studi di Cagliari

UNISS UNIVERSITÀ DEGLI STUDI DI SASSARI

CPS Creative Lab

Algero, September 16th, 2024

Andrés Otero, Juan Encinas, Francesco Ratto, and Claudio Rubattu

SEPTEMBER 16-20 2024 – ALGHERO, SARDINIA, ITALY

CYBER PHYSICAL SUMMER SCHOOL SYSTEMS 2024

UNIVERSIDAD POLITÉCNICA DE MADRID

POLITÉCNICA

Who We Are?

Andrés Otero
joseandres.otero@upm.es

Juan Encinas
juan.encinas@upm.es

POLITÉCNICA | UNIVERSIDAD POLITÉCNICA DE MADRID

CEI UPM

Francesco Ratto
francesco.ratto@unica.it

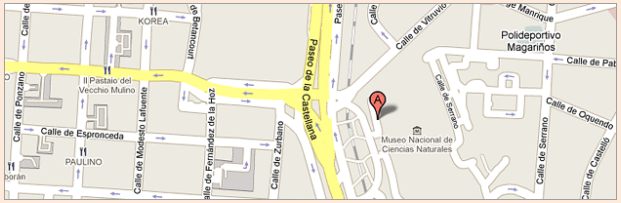
Claudio Rubattu
crubattu@uniss.it

Università degli Studi di Cagliari

UNISS UNIVERSITÀ DEGLI STUDI DI SASSARI

Centro de Electrónica Industrial (CEI) at UPM

Placed at E.T.S.I. Industriales of the UPM



Located in the City Center of Madrid

CEI UPM | Centro de Electrónica Industrial

Some Numbers

- 46 Full-time researchers
  - 19 Doctors (Faculty, Contracted researchers)
  - 30 Full-time Students (18 PhD, 20 MSc)
- 19 Part time, sponsored students (Bachelor)
- 3 Administration & Technicians

Power Electronics

- Power Converters
- Power Supply Architectures
- Modeling & Simulation
- Smart Grids
- Energy Harvesting

Digital Embedded Systems

- Reconfigurable Embedded Systems
- *Open HW* and *RISC-V*
- Internet of Things
- Machine Learning at the Edge
- Post-Quantum Cryptography

The Creative Lab at CPS 2024

The Creative Lab is a **highly practical hands-on experience** meant to promote the exchange of ideas between PhD students and young researchers with the spirit of creating a new business culture.

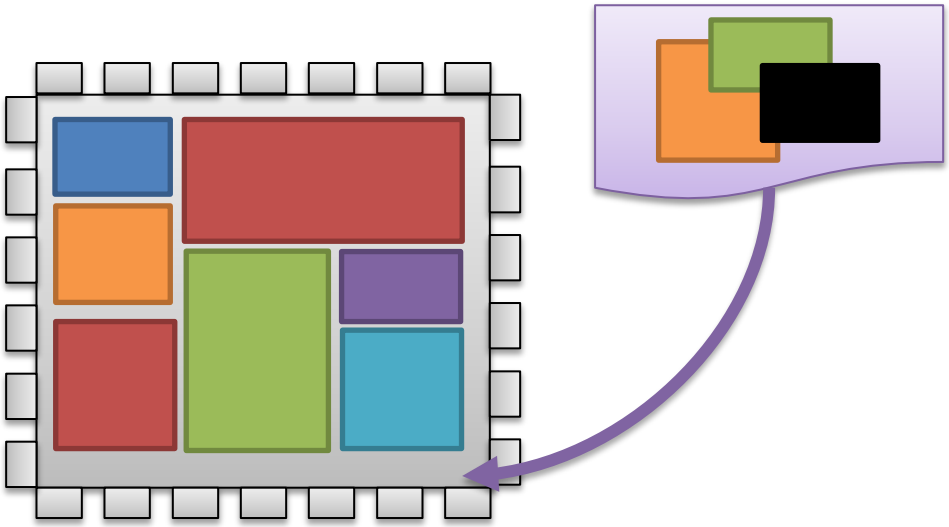
This year's topic:

Smart cyber-physical edge systems

Learning by doing!

What is Reconfigurable Computing?

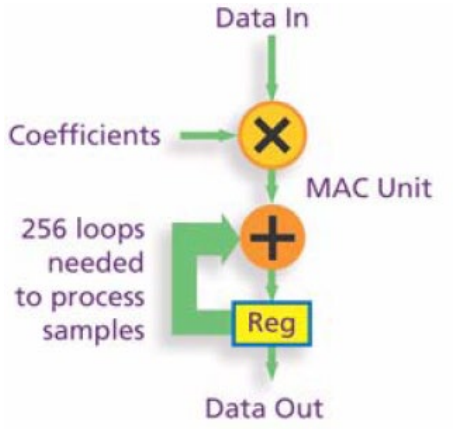
Reconfigurable Computing is a computer architecture combining some of the **flexibility** of software with the **high performance** of hardware by processing with very flexible high speed computing fabrics like field-programmable gate arrays (FPGAs). The principal difference when compared to using ordinary microprocessors is **the ability to make substantial changes to the data-path** itself in addition to the control flow.



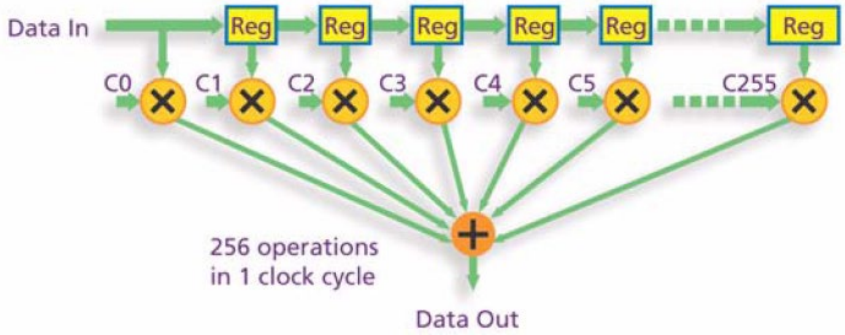
Performance in RC: Parallel Processing and Custom Computing

Consider a convolution filter… $Y_{(n)} = \sum_{i=0}^{i=N-1} k_i . X_{(n-i)}$

Sequential Implementation (Temporal Mapping)

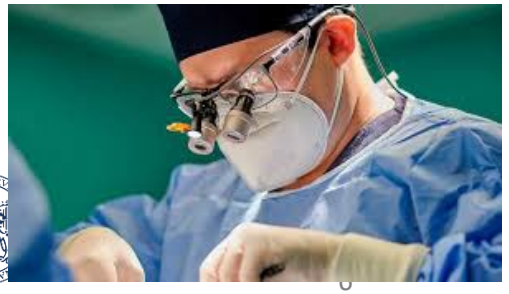


Parallel Implementation (Spatial Distribution)



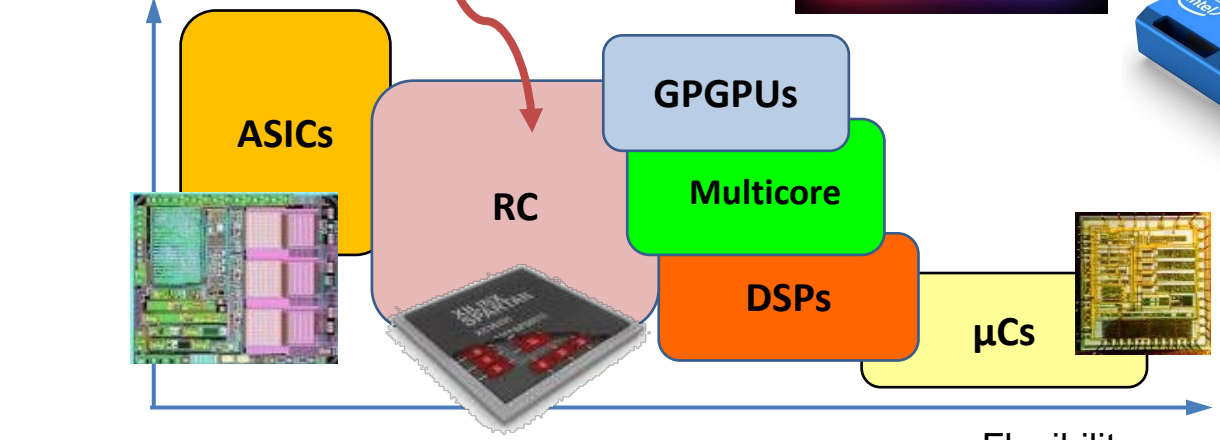
General Purpose Architectures are able to do everything

