CPS Creative Lab

Motivation, Technology, and Applications

Cyber-Physical Systems (CPSs) Summer School

Aula Nivola, Universitá di Sassari, Algero (SS)

September 18th, 2023

Alberto Zeni <alberto.zeni@polimi.it>
Davide Conficconi <davide.conficconi@polimi.it>

POLITECNICO MILANO 1863

SEPTEMBER 18-22 2023 – ALGHERO, SARDINIA, ITALY

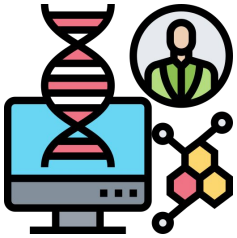
CYBER PHYSICAL SYSTEMS SUMMER SCHOOL 2023

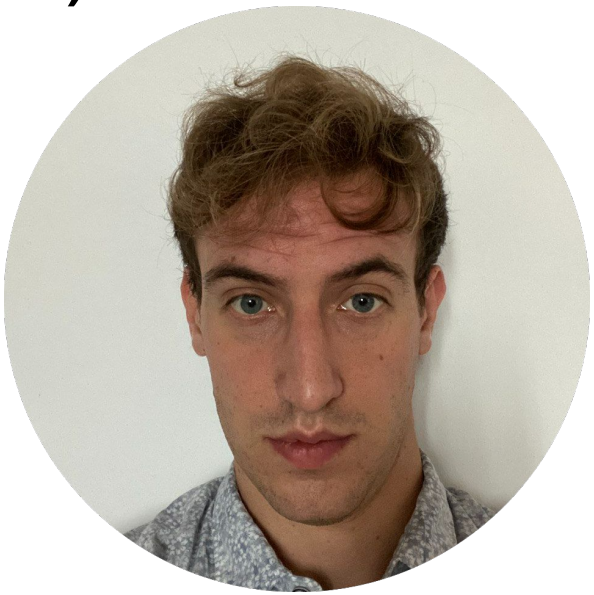
POLITECNICO MILANO 1863 NECST laboratory

Who We Are (1/2)

Ph.D. Candidate @Politecnico Di Milano

My research focuses on HPC applications and genomics

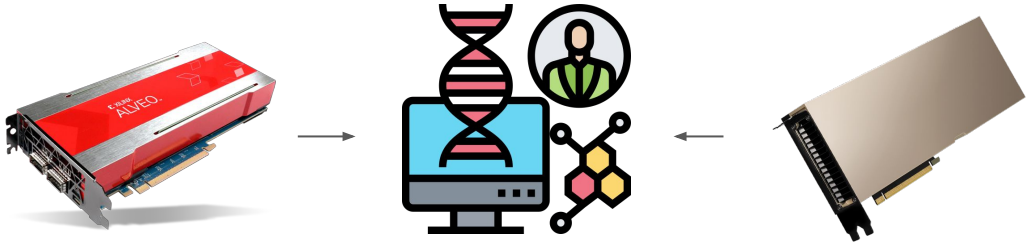




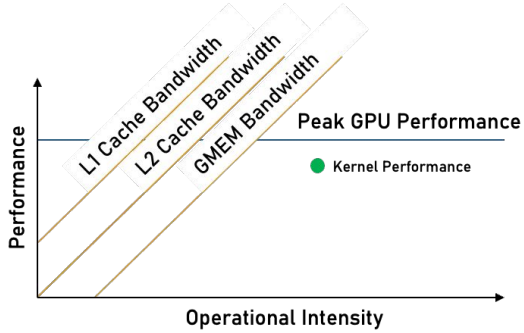
Who We Are (1/2)

Ph.D. Candidate @Politecnico Di Milano

My research focuses on HPC applications and genomics



- T.A. of for Computer Science 101, GPUs & Heterogeneous Systems
- Lecturer and T.A. of GPU101 Academy Passion In Action@Polimi



Intern/Visiting Student at

MIT/DFCI (‘23), NVIDIA(’22,’23), AMD Xilinx (’20,’21), LBNL(’19)

Dana-Farber Cancer Institute

Who We Are (2/2)

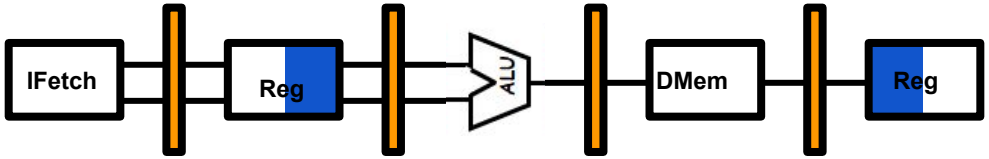
Post Doctoral Researcher @ Politecnico di Milano

Domain-Specific Reconfigurable Architecture Computer Organization

Design Methodologies **Automation** **Usability**

First steps on Quantum Computing

Adjunct Prof. for Bachelor CS101 and Advanced Computer Architecture



@ POLITECNICO MILANO 1863

Lecturer of FPGA101 Elective Course

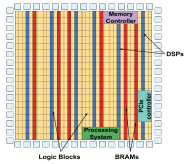


image: Flaticon.com

Who We Are (2/2)

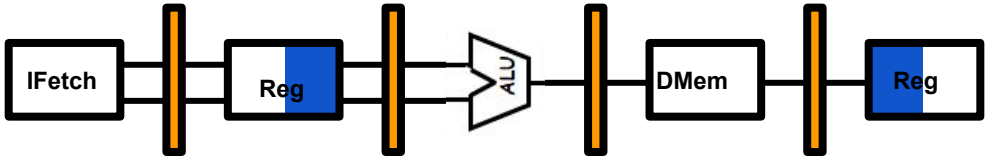
Post Doctoral Researcher @ Politecnico di Milano

Domain-Specific Reconfigurable Architecture Computer Organization

Design Methodologies **Automation** **Usability**

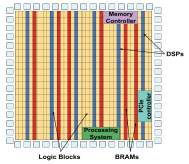
First steps on Quantum Computing

Adjunct Prof. for Bachelor CS101 and Advanced Computer Architecture



@ POLITECNICO MILANO 1863

Lecturer of FPGA101 Elective Course



Intern at research teams of IBM (21/22), Xilinx (18/19) Oracle (18)

IBM **Research** | Zurich XILINX ORACLE®

image: Flaticon.com

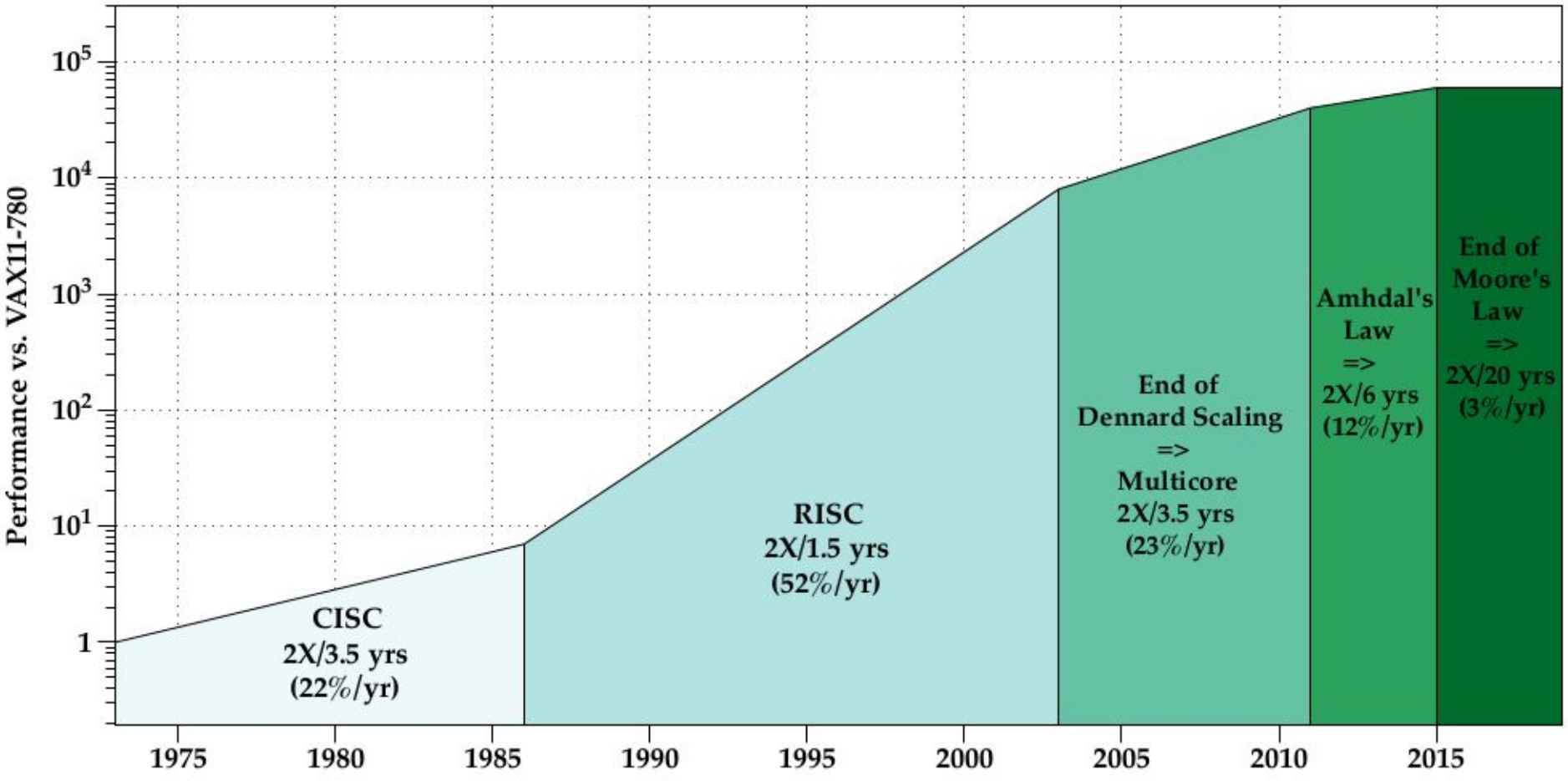
The Creative Lab at CPS 2023

It's 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.

Creative theme:

Smart Cyber-Physical Edge Systems

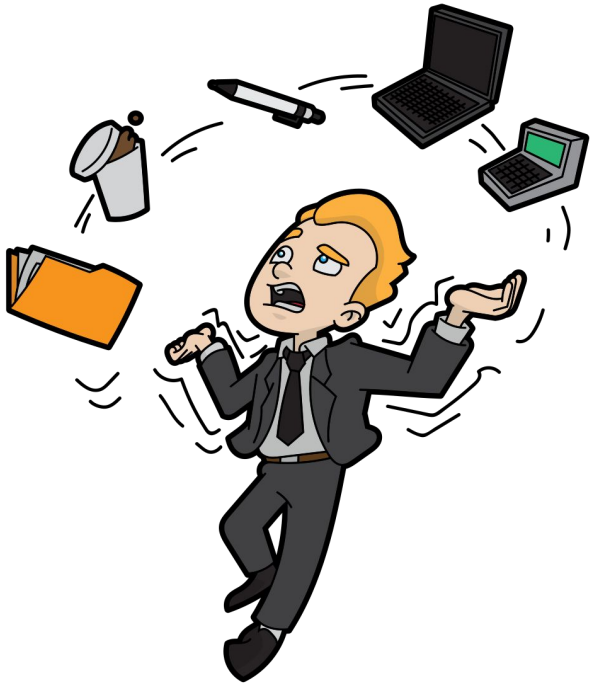
Creative Lab Theme: Motivations



Adapted from Lubos Sozzo. On how to effectively target fpgas from domain specific tools. 2019.
Data from: J. L Hennessy and D. A Patterson. Computer architecture: a quantitative approach 6th edition. Elsevier, 2018.

Generality vs Specialization in Computer Architectures

Do everything



High Flexibility

Less Performance

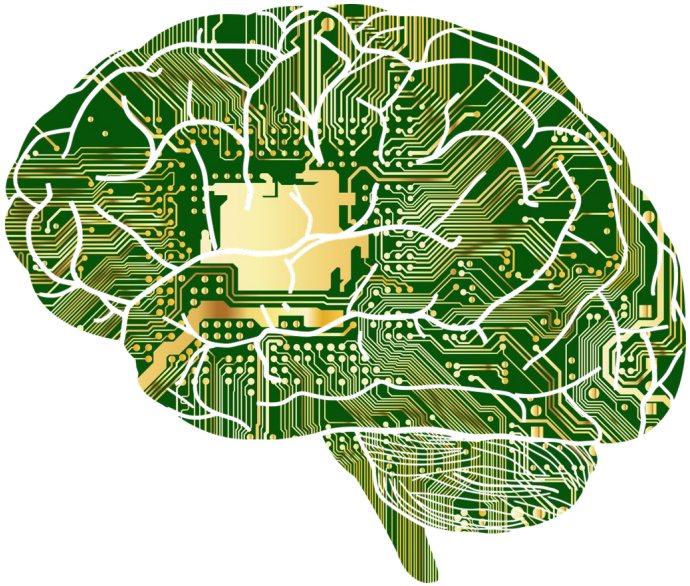
Do few tasks but extremely well

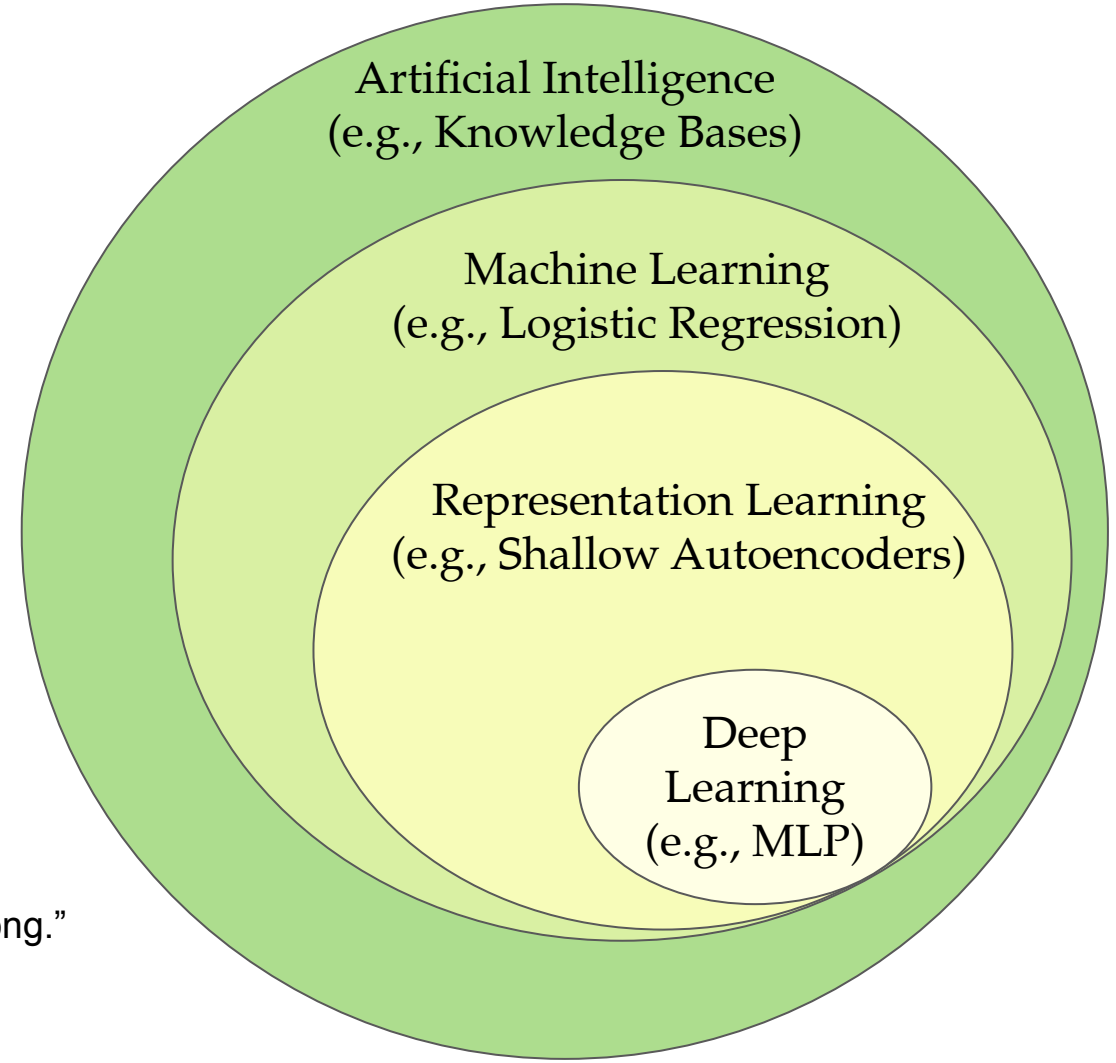
No Flexibility (frozen)

Extremely High Performance

Image source:
Report icons created by Freepik - Flaticon

The Case of Artificial Intelligence





"being too early is the same as being wrong."

