

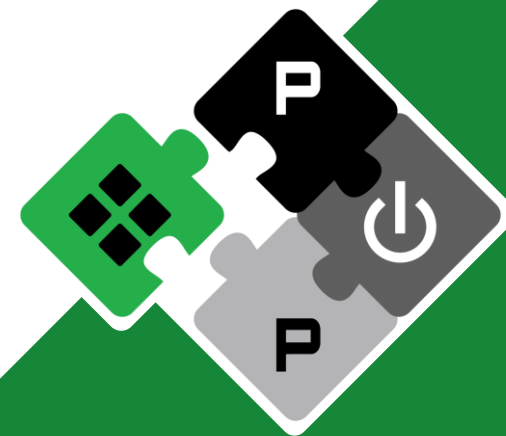
# Paving the Road of RISC-V Embedded Parallel Power Controllers for Next-Generation HPC Processors

Integrated Systems Laboratory (ETH Zürich)  
University of Bologna

Alessandro Ottaviano, Robert Balas, Giovanni Bambini,  
Corrado Bonfanti, Simone Benatti, Davide Rossi, Luca  
Benini and Andrea Bartolini

**PULP Platform**

Open Source Hardware, the way it should be!



@pulp\_platform 

<http://pulp-platform.org> 

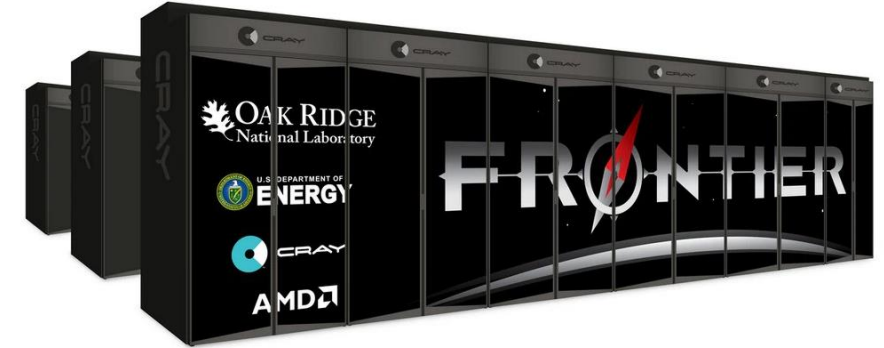
[https://www.youtube.com/pulp\\_platform](https://www.youtube.com/pulp_platform) 

# Introduction



- **Energy-efficiency is getting growing attention in the HPC field**

- Top500 leading supercomputer ORNL's Frontier leads Green500
- 150x energy-efficiency improvement in the last 15 years
- Cutting-edge cooling systems and advanced power management



- **System-level power and thermal management for HPC**

- Crucial matter in the many-core era of computing systems
- Need for fine-grained, workload-aware dynamic power management



# Pervading the RISC-V growing ecosystem



- RISC-V HPC processors development is approaching at fast pace
  - Examples: [Monte Cimone](#), [Ventana Veyron V1](#)
  - HPC requires **dynamic power management** to mitigate the drawbacks of the end of Moore's law (the *power wall*)
  - Mostly an orthogonal problem to the ISA, but RISC-V needs its own power management strategy, infrastructure and services

**What can be done to support the transition of RISC-V to HPC from a power management point of view?**





- **Existing commercial power controller are changing to account for scalability**

- To face the ever-increasing number of controlled cores in single-die and chiplet-based modern chip
- With Alder Lake, Intel integrates one PCU per P-core
- Since Ryzen, AMD has a master-slave design with one slave per core
- IBM has one OCC per Tile

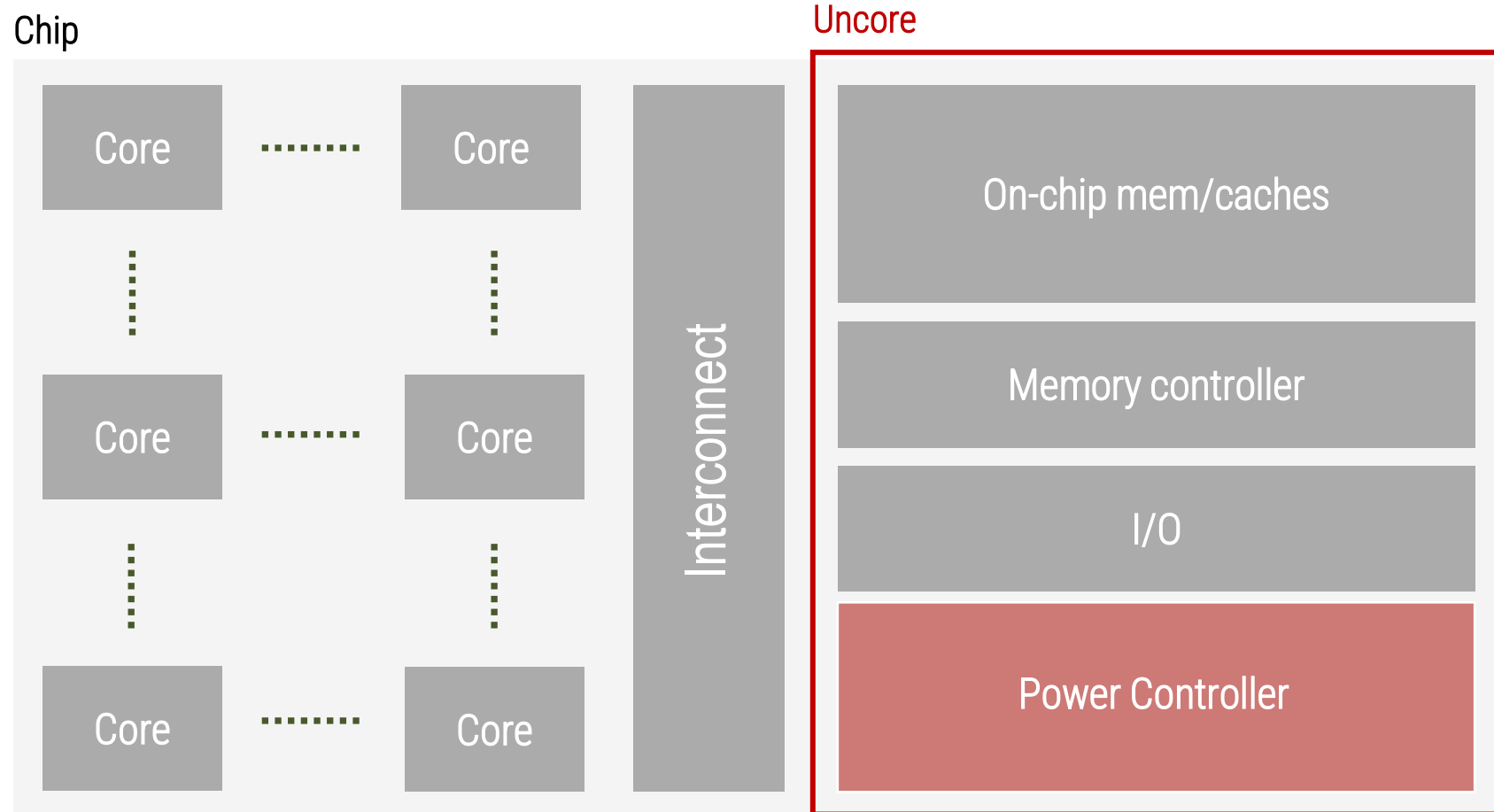
## Need for a scalable power controller solution

