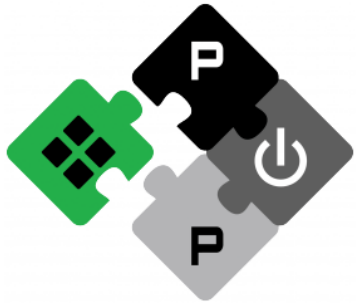


PULP PLATFORM

Open Source Hardware, the way it should be!

RISC-V based Power Management Unit for an HPC processor

Andrea Bartolini <a.bartolini@unibo.it>

Alessandro Ottaviano <aottaviano@iis.ee.ethz.ch>

ETH zürich

 http://pulp-platform.org  @pulp_platform  https://www.youtube.com/pulp_platform

Outline

Power Management in HPC

ControlPULP Hardware and Software Architecture

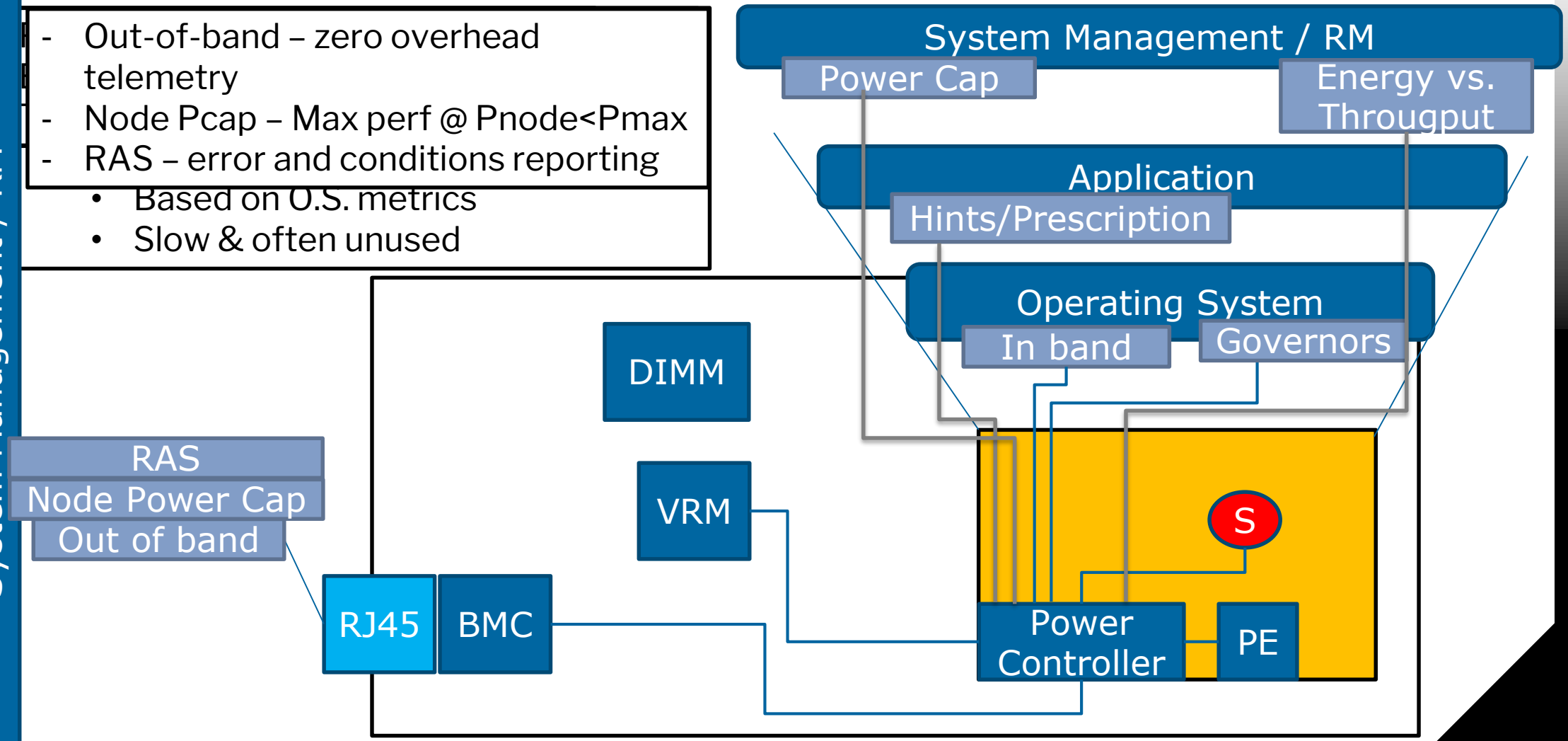
ControlPULP Validation

ETH zürich

HPC Power Management

System Management / RM

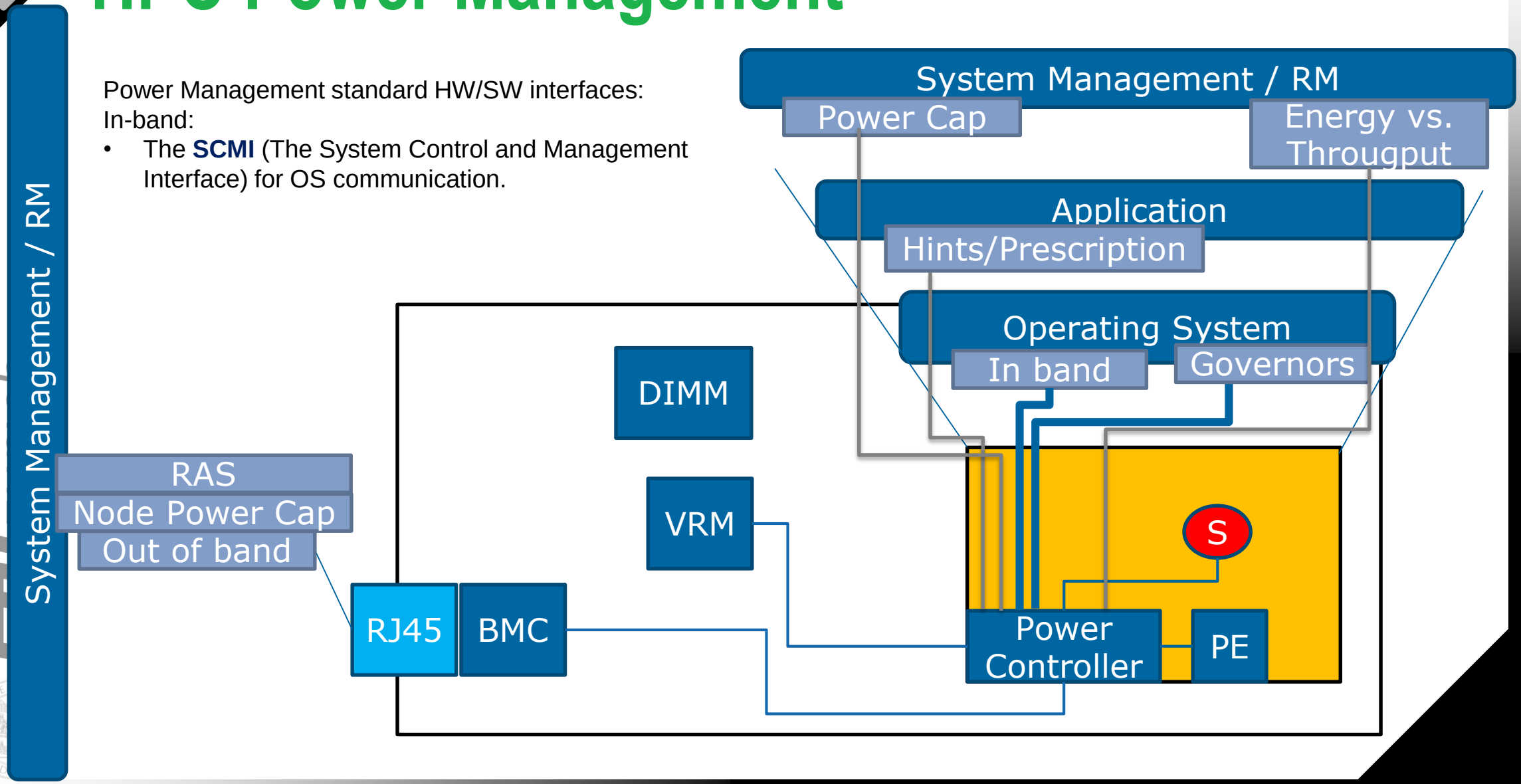
- Out-of-band – zero overhead telemetry
- Node Pcap – Max perf @ Pnode<Pmax
- RAS – error and conditions reporting
  - Based on O.S. metrics
  - Slow & often unused



HPC Power Management

Power Management standard HW/SW interfaces:
In-band:

- The **SCMI** (The System Control and Management Interface) for OS communication.



