400 Gb/s Silicon Photonic Transmitter and Routing WDM technologies for glueless 8-socket Chip-to-Chip interconnects

S. Pitris, C. Mitsolidou, M. Moralis-Pegios, K. Fotiadis, Y. Ban, P. De Heyn, J. Van Campenhout, J. Lambrecht, H. Ramon, X. Yin, J. Bauwelinck, N. Pleros and T. Alexoudi

***Abstract* — Arrayed Waveguide Grating Router (AWGR)-based interconnections for Multi-Socket Server Boards (MSBs) have been identified as a promising solution to replace the electrical interconnects in glueless MSBs towards boosting processing performance. In this paper, we present an 8-socket glueless optical flat-topology Wavelength Division Multiplexing (WDM)-based point-to-point (P2P) interconnect pursued within the H2020 ICT project ICT-STREAMS and we report on our latest achievements in the deployment of the constituent silicon (Si)-photonic transmitter and routing building blocks, exploiting experimentally obtained performance metrics for analyzing the 8-socket chip-to-chip (C2C) connectivity in terms of throughput and energy efficiency. We demonstrate an 8-channel WDM Si-photonic microring-based transmitter (Tx) capable of providing 400 (8×50) Gb/s non-return-to-zero (NRZ) Tx capacity and an 8×8 Coarse-WDM (CWDM) Si-AWGR with verified cyclic data routing capability in O-band. Following an overview of our recently demonstrated crosstalk (XT)-aware wavelength allocation scheme, that enables fully-loaded AWGR-based interconnects even for typical sub-optimal XT values of silicon integrated CWDM AWGRs, we validate the performance of a full-scale 8-socket interconnect architecture through physical layer simulations exploiting experimentally-verified simulation models for the underlying Si-photonic Tx and routing circuits. This analysis reveals a total aggregate capacity of 1.4 Tb/s for an 8-socket interconnect when operating with 25 Gb/s line-rates, which can scale to 2.8 Tb/s at an energy efficiency of just 5.02 pJ/bit by exploiting the experimentally verified building block performance at 50 Gb/s line. This highlights the perspectives for up to 69% energy savings compared to the standard QuickPath Interconnect (QPI) typically employed in electronic glueless MSB interconnects, while scaling the single-hop flat connectivity from 4- to 8-socket interconnection systems.**

Manuscript received XXXXXXX XX, XX; revised XXXX XX, XX; accepted XXXX XX, XX. This work was supported by the European Commission through the H2020-ICT-STREAMS project (No. 688172)

S. Pitris, C. Mitsolidou, M. Moralis-Pegios, K. Fotiadis, N. Pleros and T. Alexoudi are with the Dept. of Informatics and Center for Interdisciplinary Research and Innovation, Aristotle University of Thessaloniki, 57001, Greece (e-mail: skpitris@csd.auth.gr; cvmitsol@csd.auth.gr; mmoralis@csd.auth.gr; theonial@csd.auth.gr; kfotiadi@csd.auth.gr; npleros@csd.auth.gr).

Y. Ban, P. De Heyn, and J. Van Campenhout are with IMEC, Kapeldreef 75, B-3001, Leuven, Belgium (e-mail: Yoojin.Ban@imec.be; Peter.DeHeyn@imec.be Joris.VanCampenhout@imec.be).

J. Lambrecht, H. Ramon, X. Yin and J. Bauwelinck are with the ID Laboratory, Department of Information Technology, Ghent University–IMEC, 9052 Ghent, Belgium (email: Joris.Lambrecht@ugent.be; hannes.ramon@ugent.be; xin.yin@ugent.be; johan.bauwelinck@ugent.be).

***Index Terms*— Optical transmitters, optical interconnections, photonic integrated circuits, silicon photonics, computing architectures, AWGR, AWGR-based interconnections, MSBs.**

I. Introduction

The ever-increasing amount of data generated nowadays from High-performance Computing, Internet-of-Things and 4G/5G applications is currently generating vast amounts of traffic within the Data Centers (DCs) [1] that need to embrace fundamental changes in their infrastructure to keep up with the ever-increasing computational needs at a cost-effective and energy-efficient manner. The energy restrictions in system processing performance advances can be outlined in the example of the HPC performance evolution during the past 10 years: Fig. 1 depicts the HPC performance predictions established in 2011 [2], which were building upon the respective progress and performance metrics of the top HPC machines witnessed until 2011 in order to forecast the Exaflop performance target by 2020, assuming that advances will continue to rely on the 10x performance improvement factor per 3.5-4 years. However, reality has been proven to be very different and the performance of HPC machines from 2011 until today has slowed down dramatically, improving by 10x every 6 years and reaching now the 200 Pflop/s target in world's No. 1 IBM Summit HPC machine at an energy envelope of already 10 MW. Inverting this performance slow-down can only be accomplished by increasing processing

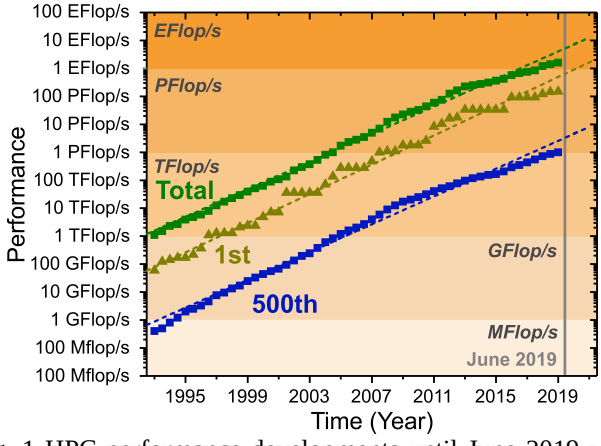


Fig. 1 HPC performance developments until June 2019 with the dashed lines depicting the performance evolution predictions made in 2011 [2]

performance at a lower energy consumption framework, which in turn has to rely on the energy-efficient synergy of the most powerful high-performance multi-core engines.

Increasing the number of high-performance cores within the same multi-core processor die has already stalled as a solution due to real-estate and pin-out limitations [3], with glueless multi-socket server boards (MSBs) relying on Chip-to-Chip (C2C) electrical interconnects between processor sockets forming nowadays the only promising alternative for synergizing a high number of high-performance cores into high computational density setups [4]. Typically, the MSB interconnects exploit the dominant point-to-point (P2P) Intel QuickPath Interconnect (QPI) and are formed in 4-socket layouts [5]. However, electrical MSB interconnects face practical scalability issues when 8- or >8-socket configurations are targeted: "glueless" 8-socket layouts can only be formed when a certain number of dual-hop connections is employed yielding increased latency values [3], while "glued" >8-socket interconnects can only be implemented by employing active switches in-between glueless QPI 4- or 8-socket islands, which inevitably increases energy consumption in addition to the higher latency values.

To face these limitations, Arrayed Waveguide Grating Router (AWGR)-based MSB optical interconnections have been identified during the last few years as a promising solution for replacing the electrical interconnects in MSBs, providing an all-to-all point-to-point (P2P) optical interconnection scheme with direct connectivity between all sockets [3][7][8]. With Si-photonics being now the most attractive photonic integration platform for datacom and computercom applications due to its energy-efficient, low-cost and CMOS-compatibility credentials, the synergy of AWGR-based interconnects with Si-photonic components can potentially boost the future MSB systems. This layout has been firstly proposed in [8] where the MSB execution time and latency performance benefits have been addressed by cycle-accurate simulations over a C-band AWGR-based interconnect scheme performing at 10 Gb/s line-rates. However, its first experimental deployment using Si-based transceiver (TxRx) and AWGR circuitry on the same Si-photonic chip layout demonstrated only a few Tx-Rx link combinations and was capable of performing at a rather low data-rate of only 0.3 Gb/s [7].

