Chip-to-Chip Interconnect for 8-socket direct connectivity using 25Gb/s O-band integrated transceiver and routing circuits

Miltiadis Moralis-Pegios[1], Stelios Pitris[1], Theonitsa Alexoudi[1], Joris Lambrecht[2], Xin Yin[2], Johan Bauwelinck[2], Yoojiin Ban[3], Peter de Heyn[3], Marianna Pantouvaki[3], Joris Van Campenhout[3] and Nikos Pleros[1]

[1] Department of Informatics, Center for Interdisciplinary Research and Innovation, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece, mmoralis@csd.auth.gr
[2] Ghent University imec, IDLab, Department of Information Technology, B-9052, Ghent, Belgium
[3] imec, Kapeldreef 75, B-3001, Leuven, Belgium

Abstract *We present an O-band Chip-to-Chip Interconnect for 8-socket direct connectivity exploiting a Si-based Ring Modulator and a packaged PD-TIA connected over a Si-based 8×8 AWGR routing module. Eight routing scenarios are experimentally demonstrated at 25Gb/s revealing error-free operation.*

Introduction

To cope with the compute and networking requirements in today's East-West DataCenter (DC) communication patterns[1], the research community has focused on novel direct chip-to-chip (C2C) interconnect technologies that can allow for low latency and energy efficient communication. Nevertheless, offered C2C solutions so far face an inevitable trade-off: Peer-to-peer schemes, like Intel's Quick Path Interconnect (QPI)[2], offer low-latency and direct any-to-any communication for a maximum number of 4 sockets and can scale to 8-node interconnection only through dual-hop and, as such, increased latency setups. Scaling to >8-socket connectivity schemes can only be accomplished through switch-based layouts like Bixby[3] and PCIe[4] that connect a high number of QPI-islands at the cost of increased energy and latency.

Novel optical interconnections schemes hold the potential to overcome this trade-off and offer high-bandwidth and low latency any-to-any connectivity among more than 4 nodes[5], leveraging the cyclic properties of Arrayed Waveguide Grating Routers (AWGR)[6]. The indisputable advantage offered by AWGR-based schemes for direct connectivity between a high number of nodes has already been shown via simulations to offer significant latency and energy benefits when performing at 10Gb/s[5]. At the same time, bit-parallel schemes allow for additional energy gains[7] and recent broadcast-friendly architectures reveal a great potential to serve the cache coherency needs in multisocket setups with energy improvements of up to 22%[8]. Realizing the Flat-Topology AWGR-based architecture with integrated photonics can potentially trigger new advances in multisocket server boards, however the only experimental demonstration where integrated transceiver and routing devices were employed performed so far at a rather low data-rate of 0.3Gb/s[9].

In this paper, we present experimentally a flat-topology O-band AWGR-based 8×8 C2C interconnection scheme at 25Gb/s line-rates exploiting integrated transceiver and routing elements. Eight different routing scenarios have been experimentally demonstrated to perform error-free at 25Gb/s with <2dB power penalty using a carrier-depletion silicon-based ring-modulator[10] (RM) as the transmitter (Tx), a packaged photodiode (PD) with its SiGe transimpedance amplifier (TIA) as the receiver[11] (Rx) and a Si-based 8×8 AWGR[12]. Taking into account the power requirements of state-of-the-art integrated RM drivers[13] and assuming a laser wall-plug efficiency of 10%, the proposed 25 Gb/s silicon C2C photonic link can allow for an overall power consumption of 9.48 pJ/bit, offering a significant 41% reduction compared to the 16.2 pJ/bit of Intel's QPI link[2].