- **ControlPULP: first RISC-V based, multi-core power controller**

- Complete HW/SW platform, to be open-sourced soon
- EPI Choice -> freedom



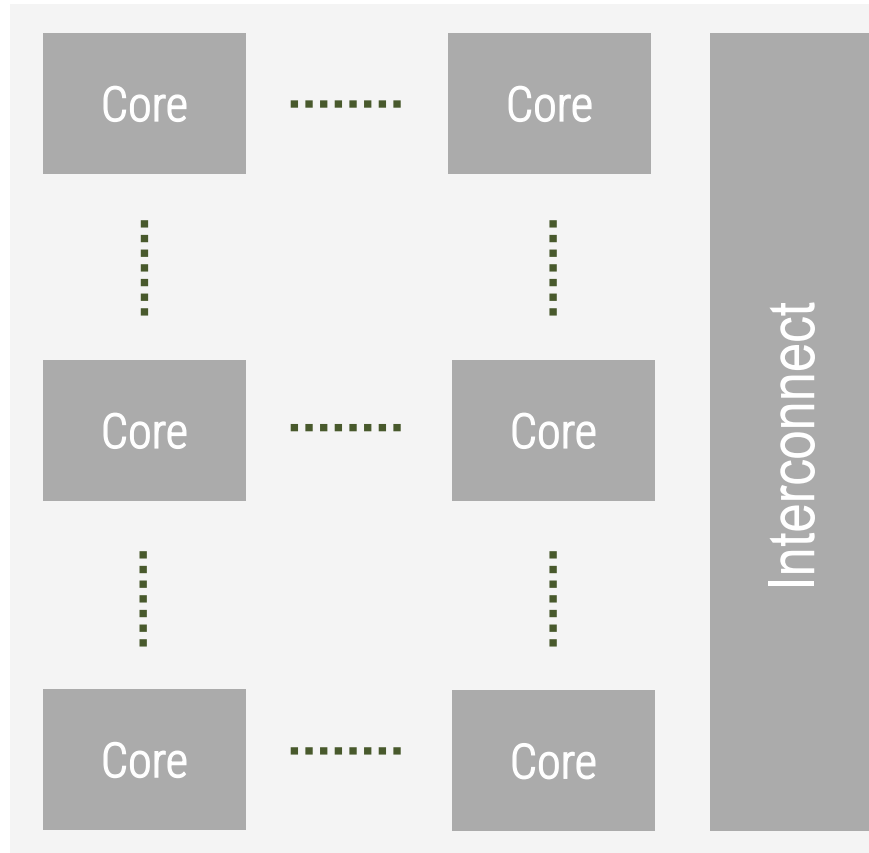
# Introduction



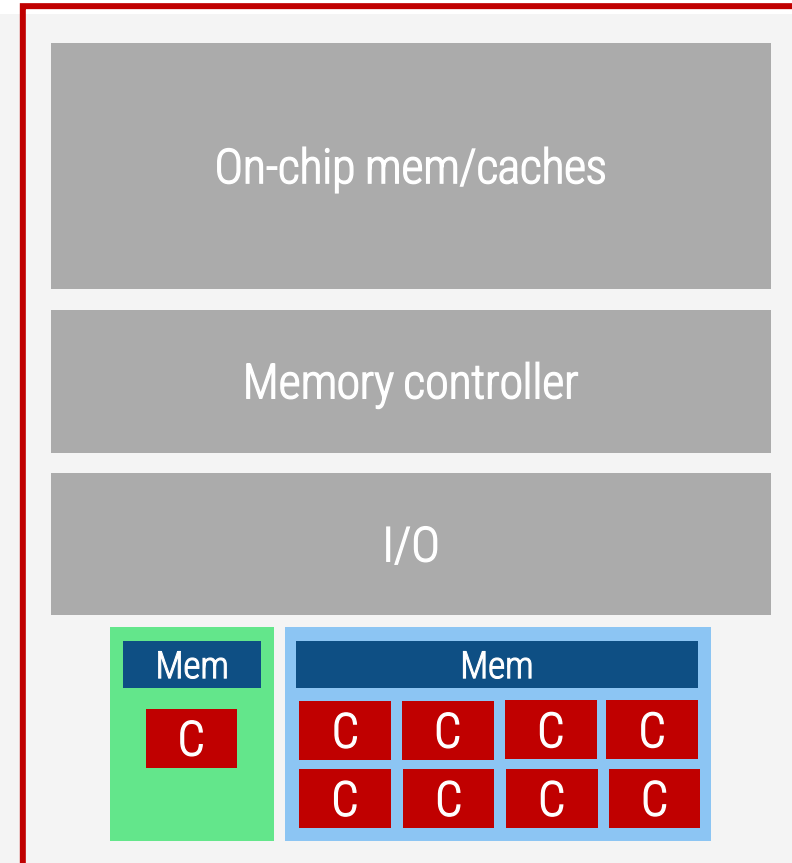
# Introduction



Chip



Uncore

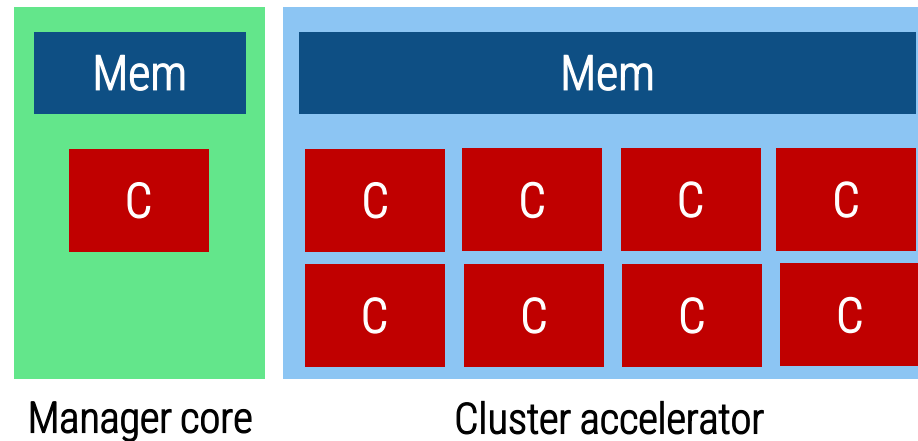


# ControlPULP



- **Scalable architecture:**

- Based on PULP<sup>1</sup> project (Parallel Ultra Low Power)
- **Multi-core** cluster with private FPU (float16, bfloat, float32)
- DMA for 2-D strided access from PVT sensor registers

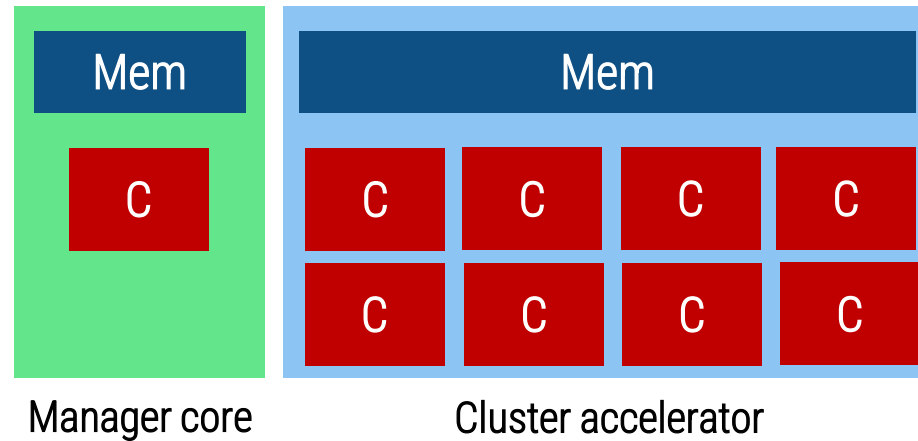


# ControlPULP



- **Predictable architecture:**

- Cache-less system
- Constant access time to scratchpad memories
- Manager core has dedicated banks for data and instructions
- Platform level interrupt controller (RISC-V PLIC)