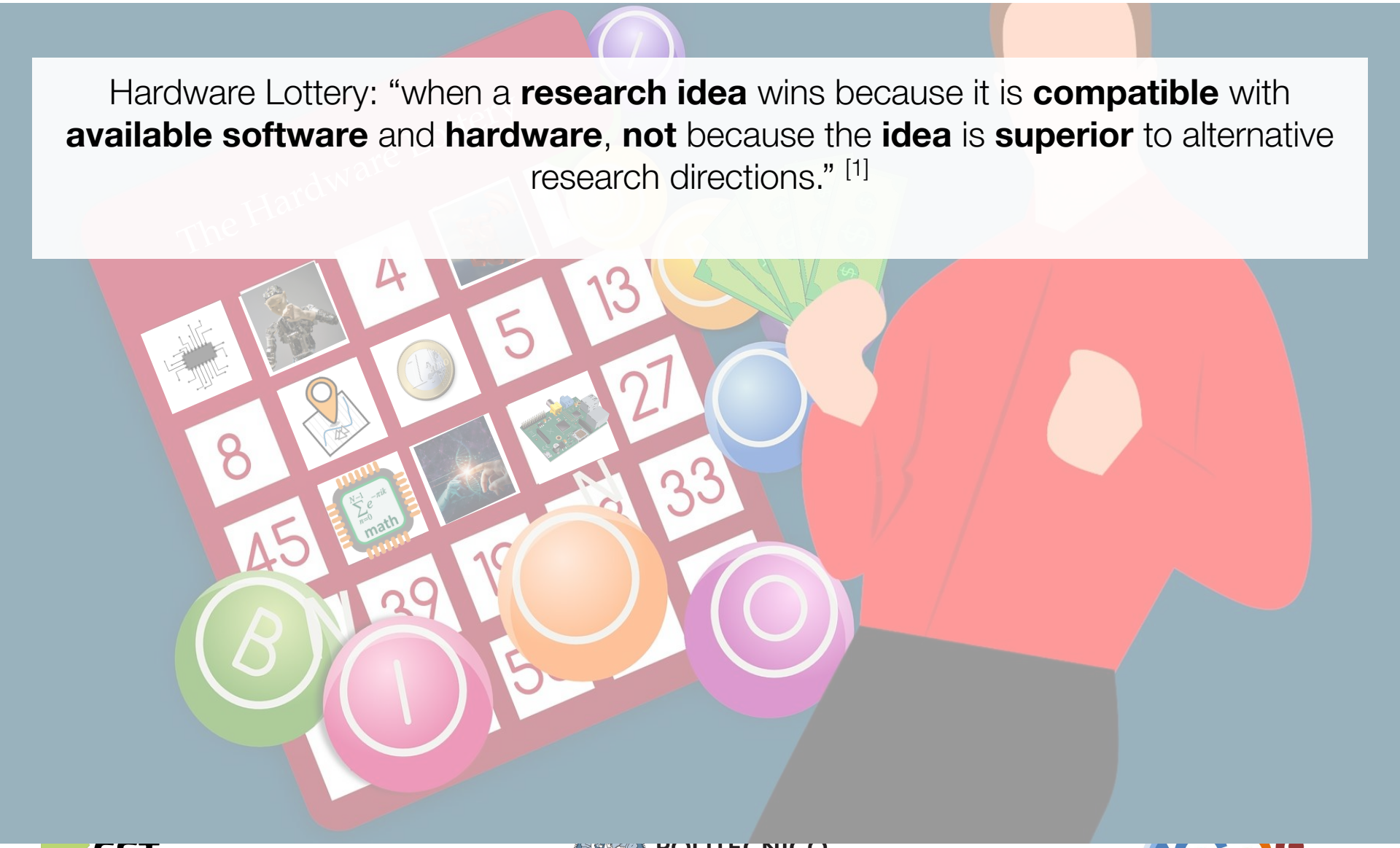
Images from https://commons.wikimedia.org/wiki/File:Anatomy-1751201_1280.png#/media/File:Anatomy-1751201_1280.png; https://www.deeplearningbook.org/contents/intro.html

The Hardware Lottery

[1] Hooker, Sara. "The hardware lottery." Communications of the ACM 64.12 (2021): 58-65.

The Hardware Lottery

Hardware Lottery: “when a **research idea** wins because it is **compatible** with **available software** and **hardware**, **not** because the **idea** is **superior** to alternative research directions.” [1]



[1] Hooker, Sara. "The hardware lottery." Communications of the ACM 64.12 (2021): 58-65.

The Hardware Lottery

Hardware Lottery: “when a **research idea** wins because it is **compatible** with **available software** and **hardware**, **not** because the **idea** is **superior** to alternative research directions.” [1]

“[...] a crucial **paradox**: **machine-learning researchers** mostly ignore **hardware despite** the **role** it plays in determining which ideas succeed.” [1]

“[...] Hardware design has **prioritized** delivering on **commercial use** cases, while built-in **flexibility** to accommodate the **next generation** of ideas remains a **secondary** consideration.” [1]

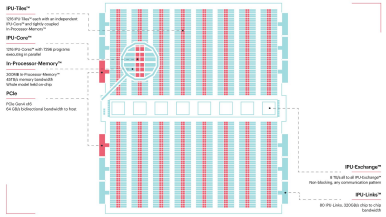
“Any attempt to avoid future hardware lotteries must be concerned with making it **cheaper** and **less time consuming** to explore **different hardware/software/algorithm** combinations.” [1]

[1] Hooker, Sara. "The hardware lottery." Communications of the ACM 64.12 (2021): 58-65.

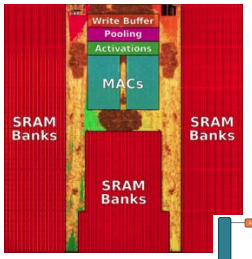
Some AI Accelerators

Neuromorphic (e.g., IBM TrueNorth)

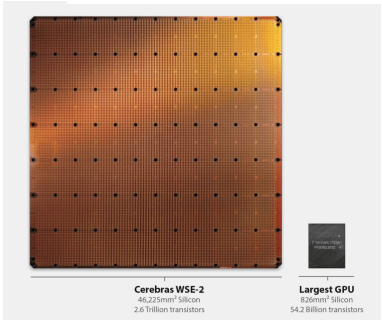
Graphcore Intelligent Processing Unit (IPU)



Tesla Neural Processing Unit (NPU) Chip

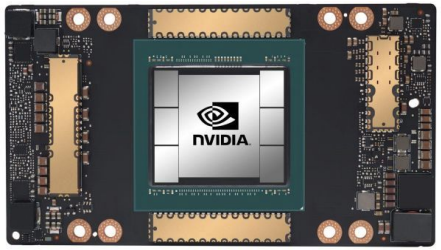


Cerebras Wafer-Scale Engine (WSE)

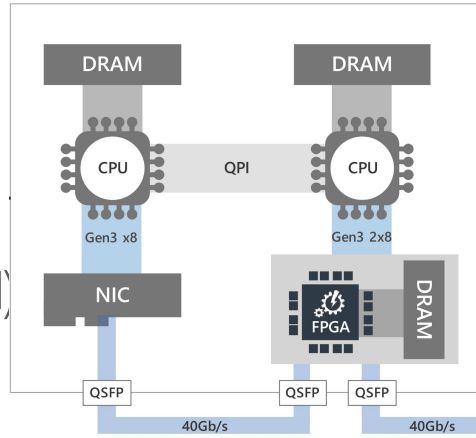


Sambanova Dataflow Engine

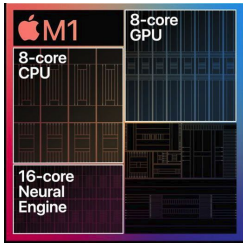
GPGPUs



MSR Brainwave NPU (FPGA-based)

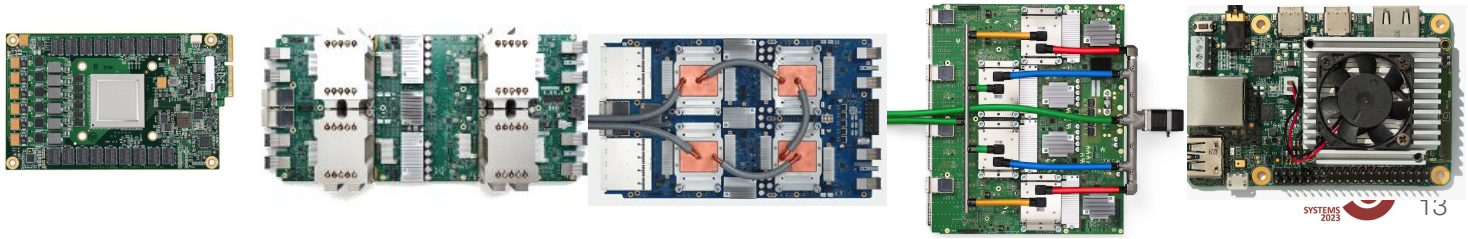


Apple M1 Neural Engine



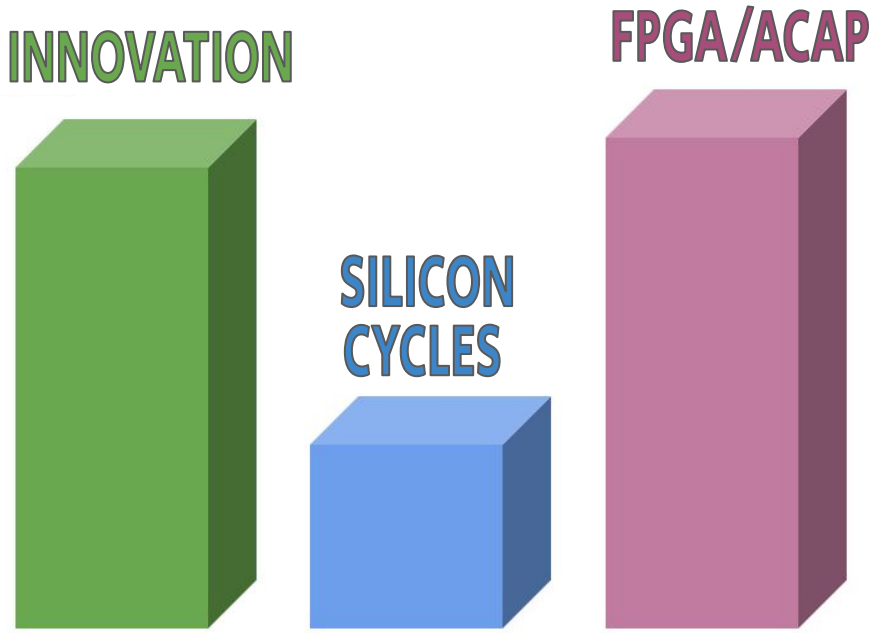
Google Tensor Processing Unit (TPU)

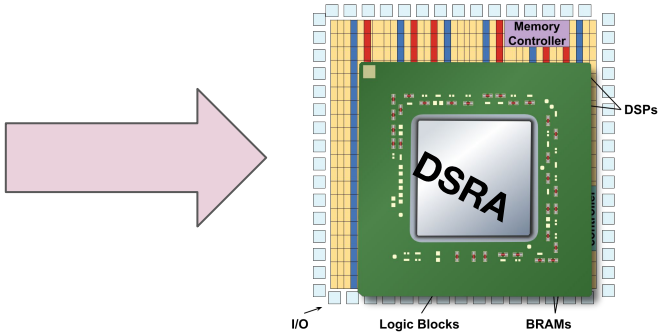
v1 v2 v3 v4 Edge



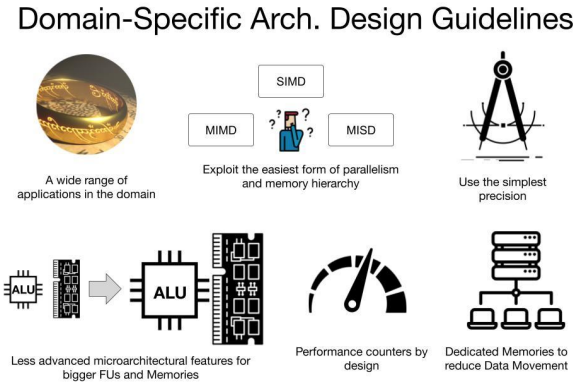
Images from https://en.wikichip.org

Evolving Rapidly → Adapt to Workload





SW-Programmable DSRA



Images from https://github.com/Xilinx/Vitis-AI; Xilinx Website

Field-Programmable Gate Arrays (FPGAs)

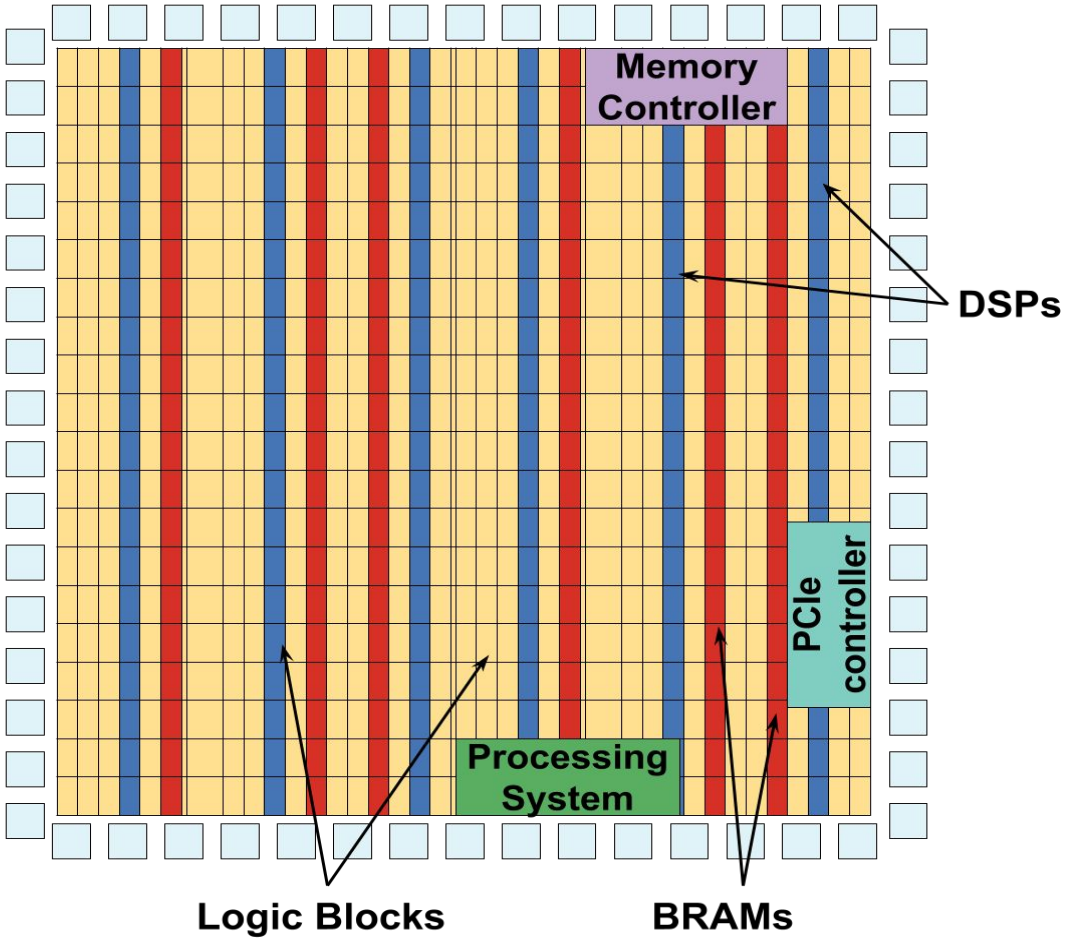
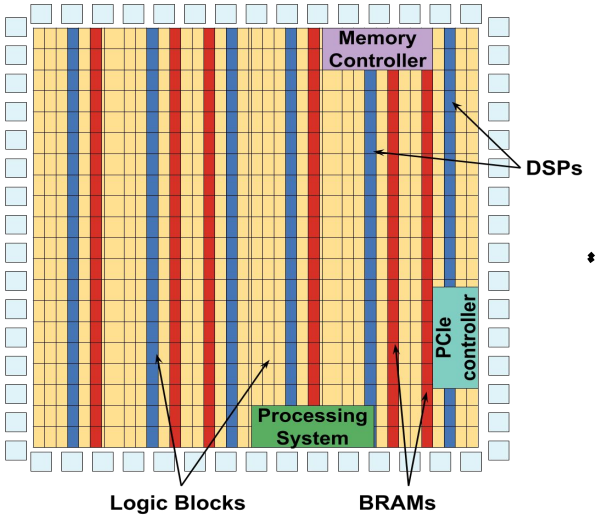