Concept and Experimental Setup

Fig. 1(a) depicts the typical 4- and 8-socket QPI C2C interconnect topologies, with the yellow- and red-coloured nodes being at 1- and 2-hop distance, respectively, from the reference blue-coloured node. Fig.1(b) shows the respective 8-node connectivity scheme when leveraging the cyclic routing properties of an 8x8 AWGR, revealing that all nodes can now be accessed at a single-hop distance. Fig.1(c) illustrates the C2C interconnect architecture as this is envisioned when using integrated photonics technology, where a Si-based RM preceded by a tunable laser source (TLS) comprises the Tx at the socket interface, an 8x8 Si-based AWGR performs the routing function and a PD and SiGe TIA is employed as the receiver at every socket . Every socket can connect to any of the remaining 7 sockets by tuning its TLS to a different
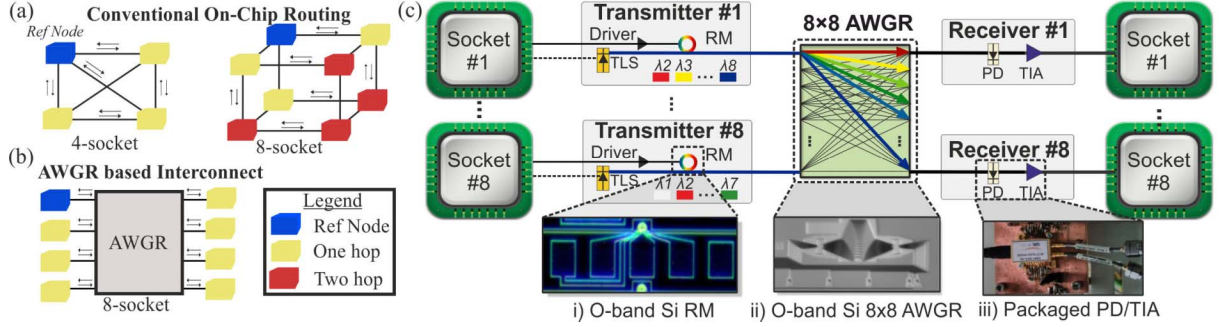
Fig. 1: (a) Conventional On-Chip Routing for 4 and 8 sockets chip interconnects (b) Optical AWGR based chip interconnect (c) Schematic of any-to-any 8×8 AWGR-based interconnection architecture utilizing integrated RMs, 8×8 AWGR, PD-TIAs . (i)-(iii) Close up views of the chips utilized for the 25 Gb/s routing demonstration

wavelength that matches a RM and an AWGR resonance, exploiting the cyclic wavelength-routed properties of the AWGR. The RM, AWGR and co-packaged PD-TIA circuits used in the experimental demonstration comprised three discrete chips and are shown as insets in Fig. 1(c):

a. **Silicon RM**: A silicon O-band micro-ring pn junction modulator has been used, shown in Fig. 1(c(i)). The silicon device is an all-pass ring resonator fabricated on imec's ISIPP200 platform. Its operation is based on the carrier-depletion mechanism and has a 39 fJ/bit energy efficiency and up to 56Gb/s modulation speed[10]. Moreover, coupled high a recently developed low-power driver, the RM can be employed consuming only 1.6 pJ/bit at 25 Gb/s[13].

b. **Silicon 8×8 AWGR**: For the routing platform, the architecture relied on an integrated silicon 8×8 passive AWGR device, depicted in Fig. 1(c(ii)), with cyclic-frequency operation in the O-band. The AWGR relied on imec-ePIXfab silicon photonics passive technology, exhibiting 10 nm-channel spacing, a 3-dB bandwidth of 5.5 nm and a free spectral range (FSR) of 80 nm[12].

c. **Co-packaged PD-TIA**: At the Rx side, a uni-traveling InGaAS/InP PIN photodiode (PD) connected with a low-power transimpedance amplifier (TIA) implemented in 0.13 μm SiGe BiCMOS was employed as a packaged and fiber-pigtailed module and is depicted in Fig.1(c(iii))[11]. Its sensitivity was -10.6dBm at 25Gb/s consuming a power of 158mW and yielding an energy efficiency of 6.32pJ/bit.