# ControlPULP



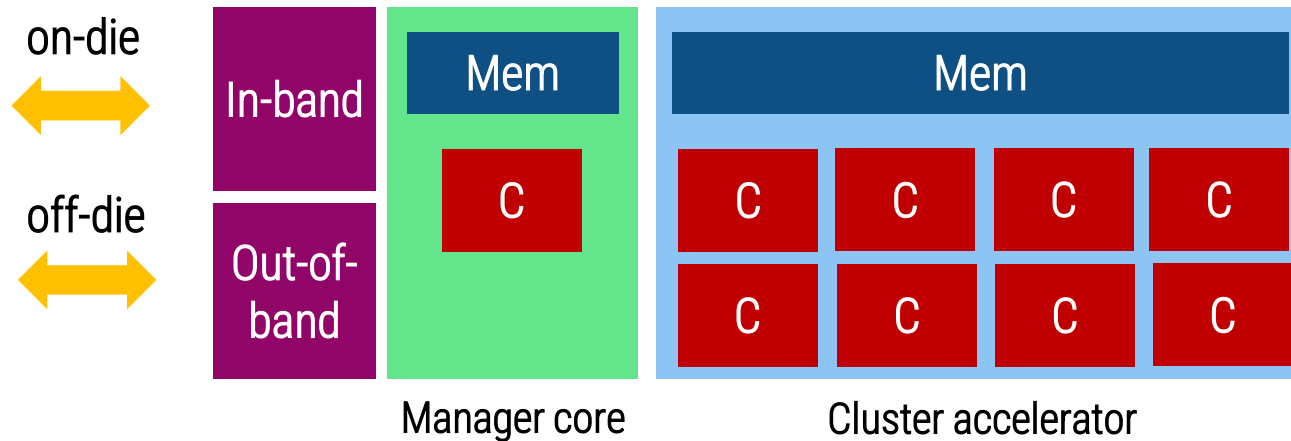
- **Industry standard power management interfaces:**

- PMBUS/AVSBUS: Voltage Regulators control
- SPI: Inter-socket communication (Multi ControlPULP)
- ACPI/MCTP: Motherboard/BMC interface (OpenBMC)

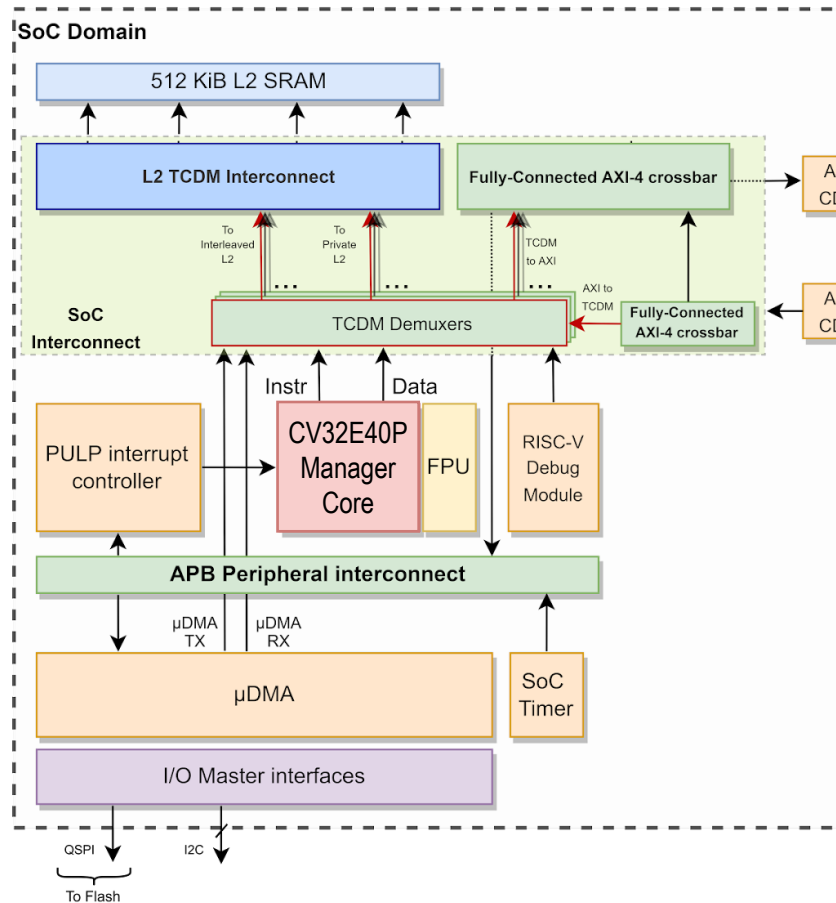
} Out-Of-Band

- SCMI: OS governors and telemetry

} In-Band



# Architecture



## Single-core subsystem

- 32-bit CV32E40P
- 512 KiB scratchpad memory
- Executes the main control policy routine
- Offloads tasks to accelerator cluster

