A RISC-V Heterogeneous SoC for Embedded Devices

Luca Valente*, Mattia Sinigaglia*, Yvan Tortorella*, Davide Rossi*, Luca Benini*†
*DEI, University of Bologna, Italy** *IIS lab, ETH Zurich, Switzerland†*
{luca.valente, mattia.sinigaglia5, yvan.tortorella, davide.rossi}@unibo.it {lbenini}@iis.ee.ethz.ch

***Abstract*—Heterogeneous embedded systems on chip (HESoCs) are used in various application domains to combine general-purpose computing with domain-specific, efficient processing capabilities. Such architectures present dedicated hardware accelerators along with a general-purpose host processor. Programmable manycore accelerators (PMCAs) are often utilized in HESoCs as they offer high energy efficiency and ease of use at the same time. Indeed, leading companies continue to advance their heterogeneous products [1][2]. On the other hand, HESoCs with PMCAs have been introduced by researchers to explore the design space [3][4] or as real SoCs [5]. This work presents a fully RISC-V heterogeneous SoC with a 64-bit Linux-capable host CPU and a PMCA composed of 8 32-bit RISC-V cores enhanced with ML and DSP extension.**

I. Motivation and methodology

Embedded devices have significant computational capabilities nowadays. They run intensive workloads and process significant amount of data gathered from multiple high-bandwidth sensors. IoT processors need to feature flexible accelerators to achieve high energy efficiency in multiple applications. At the same time, these devices interact with the physical word and hence require precise, predictable control. Thus, intensive data processing and real-time control must coexist on the same platform. Therefore, different kinds of processors are needed. Starting from these requirements, we developed a HESoC with a host CPU with four privilege levels and memory virtualization support that can run full-featured OS assisted by a PMCA for computationally intensive workloads.

This work has been made possible only thanks to open-source hardware and software. On the hardware side, the HeSoC is composed of well known open source IPs, such as CVA6 [6] as host core and the Parallel Ultra Low Power cluster [7] as accelerator. On the software side, we will leverage the HERO project [4]. CVA6 will run a lightweight Linux version and NuttX on top of the BAO hypervisor. User-level applications will be able to offload computation on the cluster when needed.

The project's goal is to provide an open-source IoT processor that combines full-fledged Operating Systems with a low power programmable accelerator to achieve high energy efficiency while at the same time offering a standard Linux-based software environment.

II. Current development

The project kicked off in May 2021. Mainly, two methodological approaches aided the development: RTL simulation and FPGA prototyping. We had an initial version of the system working on the Xilinx VCU118 board by the end of July. From that point on, we carried out a massive testing and integration effort that will end up with a $9mm^2$ tape-out in Global Foundries 22-nanometer FDSOI technology in June. The implemented chip will have four frequency domains: CVA6, the cluster, the host domain and the peripherals. CVA6 is expected to work up to 900 MHz, while the cluster and the host domain will work around 400 MHz. Such a configuration will allow us to reach more than 1GOPs.

To validate the RTL, we set up a proper Linux boot flow on the VCU118 and integrated many machine-level tests stressing the cluster. Two ongoing primary efforts are tape-out-oriented modifications and software development for lightweight host to cluster offload and data sharing.

III. Conclusion and future work

In conclusion, even though the project is still in the early development stages, it can contribute to research and the open-source hardware landscape. After the tape-out, we will focus on the software stack and acceleration of the targeted applications, leveraging the communications overheads between host and cluster and possibly enhancing the chip capabilities with on-chip accelerators.

References

[1] "NVIDIA's Tegra K1," 2022. [Online]. Available: https://developer.nvidia.com/embedded/buy/tegra-k1-processor.

[2] "Snapdragon 8 Gen 1 Mobile Platform," 2022. [Online]. Available: https://www.qualcomm.com/products/snapdragon-8-gen-1-mobile-platform.

[3] J. Balkind *et al.*, "Byoc: A "bring your own core" framework for heterogeneous-isa research," in *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 699–714, ISBN: 9781450371025. [Online]. Available: https://doi.org/10.1145/3373376.3378479.

[4] A. Kurth *et al.*, "Herov2: Full-stack open-source research platform for heterogeneous computing," *CoRR*, vol. abs/2201.03861, 2022. arXiv: 2201.03861. [Online]. Available: https://arxiv.org/abs/2201.03861.

[5] A. Pullini *et al.*, "A heterogeneous multicore system on chip for energy efficient brain inspired computing," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 8, pp. 1094–1098, 2018. DOI: 10.1109/TCSII.2017.2652982.

[6] "CVA6," 2022. [Online]. Available: https://github.com/openhwgroup/cva6.

[7] "PULP cluster," 2022. [Online]. Available: https://github.com/pulp-platform/pulp_cluster.