The experimental setup used for the proof-of-concept demonstration at 25 Gb/s is shown in Fig. 2. A TLS was used to produce a CW signal at λ1=1278.76 nm. The RM chip was optically probed with single-mode fibers through TE-polarization grating couplers (GCs). An RF source was used to generate a pseudo-random binary sequence (PRBS7) at 25 Gb/s that was amplified by a driver and applied on the RM through an RF probe. A DC voltage was applied as a reverse bias to the RM assuring operation in the depletion region. The resulting modulated signal was then launched into one of the 8 input ports of the AWGR. A 16-channel fiber array was used to couple the signal in and out of the respective AWGR input/output ports through TE-polarized GCs. Instead of changing the input RM wavelength, the RM was sequentially coupled into different input ports of the AWGR using the same wavelength, so that the signal exits the AWGR through a different output port for every input port connectivity scheme. By sequentially connecting the input signal to all 8 AWGR input ports, routing to all 8 AWGR output ports was obtained corresponding to an 8-node connectivity scenario. After exiting the AWGR, the signal was launched to the PD-TIA. Polarization controllers (PC) were used at different stages of the setup to maintain proper signal polarization. 1300nm Semiconductor Optical Amplifiers, SOA1 and SOA2 were used between the RM and the AWGR and between the AWGR and the PD-TIA to compensate for GC losses that were 9dB for the Silicon RM chip and 9B for each of the two AWGR in/out coupling stages. Signal quality monitoring before and after SOA1 was obtained
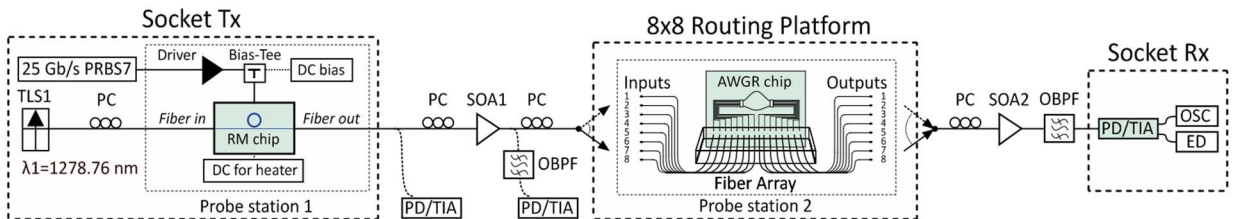

Fig. 2: Experimental setup used for the 25Gb/s demonstration employing the silicon RM, the silicon 8×8 AWGR chip and the PD-TIA module
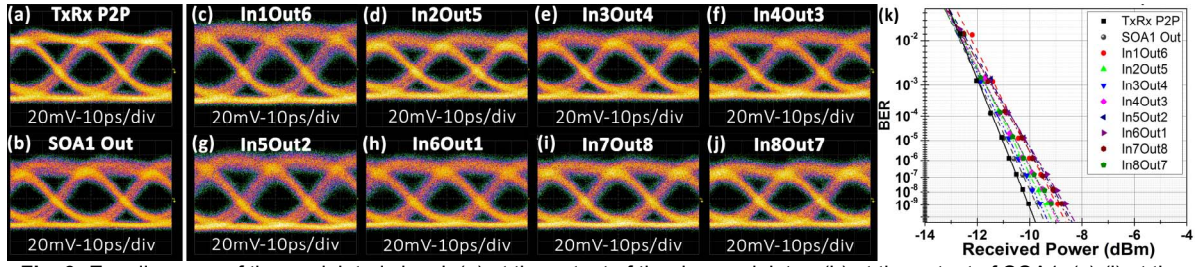
Fig. 3: Eye diagrams of the modulated signal: (a) at the output of the ring modulator, (b) at the output of SOA1, (c)-(j) at the Socket Rx PD-TIA output for all 8 possible AWGR in/out port combinations, (k) BER measurements

by connecting the socket Rx stage directly at the RM chip and the SOA1 output, respectively. SOA noise filtering was accomplished using an optical band-pass filter (OBPF) with a 2.5 nm 3-dB bandwidth.