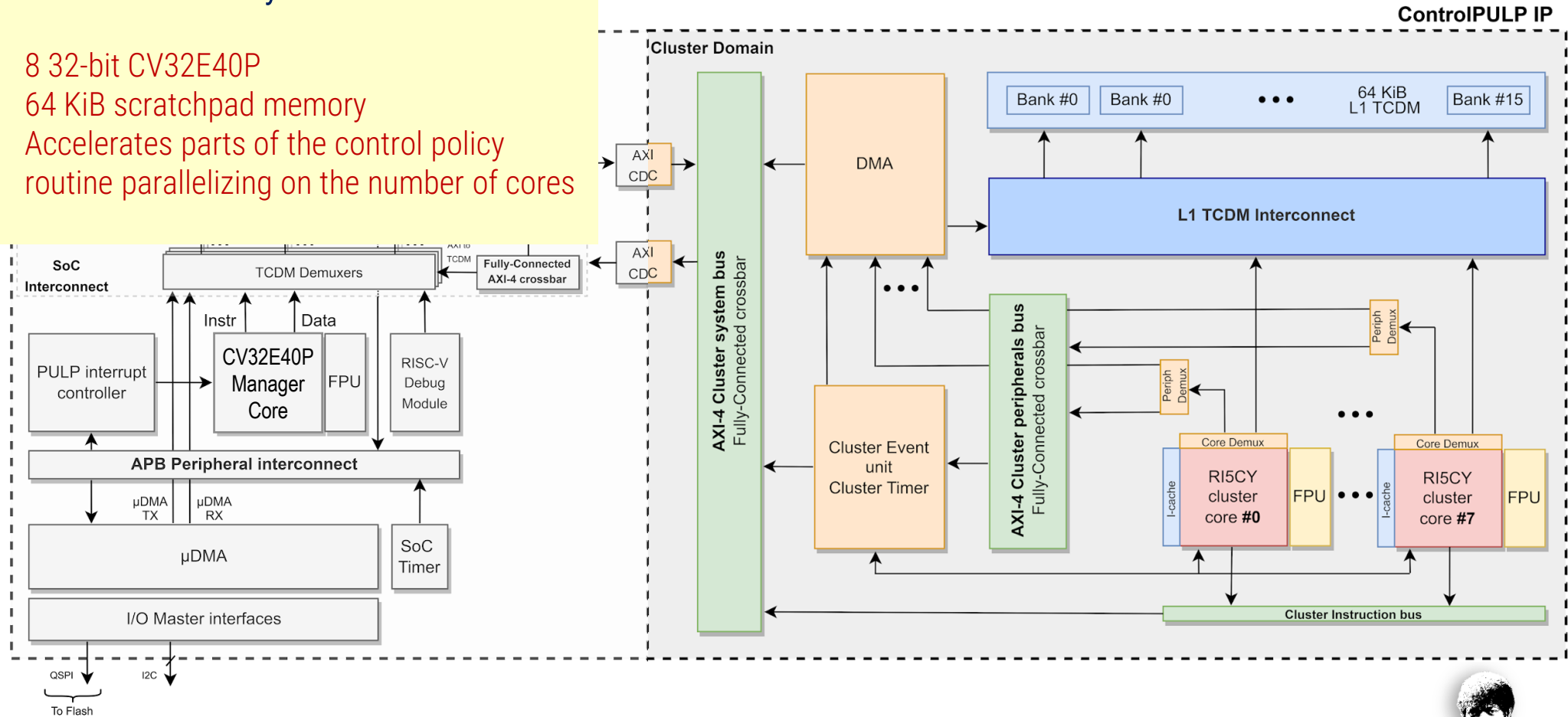


# Architecture



## Multi-core cluster subsystem

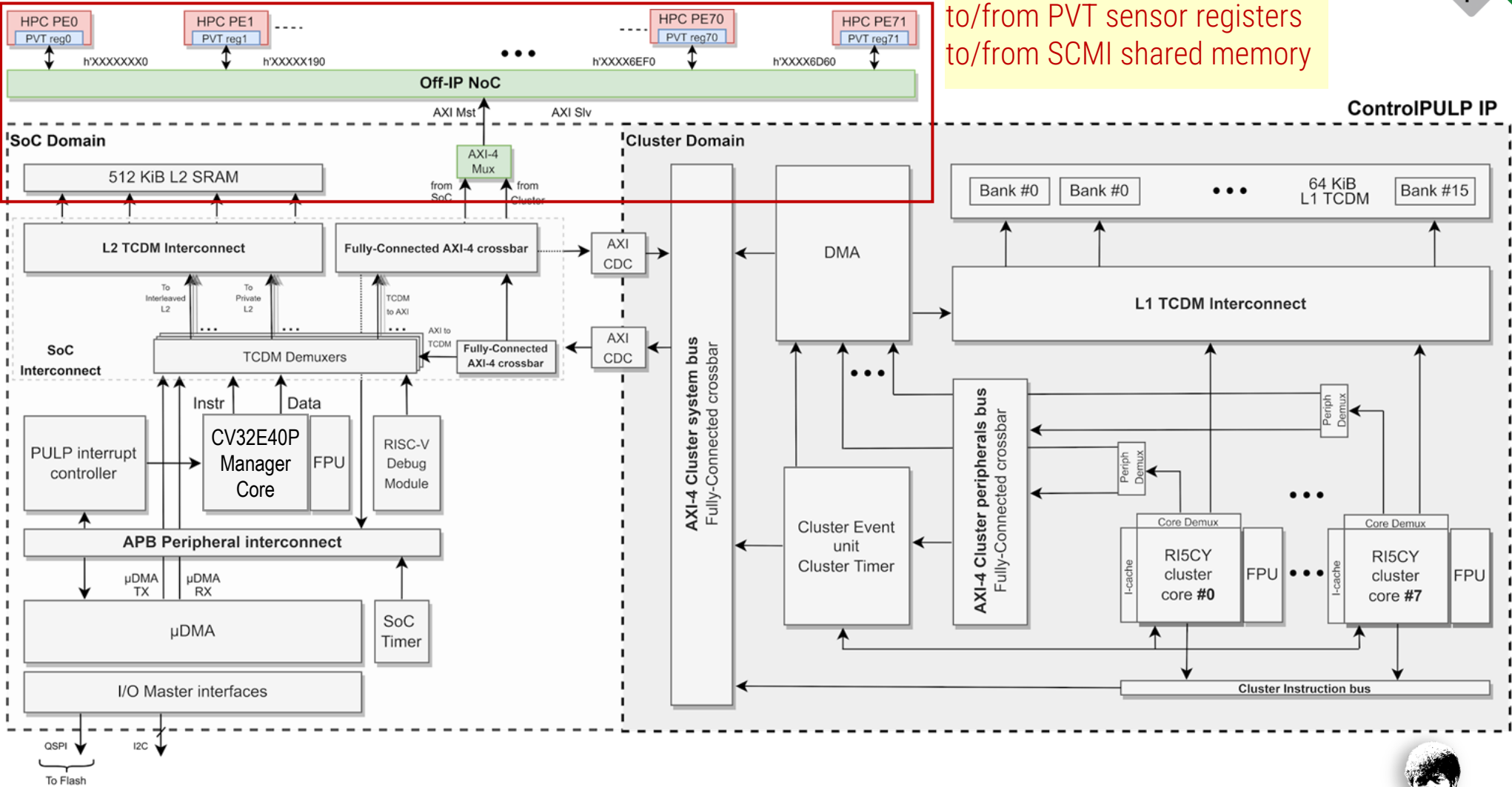
- 8 32-bit CV32E40P
- 64 KiB scratchpad memory
- Accelerates parts of the control policy routine parallelizing on the number of cores