Nevertheless, facing the real estate limitations of multiple sockets residing on the same Si-chip, the scalability of optical MSBs in terms of number of sockets and line-rate performance can be possibly achieved by following a disintegration approach where the transceivers and the AWGR router module reside in separate chips. The sockets can communicate optically with C2C optical links via an Electro-Optic Printed Circuit Board (EOPCB), relaxing in this way the real estate limitations of closely-placed sockets. In this scenario, the O-band (1260-1360 nm) spectral region offers significant advantages compared to C-band, since optical polymer waveguides embedded in EOPCBs typically offer a much lower waveguide loss figure in this spectral region [9], while at the same time ensures maximum compatibility with the standard datacom and computercom fiber-based infrastructure where O-band is the dominant operational spectral regime. Towards this goal, many recent novel demonstrations of Si-photonic high-speed TxRxs [10]-[13], Si-photonic cyclic-AWGR routing elements [14] and low-loss polymer EOPCBs [9] have already indicated the availability of all necessary building blocks required for realizing high-speed O-band AWGR-based optical MSB interconnections.

The Si-based deployment perspective of optically-enabled multi-socket interconnects has to ensure also certain performance requirements for the constituent photonic building blocks in order to allow for a successful full-scale interconnect. AWGR-based MSB interconnects have to support fully-loaded AWGR routing schemes where all AWGR input ports will be loaded with wavelength division multiplexing (WDM) data traffic, necessitating in this way low XT values for the AWGRs [15] that cannot be, however, still accomplished in Si-integrated AWGR modules [7][8][14][16]. This reality creates an extra need for architecture- and system-driven solutions towards transforming the available AGWR silicon-integrated circuitry into successful fully-loaded engines for MSB interconnection.

In this paper, we present our latest achievements accomplished within the H2020-project ICT-STREAMS in Si-photonic O-band WDM Tx and routing circuitry for optically-enabled 8-socket interconnect systems, revealing a viable perspective for the realization of glueless 8-socket MSBs in the optical domain. We demonstrate a novel WDM 400 Gb/s 8-channel Si-photonic ring modulator (RM)-based transmitter that is aligned with the latest roadmap for optical GbE transceivers [18] and can be utilized as the Si-based socket Tx interface in an 8-socket arrangement requiring only 1.04 pJ/bit (excl. the laser source & electronics), performing also at the highest non-return-to-zero (NRZ) line-rate among 400 Gb/s-capable RM-based Si WDM Txs. We review our recent advances in passive 8×8 Si-photonic cyclic-AWGR modules for cyclic-operation in O-band [14] and its performance when configured in a novel crosstalk (XT)-aware flat-topology interconnect scheme for fully-loaded 8-socket AWGR-based MSB configurations even with a moderate XT value of just -11 dB [17]. Finally, we utilize the experimentally obtained performance metrics of the Si-based Tx and AWGR circuits towards deploying reliable and experimentally validated physical layer simulation models and we present a fully-loaded 8-socket optical interconnect via physical layer simulations, revealing an aggregate 8-socket MSB throughput of 1.4 Tb/s when performing at 25 Gb/s that can scale to 2.8 Tb/s when exploiting the 50 Gb/s line-rate-capable photonic circuits. This highlights the credentials of Si-based AWGR-based interconnects to allow for single-hop connectivity even in 8-socket topologies, significantly contributing towards lower latency processing compared to electronic QPI glueless setups, while offering an energy efficiency of 5.02 pJ/bit that is close to 69% reduced compared to the 16.2 pJ/bit energy consumed by QPI.

The paper is organized as follows: We present our proposed 8-socket MSB interconnection. Then, we present the 8-

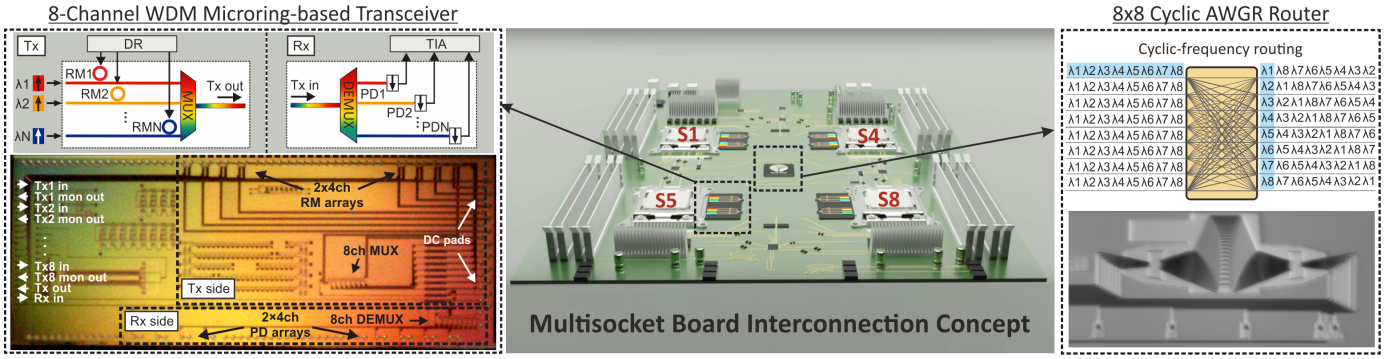


Fig. 2 Optical MSB interconnect concept with insets of WDM TxRx/Router layouts and integrated Tx and AWGR [14] chips.

channel WDM 400 Gb/s Tx and the 8×8 cyclic-AWGR. Next, we present our XT-aware scheme, which we validate through a proof-of-concept experiment, and we successfully verify the performance of a fully-loaded AWGR-based 8-socket silicon photonic interconnect through physical layer simulations. Finally, we conclude this article.

II. Flat-Topology All-to-All MSB Interconnection

Fig. 2 illustrates an artistic perspective of the envisaged optical P2P MSB architecture when deployed onto a hosting polymer EOPCB for the case of 8 interconnected processor sockets, with the corresponding inset figures depicting the integrated individual subsystems of the architecture. For easier illustration purposes, Fig. 2 includes only 4 sockets out of 8 sockets considered in the presented work. Eight different sockets (S1-S8) are attached to respective mid-board WDM Si-Pho TxRx interfaces that convert the electrical data of the processor socket to optical and launch them into the optical polymer waveguides embedded into the EOPCB board through Si-to-polymer adiabatic coupling (ADIAC) scheme [20]. As it shown in the detailed layout of the TxRx left inset in Fig. 2, the TxRx harness the low-power and low-dimension credentials of micro-ring resonator (MRR) structures both in the Tx side and the Rx side. The Tx comprises 4 RMs driven by a multi-channel electronic driver (DR) powered by external laser diodes (LD) and a 2nd-order MRR-based multiplexer (MUX) to multiplex the signals in the common WDM output. The Rx side comprises a 2nd-order MRR-based demultiplexer (DEMUX) that splits the incoming WDM signals leading to the photodiodes (PDs) which are driven by transimpedance amplifiers (TIAs) that convert and amplify the optical data to electrical so that it can be received by the socket, respectively. The all-to-all communication is ensured via a mid-board Si-Pho AWGR passive router chip plugged at the center of the EOPCB that connects all the 8 sockets together in a passive and latency-free way. Our recently-developed integrated 8-channel Si-photonic WDM TxRx can be seen in the microscope photo of the left inset in Fig. 2. The right inset in Fig. 2 depicts the wavelength mapping of the incoming WDM signals from all input ports to the output ports of the AWGR following the cyclic-frequency routing principle [14], i.e. the wavelength set launched at a specific input of the AWGR will be routed to all AWGR outputs with the signals at different input wavelength sets emerging at the outputs shifted compared to the wavelength set launched in their neighboring input. The microscope photo in Fig. 2 shows the integrated Si-photonic 8×8 AWGR.