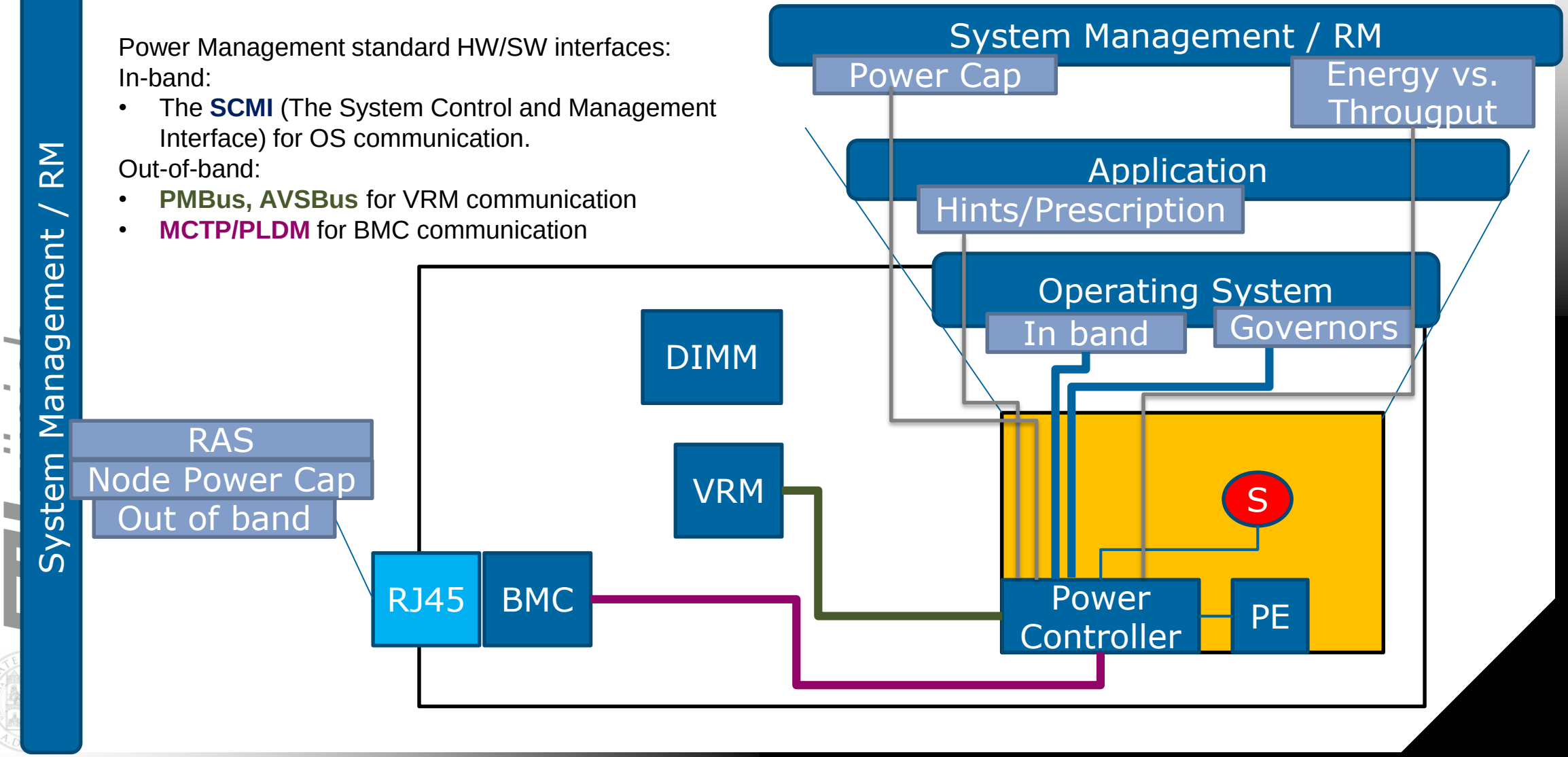
HPC Power Management

Power Management standard HW/SW interfaces:
In-band:

- The **SCMI** (The System Control and Management Interface) for OS communication.

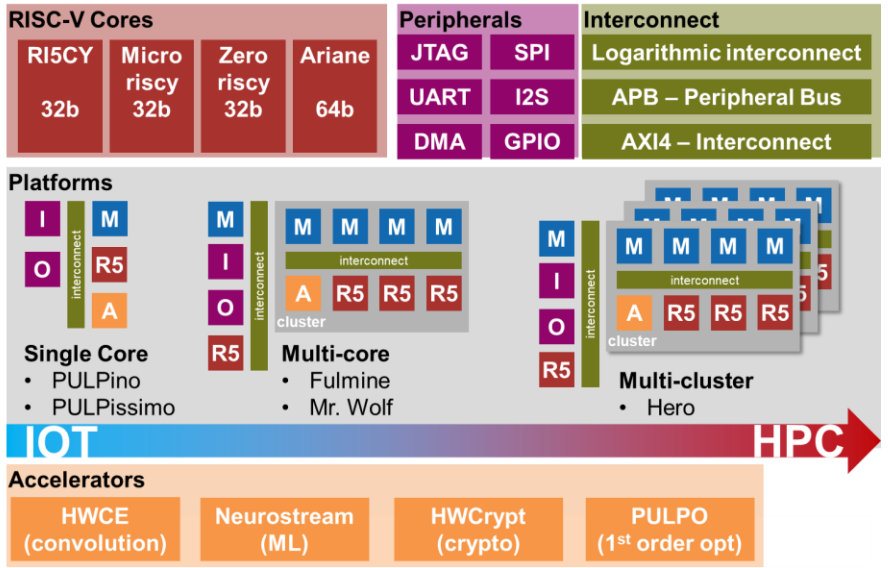
Out-of-band:

- **PMBus, AVSBus** for VRM communication
- **MCTP/PLDM** for BMC communication



On-Chip Power Controller

- Integrated **Power Controller Subsystem (PCS)** for HPC processors
- RISC-V based & open-* (PULP Platform-based), extended to **support standard power management interfaces**
- To be **integrated within Rhea**, EPI first-generation chip family.