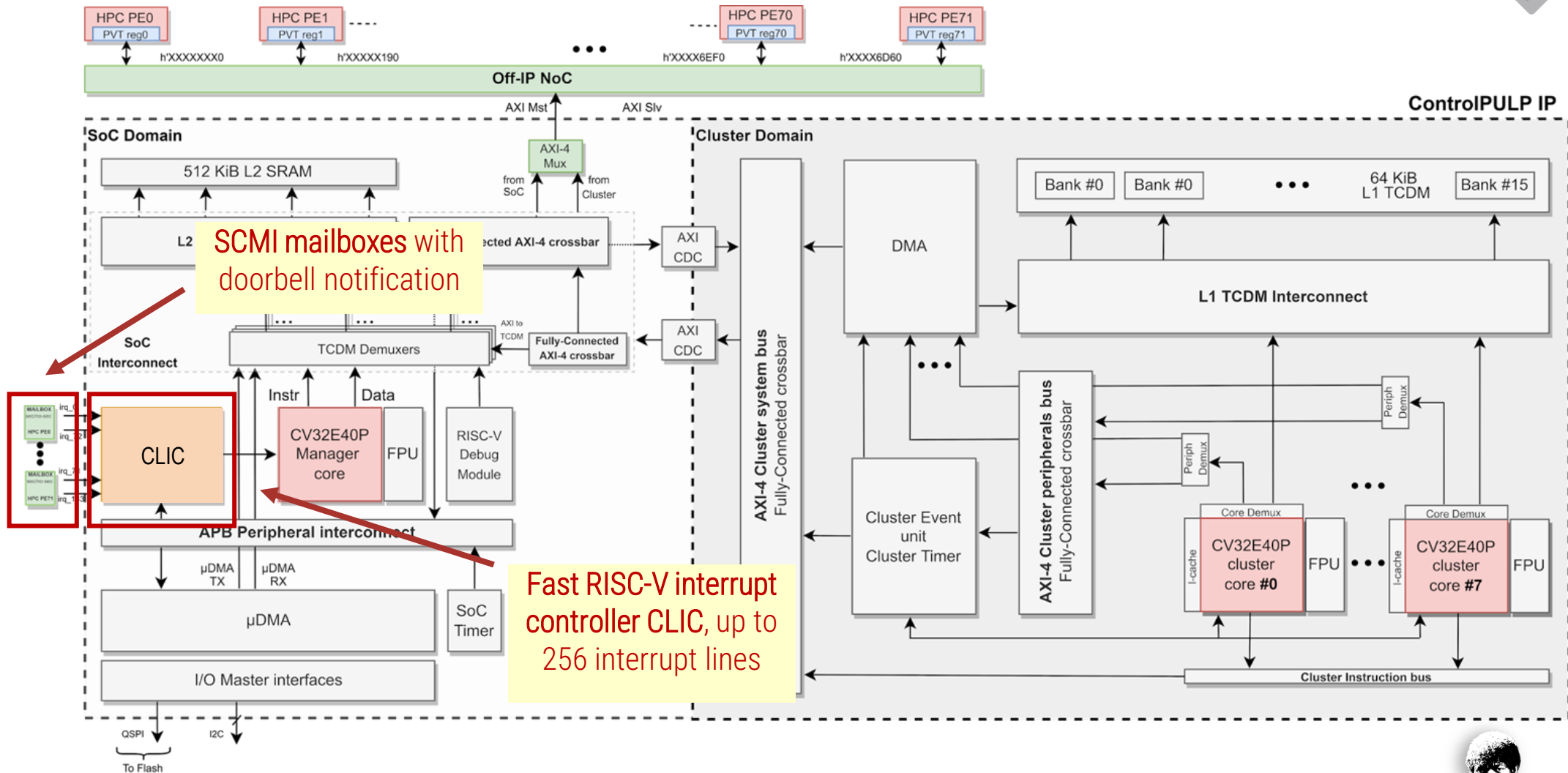
# Architecture



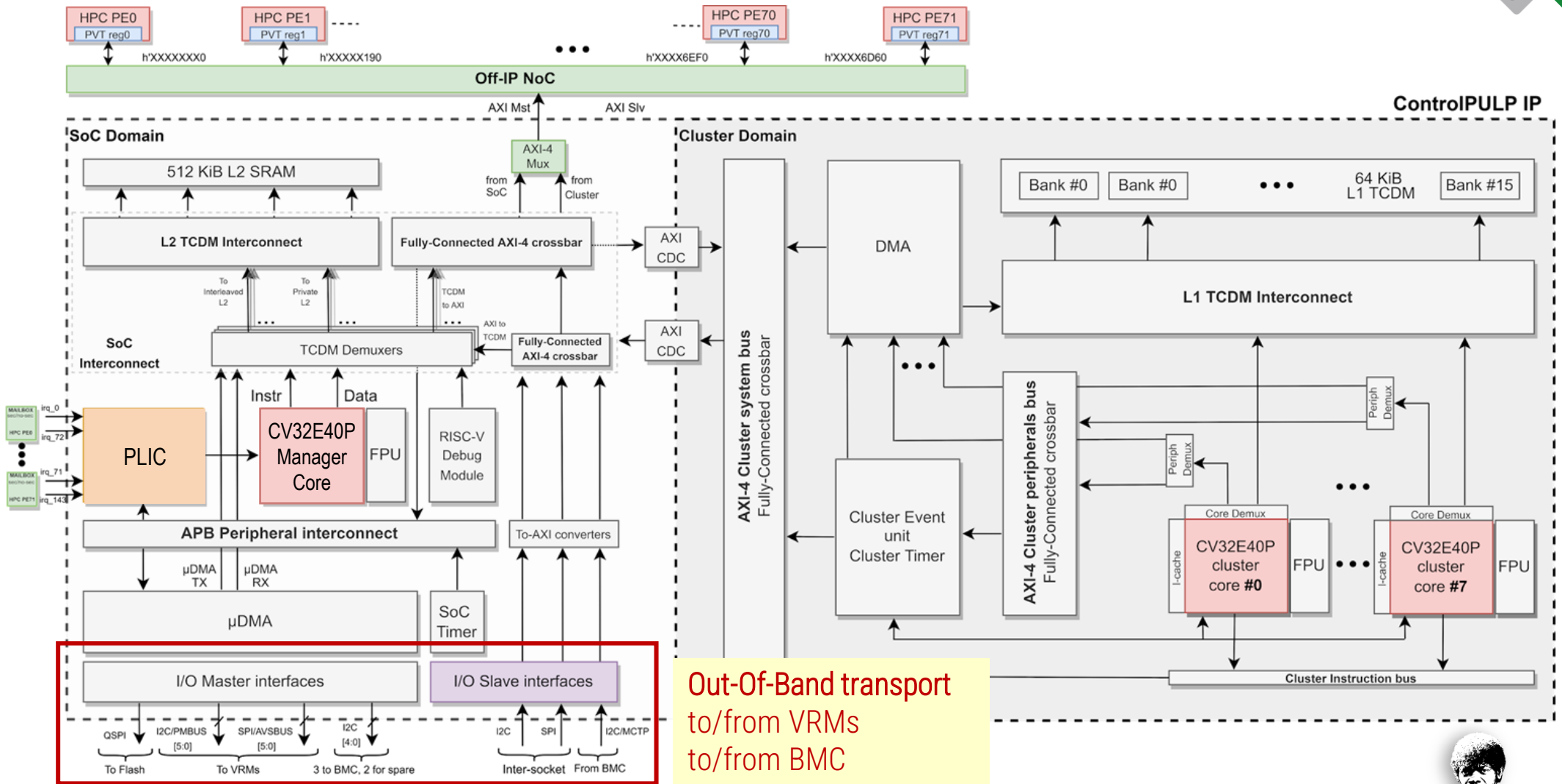
In-Band transport  
to/from PVT sensor registers  
to/from SCMI shared memory