III. 400 (8×50) Gb/s Si-Photonic Ring-based Tx

The purpose of the Si-Pho WDM mid-board TxRx attached to every socket is to optically interface the processor with the EOPCB. Following our recent work on a 4-channel O-band WDM Tx co-packaged with its electronic DR-TIA chips operating at 160 (4×40) Gb/s [10] and up to 200 (4×50) Gb/s capacities [21], we present here its extended WDM Tx version, i.e. an 8-channel O-band WDM TxRx capable to provide up to 400 (8×50) Gb/s aggregate Tx capacities and we report on its characterization and evaluation.

A. Device Fabrication & Description

The Si-Photonic TxRx (Tx area: 6400×2900 μm^2), which can be seen in Fig. 2, was fabricated in IMEC's ISIPP50G platform [22]. The Tx chip is equipped with TE-polarization grating couplers (GC) with a peak wavelength at 1315 nm comprising the I/O ports, namely *Tx#1-8* in and *Tx out*, that refer to the 8 Tx inputs and the combined Tx output, respectively. The Tx comprises two arrays of four carrier-depletion RMs designed with a free-spectral range (FSR) of 9 nm (1.6 THz), corresponding to RMs of Tx channels #1-8, respectively, The RMs exhibit a capacitance of ~30 fF, a Q factor of ~5500 and a modulation bandwidth (BW) of 33 GHz @ 0 V. Thermal tuning of each RM resonance is achieved via a dedicated tungsten heater on top of the RM pn junction that is controlled by respective DC pads at the east side of the chip. 3 dB-MMI couplers are incorporated in the Tx circuitry after each RM of leading to the waveguides connected with the *Tx mon* output GCs for intermediate monitoring purposes of the signals. The modulated signals from all 8 RMs are multiplexed into a common output via an 8×1 MUX unit that exploits 2nd-order MRRs and that was designed with a channel spacing of 1.13 nm (200 GHz) and an FSR of 9 nm (1.6 THz). The MUX can also be thermally tuned by a collective heater implemented at the sides of the 2nd-order MRR-based MUX structure for simultaneous thermal tuning of all MUX channels. Although not evaluated and presented in this work, the Tx chip also comprises a Rx based on 2nd-order MRR-based DEMUX unit and two arrays of four high-speed Ge photodiodes.

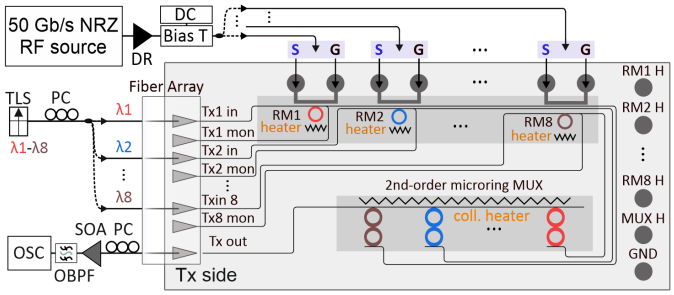


Fig. 3 Schematic of the WDM Tx and the experimental setup used for the 8×50 Gb/s NRZ evaluation.

B. Experimental Setup & Results

A schematic of the 8-channel Tx layout and the experimental setup that was used for the characterization and the evaluation of the 8×50 Gb/s NRZ operation is depicted in Fig. 3. The Tx chip was probed with a fiber array while a GS RF probe was used to access the electrical pads of the RMs #1-8 sequentially. A tunable laser source (TLS) was used to generate each time one of the 8 continuous wave (CW) signals at λ1=1308.4 nm, λ2=1309.65 nm, λ3=1310.79 nm λ4=1312.18 nm, λ5=1313.24 nm, λ6=1314.73 nm, λ7=1315.55 nm and λ8=1316.56 nm, that were sequentially launched at the input GCs corresponding to Tx1-Tx8 in, through the fiber array to reach RMs #1-8, respectively. An RF source was used to generate a non-return-to-zero (NRZ) pseudo-random binary sequence (PRBS7) that was amplified by an SHF807b high-frequency driver (DR) amplifier and applied to the RMs along with a reverse-bias DC voltage. The Tx output signal was obtained at Tx out and was then amplified in a semiconductor optical amplifier (SOA). The amplified spontaneous emission (ASE) noise of the SOA was filtered out of the amplified signal each time by a tunable narrow optical bandpass filter (OBPF) with 0.5 nm 3 dB-BW before monitored at the digital oscilloscope (OSC). Standard single mode fiber (SMF) was utilized in the experimental setup and for this reason polarization controllers (PC) were used to maintain proper signal polarization before the TE GCs of the Si-chip and the SOA.

Fig. 4(a) depicts the normalized output spectra of the 8 Tx channels. The 8-channel MUX exhibited a channel spacing of 1.17 nm (208 GHz), a channel 3 dB-BW of 0.73 nm (129 GHz) and a free spectral range (FSR) of 9.78 nm (1.73 THz), with all values above calculated on average from all MUX channels characteristics within the working spectral region (1300-1320 nm). The channel insertion losses were measured in the range of 0.9 dB to 2.7 dB, respectively. Fig. 4(b) depicts the normalized transmission spectra of RM1 as it was obtained at the Tx1 monitor output port for different reverse bias voltages applied at the RM pn junction. The modulation efficiency of RM#1 was measured to be 29 pm/V in the range of 0 V to -2 V, which was also verified for the rest of the RM structures. Fig. 4(c) depicts the normalized overlapping spectra of the RM#i and the MUX channel #i transmission for each Tx channel, respectively, without enforcing any thermal tuning at either the RM or MUX. The FSR of the RMs #1-8 resonances was measured to be 9 nm (1.6 THz) on average. The observed spectral distance of the MUX passbands from their respective Tx channel RM resonances when no thermal tuning is enforced was found to range between 6.84 nm and 7.18 nm. This indicated that an average MUX tuning of 6.98 nm is required to tune the MUX passbands with the RM resonances to achieve Tx operation. The RM resonance near 1315 nm were selected from each RM structure for the WDM Tx operation so as to reside close to the peak wavelength of the GCs for minimizing total Tx channel insertion losses. Fig. 4(d) shows the Tx output spectra for each one of the 8 Tx channels at the common *Tx out* port before and after tuning the MUX to match with the respective RM resonance for Tx operation.