Image source:
Report icons created by Freepik - Flaticon

Some AI Applications for FPGAs[2]

Healthcare



Space

Video Surveillance

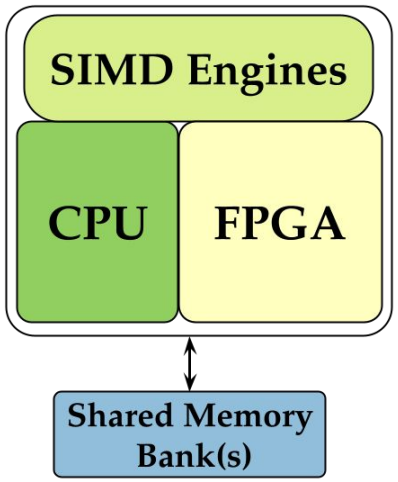
Particle Accelerator

Autonomous Driving

[2] https://www.xilinx.com/registration/subscribe/ai-ebook-signup.html

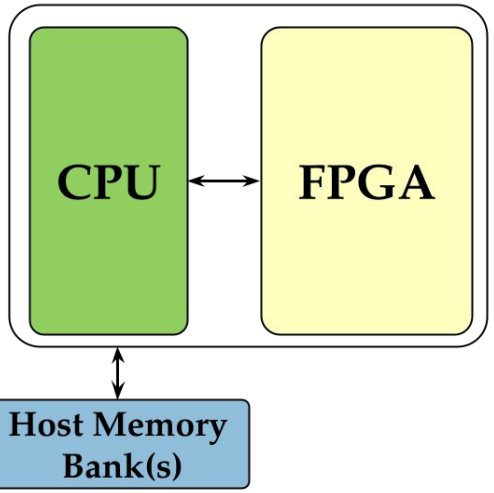
Image source: Report icons created by Flaticon.com

Reconfigurable Systems: a (possible) Taxonomy



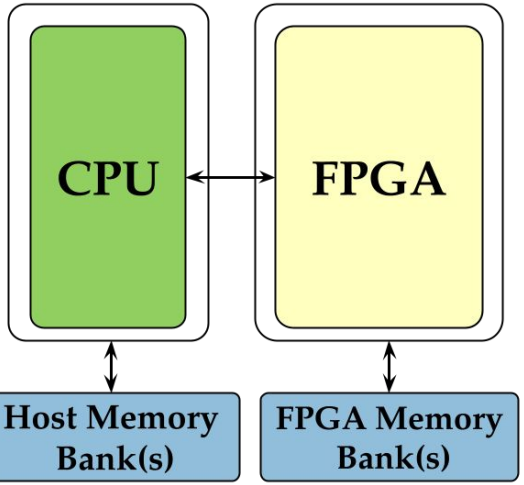
System on Chip

Ⓐ



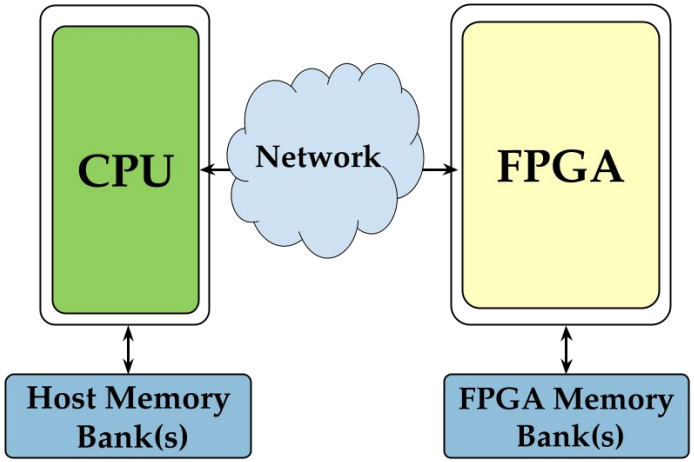
System in Package

Ⓒ



Loosely-coupled System

Ⓑ



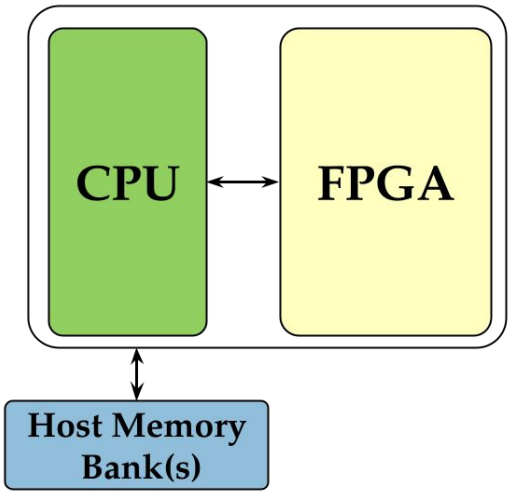
Network Attached

Ⓓ

Reconfigurable Systems: a (possible) Taxonomy

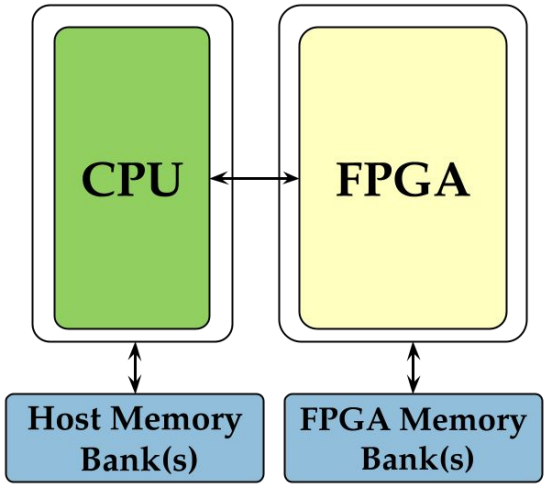
System on Chip

Ⓐ



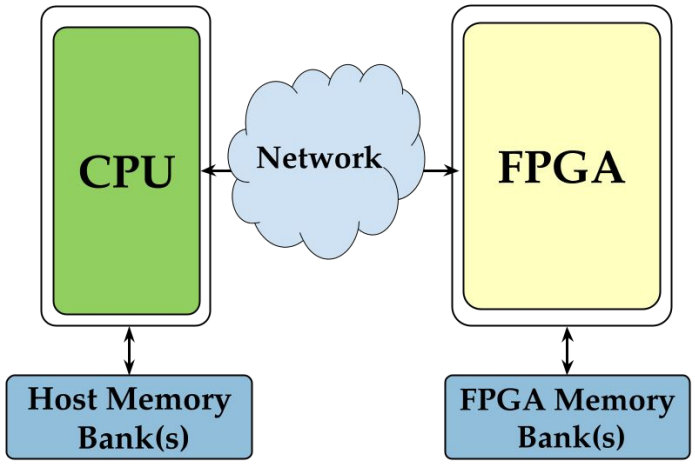
System in Package

Ⓒ



Loosely-coupled System

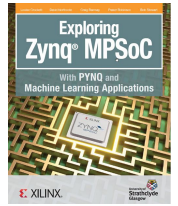
Ⓑ

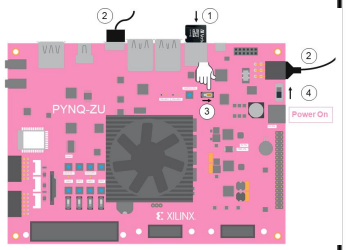


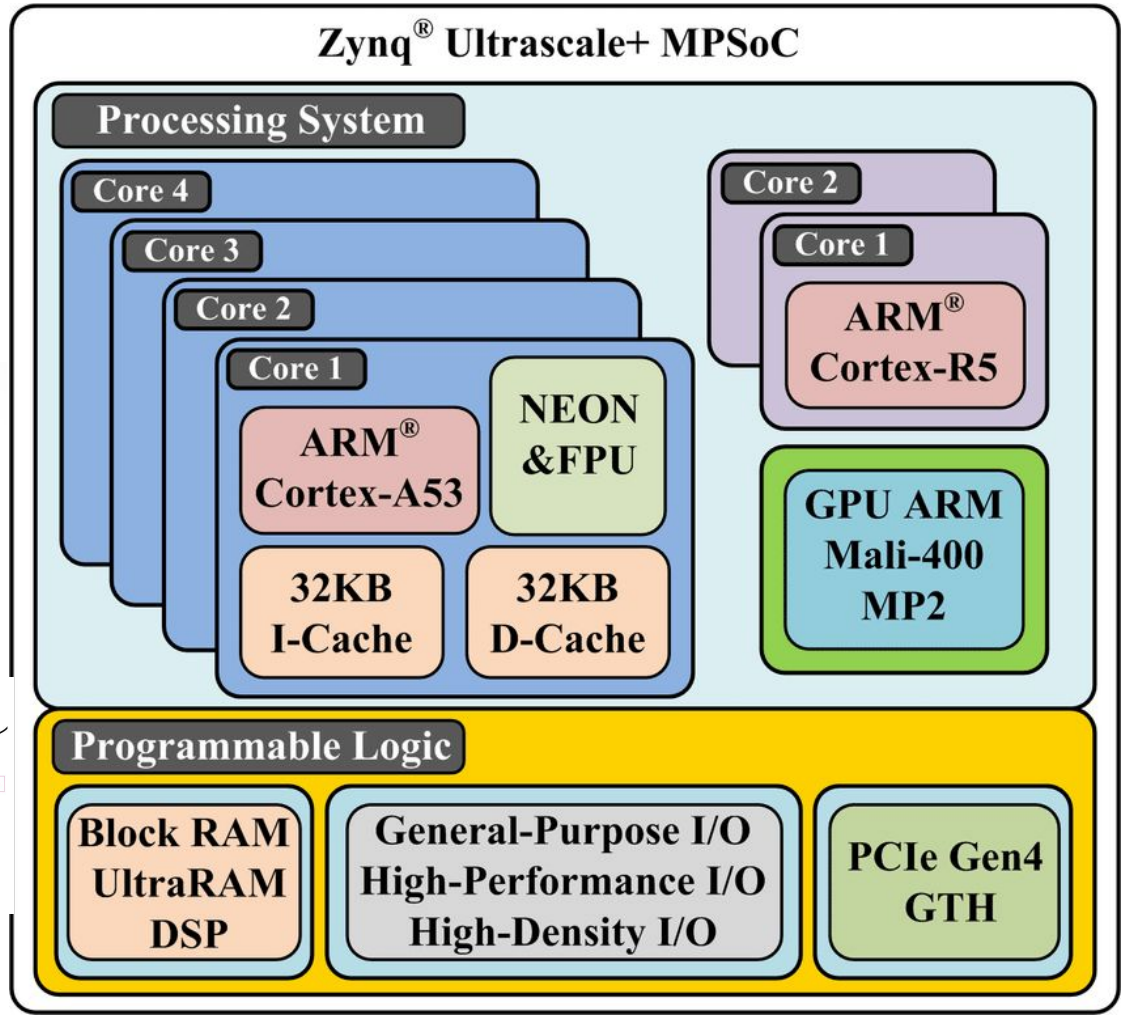
Network Attached

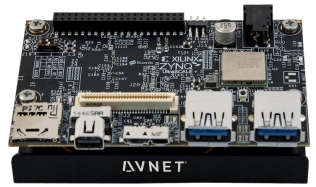
Ⓓ

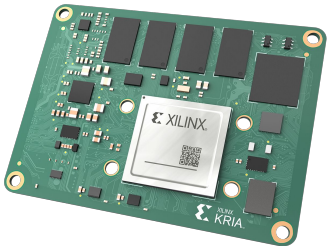
Creative Focus: ZYNQ MPSoC











Creative Essentials (1): 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

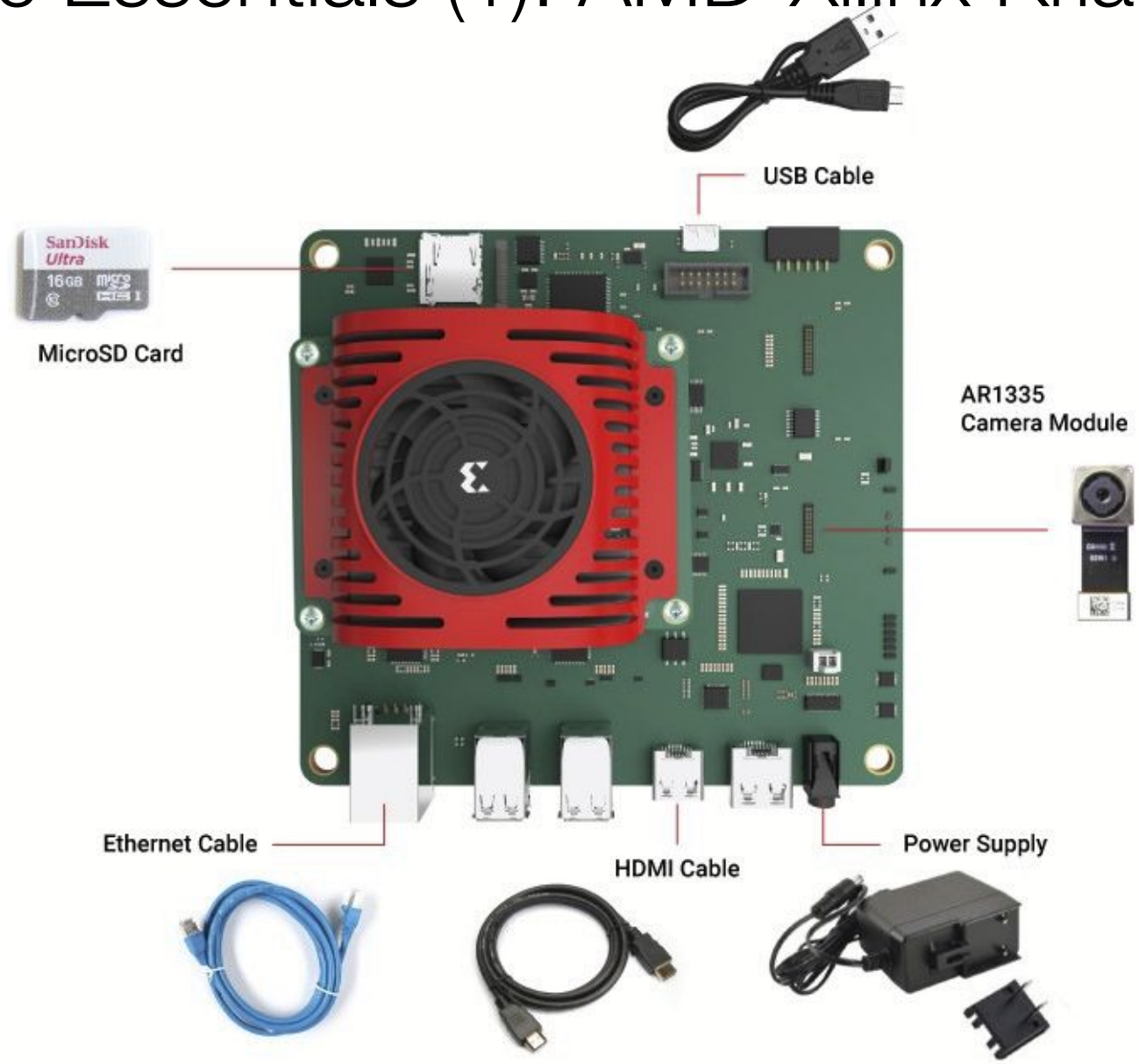
FLEXIBLE CONNECTIVITY
- 1Gb Ethernet
- USB 3.0 / 2.0