Design goals:

- Flexible Power Control Firmware (PCF)
  => Real-time support in hw/sw (low/predictable interrupt latency, FreeRTOS, …)
- Fine-grain power management w. large core count and high efficiency
  => multicore design support w. Packed-SIMD FP support.
- Support of large number of on-chip interfaces
  => Decouple on-chip transfers and computation with DMA-based data movement

Outline

Power Management in HPC

ControlPULP Hardware and Software Architecture

ControlPULP Validation

ETH zürich

Architecture

[1] https://github.com/pulp-platform/pulp

- PULP[1]-based design
- **Scalable architecture**:
  - **Multi-core** cluster with private FPU, up to float16 and bfloat precision
  - RISC-V **fast-interrupt controller: CLIC**
  - DMA for 2-D strided access from PVT sensor registers
- **Industry standard power management interfaces**:
  - **PMBUS**: Voltage Regulators control - slow/multi
  - **AVSBUS**: Voltage Regulators control - fast/p2p
  - **SPI**: Inter-socket communication (Multi ControlPULP)
  - **ACPI/MCTP**: Motherboard/BMC interface (OpenBMC)

Out-Of-Band

  - **SCMI**: OS PM governors and telemetry

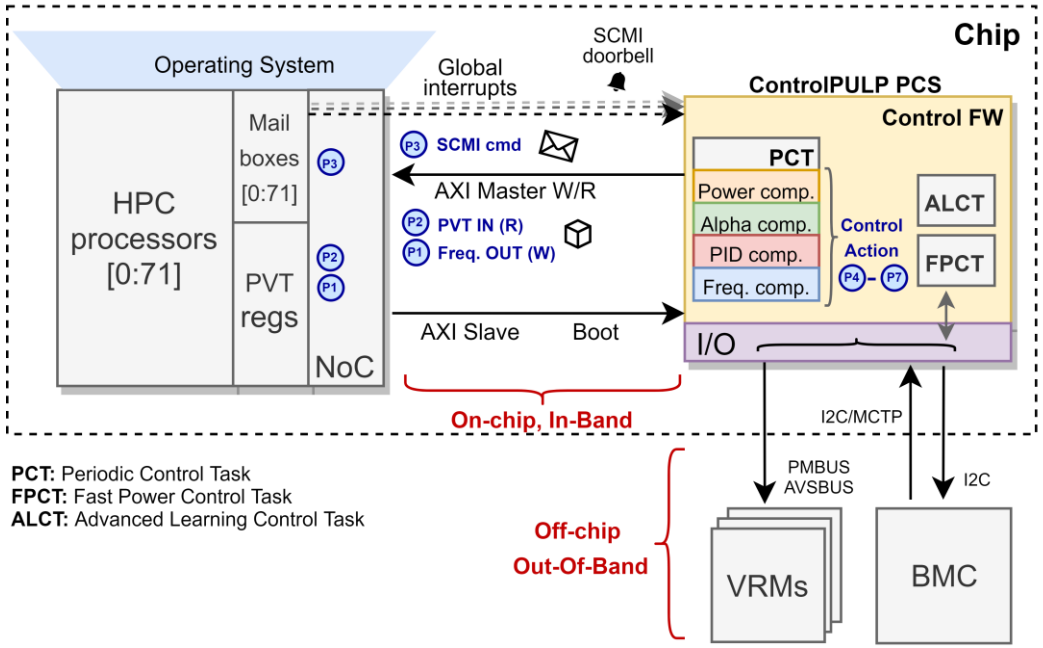
In-Band

Control Firmware

Three main control tasks [2] :

1. **Periodic Control Task (PCT)**
2. **Fast Power Control Task (FPCT)**
3. **Advanced Learning Control Task (ALCT):**

- **Control Action:** computational block
- **In-Band** transfers:
  (i) PVT data gathering- AXI4
  (ii) Doorbell-based SCMI response
- **Out-Of-Band** transfers:
  (i) VRMs power consumption – PMBUS/AVSBUS (I2C/SPI)
  (ii) BMC interaction – I2C/MTCP

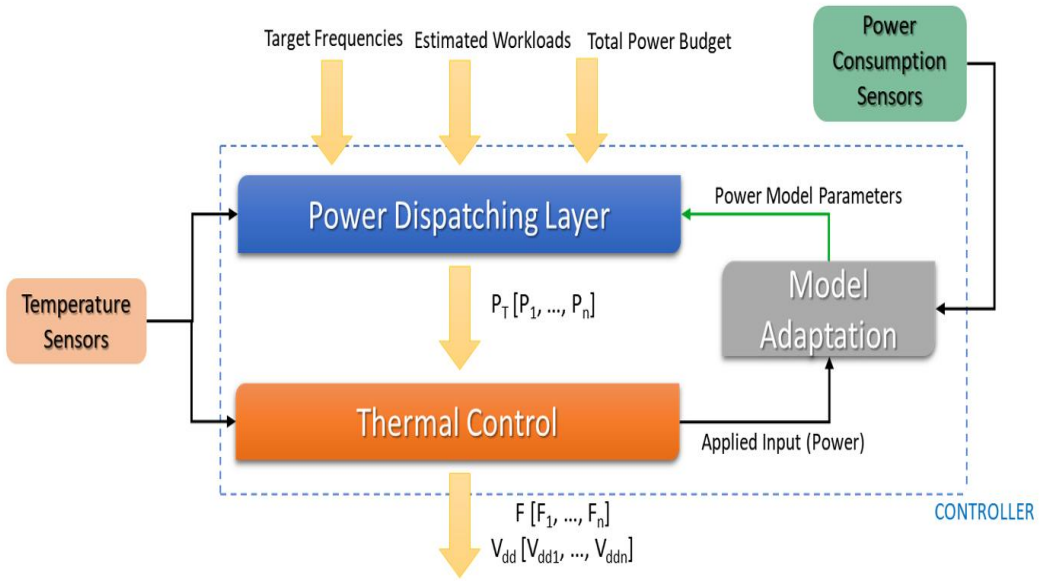


[2] G. Bambini et al., "An Open-Source Scalable Thermal and Power Controller for HPC Processors", 2020

Control Firmware

Three main control tasks [2] :

1. **Periodic Control Task (PCT)**
2. **Fast Power Control Task (FPCT)**
3. **Advanced Learning Control Task (ALCT)**

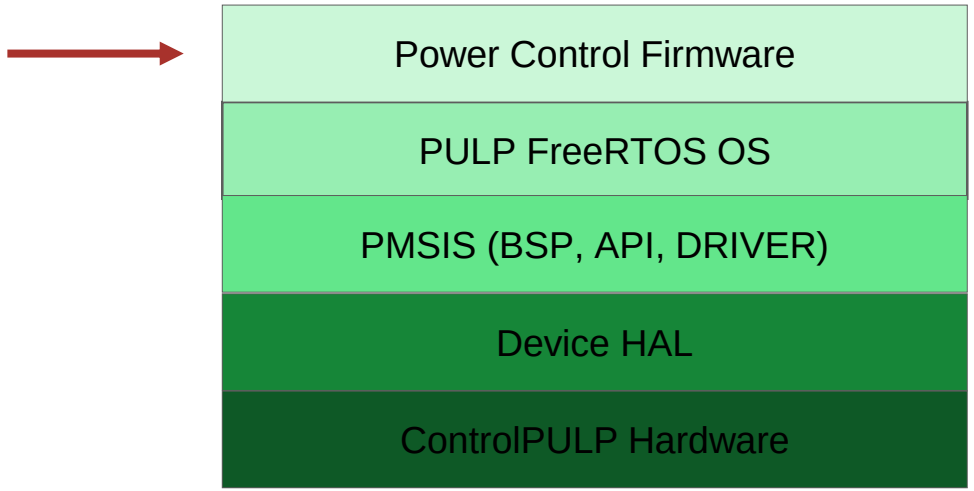


- **Control Action:** computational block
- **In-Band** transfers:
  (i) PVT data gathering- AXI4
  (ii) Doorbell-based SCMI response
- **Out-Of-Band** transfers:
  (i) VRMs power consumption – PMBUS/AVSBUS (I2C/SPI)
  (ii) BMC interaction – I2C/MTCP

[2] G. Bambini et al., "An Open-Source Scalable Thermal and Power Controller for HPC Processors", 2020