The Tx was evaluated for its high-speed modulation capabilities in an 8×50 Gb/s NRZ data generation scheme for all channels. Fig. 5(a) shows the WDM grid formed by the

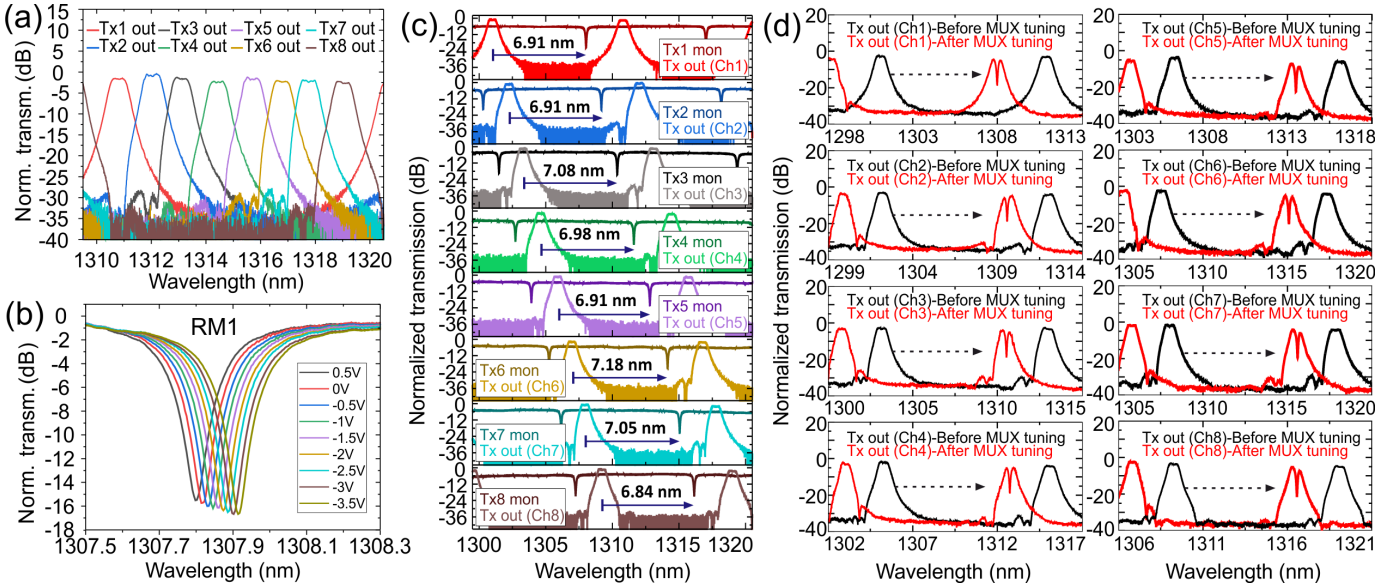


Fig. 4 (a) Tx output spectra, (b) RM1 transmission spectra for different reverse bias voltages, (c) overlapping output spectra of RMs #1-8 and MUX channels #1-8, (d) Tx output spectra before and after tuning the MUX to the RM resonances.

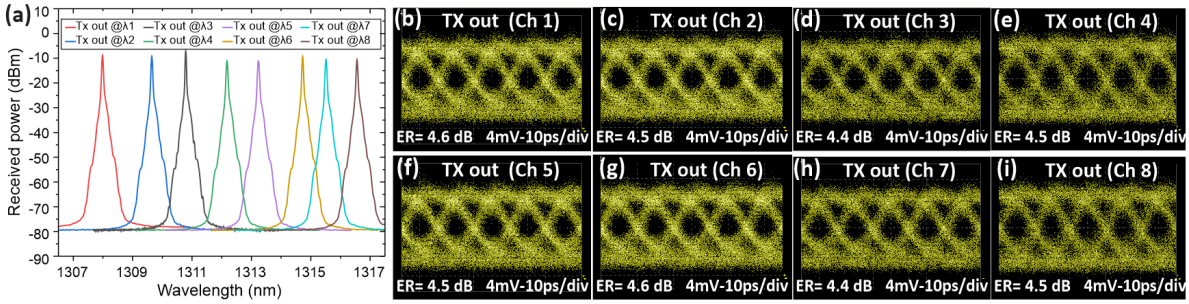


Fig. 5 (a) Superimposed output spectra of all Tx channels depicting the modulated signals at λ1-λ8 and (b)-(i) eye diagrams obtained from individual operation of each Tx channel during the 8×50 Gb/s NRZ evaluation.

superimposed output spectra of all Tx channels depicting the modulated signals at λ1-λ8 as they were obtained at Tx out from the individual operation of each Tx channel separately during the 8×50 Gb/s NRZ evaluation. Figs. 5(b)-5(i) depict the eye diagrams of the signals at 50 Gb/s NRZ as they were obtained at Tx out for each one of the Tx channels separately, exhibiting ER values in the range of 4.4 dB to 4.6 dB. corresponding to an optical modulation amplitude (OMA) values in the range of 4.5 dBm to 4.59 dBm, respectively. The 8 RMs were driven with peak-to-peak voltages of ~2.15 Vpp, respectively, while the applied DC bias reverse voltages were in the range of -1 V to -1.2 V, respectively. The optical power of the CW signals at the 8 wavelengths (λ1-λ8) at the input of the transmitter chip was 10 dBm, while the average power of the modulated signals at every one of the 8 wavelengths emerging at the output of the chip was measured to be in the range of -14.5 dBm to -9.5 dBm, respectively. The break-down of the optical losses is as follows: the GCs imposed ~7-8 dB, the monitor MMIs after RMs ~3.3 dB, the MUX ~0.9-2.7 dB depending on the channel, while the RM exhibited a transmission penalty of ~3-4 dB, depending on the Tx channel and operating wavelength. The average optical power of the signals at λ1-λ8 after the SOA and the OBPF was in the range of 3.5 dBm to 8.5 dBm, respectively. The SOA was electrically driven at 175 mA. The DC voltage applied to the MUX heater to collectively tune all of its passbands to the 8 RM resonances channels was 3.65 V corresponding to a consumed electrical power of 400 mW. No power was applied to the RM heaters during the evaluation and the entire operation was obtained at room temperature (25° C) without requiring a temperature controller. The average energy efficiency of the RM array for 8×50 Gb/s NRZ operation was estimated at 34.66 fJ/bit/RM under 2.15 Vpp drive. Taking into account the 50 mW/channel consumed on average by the MUX heater for tuning the MUX to the RMs, the energy efficiency of the Tx considering both the RM and WDM multiplexing energy requirements (excl. the laser source & electronics) was calculated to be ~1.04 pJ/bit/channel.

IV. 8×8 O-BAND SI-PHOTONIC CYCLIC AWGR ROUTER

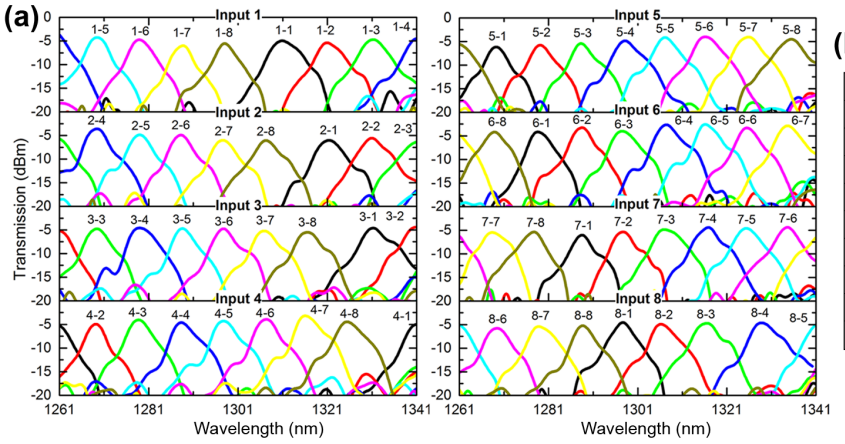
The Si-Pho AWGR-based mid-board passive router is the key component of the interconnection architecture that enables all-to-all communication between the processor sockets connected to its ports in a buffer-less and collision-less way. Given the absence of AWGR devices for operation in O-band, we developed the first cyclic-frequency integrated Si-photonic 8×8 coarse-WDM (CWDM) AWGR. Its cyclic routing properties have been demonstrated with high-speed optical signals confirming its credentials to yield P2P communication between sockets [14].

