Supporting Differentiated Services in Computers via Networking Technologies

Jiuyue Ma, Xiufeng Sui, Yupeng Li, Zihao Yu, Bowen Huang, Yungang Bao

State Key Laboratory of Computer Architecture, ICT, CAS

ABSTRACT

Contemporary data centers confront with challenges in managing the trade-offs between resource utilization and applications' quality of services (QoS). To resolve this issue, as suggested in the community white paper **21st Century Computer Architecture** [1], computer architecture needs to provide new, higher-level interfaces beyond a conventional instruction set architecture (ISA) to convey an application's QoS requirements to the hardware.

This work proposes PARD, a programmable architecture for resourcing-on-demand that provides such a new programming interface. PARD is inspired by the observation that a computer is inherently a network in which hardware components communicate via packets (e.g., over the NoC or PCIe). So we can apply networking technologies, e.g., principles of software-defined networking (SDN) [3], to this intra-computer network.

PARD addresses three major technical challenges. First, to deal with the semantic gap between high-level applications and underlying hardware packets, PARD attaches a high-level semantic tag (e.g., a virtual machine or thread ID) to each memory-access, I/O, or interrupt packet. Second, to allow a variety of hardware components to be programmed and managed via a common interface, PARD designs programmable control planes that can be integrated into various shared resources (e.g., cache, DRAM, and I/O devices) and can differentially process packets according to tag-based rules. Third, to facilitate programming, PARD abstracts all control planes as a device file tree to provide a uniform programming interface via which users create and apply tag-based rules.

With these efforts, PARD enables new functionalities like fully hardware-supported virtualization and differentiated services in computers.

More details are described in this paper [2].

BODY

A computer is inherently a network. It is promising to apply networking technologies like SDN to computer architecture.

REFERENCES

[1] Computing Community Consortium (CCC). 21st century computer architecture. *A community white paper*, 2012. http://cra.org/ccc/docs/init/21stcenturyarchitecturewhitepaper.pdf.

[2] J. Ma, X. Sui, N. Sun, Y. Li, Z. Yu, B. Huang, T. Xu, Z. Yao, Y. Chen, H. Wang, L. Zhang, and Y. Bao. Supporting differentiated services in computers via programmable architecture for resourcing-on-demand (pard). In *the 20th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2015.

[3] ONF. Software-Defined Networking.
https://www.opennetworking.org/sdn-resources/sdn-definition.