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

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

Images from https://xilinx.github.io/kria-apps-docs/kv260/main/build/html/index.html

Creative Essentials (1): AMD-Xilinx Kria KV260



Images from xilinx.com and pnq.io

Creative Essentials (1): How to reach the board

Open the terminal:

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

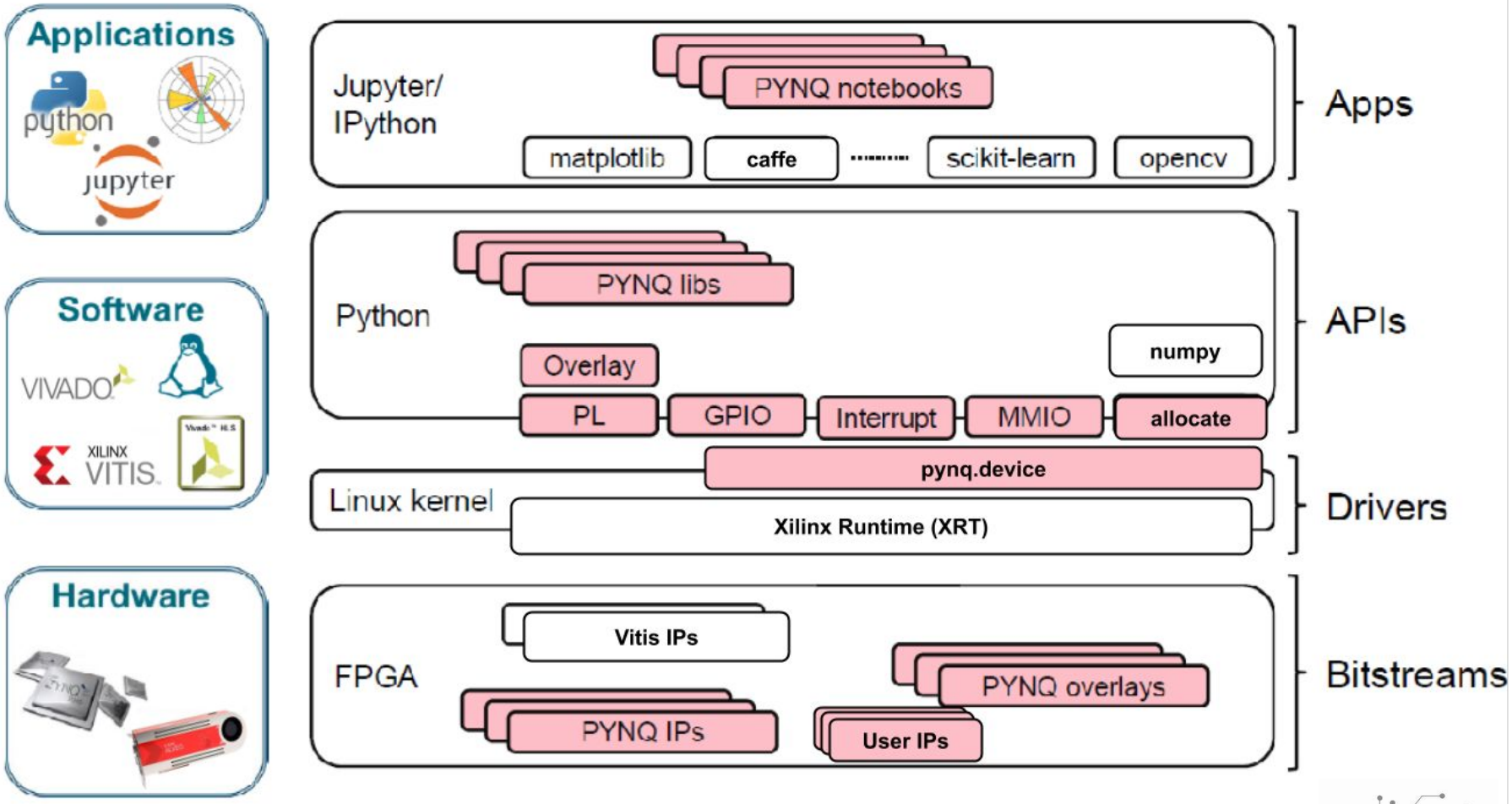
usr:pwd → ubuntu:xilinx12

via SSH (discover ip first)

$ ssh ubuntu@<my_magic_ip>

Images from xilinx.com and pnq.io

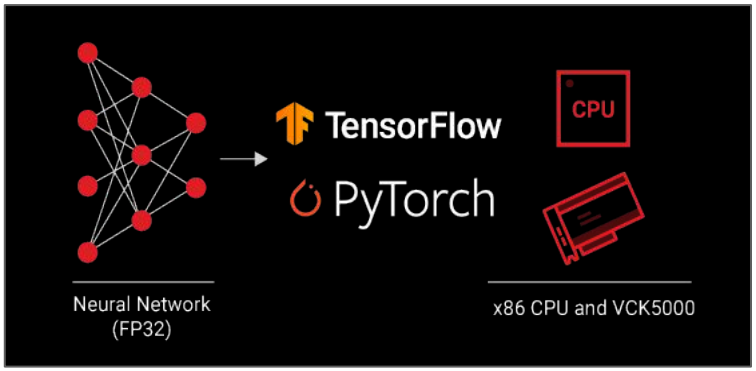
Creative Essentials (2): PYNQ™ Framework



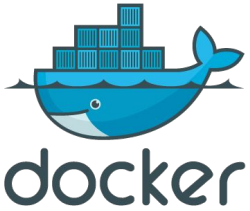
ZYNQ® ZYNQ® UltraSCALE+ ZYNQ® RFSoC XILINX ALVEO™ F1

Images from xilinx.com and pnq.io

Creative Essentials (3) : Vitis AI



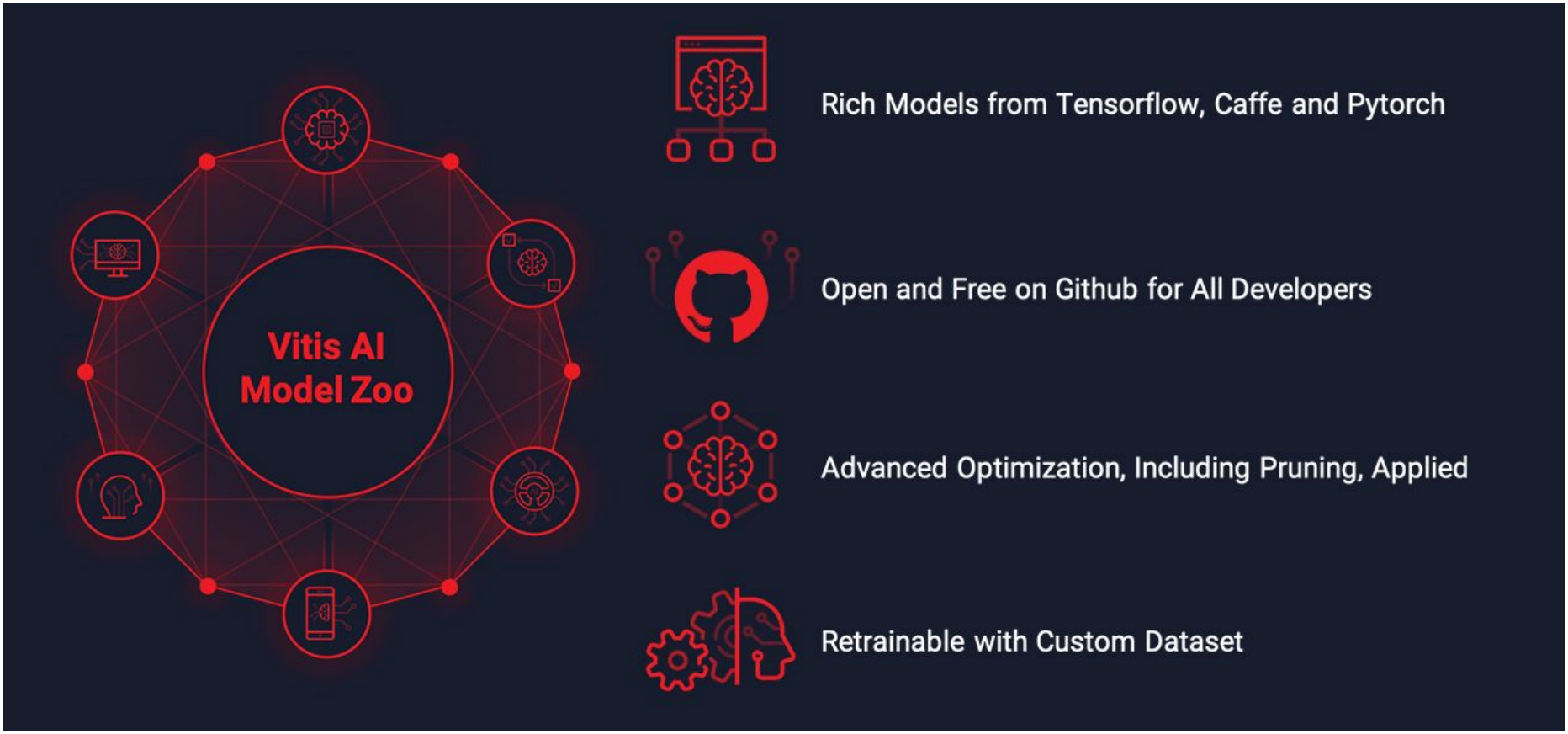
https://github.com/Xilinx/Vitis-AI



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

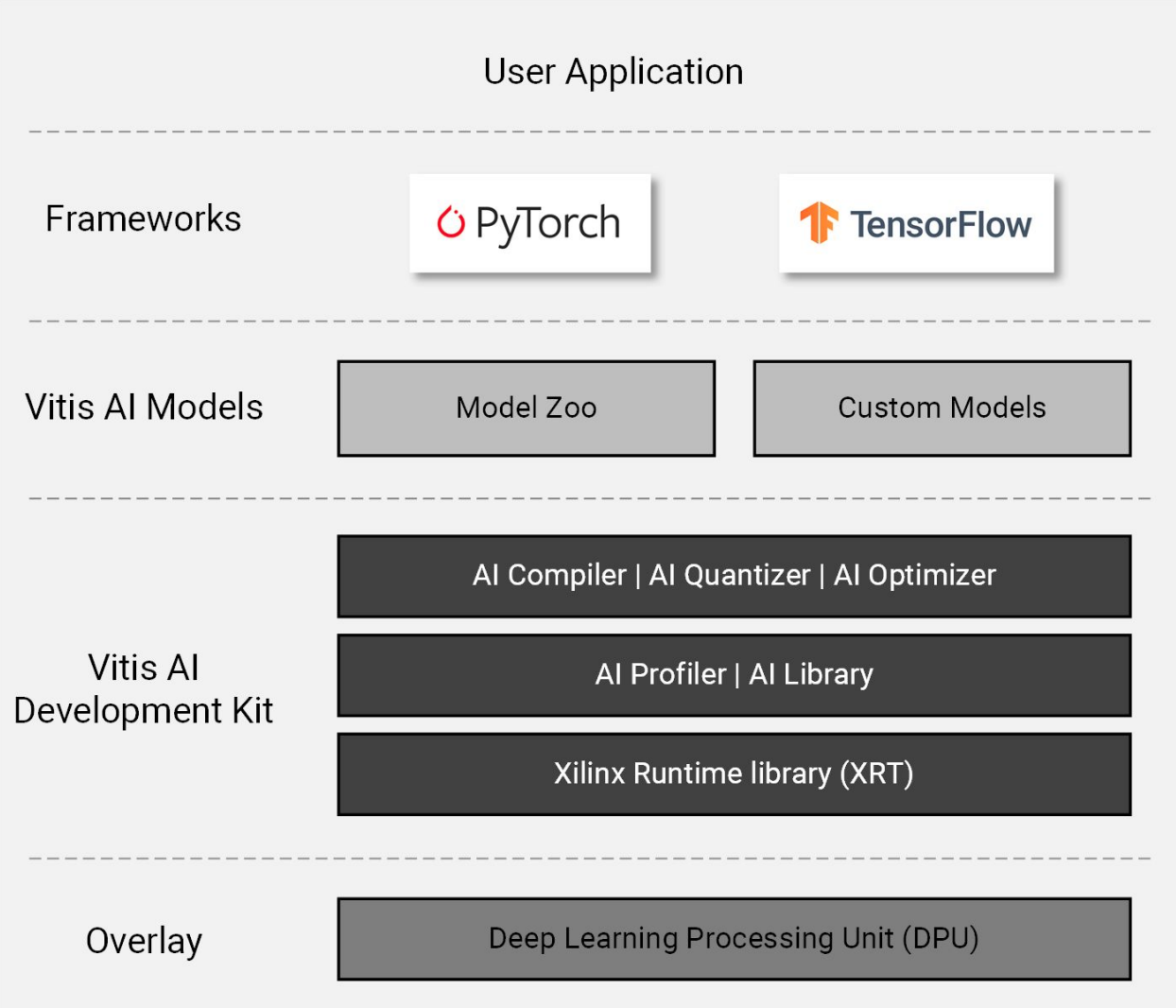
Creative Essentials (3): Vitis AI Model Zoo

https://github.com/Xilinx/Vitis-AI/tree/master/model_zoo



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

Creative Essentials (3): Vitis AI



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

Creative Essentials (3): System-View DPU

DisplayPort
USB3.0
SATA3.1
PCIe Gen2

GigE
USB2.0
UART
SPI
Quad SPI
NAND
SD

Arm Cortex-A53

Arm Cortex-R5

Processing System (PS)

AXI Interconnect

Controller

DDR

AXI Interconnect

DMA

AXI Interconnect

AXI Interconnect

demosaic

gamma

Color_ conversion

DPU

Camera

MIPI

CSI2

MIPI CSI2

Programmable Logic (PL)