Custom computing architectures do few tasks but extremely well



Reconfigurable Computing for AI Acceleration

FPGAs/MPSoCs are able to provide adaptation to the workload



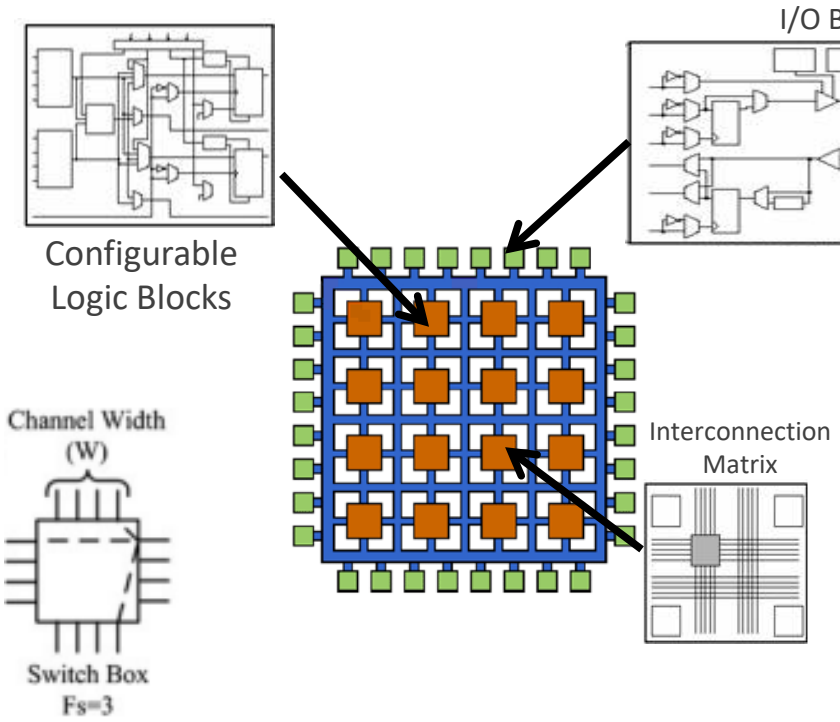
FPGA Architecture

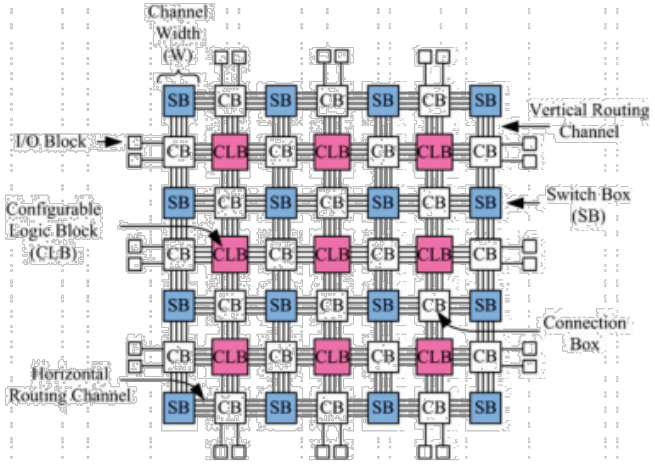
FPGAs are fine grain reconfigurable devices composed by logic elements with a functionality that can be modified as many times as needed.

CLB: Configurable Logic Block

CB: Connection Box

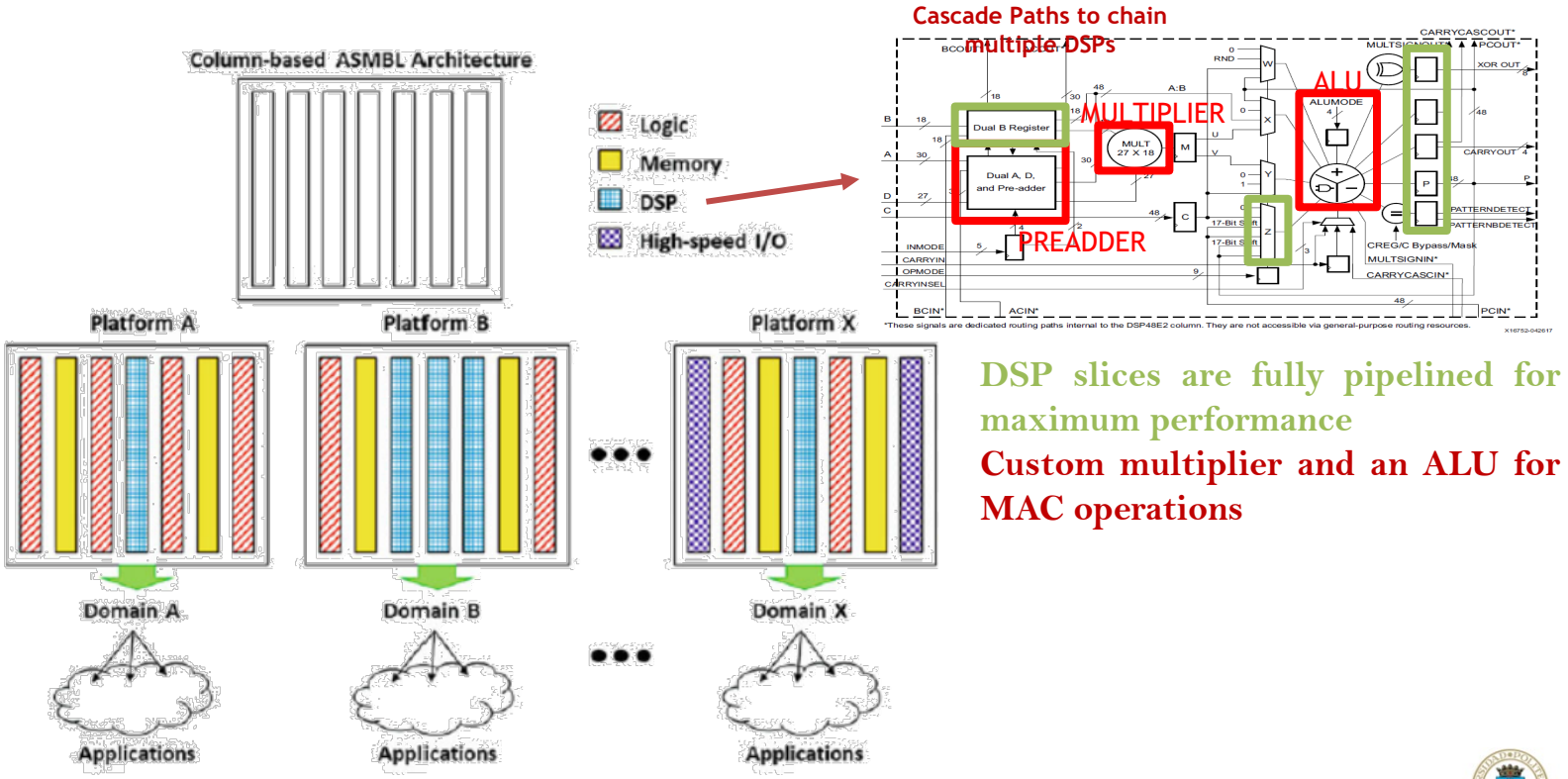
SB: Switch Box or Switch Matrix





FPGA Architecture: Heterogeneous Arrays

- To increase performance and flexibility, some reconfigurable devices have a layout with embedded heterogeneous resources, such as:
  - Custom DSP function blocks
  - Memory blocks to store frequently used data and variables on-chip for quick access, thanks to the proximity to the logic blocks that access it.