Software stack

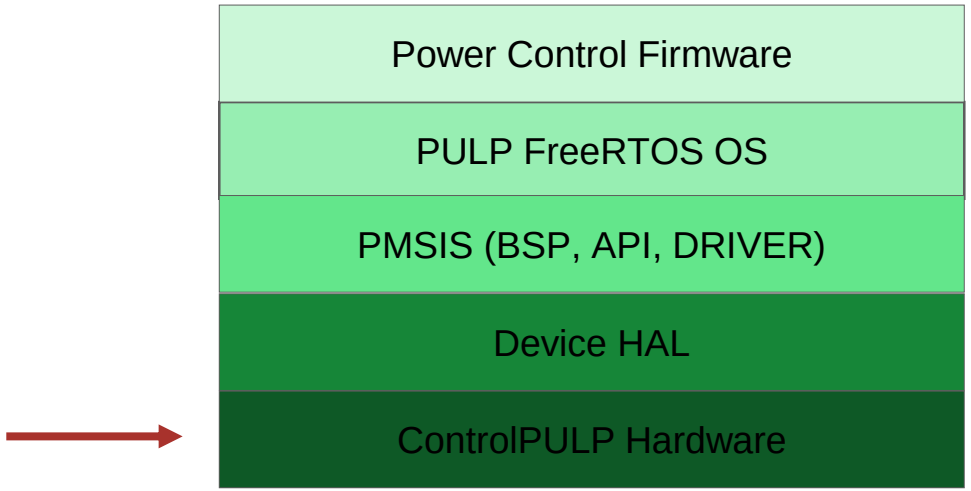
Complete software stack relying on a Real-Time operative system, **FreeRTOS**