X22329-081919

Images from Xilinx University Program Video

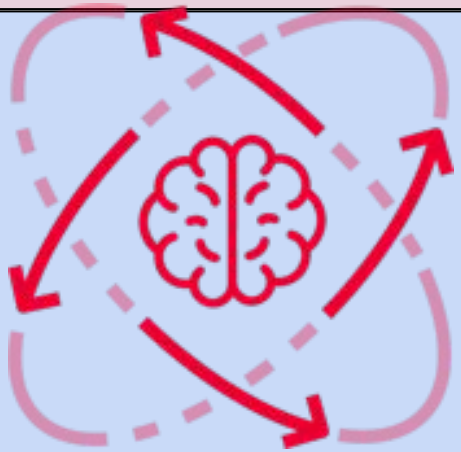
Creative Essentials (4): DPU-PYNQ

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

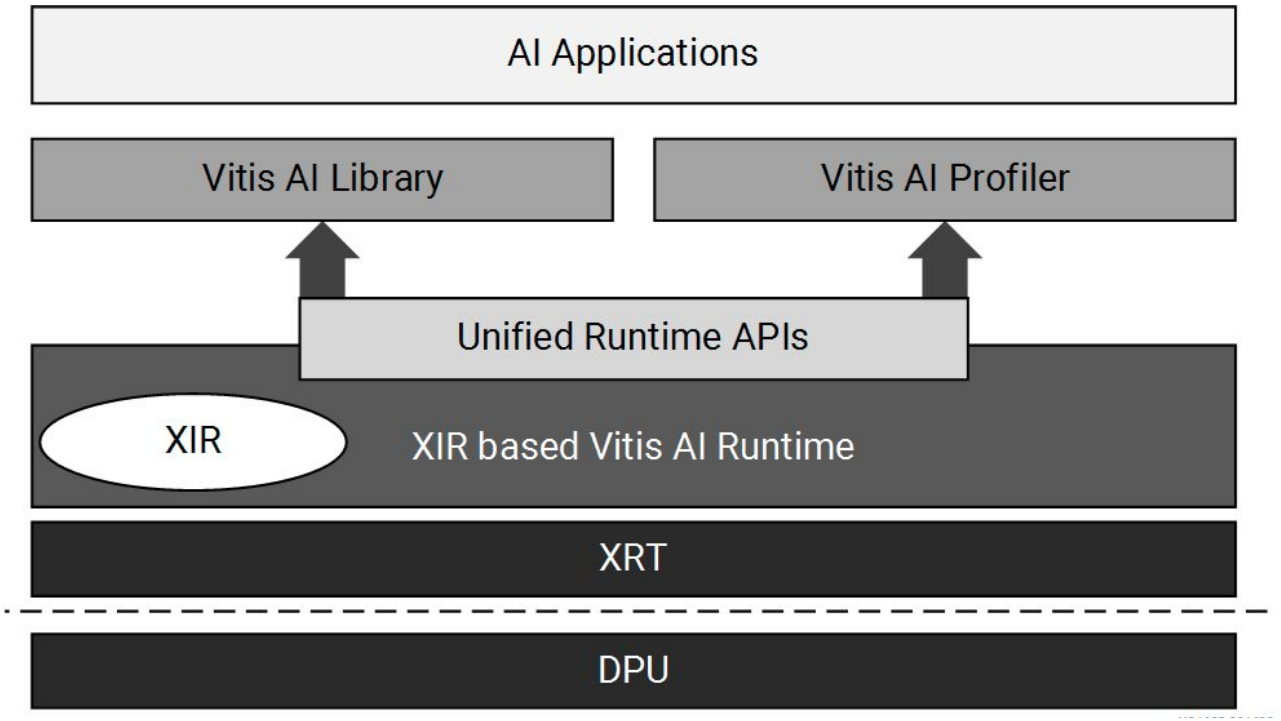
Software

PYNQ™

Hardware



Creative Essentials (4): Vitis AI Runtime (VART)

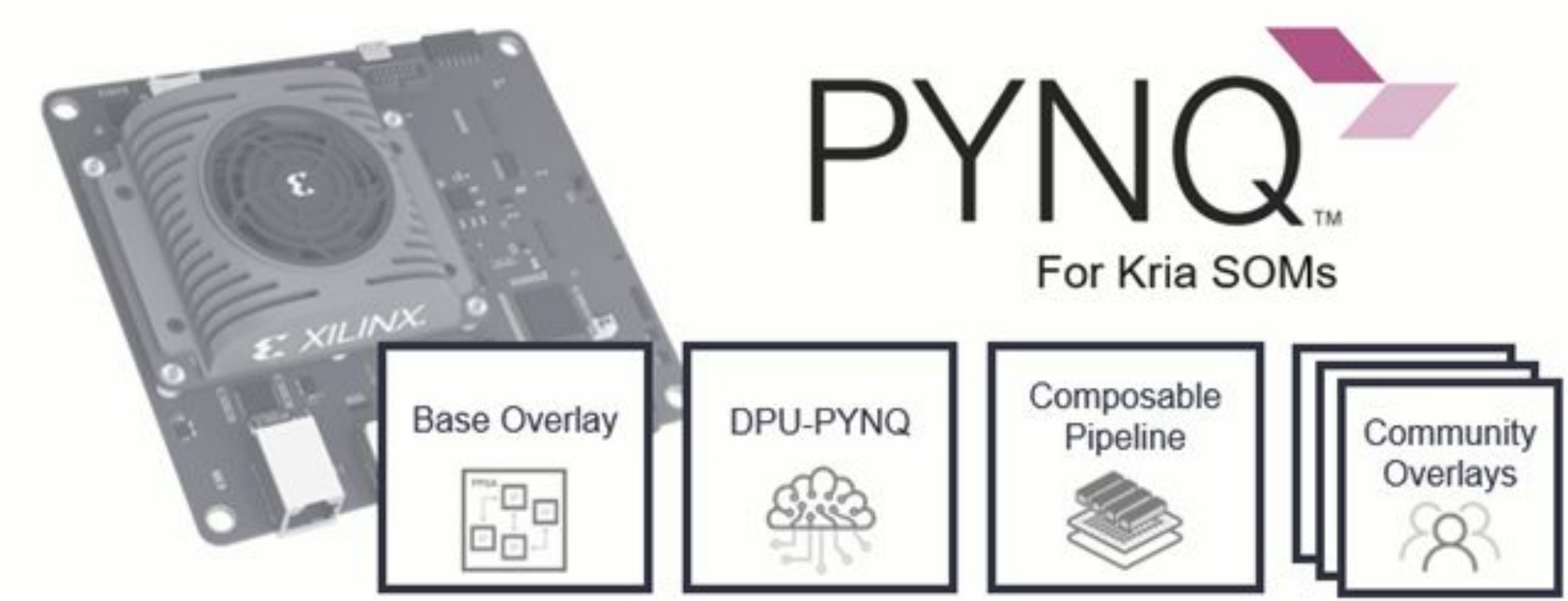


Programming the device and Debugging

- Asynchronous submission of jobs to the accelerator
- Asynchronous collection of jobs from the accelerator
- C++ and Python implementations
- Support for multi-threading and multi-process execution

Images from https://docs.xilinx.com/r/en-US/ug1414-vitis-ai

Creative Essentials (5): PYNQ on KV260

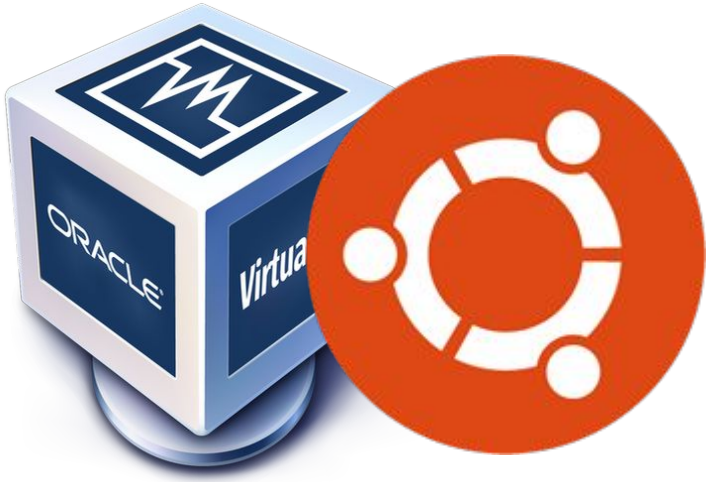


Now you can access JupiterLab via browser:

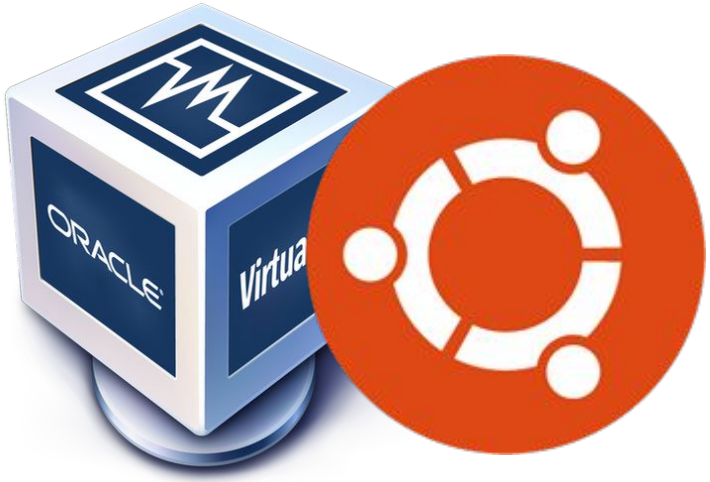
- <ip_address>:9090/lab or
- kria:9090/lab.

The password is: **xilinx**

Creative Essentials (6): The VM



Creative Essentials (6): The VM



IF YOU NEED VITIS AI

On VM Script to install Docker and setup Vitis AI

AI at the Edge CPS Applications on KV260

Smart Camera

Face Detection + Network & Display

AI Box with ReID

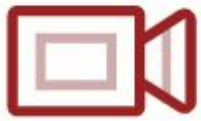
Multi-Stream Face & Pedestrian Detection

NLP SmartVision

Keyword Based Switching Between Vision Tasks

Defect Detection

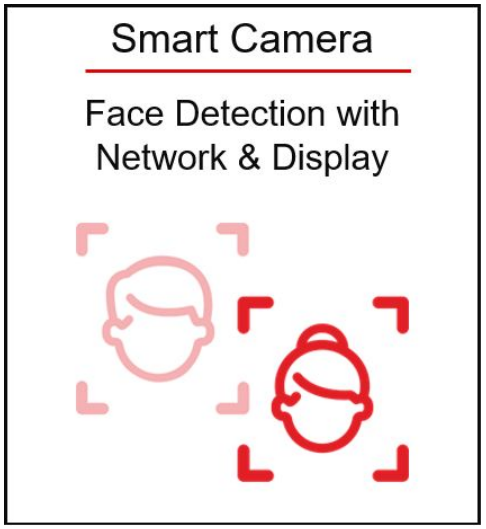
Machine Vision for Identifying Defects



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

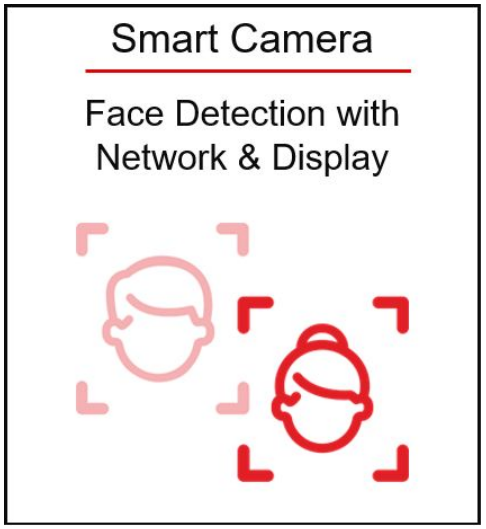
Image source: Report icons created by Flaticon.com

Creative Essentials (7): The Demos



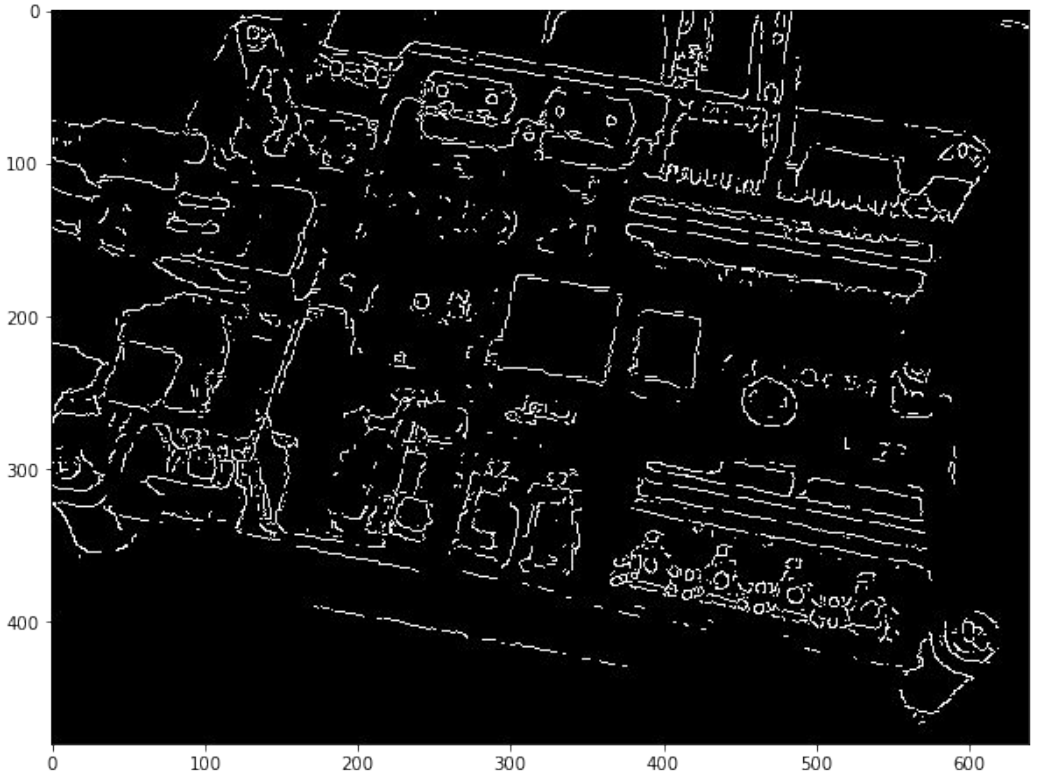
AMD Demo based on MIPI-Camera acquisition (VART-based)

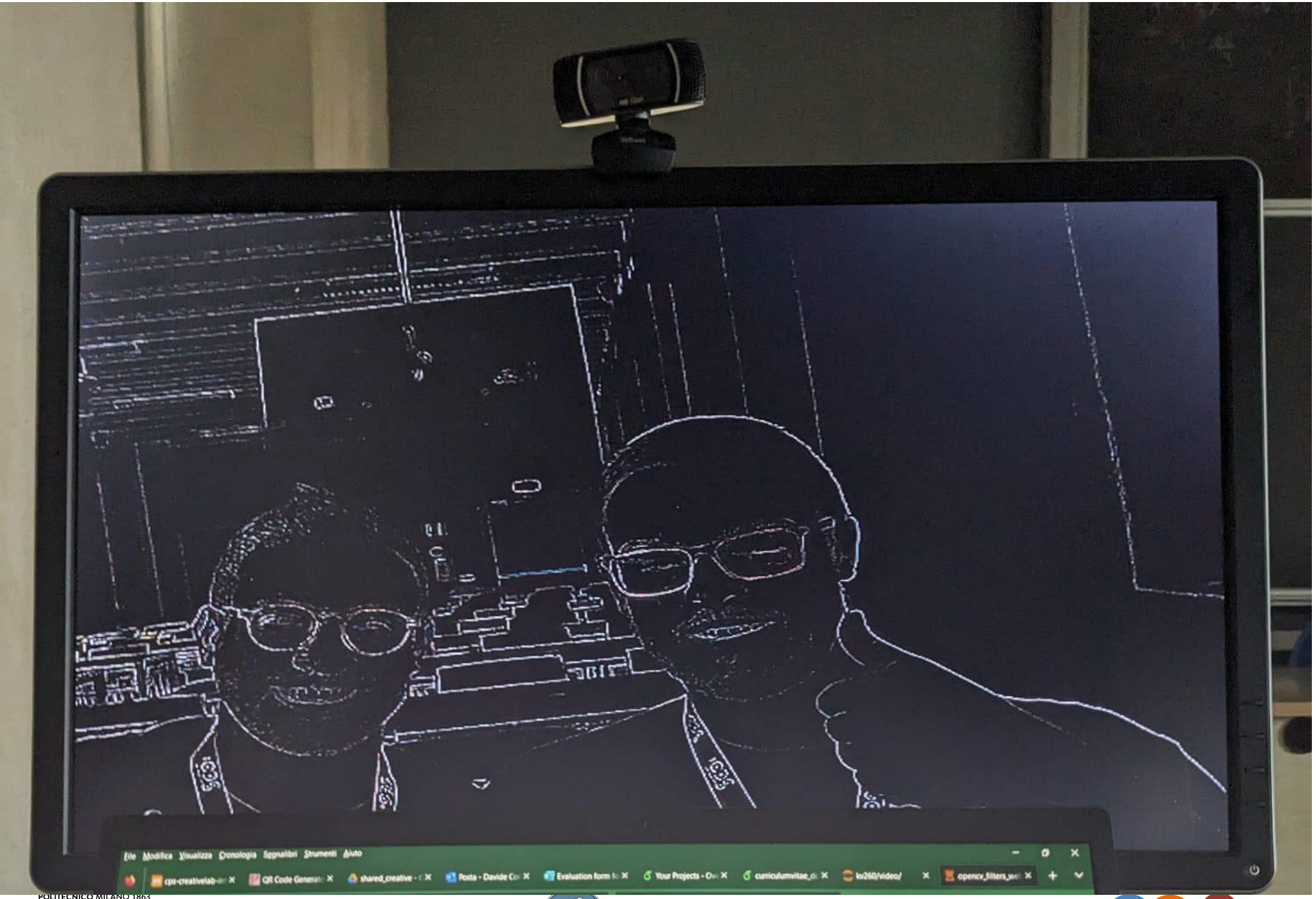
Creative Essentials (7): The Demos



AMD Demo based on MIPI-Camera acquisition (VART-based)