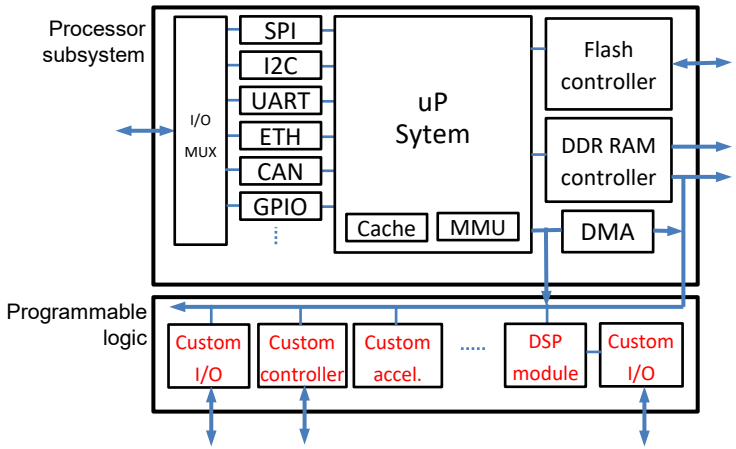


DSP slices are fully pipelined for maximum performance

Custom multiplier and an ALU for MAC operations

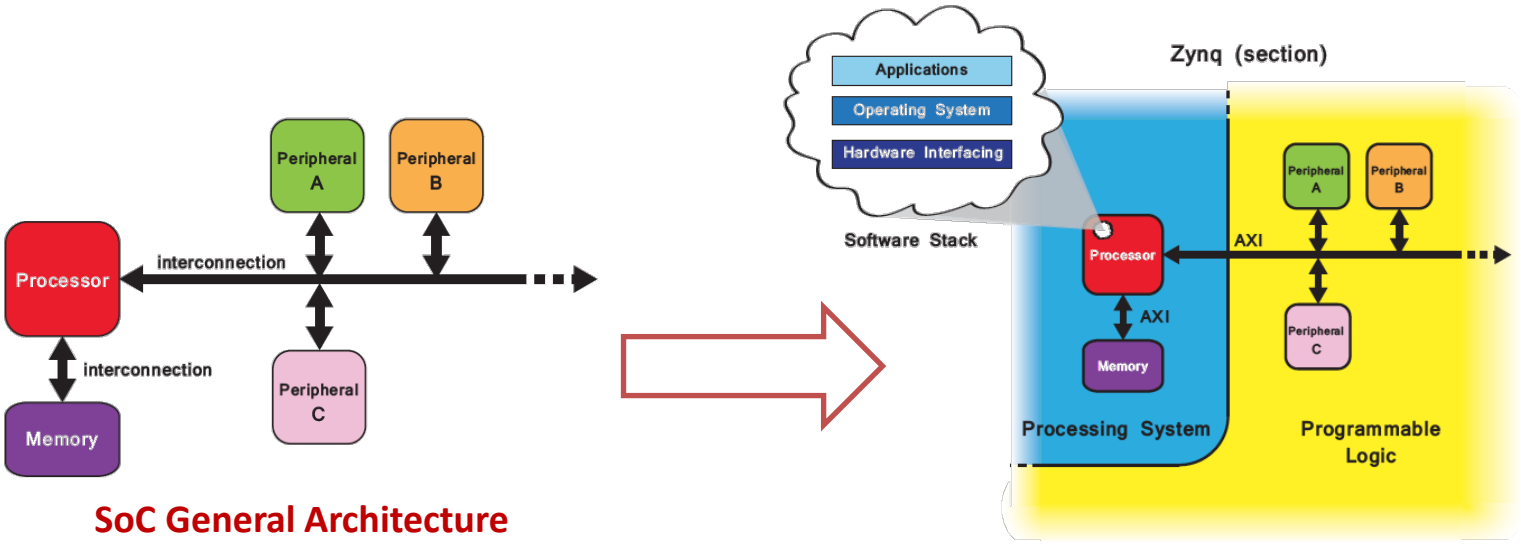
Mixed Software / Hardware Systems

- Best-of-both-worlds approach
  - Microprocessor(s) for sequential application execution
  - FPGA logic for massively parallel computation
- Different approaches
  - Traditional: System on Programmable Chip (SoPC)
    - Microprocessors embedded in FPGA logic
      - Hard cores: silicon with fixed functionality (PowerPC)
      - Soft cores: implemented using FPGA logic resources (MicroBlaze, Nios II)
  - Current: System on Chip (SoC) + FPGA
    - Silicon is split in ASIC-like CGRA and FPGA parts
    - Microprocessors can operate without enabling the FPGA



System-on-Programmable Chips (SoPCs)

System-on-Programmable-Chips (SoPCs)
Architectural concept resulting from the integration of 'hard' CPU cores with an FPGA in the same chip



SoC General Architecture

Embedded System Mapped on Zynq

"The PL section is ideal for implementing high-speed logic, arithmetic and data flow subsystems, while the PS supports software routines and/or operating systems"

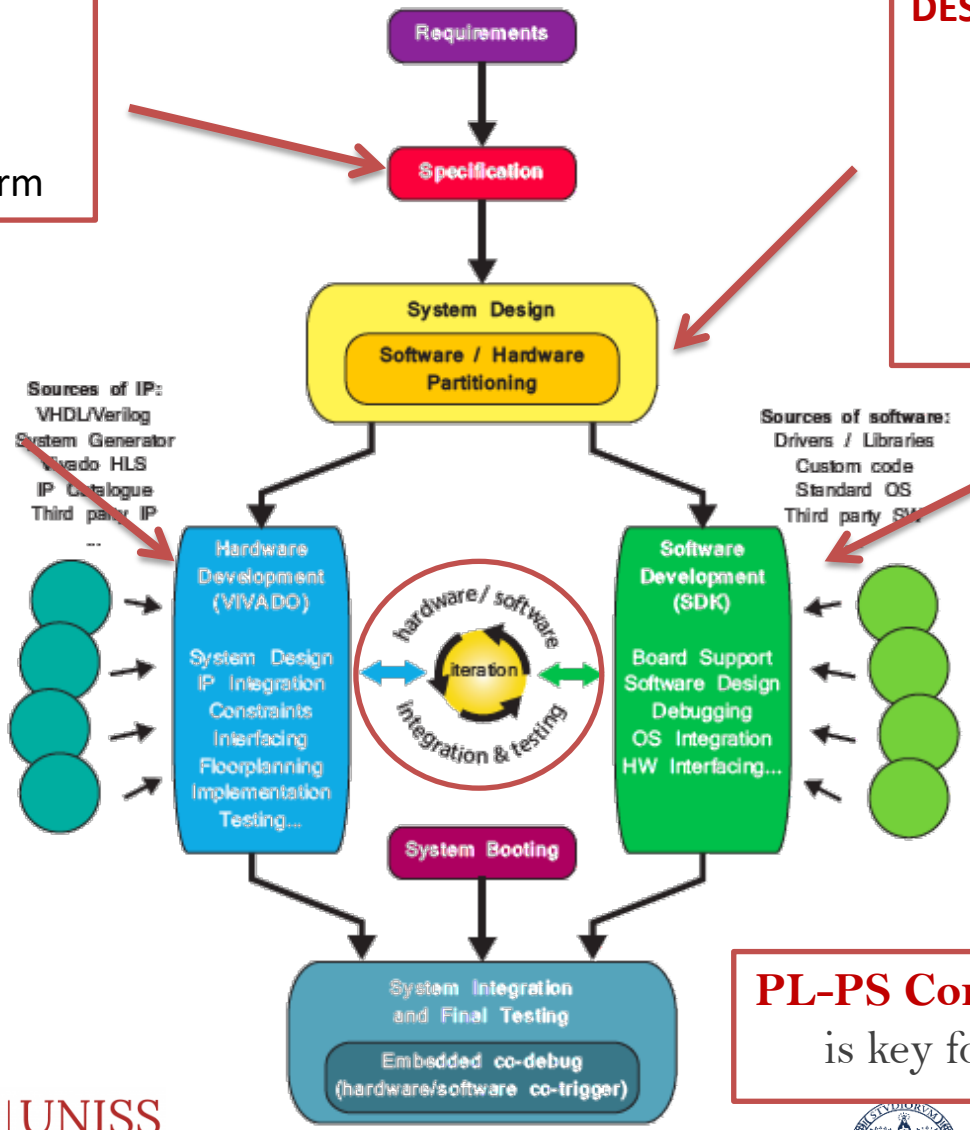
(from The Zynq Book)

SoC Based Design for Zynq (Vivado & SDK)

Functionality
Interfaces
Performance Criteria
Target Device or Platform

DESIGN ENTRY METHODS
VHDL, Verilog
System Verilog
C, C++
System C
Matlab
Simulink

PL: Computational Intensive Dataflow
Coprocessor model
Which tasks can be parallelized?