A. Device Fabrication & Characterization

The Si-photonic 8×8 AWGR was designed with the Bright Photonics *BrightAWG* toolkit with operation in the O-band targeting a center wavelength at 1301 nm, a 10 nm-channel spacing, a 3-dB channel bandwidth of 5.7 nm, an FSR of 80 nm. The fabrication of the integrated AWGR relied on the imec-ePIXfab silicon photonics passives technology platform. The fabricated device can be seen in the microscope image of Fig. 2 right inset. For its optical I/O the AWGR employed grating couplers (GC) with a peak wavelength at 1285 nm which were arranged with 250 μm-pitch for optical probing through a 16-channel Fiber Array (FA). The dimensions of the fabricated AWGR were 700×270 μm^2.

The obtained spectral response for all 8×8 port combinations can be found in Fig. 6(a) while a summary of the measured characteristics of the AWGR while of the AWGR is shown in Fig. 6(b). The measured transmission graphs shown in Fig. 6(a) were produced after extracting all the estimated fiber losses, waveguide transmission losses and grating coupler losses for the transmission link. The spectral response of the AWGR channels revealed proper cyclic-frequency operation that was further verified by the alignment of the same-colored output responses for each one of the different input ports. The measured channel peak wavelengths of the device were on average at 1260.15 nm, 1269.15 nm, 1278.71 nm, 1288.46 nm, 1297.76 nm, 1307.49 nm, 1317.55 nm and 1328.06 nm and exhibited a standard deviation of 0.177, 0.385, 0.439, 0.185, 0.200, 1.428, 2.245, 2.801 nm, respectively. The AWGR device exhibited a channel 3 dB-bandwidth of 5.5 nm, a channel XT of 11 dB, minimum channel losses of 2.5 dB, maximum channel losses of 6.045 dB revealing a loss non-uniformity of 3.545 dB among all 64 channels.

The 8×8 AWGR was also evaluated in data routing experiments with 25 Gb/s optical data signals and for all possible 64 port-combinations [14]. The evaluation revealed error-free (10^{-9}) operation with a maximum 2.45 dB power penalty (PP) for all signals transmitted at AWGR channel peak wavelengths, successfully validating its wavelength-routing interconnectivity credentials.



(b)

Measured characteristics

Port count	8×8
Center Wavelength	1301 nm
Channel spacing	10 nm
FSR	80 nm
Dimensions	700×270 μm²
Channel 3dB bandwidth	5.5 nm
Min channel losses	2.5 dB
Max channel losses	6.045 dB
Loss non-uniformity	3.545 dB
Channel XT	11 dB

Fig. 6 (a) Spectral response and (b) measured characteristics of the 8×8 AWGR.

V. Fully-Loaded 8-Socket AWGR-based Interconnect

Although AWGRs have successfully confirmed their cyclic wavelength routing properties in multisocket board interconnects [3][7][8], fully-loaded routing schemes allowing for the simultaneous all-to-all communication have not been demonstrated yet due to the constraints arising by the in-band XT (IXT) effects in AWGR structures [15]. The IXT can significantly impair the system operation when multiple signals at the same wavelength are launched simultaneously at many AWGR's input ports. In this case, the interference components from the other routing paths cause severe performance degradation on the received data signals at due to beat noise that happens within the bandwidth of the Rx PD bandwidth. As described in [15], an AWGR interconnection requires an AWGR device featuring an IXT value of -34 dB and -36 dB to achieve fully-loaded 8×8 and 16×16 connectivity, however the reported AWGRs exhibit significantly higher XT values [7][8][14][16], preventing their employment in fully-loaded configurations. To overcome this limitation, we have developed a XT-aware wavelength-selection routing scheme that allows for fully-loaded AWGR-based interconnects even when AWGRs with low IXT characteristics are employed, as this has been typically the case for the integrated AWGRs reported so far.

A. XT-Aware Wavelength Allocation Scheme

Fig. 7(a) depicts the proposed XT-aware wavelength allocation scheme where a CWDM N×N AWGR (here $N=8$) is employed for any-to-any communication between N sockets. Each socket is interfaced with a WDM Tx and Rx, with each Tx relying on N lasers, while the Rx employs an 8-channel CWDM DEMUX and N-1 narrowband optical BPFS. The Tx is powered by N-1 CWs provided by an external laser bank at the wavelengths of $\lambda_{i,j}$, with i being the index number of the Tx and j is the *"Set"* number of wavelengths. Wavelengths having the same *"Set"* number reside within the same AWGR channel spectral region but can be slightly detuned with respect to each other. The N WDM streams from the different Txs are simultaneously inserted in each of AWGR inputs to be routed to the outputs. Each output is connected to a CWDM DEMUX that de-multiplexes the WDM channels originating from different Txs. The BPFs deployed after the DEMUX filter the reference wavelengths out of XT components from signals at the detuned wavelengths that eventually emerge in the respective AWGR and DEMUX output.

Undesired signal interference can be avoided in integrated AWGRs with high IXT when the wavelength channels belonging to the same *"Set"* are imprinted at slightly different wavelengths but, at the same time, reside within the 3 BW of the same AWGR output channel. The proposed XT-aware scheme exploits a number of slightly detuned wavelengths around the nominal AWGR channel central wavelength for

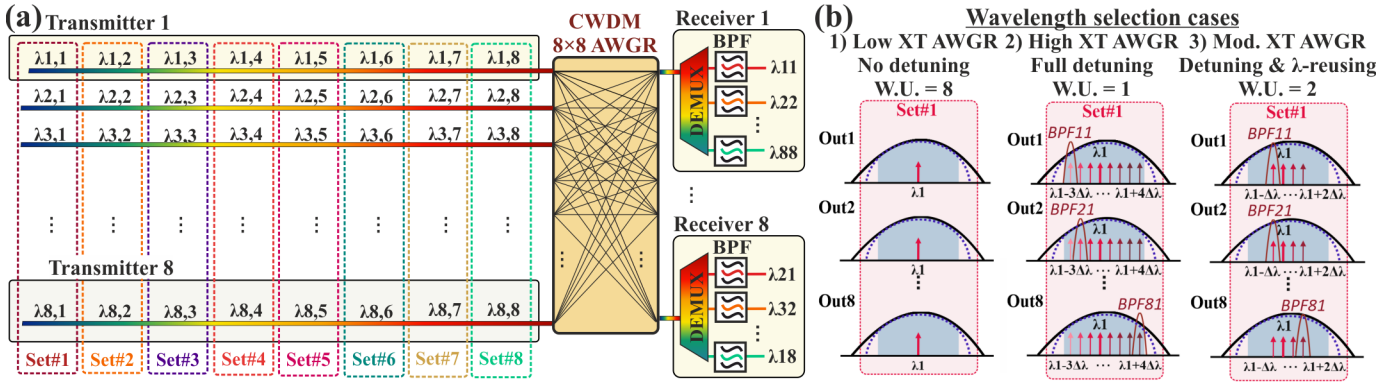


Fig. 7 XT-aware AWGR-based optical interconnection scheme for fully loaded N×N all-to-all communication, (b) Wavelength selection cases based on the XT value of the employed AWGR in the interconnection.