Results

Fig. 3(a) shows the eye diagram of the 25 Gb/s modulated signal at the RM output and Fig. 3(b) shows the respective eye diagram after SOA1 with an extinction ratio (ER) of 5.4 dB and 5.32 dB, respectively. Fig. 3(c)-Fig. 3(j) illustrate the eye diagrams obtained for the 8 routing scenarios for all possible input-output port combinations denoted as In#iOut#j, indicating successful routing with an ER of 4.8±0.3 dB in all cases. Finally, BER measurements were obtained revealing error-free operation for all 8 routing combinations at 10^{-9} BER value, as shown in Fig. 3(k). The BER curves at the output of the AWGR exhibit a variation of 1.1 dB at 10^{-9} mainly due to different insertion losses of the different routing paths that lead to different SOA2-induced noise levels. The RM was electrically driven with a peak-to-peak voltage of 2.3 Vpp while the applied DC bias was 4 V. The average optical power of the signal at the input of the RM was 8 dBm while at the output was -4 dBm. The optical power of the signal entering the AWGR was 6 dBm, while the AWGR output power was in the range of -12 dBm to -5.6 dBm, owing to different AWGR GC and channel losses depending on the input/output port combination.

Conclusions

We presented experimentally a 25Gb/s O-band Flat-Topology 8-node C2C interconnect architecture exploiting fiber-interconnected Photonic Integrated Circuits for the transceiver and passive AWGR-based routing stages. Transferring this interconnect onto the same polymer waveguide optical board[14] in order to replace GCs with low-loss adiabatic coupling structures[14] can allow for energy efficiency improvements up to 41% compared to typical QPI multisocket interconnects that can increase with higher data rates, given the proven credentials of all integrated modules to operate up to 40 Gb/s[10,11,12].

References

[1] Cisco white paper, "Cisco Global Cloud Index: Forecast and Methodology, 2015–2020," (Cisco, 2016).

[2] R. Maddox, "*Weaving High Performance Multiprocessor Fabric: Architectural Insights to the Intel QuickPath Interconnect*", Intel Press, 2009.

[3] T. Wicki et al., "Bixby: The scalability and coherence directory ASIC in Oracle's highly scalable enterprise systems,". in (HCS) 2013, pp. 1-34.

[4] J. Ajanovic, "PCI Express 3.0 Overview: A tutorial," in Proc. IEEE 21 Hot Chips Symposim (HCS) 2009.

[5] P. Grani et al., "Flat-Topology High-Throughput Compute Node With AWGR-Based Optical-Interconnects," JLT, 34, 2959-2968, 2016.

[6] N. Terzenidis et al. ," High-port low-latency optical switch architecture with optical feed-forward buffering for 256-node disaggregated data centers". *Op.Ex.*,*26*(7),pp. 8756-8766 (2018)

[7] P. Grani, et.al "Bit-Parallel All-to-All and Flexible AWGR-based Optical Interconnects," in OFC, p. M3K.4.(2017).

[8] S. Pitris, et al. "O-band Energy-efficient Broadcast-friendly Interconnection Scheme with SiPho Mach-Zehnder Modulator (MZM) & Arrayed Waveguide Grating Router (AWGR)," in OFC, p. Th1G.5 (2018).

[9] R. Yu et al, "A scalable silicon photonic chip-scale optical switch for high performance computing systems," Opt. Express 21, 32655-32667 (2013).

[10] J. Van Campenhout et al., "Silicon Photonics for 56G NRZ Optical Interconnects", in OFC, p.W1I.1, (2018).

[11] B. Moeneclaey et al, "A 40-Gb/s Transimpedance Amplifier for Optical Links" in PTL 27 (13), 1375

[12] S. Pitris et al, Silicon photonic 8× 8 cyclic Arrayed Waveguide Grating Router for O-band on-chip communication", Op. Ex,*26*(5), 6276-6284. (2018).

[13] H. Ramon, et al. "Low-Power 56Gb/s NRZ Microring Modulator Driver in 28nm FDSOI CMOS." IEEE Photonics Technology Letters 30.5 467-470 (2018):

[14] G.T. Kanellos et al, "WDM mid-board optics for chip-to-chip wavelength routing interconnects in the H2020 ICTSTREAMS," in Proc. SPIE (2017).