# Architecture

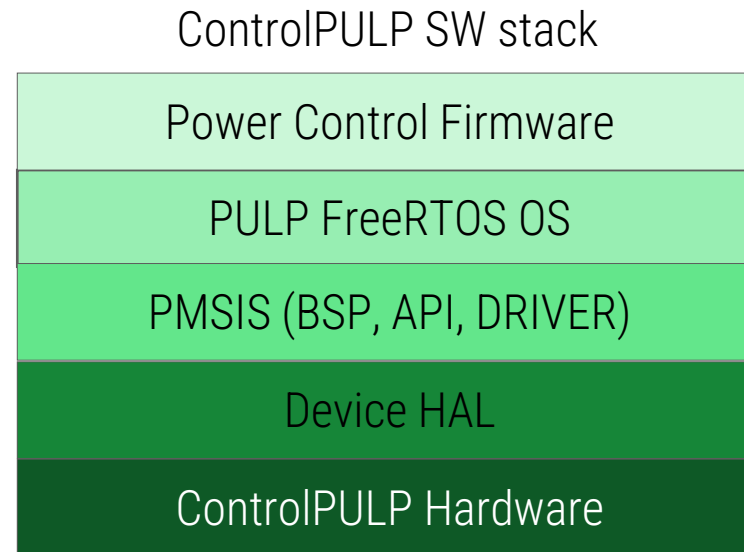


# Architecture

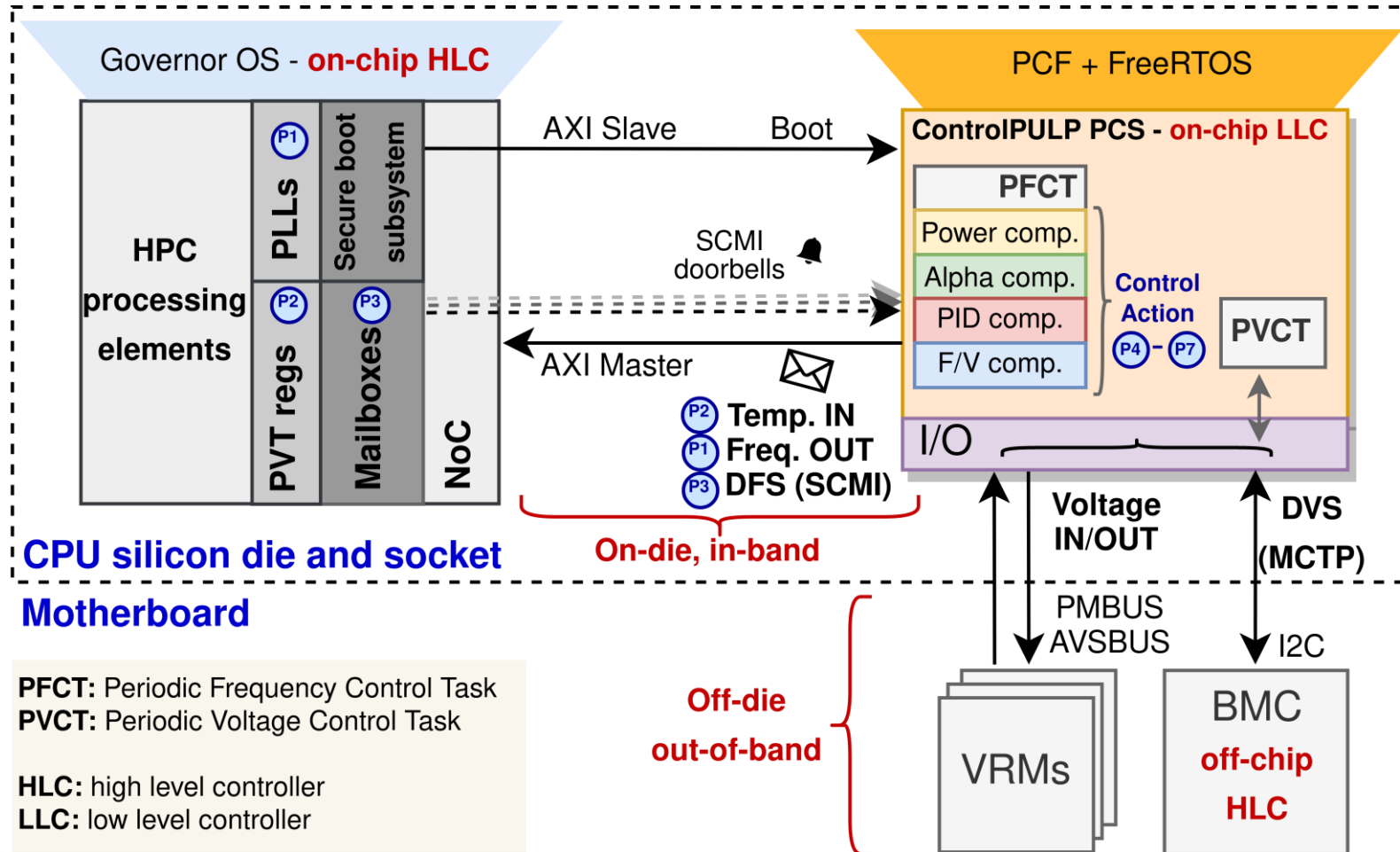


# Software stack

- Complete software stack
- With a Real-time operation system, **FreeRTOS**



# Power Control Policy





# CONTROLPULP CONTROL: NEXT STEPS

- Power-to-Frequency conversion performed with an **iterative-solving algorithm** to convert Power to a **couple** of Frequency-Voltage
  - Voltage and Frequency are reduced **together**
- Safe / Stochastic **MPC**
  - It enforces limits (power and thermal capping) automatically
  - **Optimal**
  - Relies on Model accuracy
  - **Computation** and **Memory** concerns
    - ControlPulp has a **multi-cluster** design
    - Can add MPC **hardware accelerators**
    - Can remap control into **ML** algorithms



# Experimental results



- **Platform area**

- GF22 synthesis: one manager core, one cluster, 512KiB + 64KiB @500 MHz, TT
- Total Area of 9.1 MGE
- Estimated **< 1%** of a HPC processor die in modern technology node

Table 1: ControlPULP post-synthesis area breakdown on GF22FDX technology.

Unit	Area [mm <sup>2</sup> ]	Area [kGE]	Percentage [%]
Cluster unit	0.467	2336.7	25.5
SoC unit	0.135	675.9	7.39
L1 SRAM	0.119	595.7	6.51
L2 SRAM	1.108	5542.1	60.6
<b>Total</b>	<b>1.830</b>	<b>9150.3</b>	<b>100</b>