PS: SEQUENTIAL PROCESSING TASKS
Operating System?
User Application
GUI
Networking?

PL-PS Communication latency
is key for the performance

Creative Lab Components: AMD-Xilinx Kria KV260

Kria™ KV260 Vision AI Starter Kit

VISION READY
- Mult-Camera Support: Up to 8 interfaces
- 3 MIPI sensor interfaces, USB cameras
- Built-in ISP component
- HDMI, DisplayPort outputs

FLEXIBLE CONNECTIVITY
- 1Gb Ethernet
- USB 3.0 / 2.0

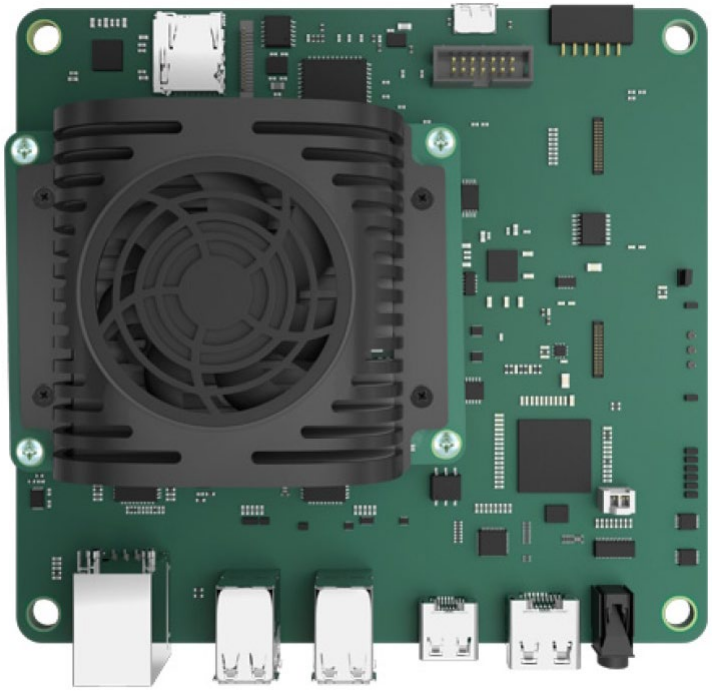
EXPANDABLE
- Extend to any sensor or interface
- Access Pmod ecosystem

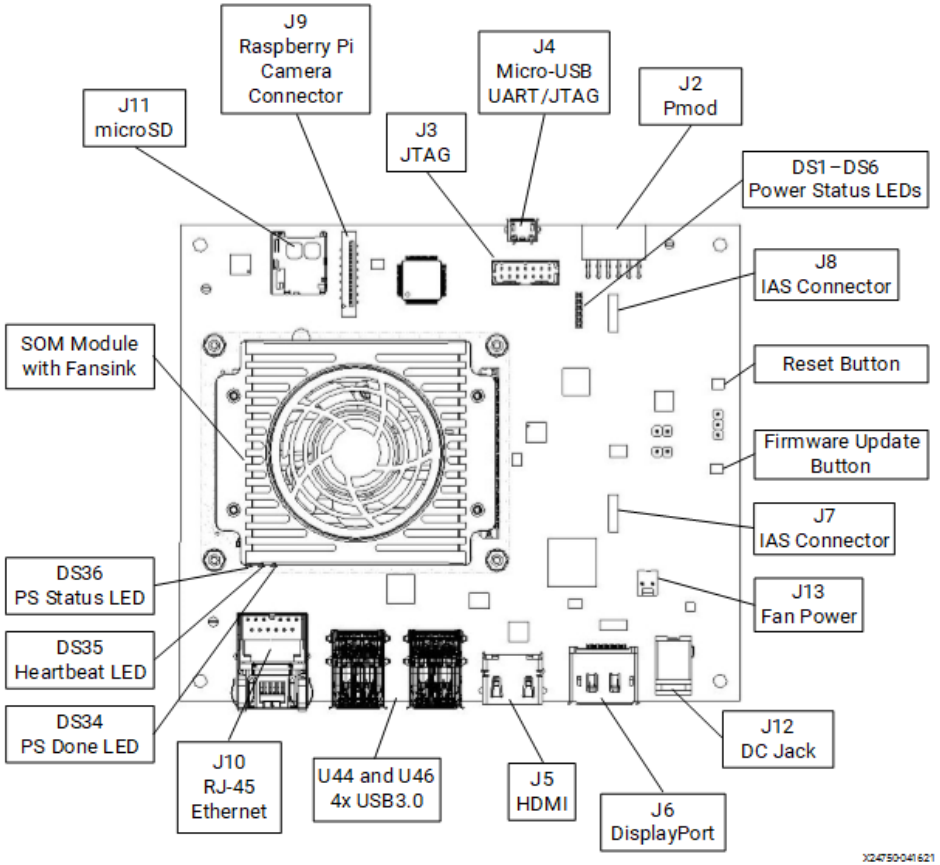
ACCESSIBLE
- Low cost, enabling design exploration
- Available from Xilinx and distributors