ControlPULP SW stack

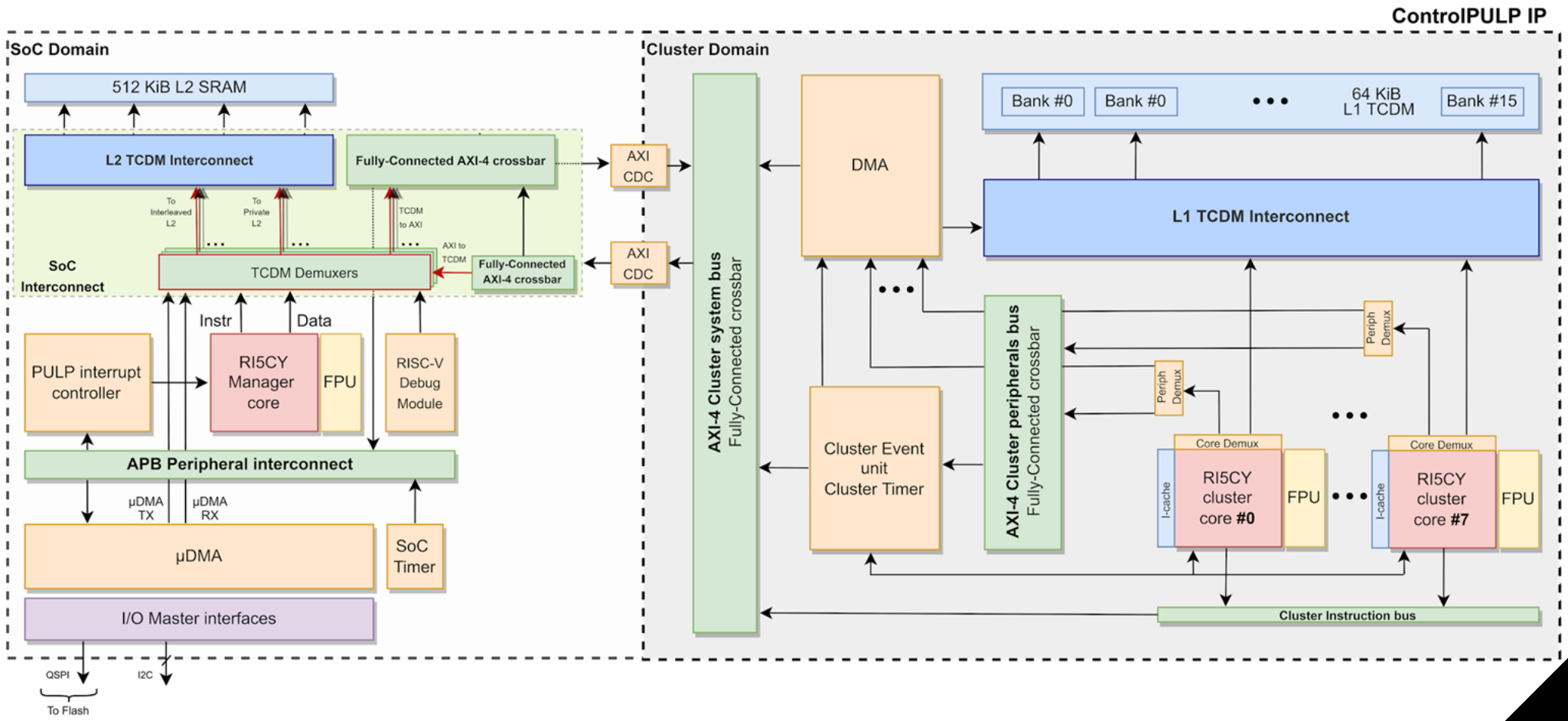
Software stack

Complete software stack relying on a Real-Time operative system, **FreeRTOS**

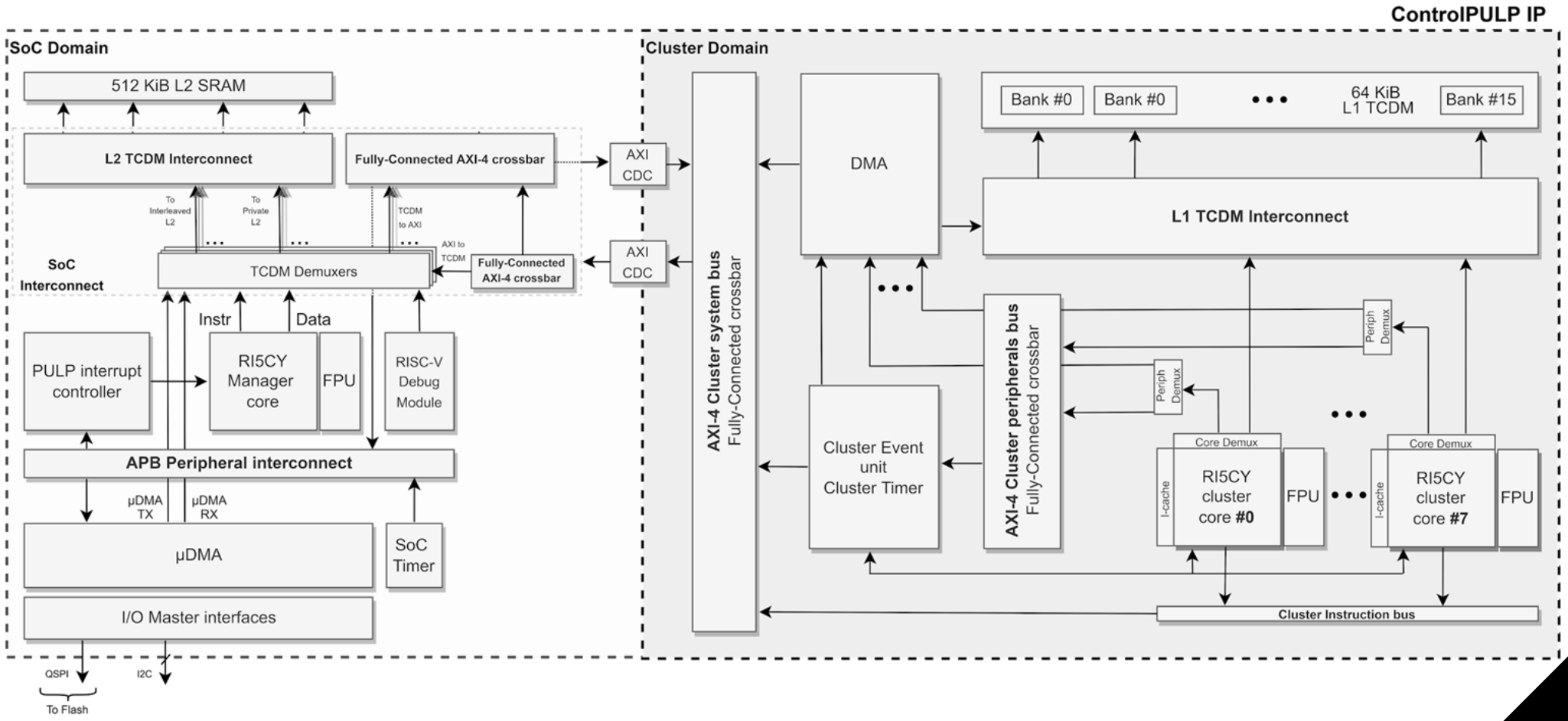


ControlPULP SW stack

Architecture

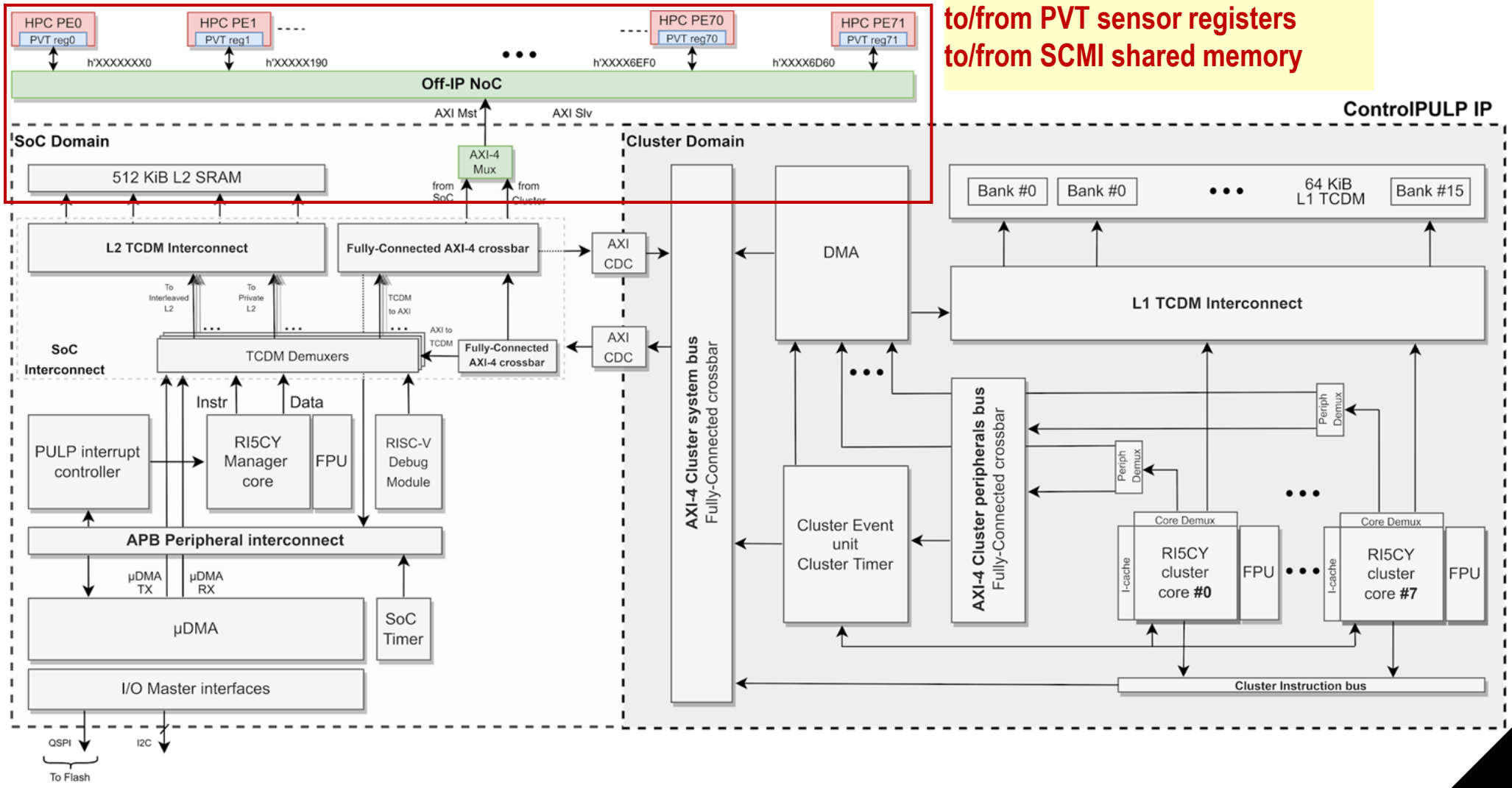


Architecture



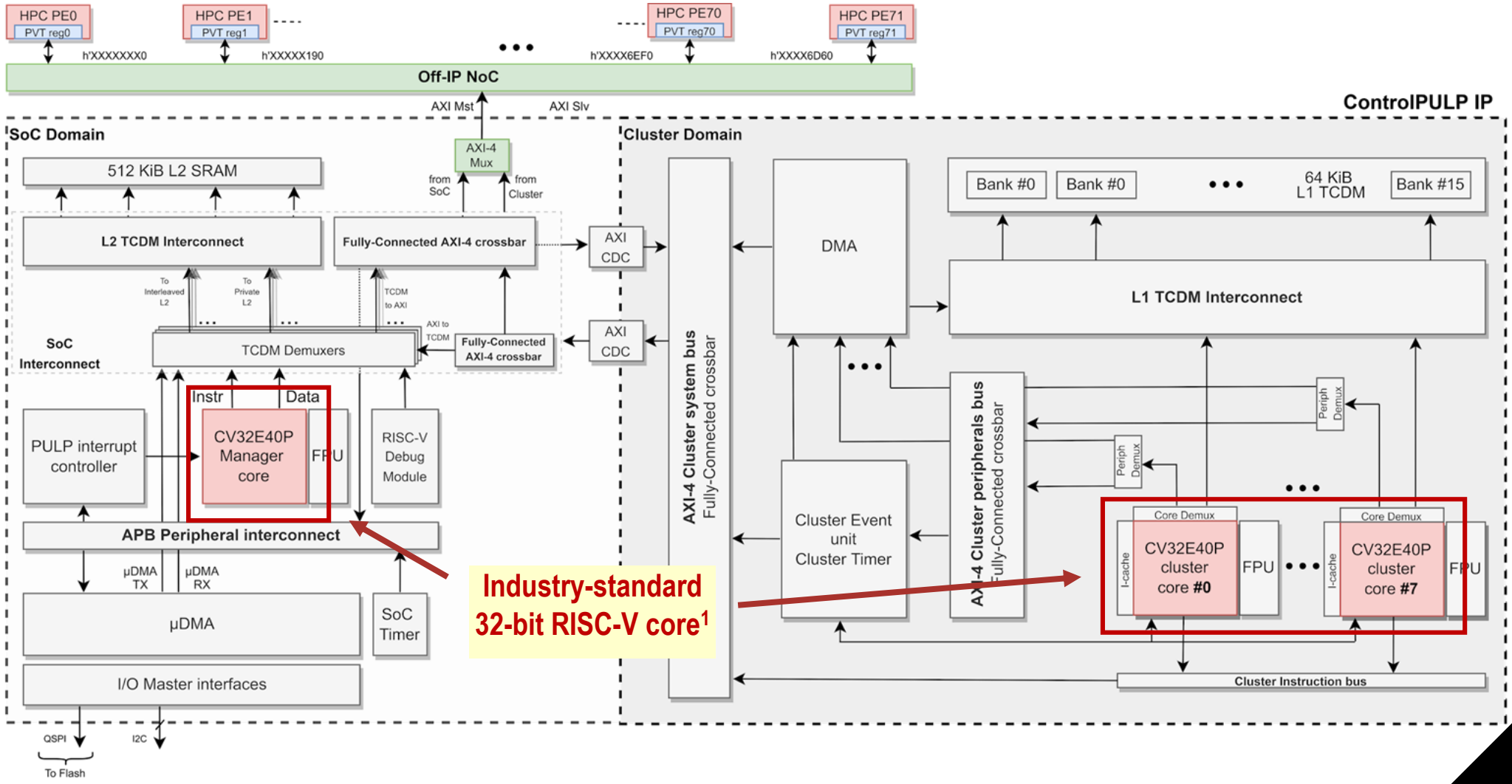
Architecture

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

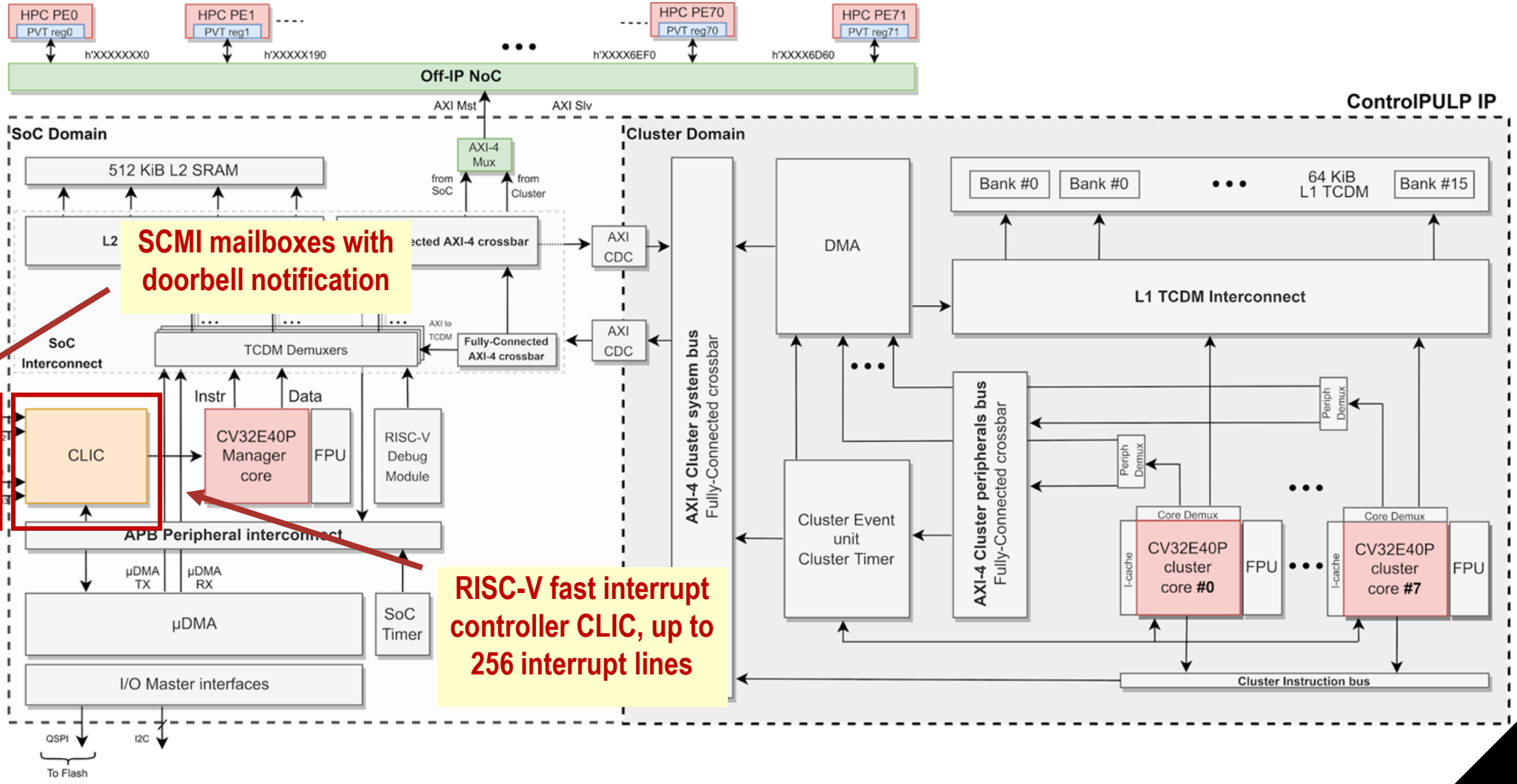