USB-camera Edge detection (PYNQ-based)




File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

Creative Essentials : Sum up

1. AMD Kria KV260 and kit
2. PYNQ Framework
3. Vitis AI and DPU
4. DPU-PYNQ & VART
5. PYNQ on KV260
6. The CPS VM :)
7. Demos:
   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

https://tinyurl.com/cps23-creative

Creative Lab Timeslots

Monday (Aula Nivola)

13:30 – 14:00 Creative Lab Presentation

Tuesday (Aula Badas)

13:30 – 14:00 Creative Lab

Wednesday (Aula Badas)

13:30 – 16:30 Creative Lab

Thursday (Aula Badas)

13:30 – 14:00 Creative Lab

Friday (Aula Badas)

9:00 –12:30 Creative Lab

12:30 – 13:30 Buffet Lunch

13:30 – 15:30 Creative Lab Pitch & Demo

Creative Lab Teams…

Creative Lab Teams: Yellow Team

Alessandro Armellin
Mahyar Pourjabar
Abolfazl Sajadi
Luca Bompani
Paola Busia

Creative Lab Teams: **Green Team**

Roberto Caviglia
Seyed Ahmad Mirsalari
Stephanie Soldavini
Nazareno Bruschi
Andrea Stanco

Creative Lab Teams: **Purple Team**

Mattia Sinigaglia
Diego Navarro-Cabrera
Sara Groppo
Lorenzo Carletti

Creative Lab Teams: **Orange Team**

Sanjay Deshpande
Luigi Pugliese
Yvan Tortorella
Roozbeh Siyadatzadeh
Francesco Ratto

Creative Lab Teams: **Gray Team**

Benedetta Mazzoni
Alberto Musa
Kyrian Maat
Michele Guagnano

Creative Lab Teams: **Blue Team**

Marius Herget
Fatemeh Mehrafrooz
Maicol Ciani
Federico Manca

Thank you for your attention

Alberto Zeni <alberto.zeni@polimi.it>
Davide Conficconi <davide.conficconi@polimi.it>

Acknowledgements

Thanks to the CPS organization for KV260 kits and VM setup

Part of this material comes from:

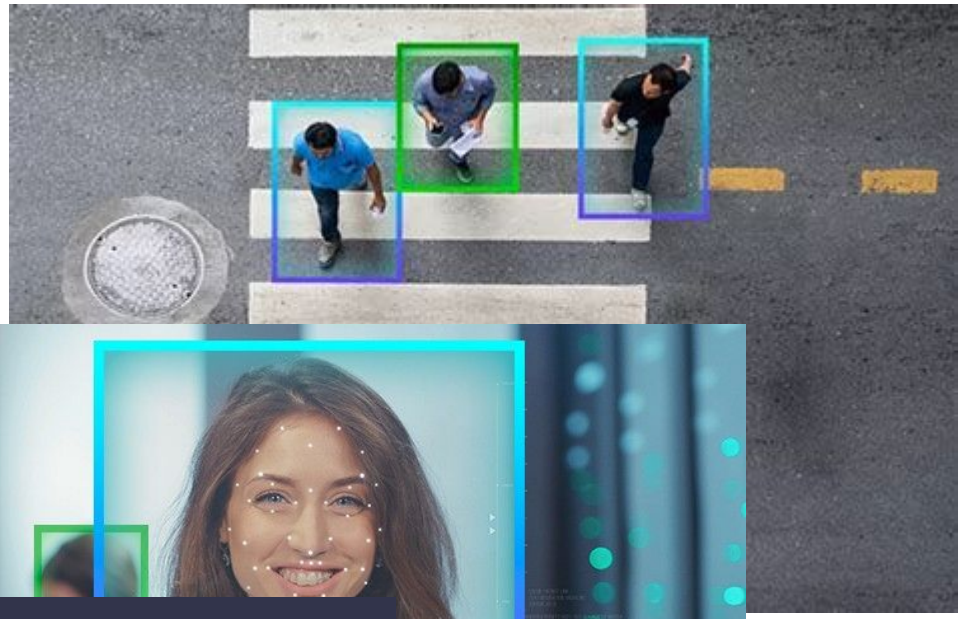
- The AMD-Xilinx websites (mainly https://github.com/Xilinx, https://xilinx.github.io/kria-apps-docs/kv260/main/build/html/index.html)
- “On the role of reconfigurable systems in domain-specific computing”, D. Conficconi, Politesi

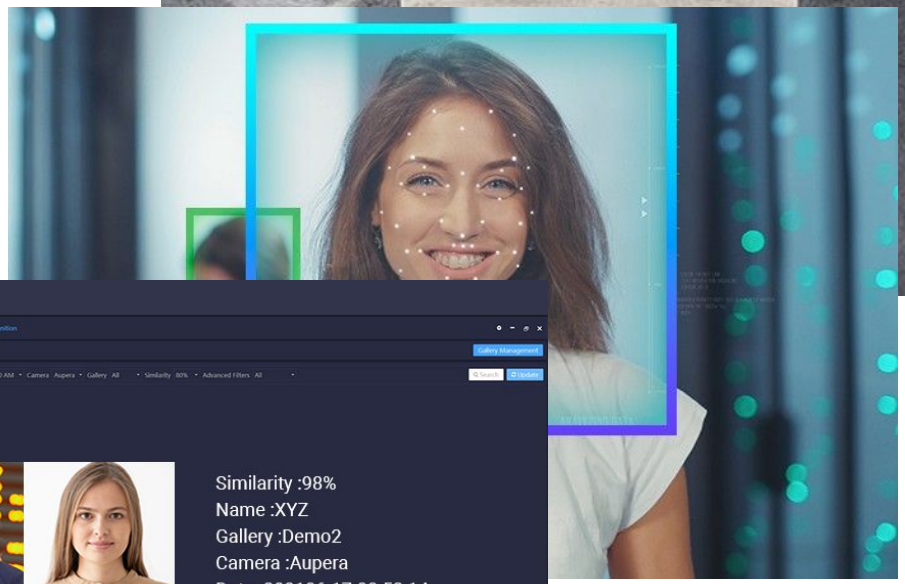
and are ***properties of their respective owners***

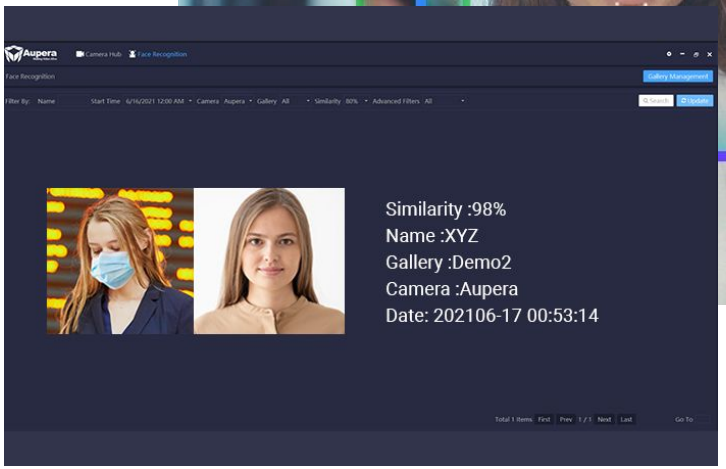
Thank you for your attention

Alberto Zeni <alberto.zeni@polimi.it>

KRIA Live Demo!



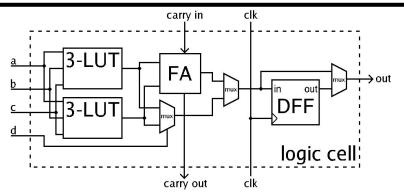




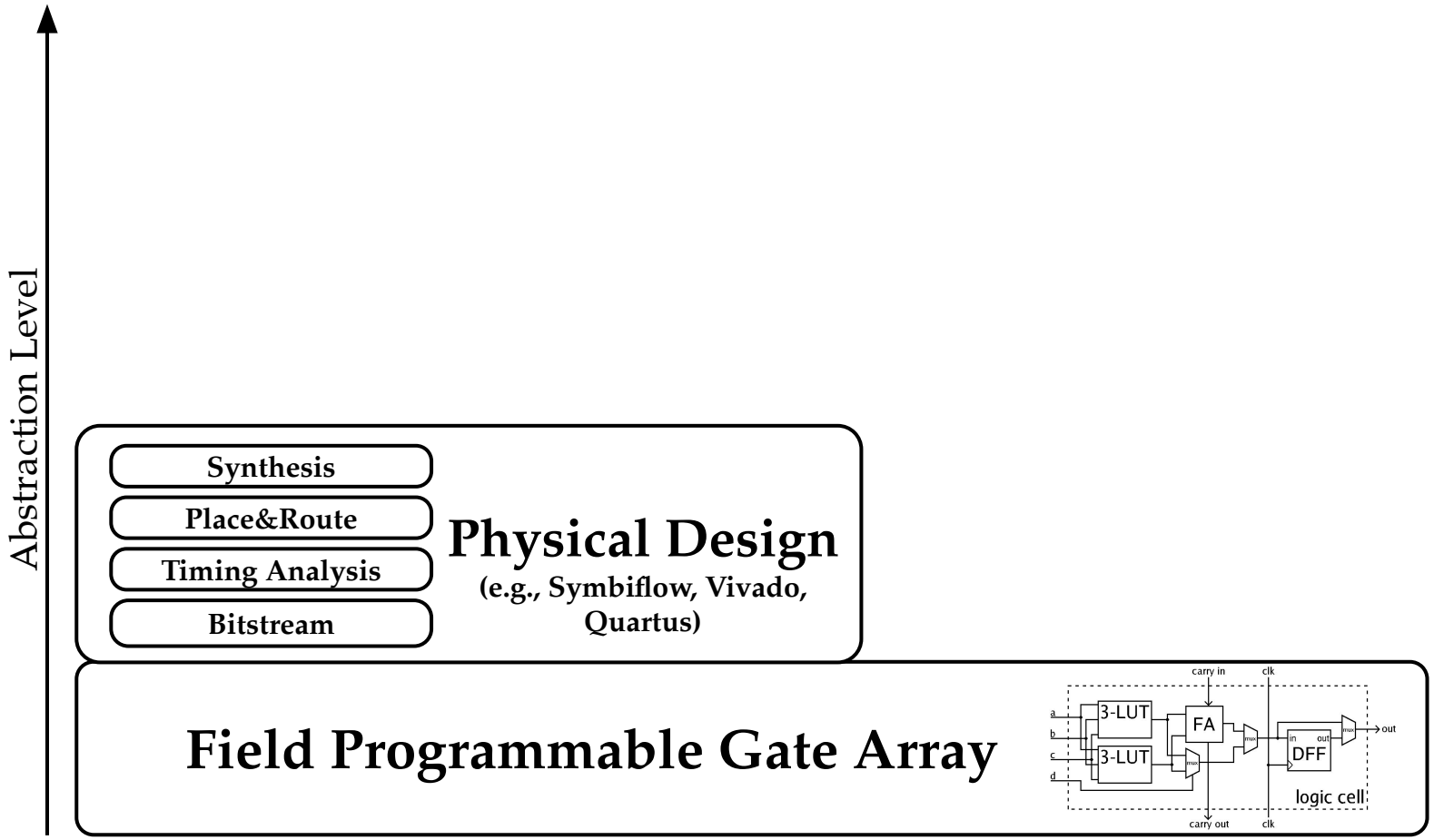
FPGAs Design Flows [https://doi.org/10.1145/3532989]

