ControlPULP: A Multi-Core RISC-V Power Controller for HPC Processors

Alessandro Ottaviano*, Robert Balas*, Giovanni Bambini†, Davide Rossi†, Luca Benini*†
Andrea Bartolini†
*ETH Zürich, †University of Bologna

After the end of Dennard's scaling, the increase in power density has become an undesired but unavoidable collateral effect of the performance gain obtained with technological scaling. This trend has made the processing elements at the heart of computing nodes energy, power, and thermally constrained. Modern high-performance processors feature a large number of cores, such as AWS Graviton 2 (64 ARM Neoverse N1 cores), Intel Alder Lake-S Xeon (16 cores, 24 threads), AMD Epyc 7003 Milan (up to 64 Zen 3 cores), SiPearl Rhea Processor (72 ARM Neoverse V1 Zeus cores) and the NVIDIA Grace CPU Superchip (144 ARM Neoverse N2 cores). Their application workload requires a dynamic trade-off between maximum performance and energy efficiency (energy-aware CPU [1]). Hence, all modern processors integrate on-die Power Controller Subsystems as dedicated hardware resources co-designed with a Power Control Firmware (PCF) implementing complex MIMO power management policies. The latter involves embedding and interleaving a plurality of activities in the PCS, namely (i) dynamic control of the CPU power consumption with short time constants to prevent thermal hazards and to meet the TDP limit (power capping), (ii) real-time interaction with inputs provided by on-die (Operating System - OS - power management interfaces and on-chip sensors) and off-die (Baseboard Management Controller - BMC -, Voltage Regulator Modules - VRMs -) units and (iii) dynamic power budget allocation between general-purpose (CPUs) and other integrated subsystems, such as graphics processors (GPUs) [1].

Existing on-die PCSs share a common design structure with an integrated single-core microcontroller [1] supported by dedicated hardware state machines or more generic accelerators [1]. Many-core power management requires fine-grained control of the operating points of the processing elements [2] to meet a given processor power consumption setpoint while minimizing performance penalties. Moreover, the control policy (Power Control Firmware - PCF) has to provide fast, reactive, and predictable responses to promptly handle the incoming requests from the OS or BMC and prevent thermal hazards. A flexible and scalable way to manage these computationally intensive operations must provide a high-quality control performance per core and support more advanced experimental control policies. This scenario suggests the need for a performant and capable PCS architecture optimized for handling a fine-grained, per-core performance state control strategy on a large number of controlled cores within the required timing deadlines.

In this work, we present ControlPULP, an end-to-end RISC-V parallel PCS architecture based on open RISC-V cores and hardware IPs. To the best of our knowledge, ControlPULP is the first fully open-source (hardware and software) PCS with a configurable number of cores and hardware resources to track the computational requirements of the increasingly complex power management policies of current and future high-performance processors. The proposed design has the following contributions:

1) ControlPULP integrates a multi-core cluster with per-core FPUs for reactive control policy, achieving 4.9x speedup on the PCF against the single-core configuration.
2) The cluster integrates a specialized DMA to accelerate the data acquisition from on-chip sensors and off-chip peripherals, with 2D strided access patterns capability.
3) We tailor ControlPULP to meet real-time power management requirements. The architecture achieves low interrupt latency thanks to a platform-level interrupt controller (RISC-V PLIC) tasked to process the global interrupts associated with OS- and BMC- driven commands and a low latency predictable interconnect infrastructure.
4) We performed a design space exploration evaluating the cost/benefit trade-off of the hardware extensions. We compared the end-to-end capabilities of ControlPULP with a case study on the control quality of the PCF against the only openly documented SoA control policy implemented by IBM on-chip controller [2], achieving 6% more precise setpoint tracking.

Future works will improve the control performance by introducing advanced control algorithms and validating the design in silicon with the integration in a real HPC processor. Furthermore, hardware and software extensions will be evaluated with FPGA-based HIL emulations.

References

[1] E. Rotem, A. Naveh, A. Ananthakrishnan, E. Weissmann, and D. Rajwan, "Power-management architecture of the intel microarchitecture code-named sandy bridge," *IEEE Micro*, vol. 32, no. 2, pp. 20–27, 2012.
[2] G. LLC, "Power management for multiple processor cores," U.S. Patent US8402290B2, Dec. 2020.

[1] https://github.com/ARM-software/SCP-firmware

[2] https://github.com/open-power