Creative Lab Components: AMD-Xilinx Kria KV260

Creative Lab Components: AMD-Xilinx Kria KV260

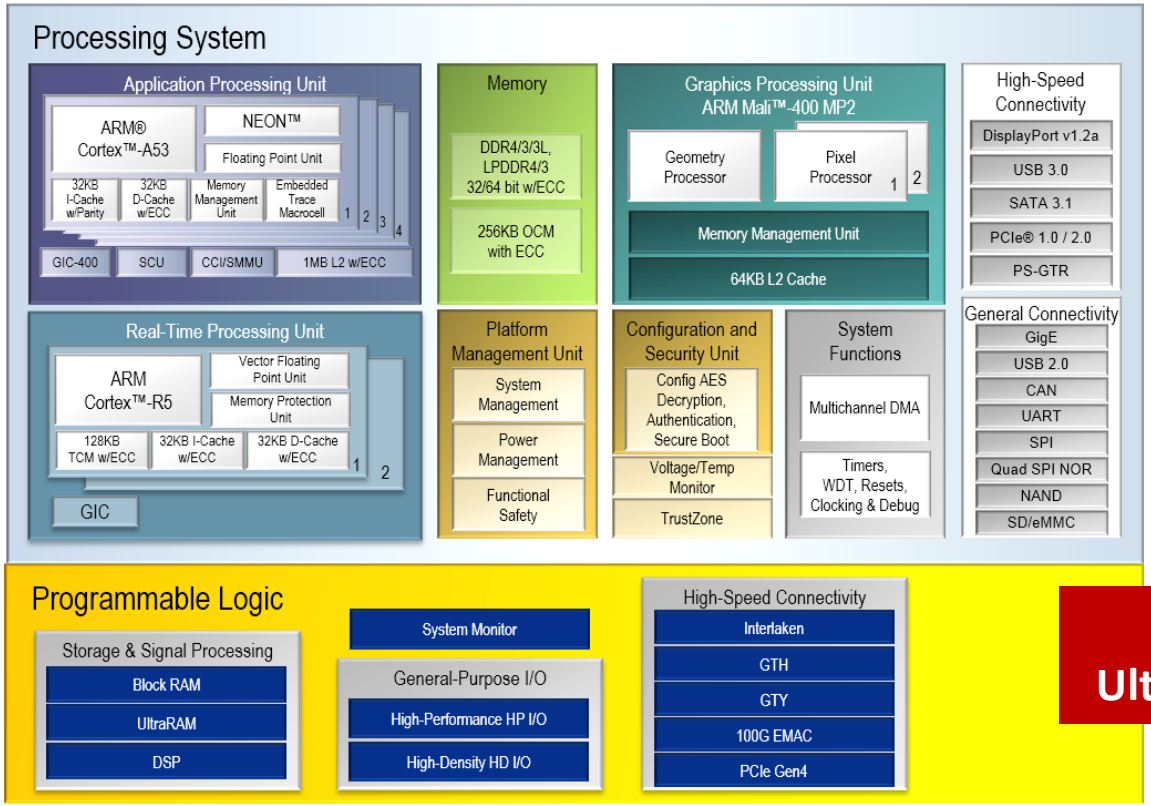
Kria™ KV260 Vision AI Starter Kit





Creative Lab Components: AMD-Xilinx Kria KV260

- *Zynq UltraScale+ EG*
  - Quad-core ARM® Cortex-A53 **processors** @1.5GHz
  - 16nm FinFET+ **FPGA**
  - Dual-core Cortex-R5 **real-time processors**
  - Mali-400 MP2 **GPU**



Xilinx Zynq UltraScale+ MPSoC

How to reach the board?

Open the terminal:

```
$ dmesg | grep tty
$ sudo putty /dev/ttyUSBXXXX -serial -sercfg 115200,8,n,1,N
$ ip a
```

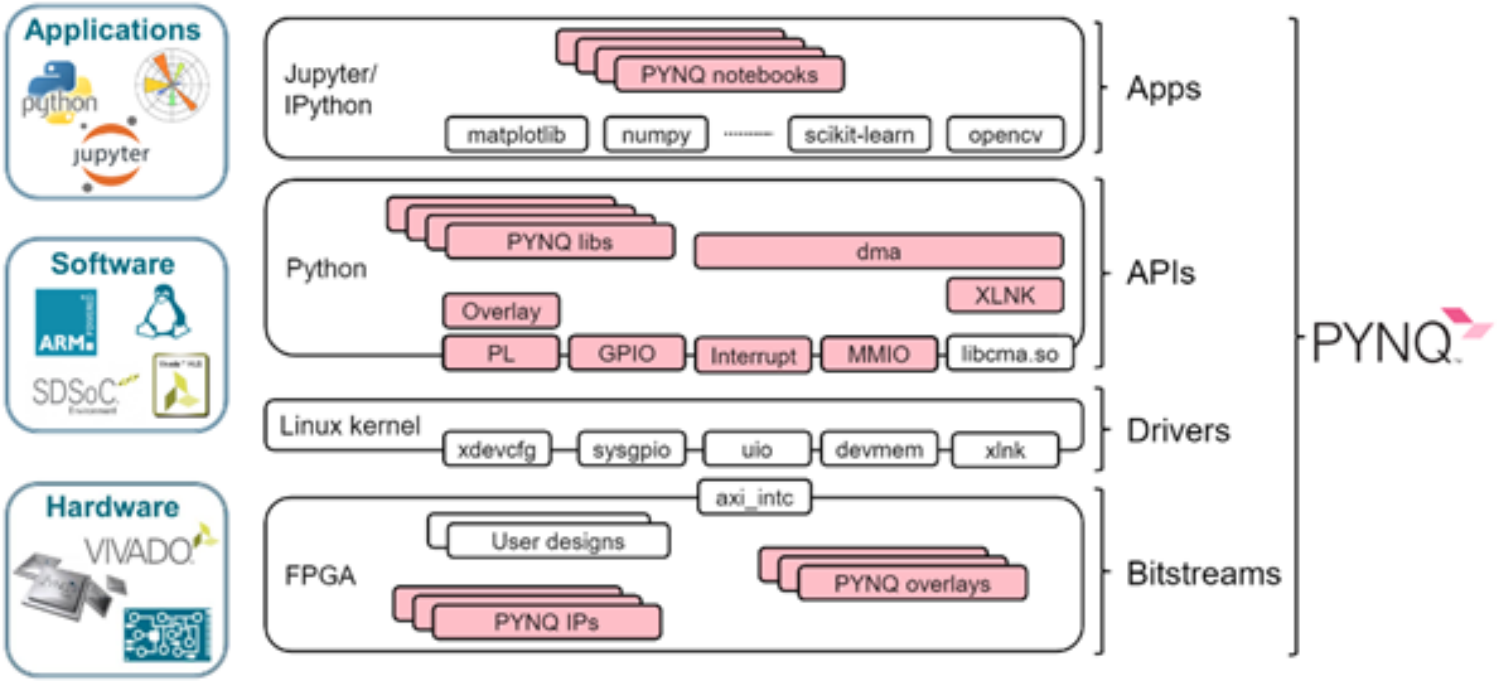
usr:pwd → ubuntu:alghero24

via SSH (discover ip first)

```
$ ssh ubuntu@<my_magic_ip>
```

Creative Lab Components: The PYNQ Framework (I)

PYNQ is an open-source project started by Xilinx, which fuses the productivity of Python with the acceleration provided by programmable logic within the Zynq / Zynq MPSoC.



Creative Lab Components: The PYNQ Framework (II)

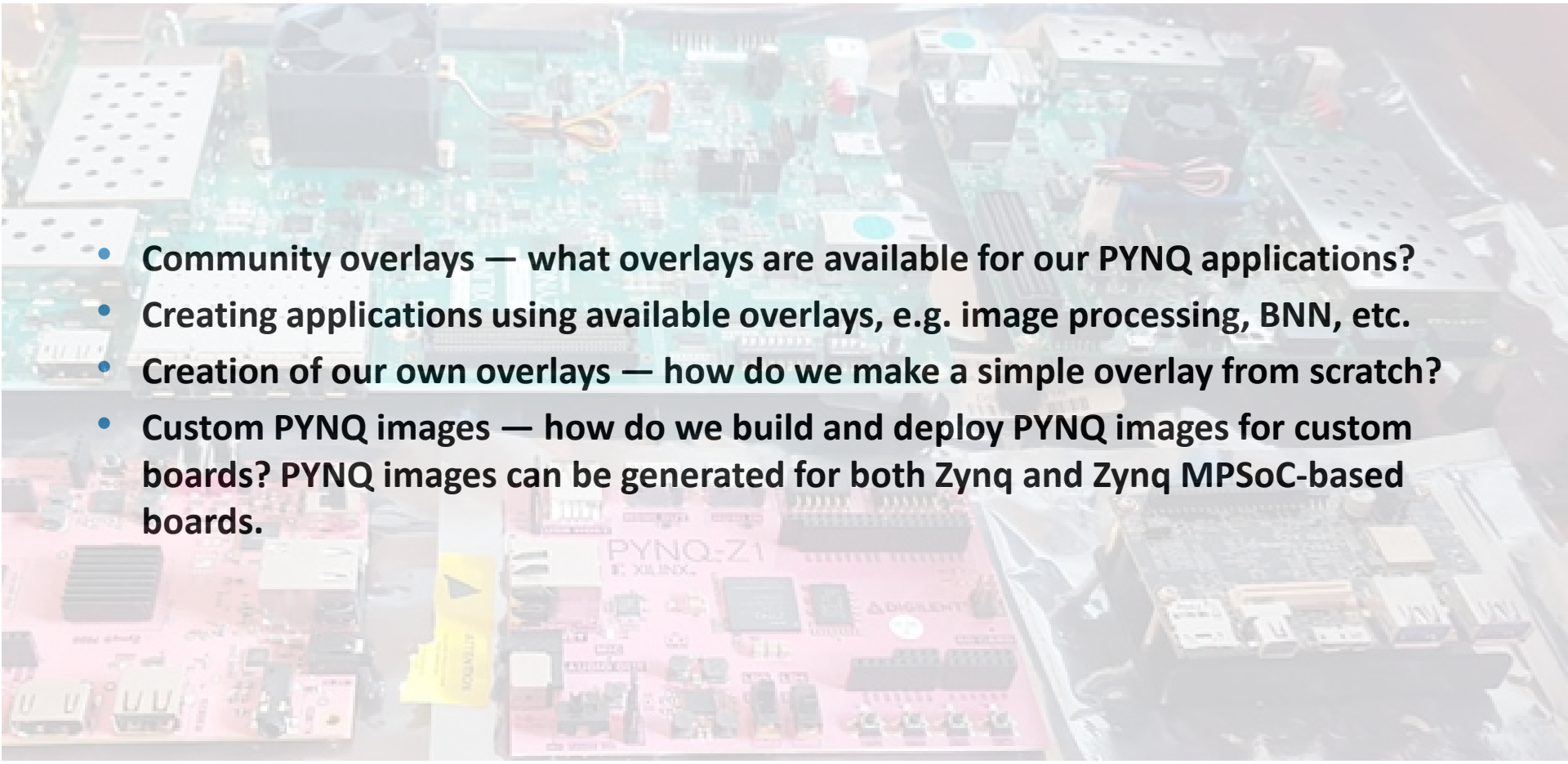
As developers, we leverage the PYNQ framework by connecting to the Jupyter server over a wired ethernet link and developing our application in a browser-based interface.