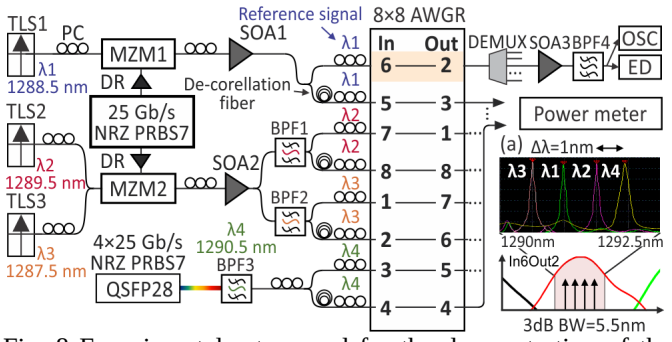


Fig. 8 Experimental setup used for the demonstration of the proposed scheme (a) the 4 detuned-λs in Δλ=1 nm spacing aligned with the AWGR channel.

each spectral band considering, however, the maximum possible wavelength utilization (WU) factor (i.e. the number of different AWGR input signals that can use the same wavelength and still obtain error-free operation in fully-loaded configuration). Thus, the number of the different detuned wavelengths within a single *"Set"* is determined as the ratio of the total number of AWGR input ports divided by the WU factor.

Fig. 7(b) depicts the wavelength selection cases for the proposed XT-aware wavelength allocation scheme. In Case #1, an AWGR with low IXT is assumed, allowing for error-free operation even for a WU equal to the number of inputs. (e.g. WU=8). In this case, the same λ can be employed for each Input*#i* of a Set*#j*, allowing for all-to-all communication with a total number of only 7 λs for all Txs (one band of the AWGR is excluded since it corresponds to the communication of the socket with itself). In Case #2, an AWGR with high IXT is assumed, allowing error-free operation only when each input λ of a Set*#j*, has a different value (WU=1). In this worst-case scenario, N-1 sets (i.e. 7 sets) of N detuned λs (i.e. 8 λs) can be employed in the interconnection in a Δλ-spaced pattern satisfying the following: (i) $(N\text{-}1) \times \Delta\lambda \leq 3\ dB\text{-}BW$ of the CWDM AWGR and (ii) $\Delta\lambda \geq B$, where B is the modulation BW of the signals in the interconnect that determines also the required 3 dB-BW of the employed BPFs. In this case, 8/1=8 different wavelengths per Set#j are distributed in the 3 dB-BW of each of the 7 spectral bands (one band of the AWGR is excluded since it corresponds to self-communication). Thus, 56 wavelengths can be employed for the 56 connections. In Case #3, where an AWGR with moderate IXT is considered, the wavelength detuning scheme can be also exploited lowering the number of deployed λs by defining the maximum WU that allows for error-free operation, and thus reduce the complexity of the system. Indicatively, in Case #3 it was assumed adequate BER performance when WU=2 is employed, resulting in N/2 (i.e. 4 when N=8) different λs per Set*#j*, distributed in the 3 dB-BW of each spectral band. Thus, by finding the maximum WU factor, i.e WU=2 in this case, a total number of 28 λs can be employed for the 56 connections, providing 50% improved λ-utilization compared to Case#2, identified as the worst-case.

The maximum number N of detuned λs can be given in any wavelength-detuning case by the relation: $N \leq (3\ dB\text{-}BW\ of\ AWGR\ /\ B) + 1$. The 3 dB-BW was selected in the proposed XT-aware scheme in order to verify adequate loss uniformity and channel shape uniformity so as to avoid signal spectral distortion for the signals transmitted through the AWGR channels. The XT-aware scheme can greatly benefit from AWGRs with "flat-top" channel design in which the entire flat spectral region of the channels can be utilized for the transmission of the detuned wavelengths avoiding any loss non-uniformity and spectral filtering.

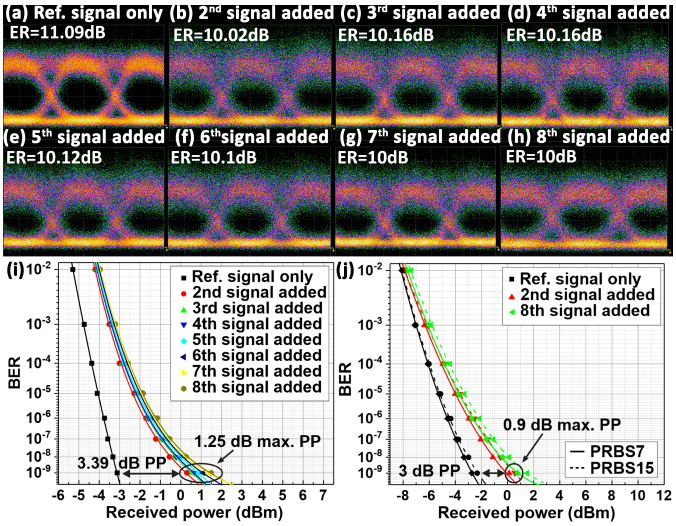
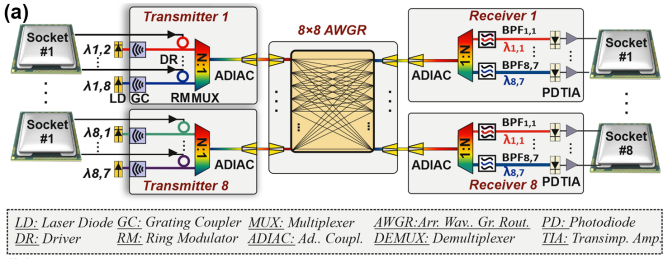


Fig. 9 (a)-(h) Eye diagrams (4mV-5ps/div) and (i) BER measurements of the ref. signal transmitted through the AWGR while more signals at detuned λs are added: (i) experimental, (j) simulations for PRBS7 & PRBS15 signals.

B. *Experimental validation*

The 8×8 Si-photonic AWGR presented in Section III was utilized for the experimental validation of the XT-aware scheme. Following the characterization of the AWGR device that revealed a moderate XT value of 11 dB for all 8×8 AWGR channels [14], the integrated device was evaluated for its capability to perform simultaneous routing of signals imprinted at the same wavelength λ_{ref}. The experimental evaluation in [17] indicated that the integrated 8×8 AWGR can tolerate up to a 2nd additional signal at the same wavelength before introducing error-floors in the transmission of a reference signal and thus the proposed XT-aware detuning scheme can be deployed with a WU=2, thus by using 8 signals imprinted in couples at 4 detuned wavelengths for simultaneous communication.