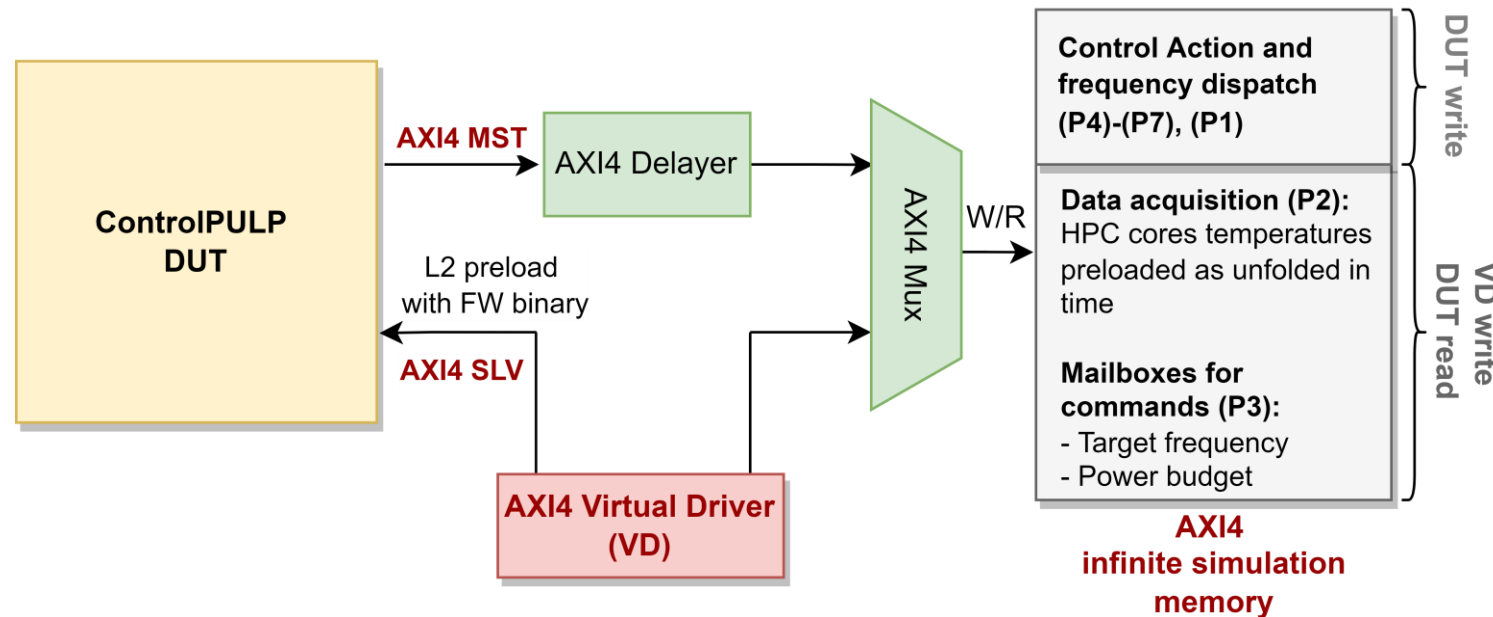


# Experimental results



- **Standalone RTL evaluation**

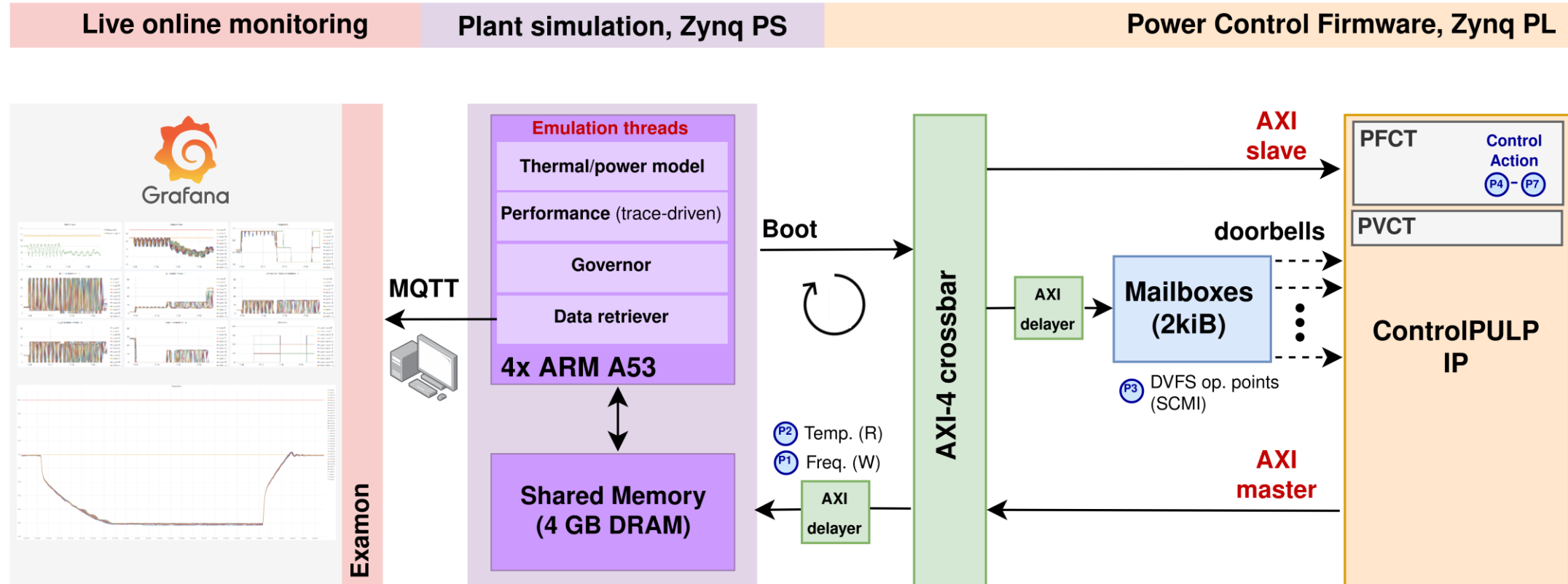
- NoC latency and controlled system are modeled in a testbench environment
- **Evaluate:** multi-core speedup (**performance**) and interrupt handling reactiveness (latency)



# Experimental results

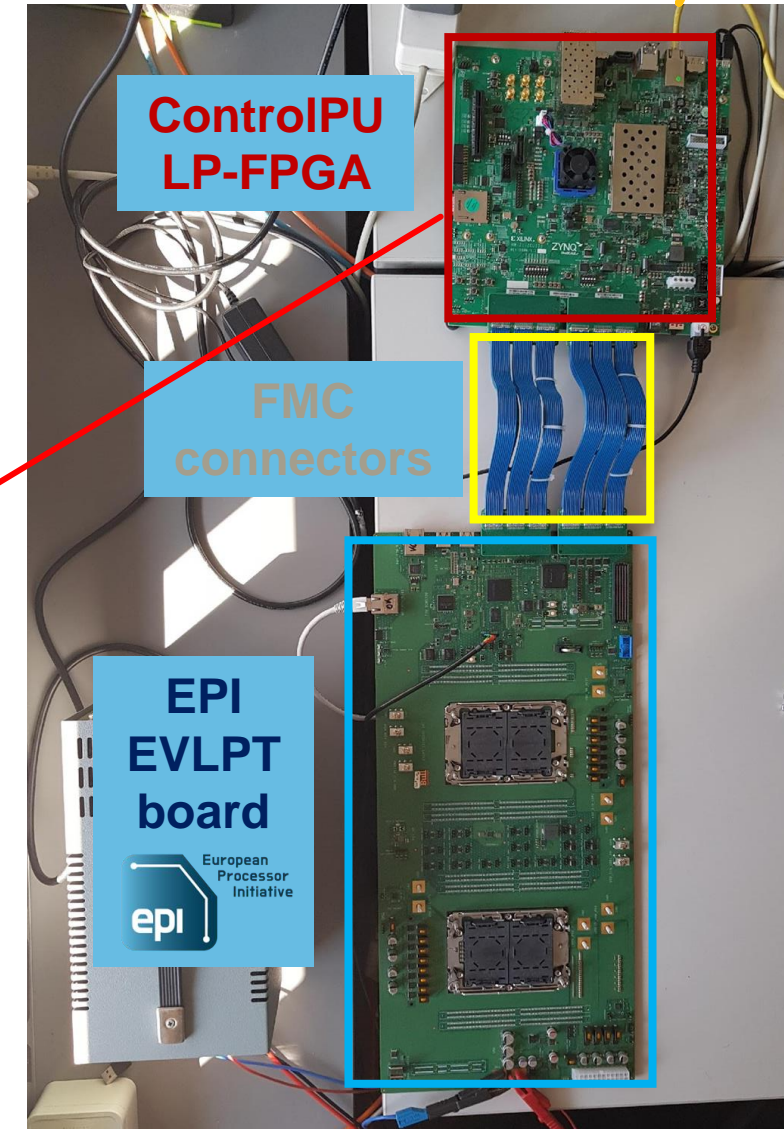
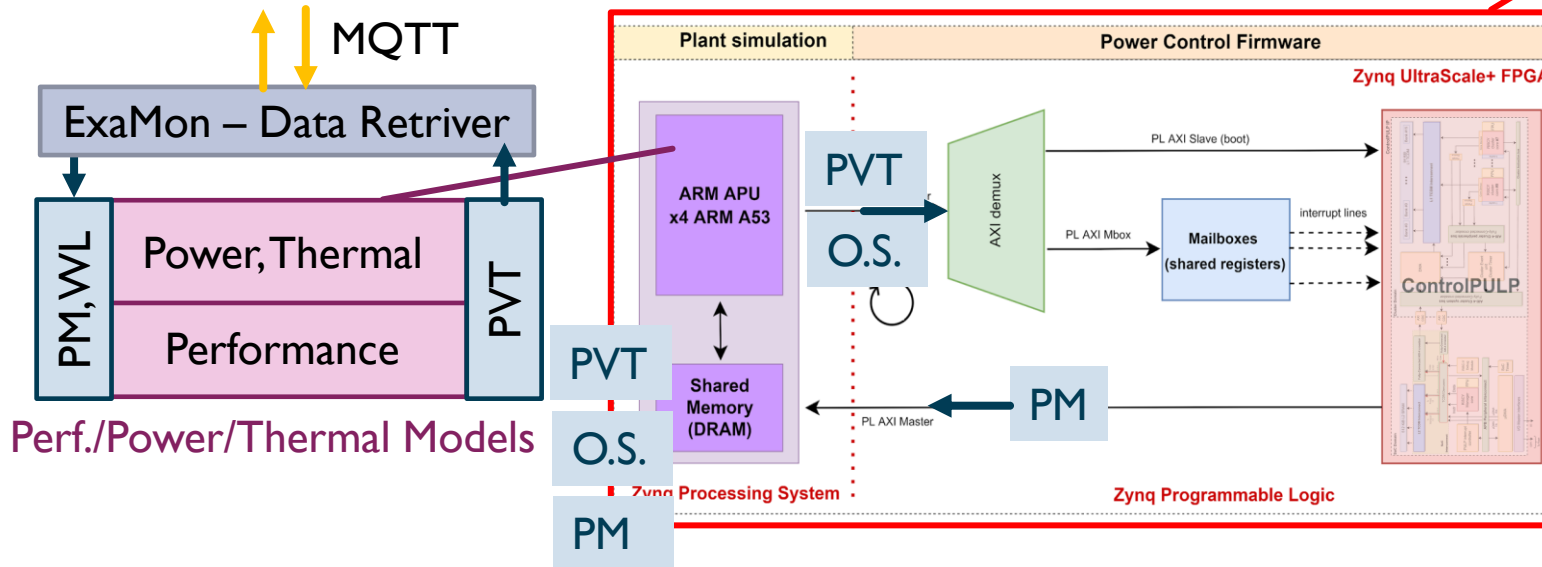
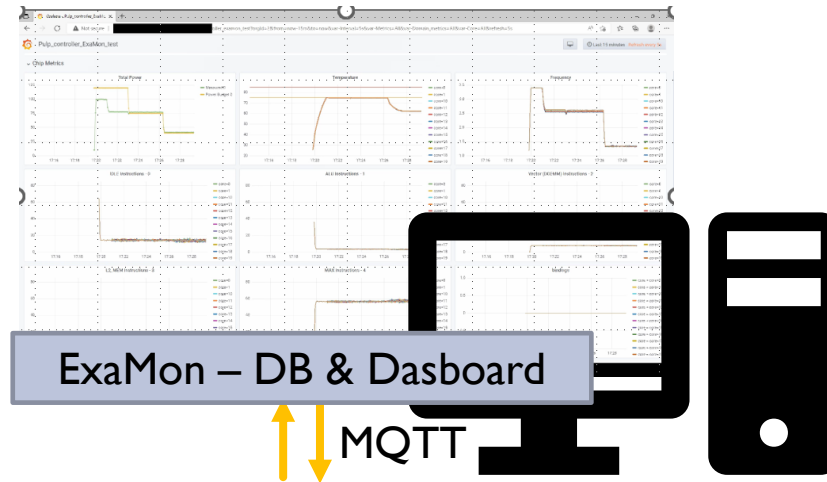