Abstraction Level

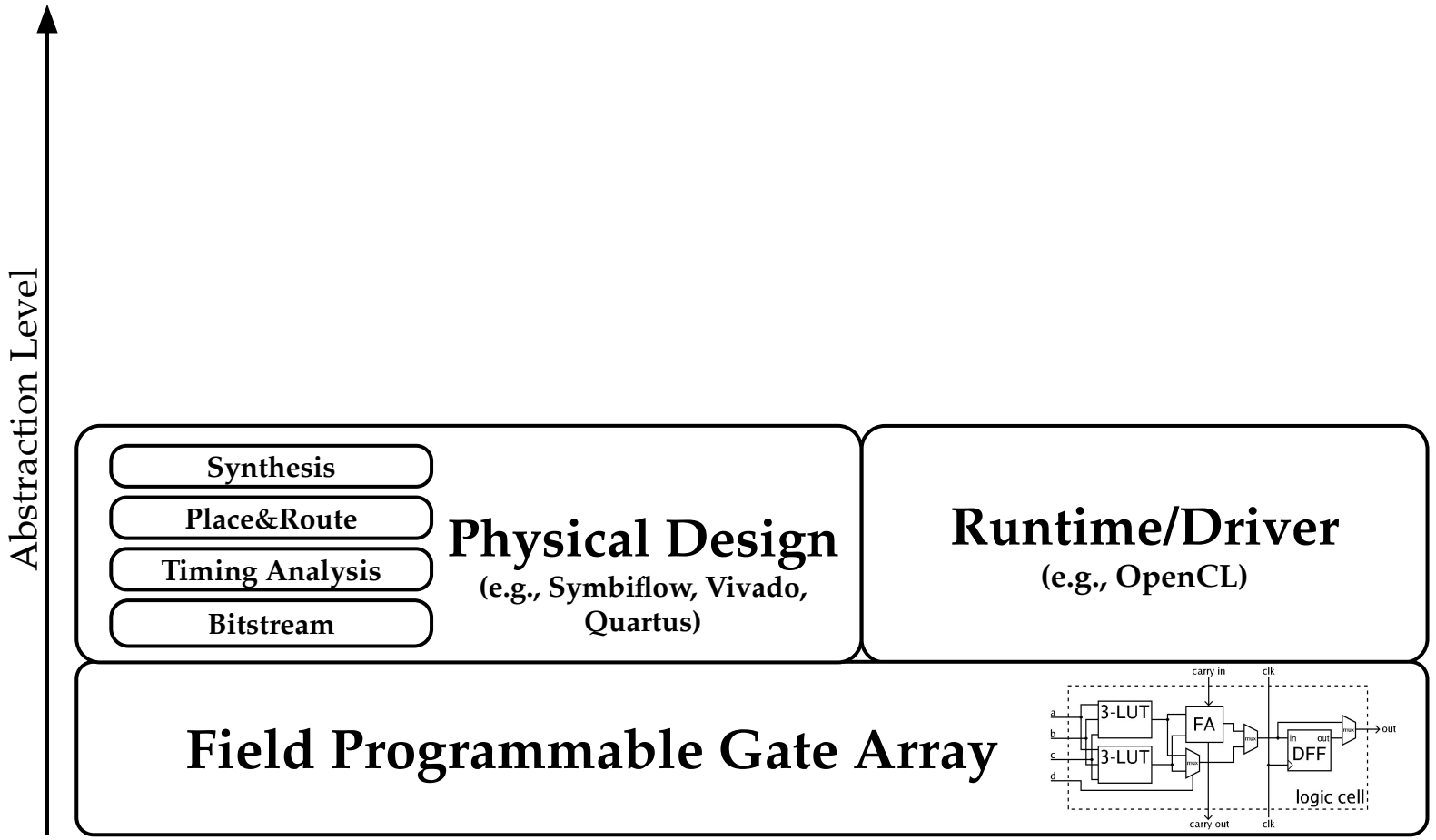
Field Programmable Gate Array



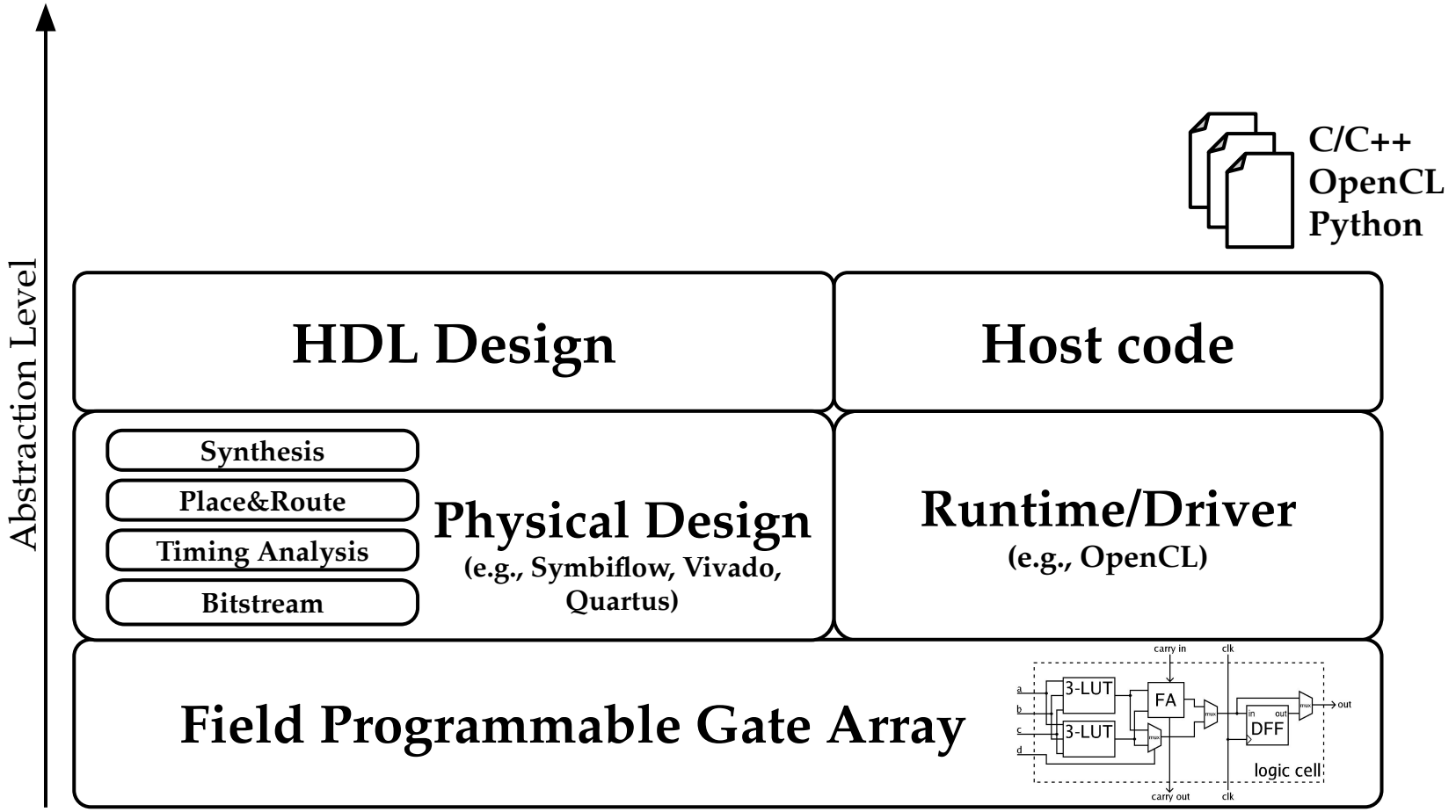
FPGAs Design Flows [https://doi.org/10.1145/3532989]



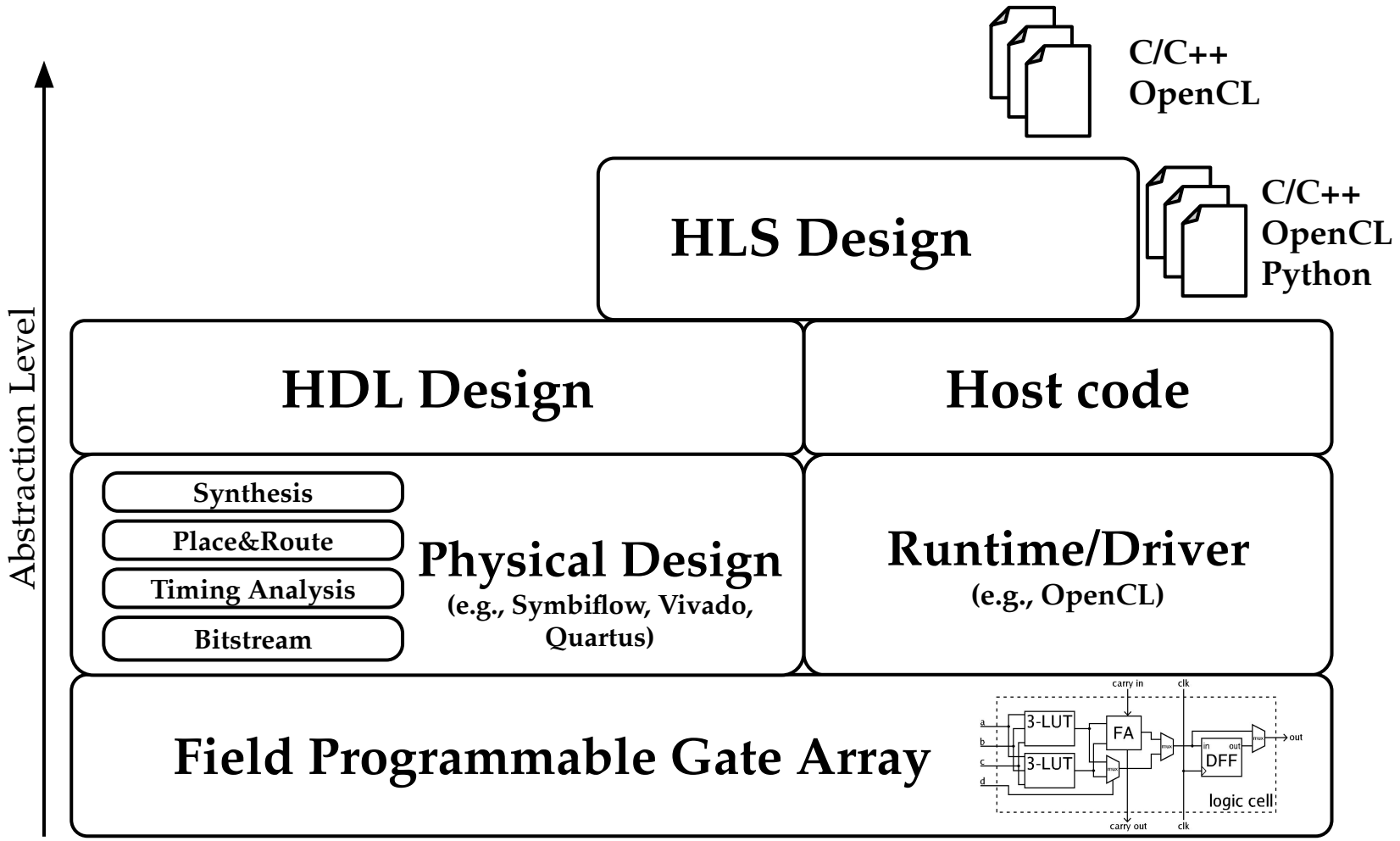
FPGAs Design Flows[https://doi.org/10.1145/3532989]



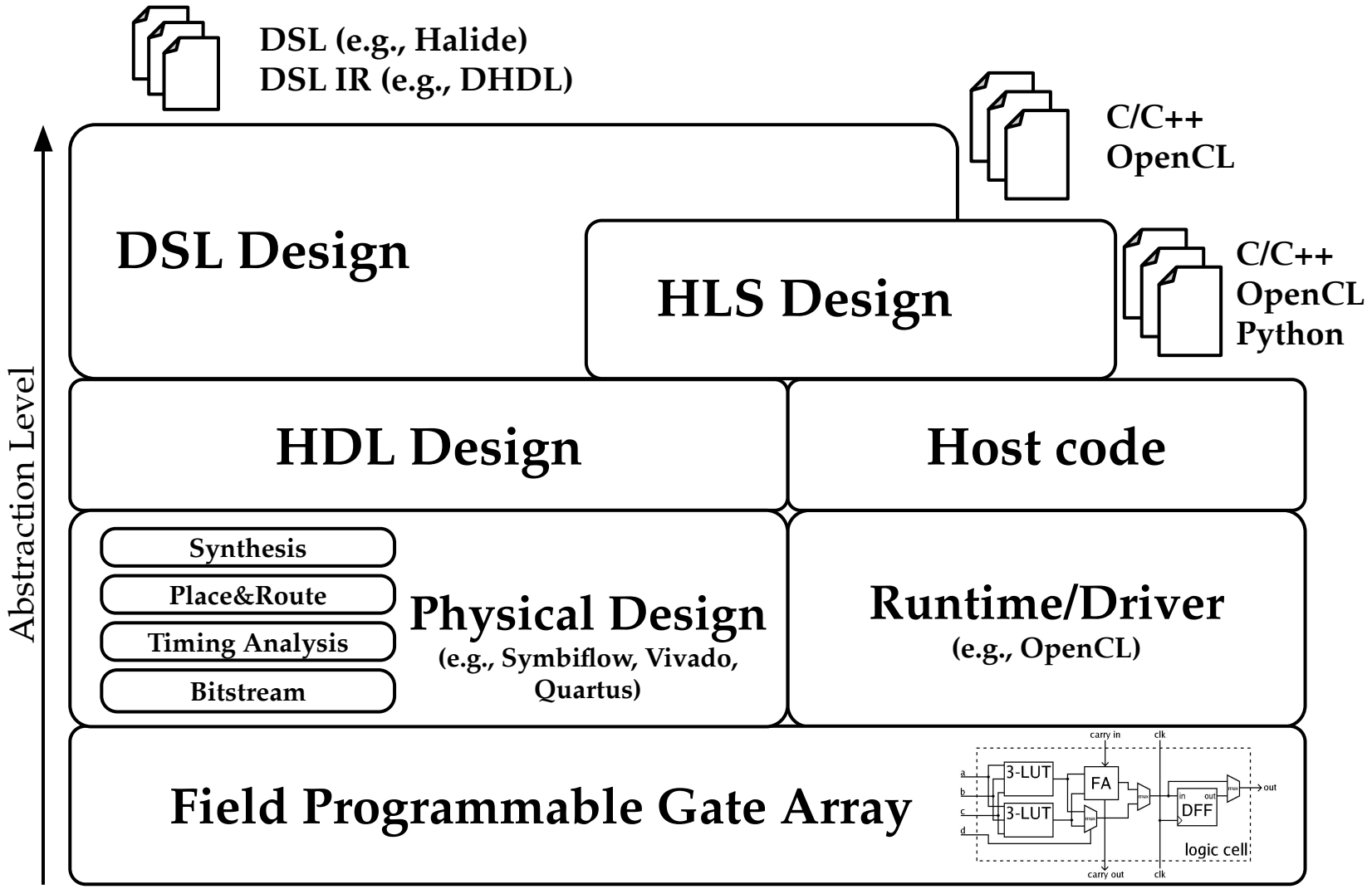
FPGAs Design Flows[https://doi.org/10.1145/3532989]



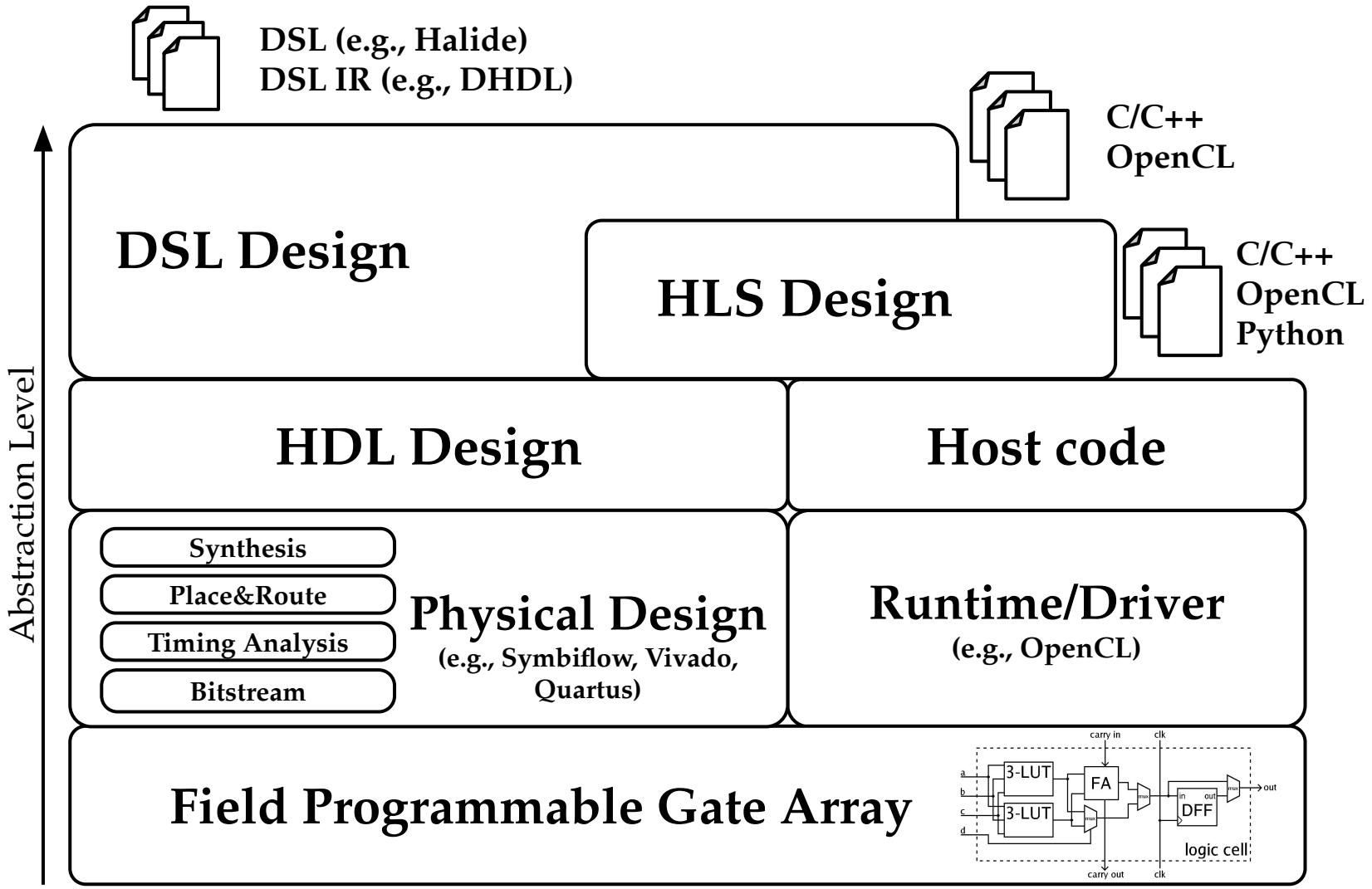
FPGAs Design Flows[https://doi.org/10.1145/3532989]