Now you can access **JupiterLab** via browser:

- <ip_address>:9090/lab or

The password is: **xilinx**

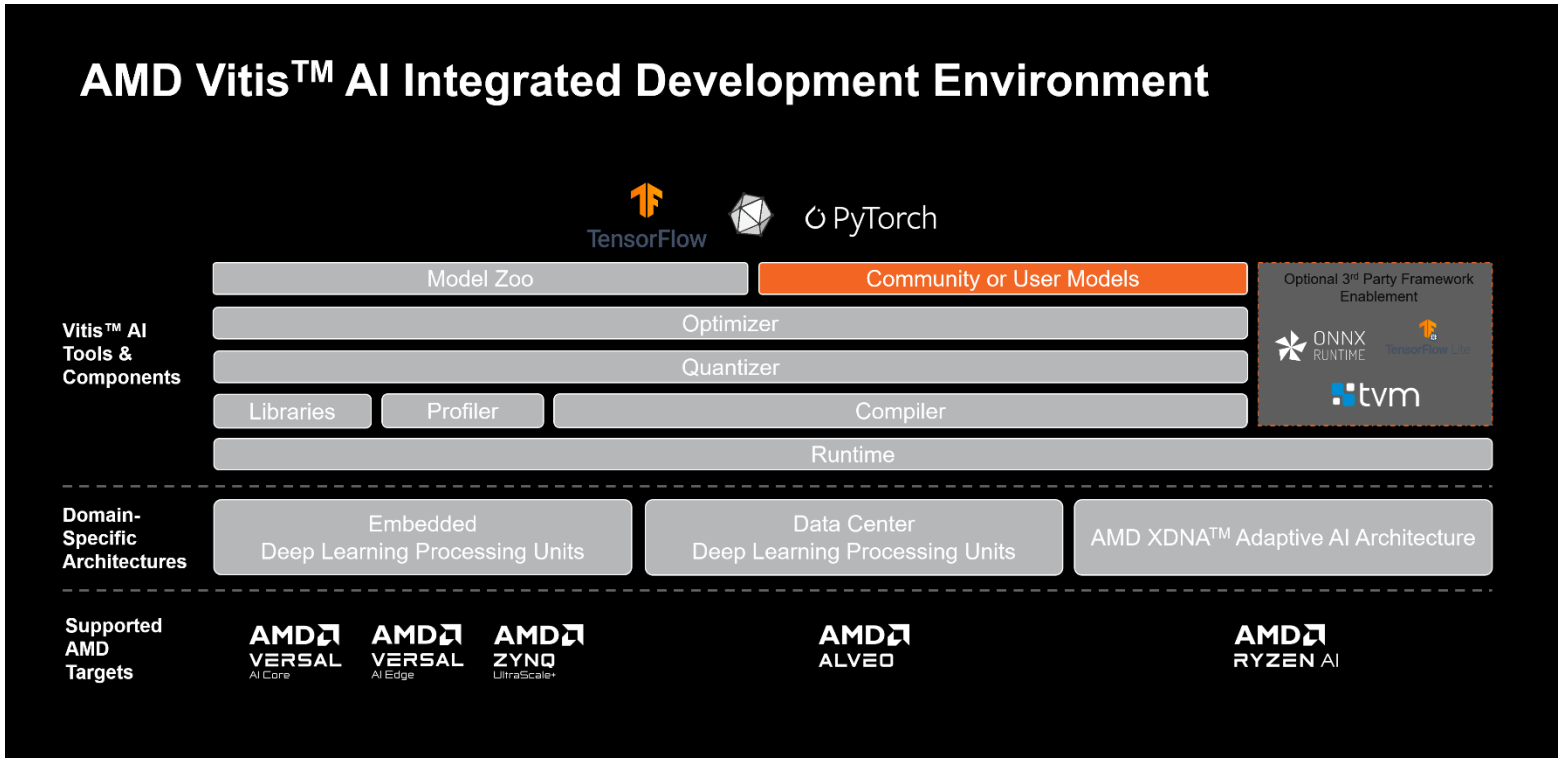
Creative Lab Components: The PYNQ Framework (III)



- Community overlays — what overlays are available for our PYNQ applications?
- Creating applications using available overlays, e.g. image processing, BNN, etc.
- Creation of our own overlays — how do we make a simple overlay from scratch?
- Custom PYNQ images — how do we build and deploy PYNQ images for custom boards? PYNQ images can be generated for both Zynq and Zynq MPSoC-based boards.

Creative Lab Components: Vitis AI

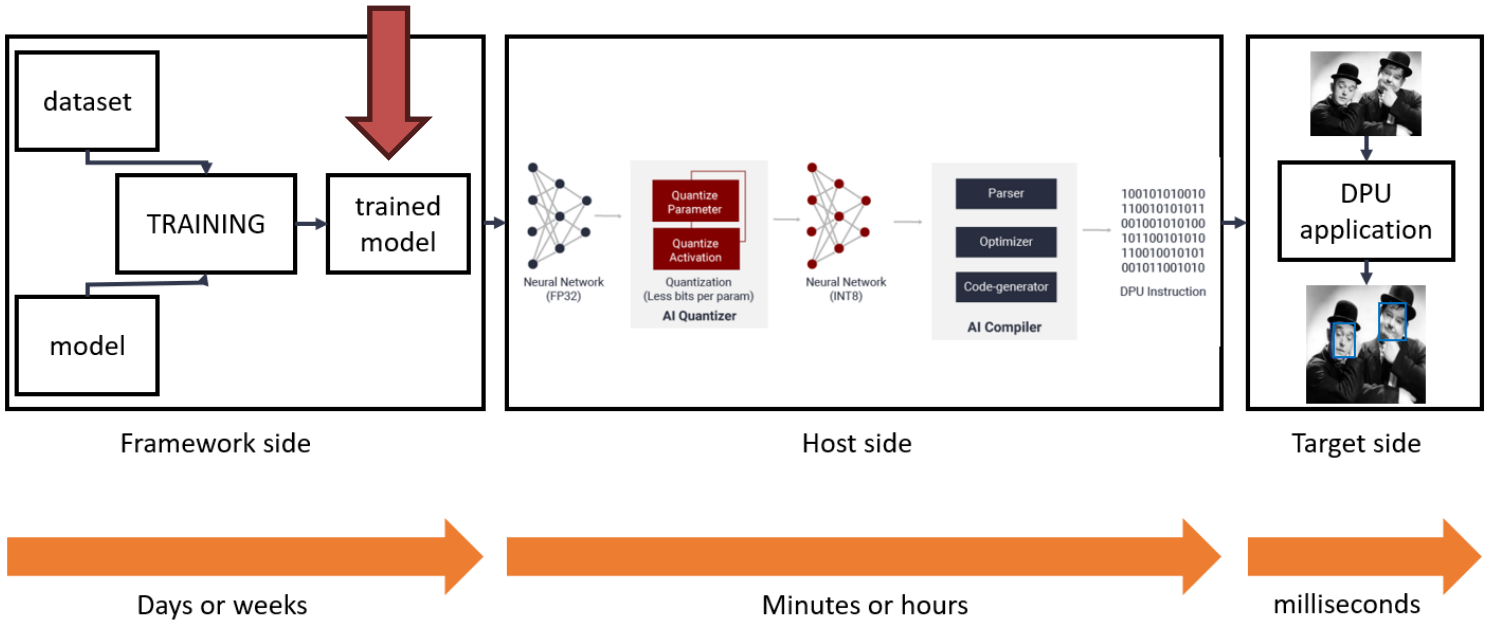
Adaptable & Real-Time AI Inference Acceleration