Fig. 8 depicts the experimental setup used for the proof of-concept demonstration of the proposed wavelength-detuning scheme introducing also the reduced WU feature. Transmission of 8 signals was achieved by using duplicates of the signals at 4 detuned wavelengths with 1 nm-spacing, rather 8 signals in <1 nm spacing. Fig. 8 shows the deployed 4-λ grid with 1 nm-spacing aligned with AWGR channel In6-Out2 residing also within its 5.5 nm 3 dB-BW. Three TLSs and a pluggable QSFP module were used to produce 3 CW signals at λ1=1288.5 nm, λ2=1289.5 nm, λ3=1287.5 nm and 1290.5 nm. The signal at λ1 was modulated by a Mach-Zehnder

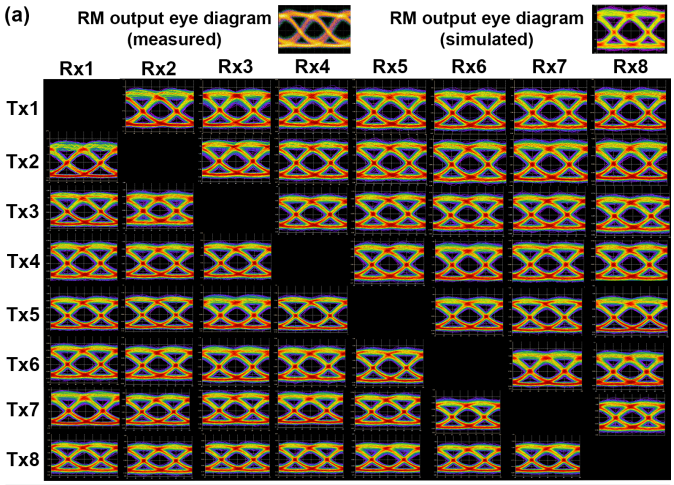


(b)

Chan. Set 1(λm: 1258nm)	Tx1-Rx4	Tx2-Rx3	Tx3-Rx2	Tx4-Rx1	Tx5-Rx8	Tx6-Rx7	Tx7-Rx6	Tx8-Rx5
AWGR peak λ (nm)	1260	1260.3	1260	1260.1	1260	1260.1	1260.2	1260.5
Detuned λ (nm)	1258.5	1261.5	1260.5	1259.5	1258.5	1259.5	1260.5	1261.5
Chan. Set 2 (λm: 1268nm)	**Tx1-Rx5**	**Tx2-Rx4**	**Tx3-Rx3**	**Tx4-Rx2**	**Tx5-Rx1**	**Tx6-Rx8**	**Tx7-Rx7**	**Tx8-Rx6**
AWGR peak λ (nm)	1269.4	1269.4		1269.2	1269.2	1268.9		1269.4
Detuned λ (nm)	1269.5	1270.5		1268.5	1269.5	1268.5		1270.5
Chan. Set 3 (λm: 1278nm)	**Tx1-Rx6**	**Tx2-Rx5**	**Tx3-Rx4**	**Tx4-Rx3**	**Tx5-Rx2**	**Tx6-Rx1**	**Tx7-Rx8**	**Tx8-Rx7**
AWGR peak λ (nm)	1278.9	1279	1278.9	1278.7	1279.1	1278.6	1277.7	1278.8
Detuned λ (nm)	1279.5	1280.5	1278.5	1277.5	1280.5	1278.5	1277.5	1279.5
Chan. Set 4 (λm: 1288nm)	**Tx1-Rx7**	**Tx2-Rx6**	**Tx3-Rx5**	**Tx4-Rx4**	**Tx5-Rx3**	**Tx6-Rx2**	**Tx7-Rx1**	**Tx8-Rx8**
AWGR peak λ (nm)	1288.7	1288.2	1288.5		1288.3	1288.5	1288.5	
Detuned λ (nm)	1288.5	1287.5	1290.5		1289.5	1288.5	1289.5	
Chan. Set 5 (λm: 1298nm)	**Tx1-Rx8**	**Tx2-Rx7**	**Tx3-Rx6**	**Tx4-Rx5**	**Tx5-Rx4**	**Tx6-Rx3**	**Tx7-Rx2**	**Tx8-Rx1**
AWGR peak λ (nm)	1298	1297.6	1297.7	1297.8	1298.1	1297.5	1297.7	1297.7
Detuned λ (nm)	1299.5	1296.5	1297.5	1298.5	1299.5	1296.5	1298.5	1297.5
Chan. Set 6 (λm: 1308nm)	**Tx1-Rx1**	**Tx2-Rx8**	**Tx3-Rx7**	**Tx4-Rx6**	**Tx5-Rx5**	**Tx6-Rx4**	**Tx7-Rx3**	**Tx8-R2**
AWGR peak λ (nm)		1307.2	1306.9	1307.3		1307.4	1306.9	1306.2
Detuned λ (nm)		1307	1306	1307		1308	1306	1305
Chan. Set 7 (λm: 1318nm)	**Tx1-Rx2**	**Tx2-Rx1**	**Tx3-Rx8**	**Tx4-Rx7**	**Tx5-Rx6**	**Tx6-Rx5**	**Tx7-Rx4**	**Tx8-R3**
AWGR peak λ (nm)	1320.9	1321.4	1316.5	1316	1316.1	1316.1	1316.9	1316.5
Detuned λ (nm)	1320	1320	1317.5	1315.5	1316.5	1315.5	1317.5	1316.5
Chan. Set 8 (λm: 1328nm)	**Tx1-Rx3**	**Tx2-Rx2**	**Tx3-Rx1**	**Tx4-Rx8**	**Tx5-Rx7**	**Tx6-Rx6**	**Tx7-Rx5**	**Tx8-Rx4**
AWGR peak λ (nm)	1331.3		1331.2	1325.4	1325.9		1325.4	1328.9
Detuned λ (nm)	1331.5		1331.5	1325.5	1326.5		1325.5	1329

Fig. 10 (a) Simulated fully-loaded AWGR interconnection utilizing the XT-aware wavelength allocation. (b) Deployed AWGR channel peaks and detuned transmission wavelengths for each AWGR Channel Set and for all Tx-Rx combinations.

Modulator (MZM) MZM1 (designated as the Reference signal) while the signals at λ2 and λ3 were modulated simultaneously by MZM2. The MZMs were both driven with two 5 Vpp 25 Gb/s NRZ PRBS7 signals. Three 2.5 nm BPFs (BPF1, BPF2 & BPF3) were used to filter the modulated signals at λ2, λ3 and λ4 after being amplified by SOA2 at MZM2 output, respectively. The modulated signals were further split via 3-dB couplers into 8 branches for inputs #1-8 that were decorrelated to avoid the same-λ signals and XT components to reach simultaneously the receiver. The 8 signals at the 4 wavelengths were launched into the respective AWGR inputs as following: λ1-λ1-λ2-λ2-λ3-λ3-λ4-λ4. The Reference signal emerging at AWGR Output2 was filtered by a narrow 0.5 nm-BPF (BPF4), connected to the OSC and the error-detector (ED). An optical spectrum analyzer was also employed in this case to obtain the optical spectrum after BPF4. The maximum transmission of all signals through the AWGR achieved for TE-polarized signals, was verified by an optical PM, corresponding to the highest contribution of the in-band and out-of-band XT components to the reference signal. Fig. 9(a)-(h) depict the obtained eye diagrams of the Reference signal at λ1 while the additional signals were launched into the AWGR input ports revealing negligible signal degradation while maintaining high ER values of approximately >10 dB for up to 8 added signals, respectively. BER measurements were obtained for the Reference signal, shown in Fig. 9(i), revealing error-free operation with a PP of 3.39 dB at an bit error-rate (BER) of 10^{-9} for the reference



(b)

	Tx B2B rec. power (dBm)	Power penalty values (dB) for fully-loaded configuration							
		Rx1	Rx2	Rx3	Rx4	Rx5	Rx6	Rx7	Rx8
Tx1	-11.9		0.1	2.8	0.9	0.7	1.3	0.1	2.2
Tx2	-11.5	4.4		6.7	1	1.1	1.1	2.6	3.1
Tx3	-11.3	2.9	0.7		0.7	2.9	0.9	1	1.1
Tx4	-11.2	1.4	1.9	1.2		0.8	4.4	2	1.5
Tx5	-11.8	0.7	0.5	4.9	3.6		0.9	0.8	1.9
Tx6	-11.1	1.4	1.6	1.9	0.9	1.7		5.9	6.1
Tx7	-11.2	1.9	0.8	2.1	6.7	2.1	2.1		0.8
Tx8	-11.5	1.4	1.1	5.4	2.1	5.7	2.3	3.1	