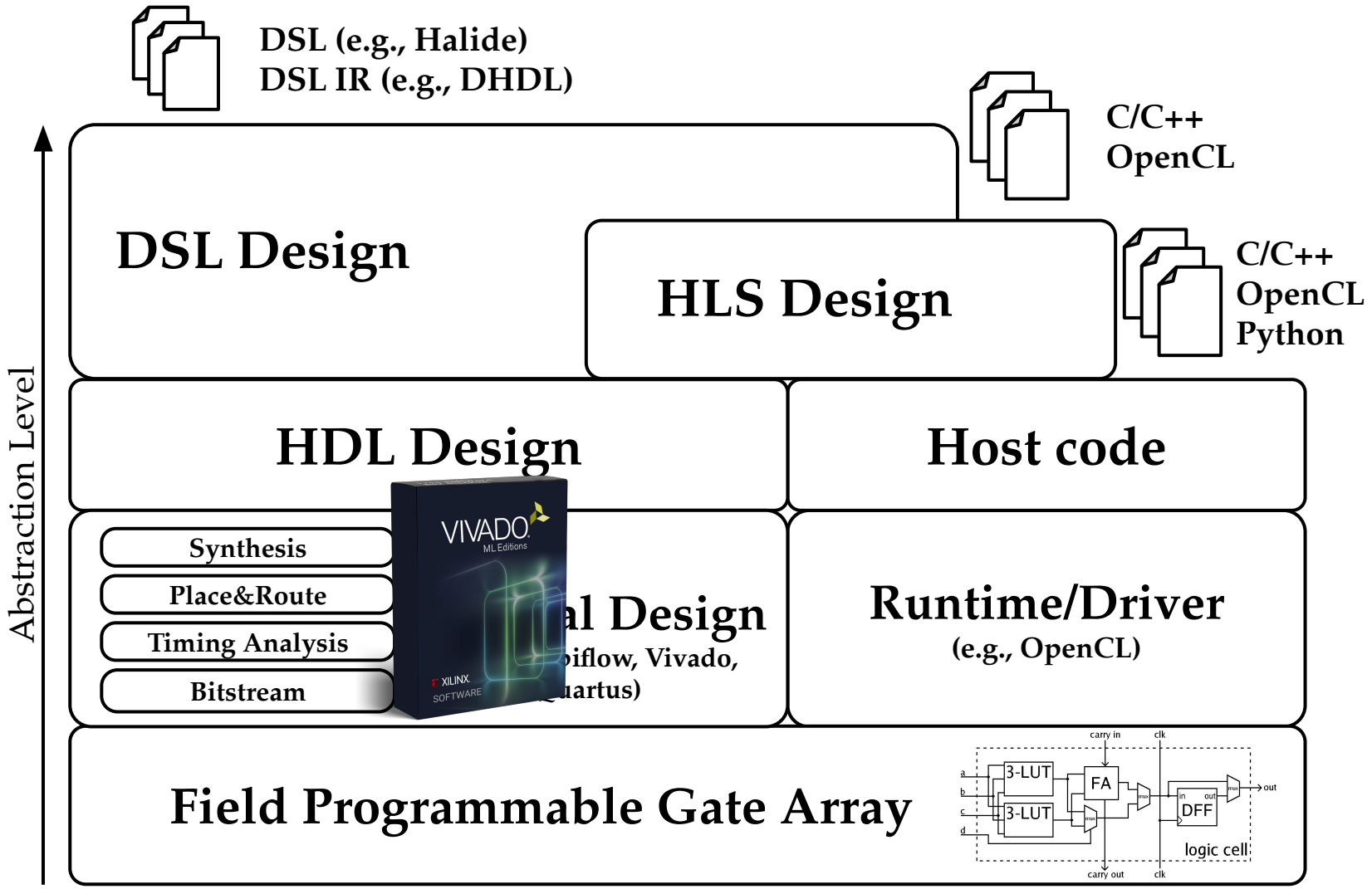
FPGAs Design Flows[https://doi.org/10.1145/3532989]



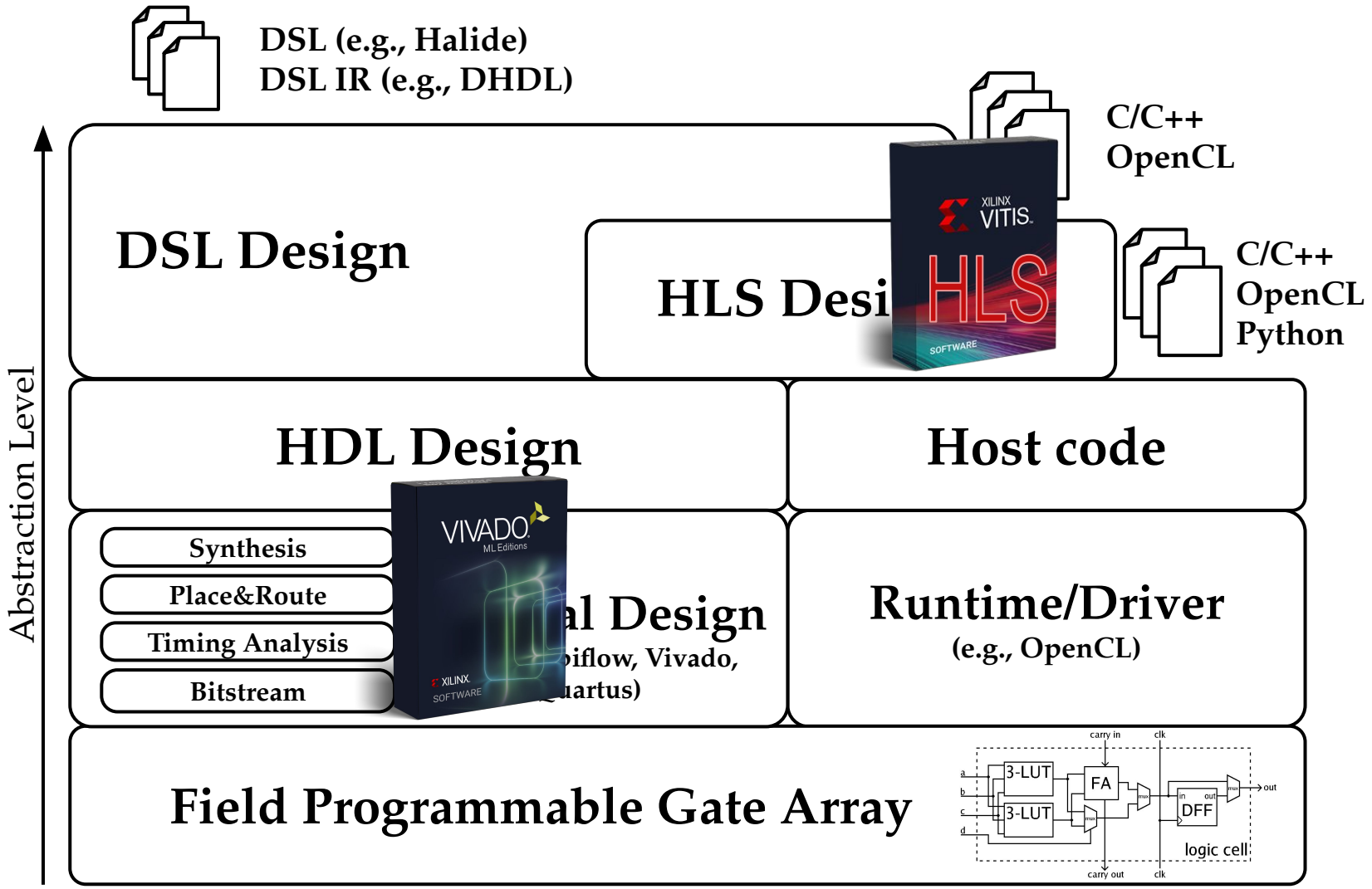
Recap FPGAs Design Flows [https://doi.org/10.1145/3532989]



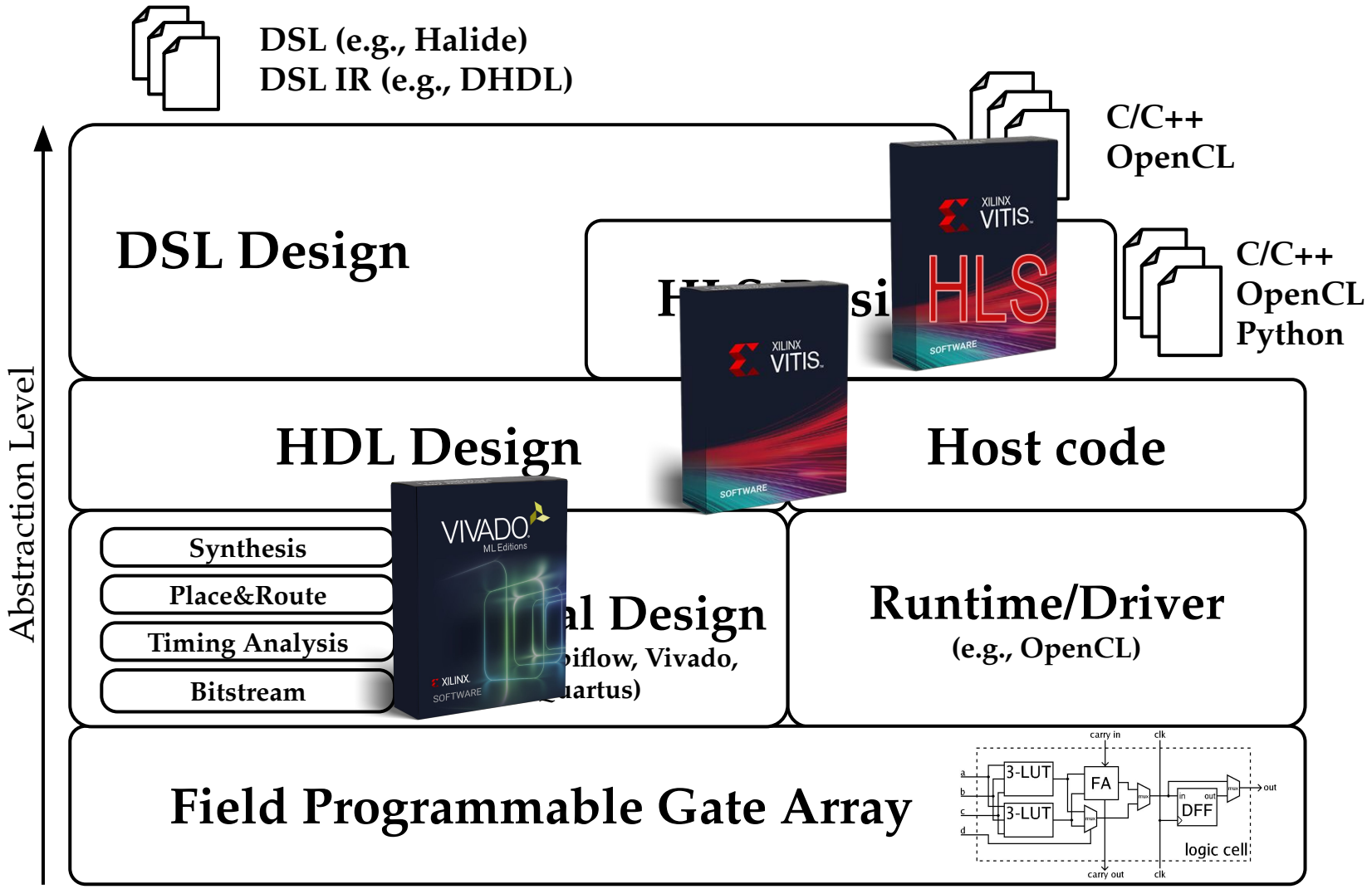
Recap FPGAs Design Flows [https://doi.org/10.1145/3532989]



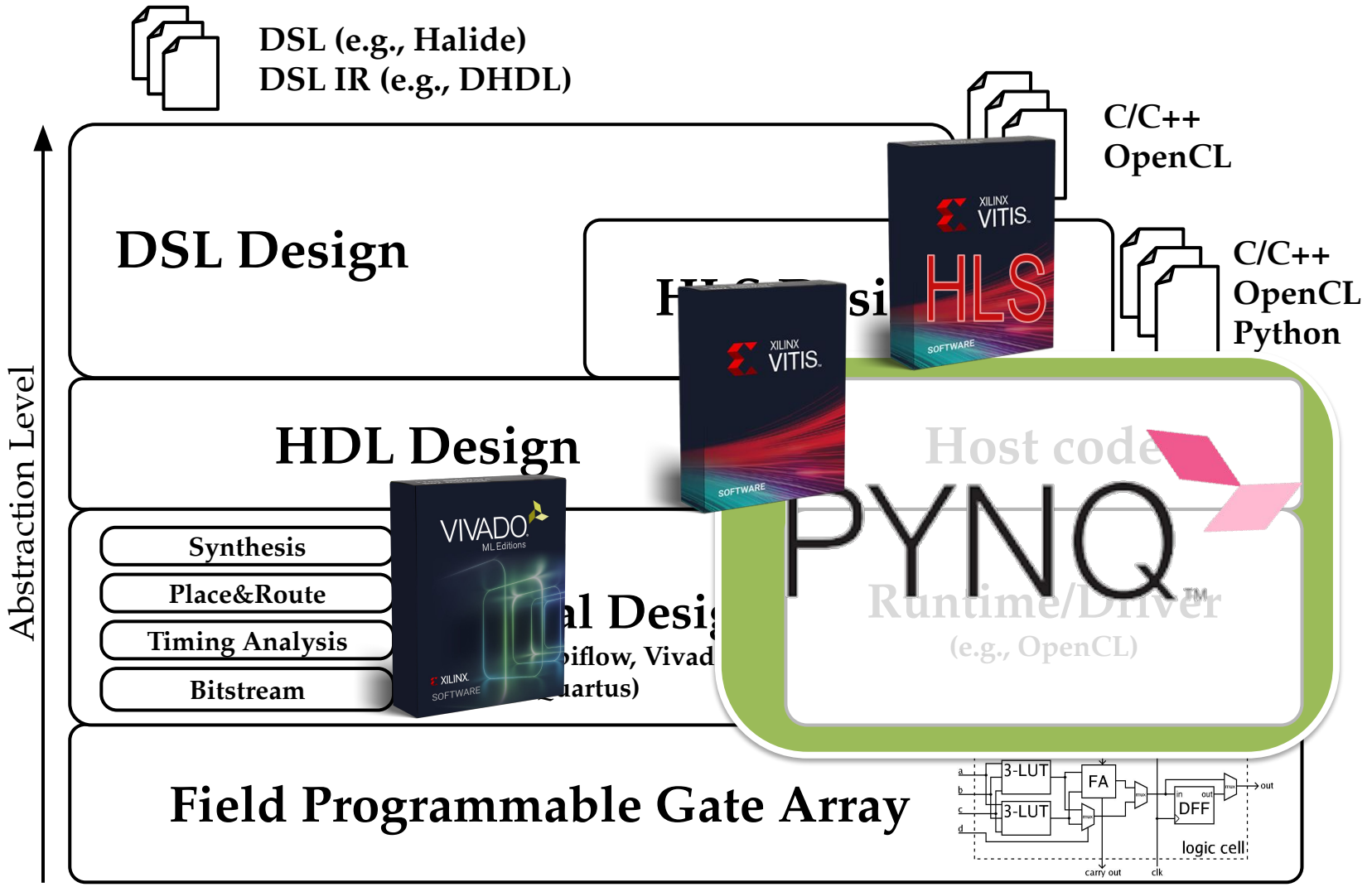
Recap FPGAs Design Flows[https://doi.org/10.1145/3532989]



Recap FPGAs Design Flows[https://doi.org/10.1145/3532989]



Recap FPGAs Design Flows [https://doi.org/10.1145/3532989]



Recap FPGAs Design Flows[https://doi.org/10.1145/3532989]