[2] https://github.com/openhwgroup/cv32e40p

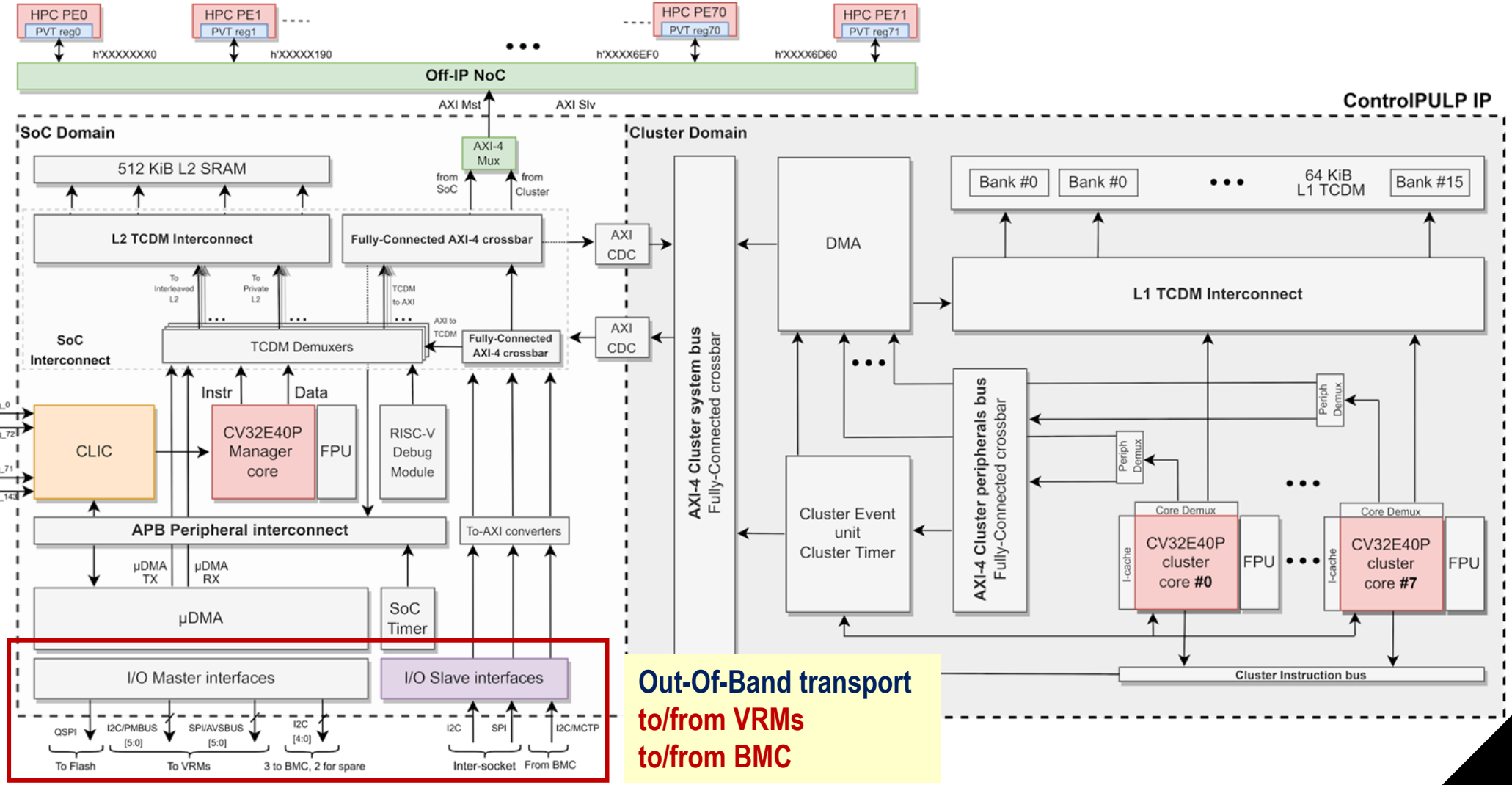
Architecture



Architecture



Architecture



Out-Of-Band transport
to/from VRMs
to/from BMC

Outline

Power Management in HPC

ControlPULP Hardware and Software Architecture

ControlPULP Validation

ETH zürich

ControlPULP validation

[4] N. Bruschi et al., "GVSoC: A Highly Configurable, Fast and Accurate Full-Platform Simulator for RISC-V based IoT Processors", 2021

