FiPS and M2DC: Novel Architectures for Reconfigurable Hyperscale Servers

Rene Griessl, Meysam Peykanu, Lennart Tigges, Jens Hagemeyer, and Mario Porrmann
Center of Excellence Cognitive Interaction Technology, Bielefeld University, Germany

Extended Abstract

1. Introduction – FiPS

Within the EU projects FiPS (FP7) and M2DC (H2020) we aim at significantly increasing the energy-efficiency of compute platforms for cloud and high performance computing. With the RECS Box System, a highly scalable heterogeneous hardware platform is developed, which seamlessly integrates CPUs, embedded CPUs, FPGAs, GPUs and many-core processors. To ease programming of the platform, FiPS is setting up a programming methodology, simplifying the usage of the heterogeneous computing devices as processing elements in a holistic integrated hardware and software server eco-system.

2. RECS|Box Server Platform

The RECS®|Box system is designed and manufactured by Bielefeld University and Christmann Informationstechnik+Medien. Utilizing the RECS®|Box platform concept enables a seamless integration of general purpose processors, embedded processors, FPGAs, GPGPUs, and multi/many-core processors into a single scalable and modular server architecture. Mixed configurations of CPUs, e.g., based on x86 or ARM architecture, can be configured and extended by PCIe-based hardware accelerators. Extensive monitoring and management functionalities are tightly integrated into hardware platform, enabling thorough evaluation of the energy efficiency achieved for the evaluated software tools and algorithms. A detailed discussion of the server platform is provided in [1]. The RECS®|Box architecture provides a heterogeneous scale out approach (horizontal scaling).

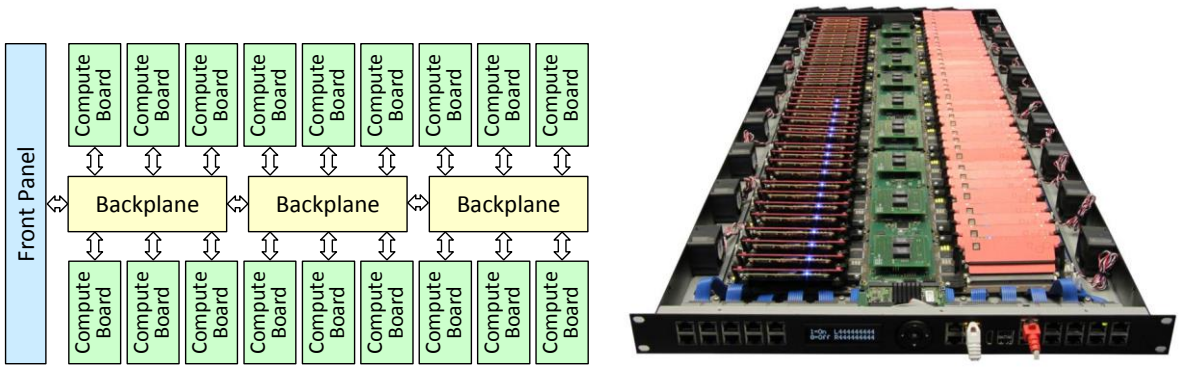

Figure 1: *Architecture of the RECS|Box Server Platform and example installation*

A single server integrates up to 18 compute boards, each equipped with up to four microserver boards in a one rack unit enclosure (cf. Figure 1). Microserver based on Intel and AMD x86/x64 CPUs, embedded and mobile ARM-based CPUs as well as FPGA-based SoCs (Xilinx Zynq 7000) are available and can be flexibly combined. All microservers are based on existing computer-on-module (COM) standards like COM Express [2] or Apalis [3]. Communication between the compute boards is facilitated via a central backplane that offers switched Gigabit Ethernet as well as a dedicated, multi-standard interconnection infrastructure with a bandwidth of up to 40 Gbit/s per compute board.

Within the FiPS project, a microserver based on the COM Express standard has been developed targeting a close integration of FPGA resources into the platform. In addition to a Xilinx Zynq-7045 SoC the microserver comprises 1 GByte DDR3 memory for the processing system (Dual-Core ARM Cortex-A9, 1 GHz) and up to 8 GByte DDR3 memory for the programmable logic. In addition to Ethernet based communication the RECS®|Box also integrates high-speed serial links for low latency communication between the Zynq-based COM Express modules, which can be utilized for a wide range of application areas. Hardware/Software implementations have been realized supporting Remote Direct Memory Access (RDMA) as well as efficient streaming between Zynq microservers, e.g., for image and video processing. Benchmarking resulted in a sustained full duplex bandwidth of 39.5 Gbps and a latency of 320 ns for communication between Zynq microservers. Being integrated into the RECS management system, the high-speed serial communication can be easily monitored enabling control and adaptation of the communication parameters at run-time.

For benchmarking performance and energy efficiency, the platform was evaluated using a sequence alignment implementation from bioinformatics. Compared to a highly optimized GPU implementation, the FPGA-based microservers have been able to double the energy efficiency although the FPGA implementation was based on a high-level synthesis flow without dedicated optimization towards the FPGA [4].

3. Outlook – M2DC

Based on the architecture and methodologies developed within FiPS, a new class of low-power TCO-optimized appliances with built-in efficiency and dependability enhancements is developed within M2DC. A novel heterogeneous server architecture will enable customization and smooth adaptation to various types of applications, which will be easy to integrate utilizing a broad ecosystem of management software and system efficiency enhancements. M2DC will especially focus on a novel high-speed low-latency communication infrastructure, combined with system efficiency enhancements to increase performance and energy efficiency of a wide range of applications.

Acknowledgments

This research was supported by the EU Seventh Framework Programme FP7/2007–2013 under grant agreement no. FP7-ICT-2013-10 (609757) and by European Union's Horizon 2020 research and innovation programme under grant agreement No 688201. This work was also funded as part of the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277), Bielefeld University.

References

[1] R. Griessl, M. Peykanu, J. Hagemeyer, M. Porrmann, S. Krupop, M. Berge, T. Kiesel, and W. Christmann, "A scalable server architecture for next-generation heterogeneous compute clusters," in 12th IEEE Int. Conf. on Embedded and Ubiquitous Computing (EUC), Aug 2014, pp. 146–153.

[2] PICMG, "PICMG COM.0 R2.1 - Com Express Module Base Specification," Available at http://www.picmg.org, accessed: 13-August-2015.

[3] Toradex, "Apalis Computer Module - Module Specification," Available at http://developer.toradex.com/hardware-resource

[4] Kierzynka , M., Kosmann, L., vor dem Berge, M., Krupop, S., Hagemeyer, J., Griessl, R., Peykanu, M., Oleksiak, A.: Energy Efficiency of Sequence Alignment Tools - Software and Hardware Perspectives. Future Generation Computer Systems. (2016).