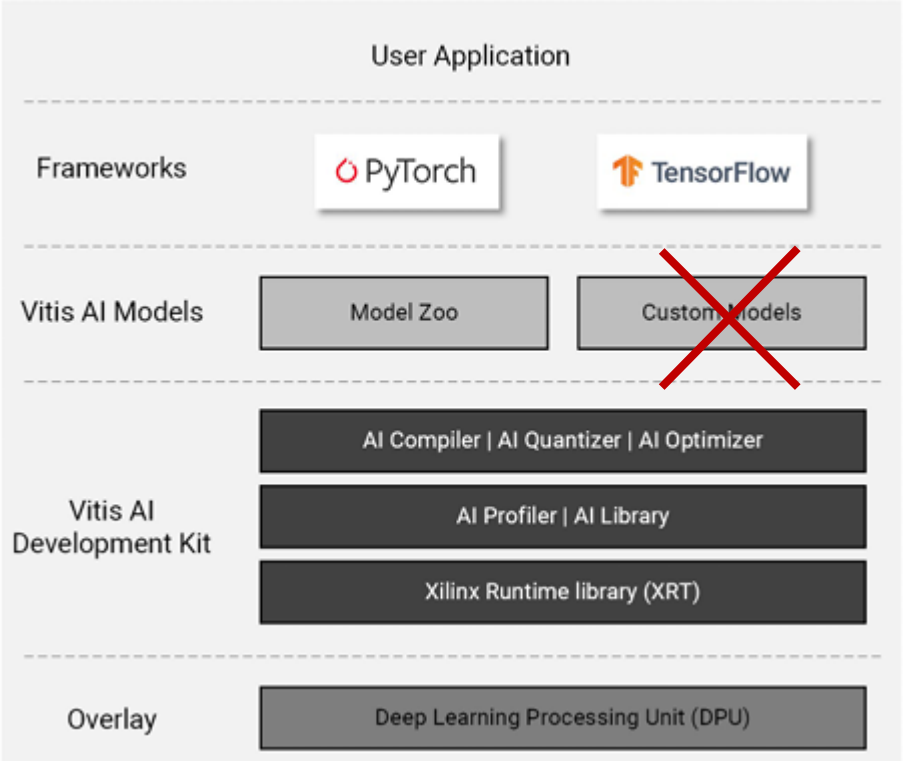
https://github.com/Xilinx/Vitis-AI

Creative Lab Components: Vitis AI

Entering here for the creative lab



Creative Lab Components: Vitis AI



Vitis AI Model Zoo

https://github.com/Xilinx/AI-Model-Zoo/

Rich Models in PyTorch, TensorFlow and ONNX

Open and Free on GitHub for All Developers

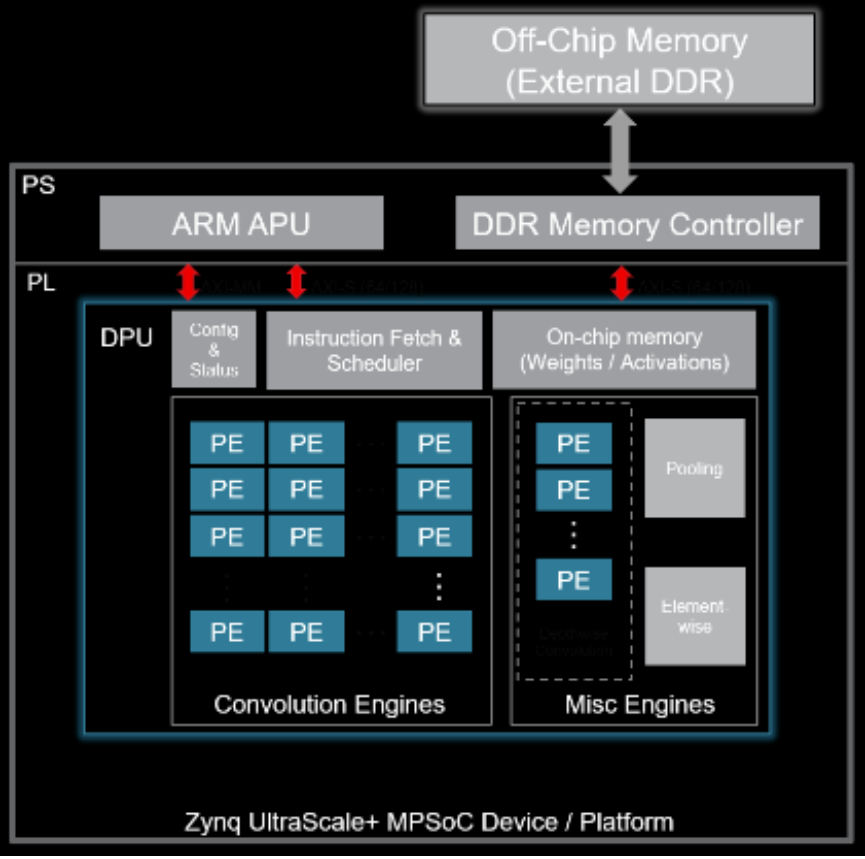
Advanced Optimization, Pruning Included

Retrainable with Custom Dataset

Creative Lab Components: DPU & run-time

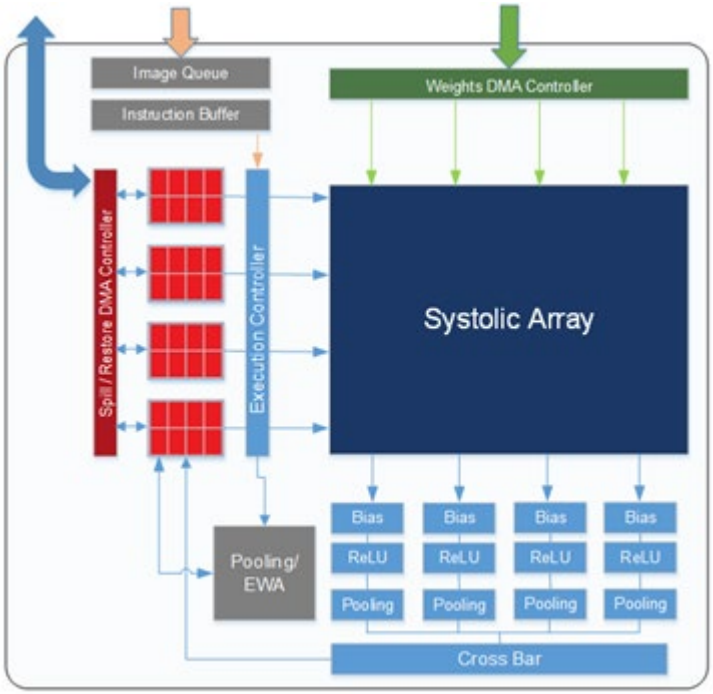
DPU for the Zynq® UltraScale+™ MPSoC

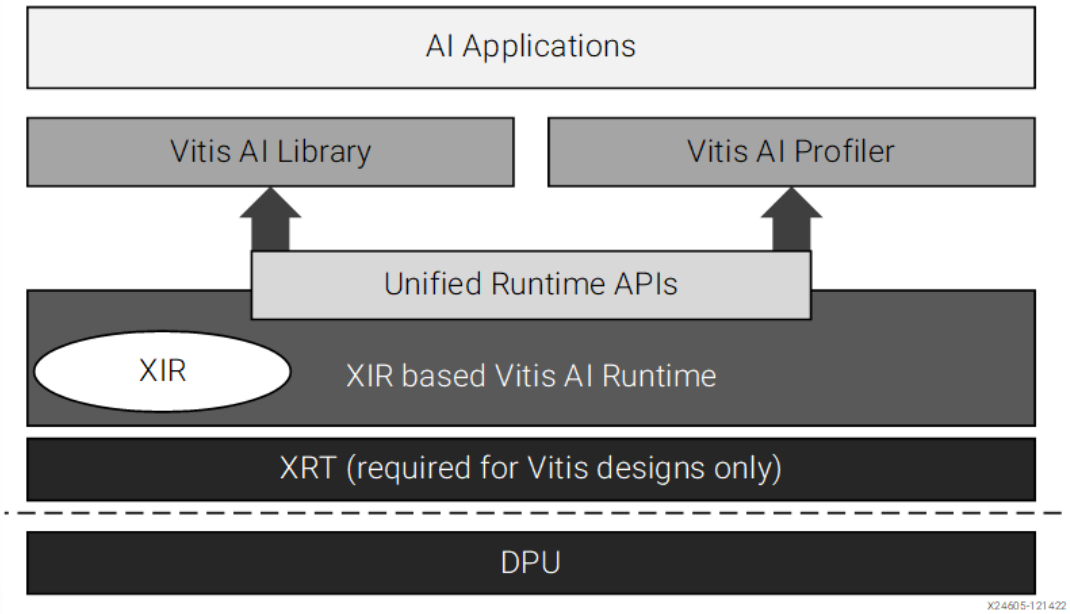
- "Deep learning Processing Unit"
  - "Matrix of Processing Engines" (MPE) architecture
    - Composed of DSPs, LUT, FFs, BRAM, UltraRAM (optional)
  - Soft-logic, micro-coded processor for INT8 deep learning inference
    - Graph compiled as micro-code, executed on DPU
      - Operator fusion, INT8 quantization, weights shuffled for optimized performance
  - Linux, QNX, and Integrity runtime enablement (QNX / Integrity support as POC, beta)
- Dedicated data movers and AXI stream interfaces for:
  - Instruction fetch
  - Weight and activation load/store
- Dedicated AXI memory-map interface for DPU configuration and status register access
- Global memory pool for buffering of intermediate activations
  - Optimizes and minimizes DDR bandwidth requirements
- Scalable from low to high parallelism (OPs)
  - 1 to 4 DPU cores per target device
  - Size DPU parallelism to fit available logic
  - Configurations for low / high DSP, RAM utilization, UltraRAM



https://github.com/Xilinx/DPU-PYNQ

Creative Lab Components: DPU & run-time





https://github.com/Xilinx/DPU-PYNQ

Creative Lab Scheduling 2024

	Monday	Tuesday	Wednesday	Thursday	Friday
	Introduction	*Modelling and Design*	*Architectures*	*Security on CPS*	*Creative Lab Day*
8:15-8:50	Registration and Opening				
8:50-9:00		Opening	Opening	Opening	Opening
9:00-10:30	Alberto Sangiovanni-Vincentelli (UC BERKELEY)	Danilo Pau (STM)	Daniel Newbrook (SOTON - ARM)	Elif Bilge Kavun (UNI PASSAU)	Creative Lab @ Work
10:30-11:00	Break	Break	Break	Break	Break
11:00-12:30	Paolo Azzoni (INSIDE)	STMicroelectronics Demo and Tutorial	MYRTUS project tutorial	SECURED project tutorial	Creative Lab @ Work