Standalone RTL validation

- Event-based RTL simulation ecosystem ☑
- GVSOC Architectural simulation[4] ecosystem ☑

Automated Continuous Integration regression check – RTL based

Pipeline Needs Jobs 38 Tests 128

Summary

128 tests | 5 failures | 0 errors | 96.09% success rate | 36445.43s

Jobs

Job	Duration	Failed	Errors	Skipped	Passed	Total
rt_soc_interconnect	92.02s	0	0	0	3	3
rt_coremark	1910.30s	0	0	0	1	1
rt_tcdm	656.33s	1	0	0	2	3
rt_mchan	10923.38s	1	0	0	9	10
rt_i2c_slv_irq	47.05s	0	0	0	1	1
rt_avs	50.43s	0	0	0	1	1
rt_sensors_rx	1648.22s	0	0	0	3	3

ControlPULP validation

[4] N. Bruschi et al., "GVSoC: A Highly Configurable, Fast and Accurate Full-Platform Simulator for RISC-V based IoT Processors", 2021

1. Standalone RTL validation

- GF22 synthesis: 500 MHz, 9.1 MGE
- **Estimated < 1%** of a HPC server processor in modern technology node

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

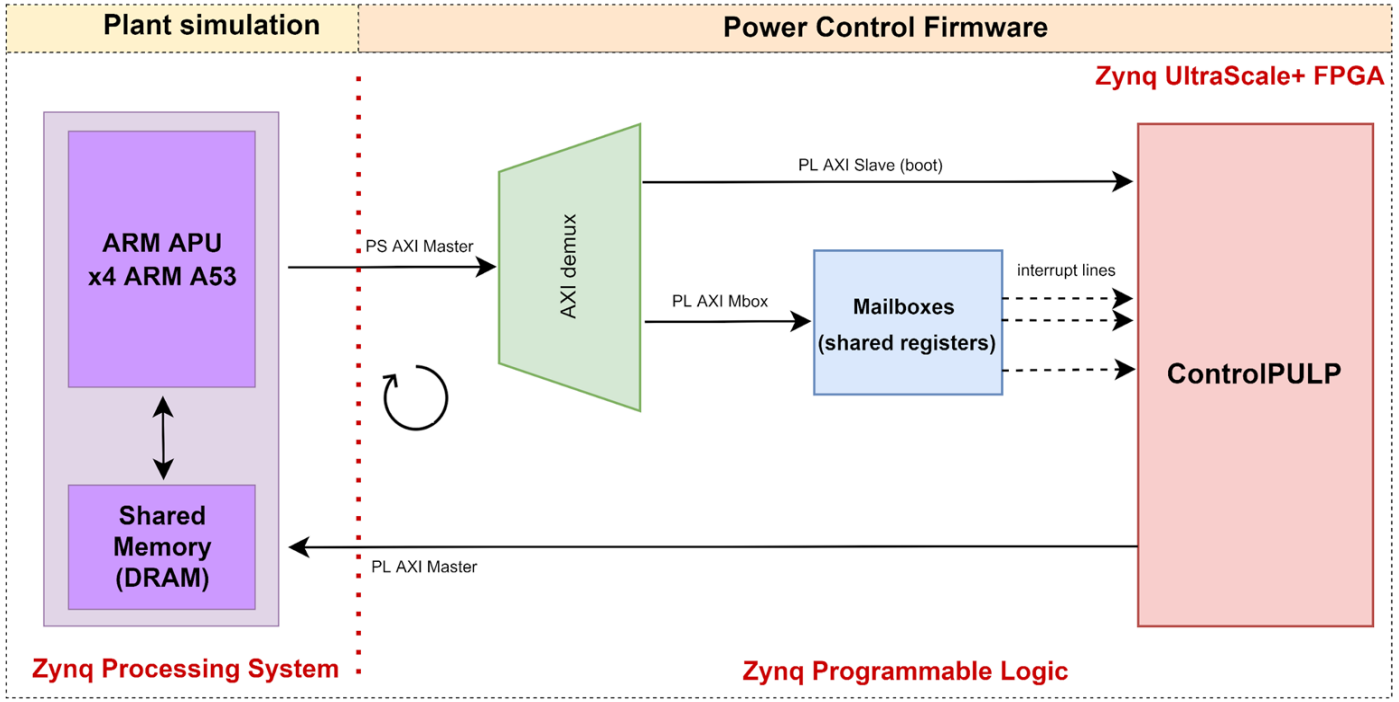
Unit	Area [mm^2]	Area [kGE]	Percentage [%]
Cluster unit	0.467	2336.7	25.5
SoC unit	0.135	675.9	7.39
L1 SRAM	0.119	595.7	6.51
L2 SRAM	1.108	5542.1	60.6
Total	**1.830**	**9150.3**	**100**

[5] https://github.com/pulp-platform/pulp/hero

ControlPULP validation

2. FPGA-based Hardware-in-the-Loop emulation

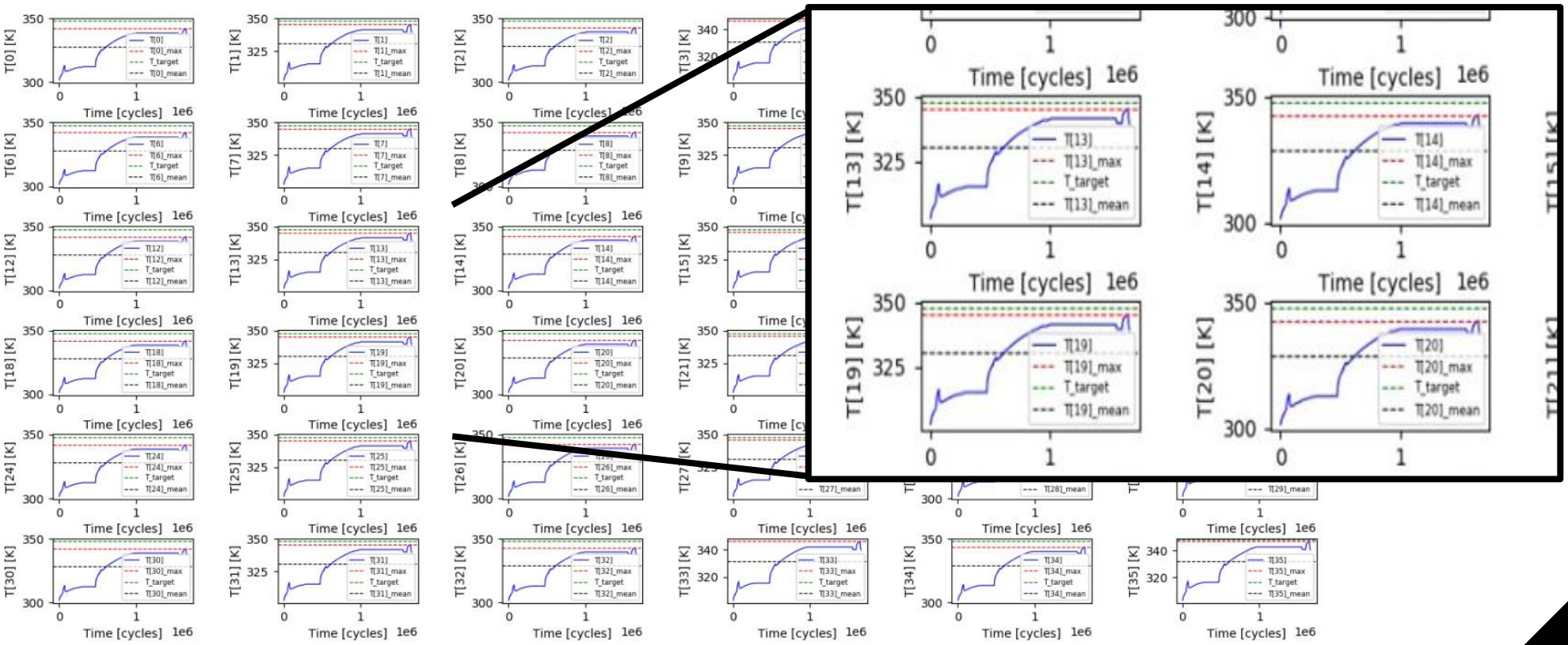
- Cycle-accurate/architectural simulators not suited for
- **Heterogeneous** approach with **FPGA HIL emulation**, based on PULP HERO[5]