- FPGA-based Hardware-in-the-loop (HIL) emulation framework

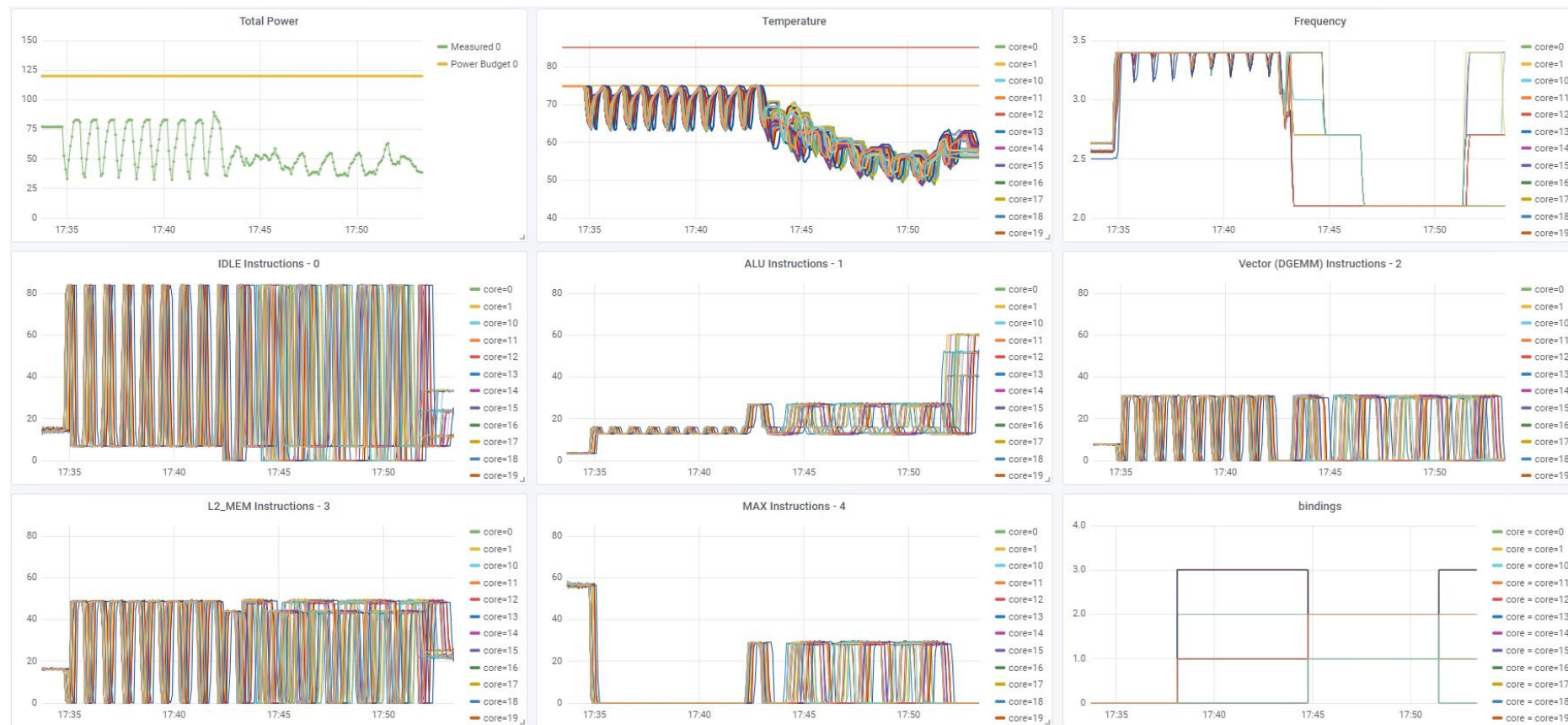


## FPGA-based Hardware-in-the-Loop emulation

- RTL + FW @ FPGA
- PLANT sim. @ A53
- ExaMon integration



# RESULTS





# ROADMAP 2023

- Open-source the Co-Design framework (HW/SW): Q3 2023
- Implement **more advanced and proactive control algorithms** leveraging cluster-based acceleration: predictive policies (e.g. **Model Predictive Control – MPC**)
- Improve emulation framework with realistic off-chip interaction
- Tapeout in TSMC 7nm for **European Processor Initiative (EPI)** - second quarter 2023: Rhea
- Lightweight version of the power controller in EPAC for EPI
- Find new Partners interested in ControlPULP
- Expand the team

# Conclusions



- **First** RISC-V power controller for HPC systems
- **Complete** HW/SW co-design platform to be open-sourced
- **Scalable** multi-core programmable accelerator with low area overhead ( $< 1\%$  of a modern HPC processor)
- Flexible **FPGA-driven thermal, power, performance emulation framework**
- **Fast RISC-V interrupt control** and custom core extensions

## Thank you for your attention