12:30-13:30	Lunch	Lunch	Lunch	Lunch	Lunch
13:30-14:00	Creative Lab Presentation	Creative Lab @ work		Creative Lab @ work	Creative Lab Pitch & Demo
14:00-15:30	CPS - WORKSHOP and Posters presentation	Jose Miranda (EPFL)		Thomas Kappelmann (INFINEON)	
15:30-16:20			Creative Lab @ work (15:00 - Till you'd like)		

Aula Badas

Aula Nivola

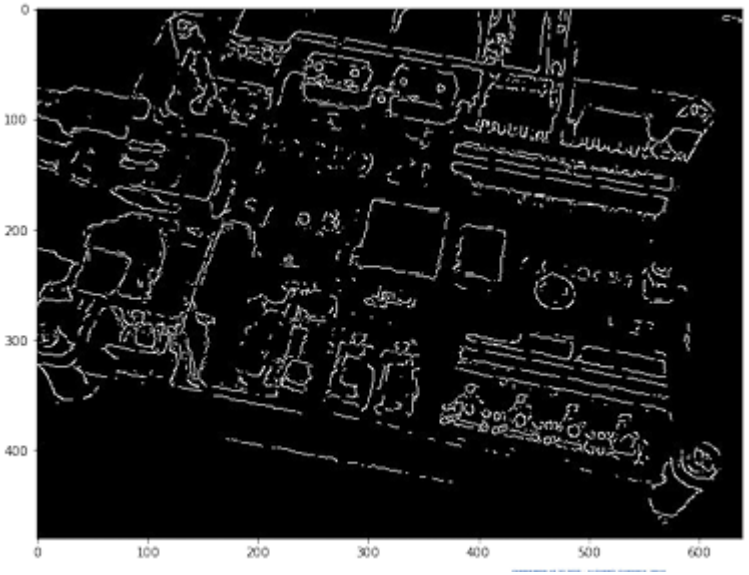
Aula Badas

Aula Badas

Aula Badas

A Demo

- AMD Demo based on USB Camera acquisition



https://xilinx.github.io/kria-apps-docs/kv260/2022.1/build/html/index.html

https://github.com/Xilinx/PYNQ/blob/master/boards/Pynq-Z1/base/notebooks/video/opencv_filters_webcam.ipynb


Creative Lab Teams

Creative Lab Teams

- **Red Team**
  - Antony Bartlett
  - Luca Martis
  - Emanuele De Luca
  - Zeinab Sadat Rabani

- **Green Team**
  - Daniele Nicoletti
  - Philip Wiese
  - Luka Macan
  - Leonardo Picchiami

Creative Lab Teams

- **Blue Team**
  - Benigno Ansanelli
  - Jiahong Bi
  - Georgios Tasopoulos
  - Afroz Mokarim
  - Alessandro Monni

- **Yellow Team**
  - Francesco Biondani
  - Marina Cordovilla
  - Marco Brohet
  - Mattia Tibaldi
  - Marco Fois

Creative Lab Teams

- **Purple Team**
  - Mario Libro
  - Angelos Dimoglis
  - Tom Slooff
  - Federico Manca

- **Orange Team**
  - Massimo Micolitti
  - Michalis Piponidis
  - Alberto Galassi
  - Gianluigi Ferrari

“Good” Creative Lab Projects

… decide “small” realistic demos.

Think big, address real problems, potential business ideas, but…

Useful Links

Setting Up the Board and Application Deployment¶
https://xilinx.github.io/kria-apps-docs/kv260/2022.1/build/html/docs/smartcamera/docs/app_deployment.html

Ready to use notebooks:
https://github.com/Xilinx/Kria-PYNQ/tree/main/kv260/base/notebooks/microblaze
https://github.com/Xilinx/Kria-PYNQ/tree/main/kv260/base/notebooks/video
https://github.com/Xilinx/DPU-PYNQ/tree/master/pynq_dpu/notebooks
https://github.com/Xilinx/PYNQ_Composable_Pipeline

Model Zoo for PYNQ-DPU based models (2.5 Vitis AI):
https://github.com/Xilinx/Vitis-AI/tree/2.5/model_zoo

Vitis AI Library
https://docs.xilinx.com/r/2.5-English/ug1354-xilinx-ai-sdk/Introduction

Vitis AI User Guide v2.5
https://docs.xilinx.com/r/2.5-English/ug1414-vitis-ai/Vitis-AI-Overview