ControlPULP validation

2. FPGA-based Hardware-in-the-Loop emulation

- **Real-Time plant emulation:** TDP budget control over 36-cores



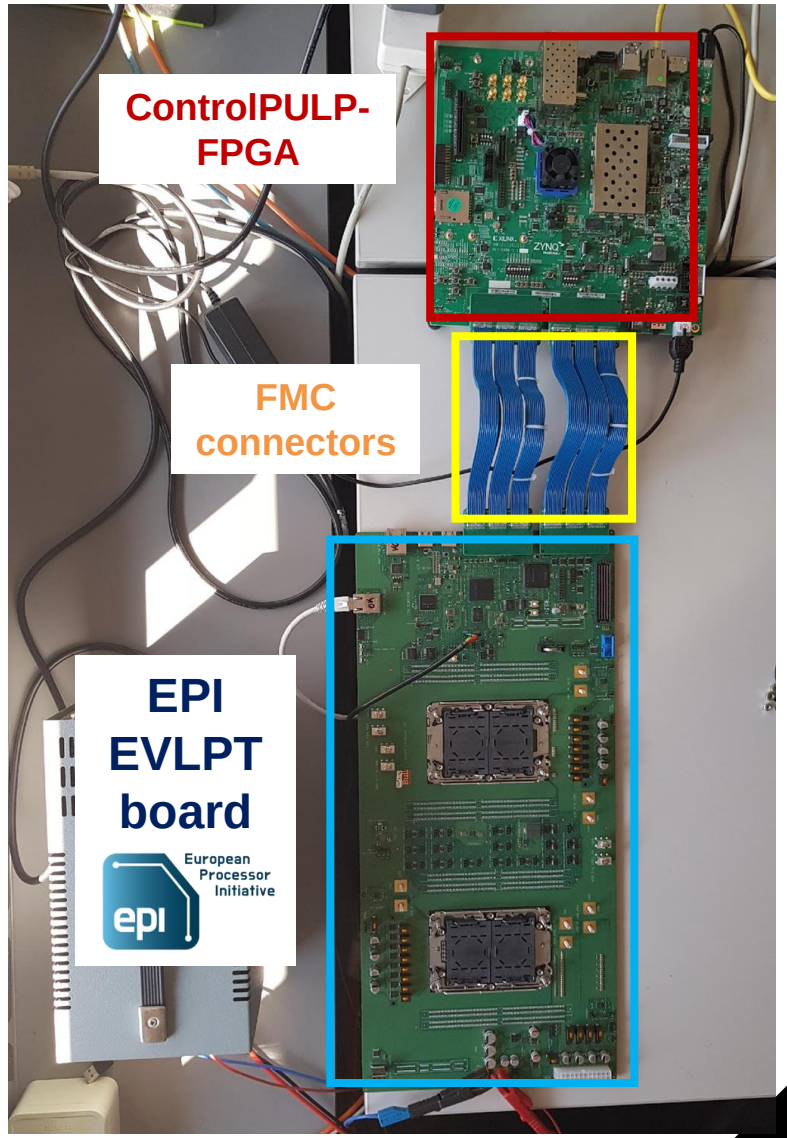
ControlPULP validation

2. FPGA-based Hardware-in-the-Loop emulation

- **EVLPT motherboard from EPI partners:**
  - **Prototype motherboard** for the future Rhea processor
  - VRMs, BMC, Intel FPGA for power sequencing

- **Test off-chip peripherals:**
  1. ACPI power sequencing test ☑
  2. PMBUS test to BMC, VRMs, IBC ☑
  3. I2C Slave (MCTP) test from BMC ☑

  4. AVSBUS test to VRMs control ☑
  5. Inter-socket (Multi ControlPULP) test ☑ **WIP**
  6. More advanced communication ☑



Conclusion

- First RISC-V Power Controller for current and future HPC processors, based on PULP
- Complete HW/SW codesign and validation platform

Roadmap

- Test chip tapeout in 65 nm to further validate the HW
- Multi-FPGA emulation for inter-socket validation
- More advanced and distributed HW/SW power management

Aknowledgment

The ControlPULP Design Team:

- Giovanni Bambini, Robert Balas Corrado Bonfanti, Antonio Mastrandrea, Davide Rossi, Simone Benatti, Luca Benini

The european-project-initiative has received funding from the European High Performance Computing Joint Undertaking (JU) under Framework Partnership Agreement No 800928 and Specific Grant Agreement No 101036168 (EPI SGA2). The JU receives support from the European Union's Horizon 2020 research and innovation programme and from Croatia, France, Germany, Greece, Italy, Netherlands, Portugal, Spain, Sweden, and Switzerland.

The European PILOT project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No.101034126. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Spain, Italy, Switzerland, Germany, France, Greece, Sweden, Croatia and Turkey.

This REGALE-project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 956560. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Greece, Germany, France, Spain, Austria, Italy.

PULP

Parallel Ultra Low Power

Luca Benini, Davide Rossi, Andrea Borghesi, Michele Magno, Simone Benatti, Francesco Conti, Francesco Beneventi, Daniele Palossi, Giuseppe Tagliavini, Antonio Pullini, Germain Haugou, Manuele Rusci, Florian Glaser, Fabio Montagna, Bjoern Forsberg, Pasquale Davide Schiavone, Alfio Di Mauro, Victor Javier Kartsch Morinigo, Tommaso Polonelli, Fabian Schuiki, Stefan Mach, Andreas Kurth, Florian Zaruba, Manuel Eggimann, Philipp Mayer, Marco Guermandi, Xiaying Wang, Michael Hersche, Robert Balas, Antonio Mastrandrea, Matheus Cavalcante, Angelo Garofalo, Alessio Burrello, Gianna Paulin, Georg Rutishauser, Andrea Cossettini, Luca Bertaccini, Maxim Mattheeuws, Samuel Riedel, Sergei Vostrikov, Vlad Niculescu, Hanna Mueller, Matteo Perotti, Nils Wistoff, Luca Bertaccini, Thorir Ingulfsson, Thomas Benz, Paul Scheffler, Alessio Burello, Moritz Scherer, Matteo Spallanzani, Andrea Bartolini, Frank K. Gurkaynak,
and many more that we forgot to mention

http://pulp-platform.org **@pulp_platform**

ControlPULP validation

[4] N. Bruschi et al., "GVSoC: A Highly Configurable, Fast and Accurate Full-Platform Simulator for RISC-V based IoT Processors", 2021

- Event-based RTL simulation ecosystem ☑
- GVSOC Architectural simulation[4] ecosystem ☑

- **Multi-core and DMA centric PCF speedup: 5x than single-core execution**