

Software-Defined Board- and Chip-level Optical Interconnects for Multi-Socket communication and Disaggregated Computing

N. Pleros¹, N. Terzenidis^{1,2}, T. Alexoudi^{1,2}, D. Syrivelis², K. Vyrsoinos¹, and G.T. Kanellos¹

email: npleros@csd.auth.gr

¹*Dptm of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece*

²*Dptm of Electrical and Computer Engineering, University of Thessaly, Volos, Greece*

Abstract: The vast amount of new data being generated is outpacing the development of infrastructures and continues to grow at much higher rates than Moore's law, a problem that is commonly referred to as the "data deluge problem". This brings current computational machines in the struggle to exceed Exascale processing powers by 2020 and this is where the energy boundary is setting the second, bottom-side alarm: A reasonable power envelope for future Supercomputers has been projected to be 20MW, while world's current No. 1 Supercomputer Sunway TaihuLight provides 93 Pflops and requires already 15.37 MW. This simply means that we have reached so far below 10% of the Exascale target but we consume already more than 75% of the targeted energy limit!

The way to escape is currently following the paradigm of disaggregating and disintegrating resources, massively introducing at the same time optical technologies for interconnect purposes. Disaggregating computing from memory and storage modules can allow for flexible and modular settings where hardware requirements can be tailored to meet the certain energy and performance metrics targeted per application. At the same time, optical interconnect and photonic integration technologies are rapidly replacing electrical interconnects continuously penetrating at deeper hierarchy levels: Silicon photonics have enabled the penetration of optical technology to the computing environment, starting from rack-to-rack and gradually shifting towards board-level communications. In this article, we present our recent work towards implementing on-board single-mode optical interconnects that can support Software Defined Networking allowing for programmable and flexible computational settings that can quickly adapt to the application requirements. We present a programmable 4x4 Silicon Photonic switch that supports SDN through the use of Bloom filter (BF) labeled router ports. Our scheme significantly simplifies packet forwarding as it negates the need for large forwarding tables, supporting at the same time network size and topology changes through simple modifications in the assigned BF labels. We demonstrate 1x4 switch operation controlling the Si-Pho switch by a Stratix V FPGA board that is responsible for processing the packet ID and correlating its destination with the appropriate BF-labeled switch output port. Moving towards high-capacity board-level settings, we discuss the architecture and technology being currently promoted by the recently started H2020 project ICT-STREAMS, where single-mode optical PCBs hosting Si-based routing modules and mid-board transceiver optics expect to enable a massive any-to-any, buffer-less, collision-less and extremely low latency routing platform with 25.6Tb/s aggregate throughput. This architecture and technology are also extended to support resource disaggregation in data centers as currently being pursued in the H2020 project dREDBox, where the any-to-any collisionless routing scheme is proposed for connecting disaggregated computing and memory bricks trying to minimize remote memory access latency and energy consumption.

This work has been supported by the European H2020 ICT projects ICT-STREAMS (Contract No. 688172) and dREDBox (Contract No. 687632)