Fig. 11 (a) Simulated eye diagrams for all Tx-Rx combinations and (b) power-penalty values for an error-rate of 10^{-9} of the signals received by all socket Rxs in the fully-loaded scenario.

signal when the 2nd signal at the same λ was added and an additional maximum PP of 1.25 dB at a 10^{-9} BER when all 8 signals are transmitted simultaneously, indicating successful proof-of-concept validation of the proposed scheme. The additional PP of 1.25 dB was due to the out-of-band XT (OXT) contributions of the signals at λ2, λ3 and λ4. The BER measurements shown in Fig. 9(j) obtained through simulations with PRBS7/15 signals in *VPI Photonics* matched with the experimentally obtained BER values. The optical signal-to-noise ratio (OSNR) of the 8 signals at the 4 detuned wavelengths before entering the AWGR was measured to be 46.26 dB, 55.9 dB, 56.1 dB and 54 dB for the signals at λ1, λ2, λ3 and λ4, respectively. The OSNR values of all 8 signals at λ1, λ1, λ2, λ2, λ3, λ3, λ4 and λ4, recorded at BPF4 output during simultaneous transmission through the AWGR, were measured to be 40 dB, 40.2 dB, 40 dB, 40.1 dB, 40.2 dB, 37.4 dB, 37.6 dB, 40.1 dB, respectively. These values were calculated after having removed the contribution of the SOA3 noise figure of 5 dB at an operating current of 170 mA. The BER evaluation was limited to PRBS7 patterns and up to 10^{-9} BER values due to equipment limitations and the simulations were performed with up to PRBS15 patterns due to computational resources constraints by the time of the evaluation. The distortion of the signals transmitted at different wavelengths through the entire 5.5. nm AWGR passband (Input6-Output2) was found to introduce up to a maximum of 0.25 dB of PP via simulations in *VPI photonics*.

C. The 8-socket fully-loaded AWGR-based interconnect

Following the proof-of-concept experimental demonstration

TABLE 1 ICT-STREAMS 8-SOCKET INTERCONNECT VS QPI

	Intel QPI	ICT-STREAMS 25 Gb/s line-rate	ICT-STREAMS 50 Gb/s line-rate
No. of Sockets	Up to 8	≥8 sockets	≥8 sockets
No. of max. hops	Up to 2 hops	1 hop	1 hop
Socket line-rate	9.6 Gb/s	25 Gb/s	50 Gb/s
Sock. capacity	307.2(102.4×3) Gb/s	175 (7×25) Gb/s	350 (7×50) Gb/s
MSB capacity	2.45 Tb/s	1.4 Tb/s	2.8 Tb/s
Link EE	16.2 pJ/bit	10.04 pJ/bit	5.02 pJ/bit

with 8×25 Gb/s data channels, we present here the simultaneous fully-loaded all-to-all operation of the proposed XT-aware scheme where all 8 Txs send 7×25 Gb/s data signals to the remaining 7 Rxs, corresponding to an aggregate-capacity of 1.4 Tb/s. The architecture in fully-loaded configuration was validated through physical layer simulations in *VPI Photonics* software using accurate simulation models for the underlying circuits with their responses following closely the experimental behavior obtained from the fabricated devices, i.e. a Si-photonic RM [3] and the 8×8 AWGR [14]. Fig. 10(a) depicts the simulated setup in VPI Photonics comprising all of the building blocks of the envisaged MSB interconnection architecture hosted on a polymer EOPCB, i.e the LD, the RM and the 2nd-order MRR (DE)MUX, the ADIACs, the AWGR, the BPFs and finally the PDs and TIAs. The simulations of the fully-loaded AWGR configuration relied on transmissions on duplicates of the signals at 4 detuned wavelengths with 1 nm-spacing, following the same rationale as the experimental validation. Fig. 10(b) shows the deployed detuned wavelengths for all 56 transmitted signals grouped together considering their AWGR Channel Set, i.e. the same spectral region where their detuned wavelength sub-grid resides, along with their respective AWGR Channel Set mean wavelength (λm) and the AWGR channel peak wavelength. The grey-highlighted areas in Fig. 10(b) correspond to links from each socket to itself which are not utilized in the interconnection. Fig. 11(a) shows on top the eye diagram at 25 Gb/s PRBS7 NRZ operation of the signal generated by a previously evaluated high-speed RM [3] exhibiting an ER value of 5.4 dB that was modeled in *VPI Photonics* by means of its experimentally-obtained Electro-optic response generating ultimately an eye diagram with the same ER that matches with the measured. The simulations of the fully-loaded interconnection with 25 Gb/s PRBS7 NRZ signals revealed clearly-open eye diagrams with an ER between 4.6 dB to 4.9 dB, which can be seen Fig. 11(a) at the bottom, depicts the eye diagrams obtained for all 56 signals transmitted through the respective AWGR channels and as obtained by the Rx of all sockets. BER measurements were also obtained in the simulations for all 56 transmitted signals and can be seen in Fig. 11(b). The Back-to-Back (BtB) values in Fig. 11(b) correspond to the received power (in dBm) of the signals at the respective Txs outputs by the same socket Rx without the AWGR routing stage for an error-rate at 10^{-9}. The PP values (in dB) of all 56 signals transmitted through the AWGR and received by the respective Rx during full-scale operation versus the B2B transmission received power for an error-rate of 10^{-9}, can be seen also found in Fig. 11(b). Error-free operation of all signals at 10^{-9} error-rate was achieved in the simulations for all TxRx combinations in the fully-loaded interconnection scheme with an average PP of 2.18 dB compared to the B2B transmissions. The reason behind the high PP values (≥4.4 dB) obtained for some of the TxRx combinations is the increased neighboring XT values for each one of the given transmitted signals. The following characteristics were also considered for the component models in the *VPI* setup: LD relative intensity noise = -130 dB/Hz & linewidth = 1 MHz, RM E/O BW = 16 GHz & Q-factor = 9900, PD responsivity = 0.75 A/W & dark current = 5 nA, TIA input equivalent noise = 3 μA, Rx BW = 23 GHz. The losses related to the LD-to-Si coupling, the RM transmission penalty, the adiabatic couplers and the MUX/DEMUX stages were simulated as constant attenuators in the VPI setup with 1.5 dB, 3 dB, 0.5 dB and 1.5 dB loss values, respectively. The MUX/DEMUX modules and the OBPFs on the Rx side were simulated with 100 GHz and 35 GHz 3 dB-channel BW, respectively. The successful evaluation via simulations of the fully-loaded configuration significantly highlighted the capability of our proposed XT-aware scheme to enable simultaneous all-to-all interconnections even AWGR with moderate IXT are deployed.

D. Throughput and Energy analysis of glueless 8-socket AWGR-based interconnects and comparison to QPI

The energy efficiency of the proposed Si-photonic interconnect can be estimated by summing the power consumption (PC) of all active components employed in the Tx-Rx link: the PC of the LD, the PC of the RM heater and its DR and the PC of the Rx TIAs. The 8-channel Si-TxRx presented in this work was designed as a part of a TxRx assembly to be co-packaged with the same electronic DR and TIA chips as the 200 (4×50) Gb/s TxRx assembly in [21] by utilizing two 4-ch DRs and two 4-ch TIAs, with each one connected to one of the 4-channel RM-arrays and PD-arrays, respectively. The PC values considered in this analysis are the following: 50 mW/ch for the RM's heater (this work), 61 mW/ch for the RM's DR [21] and 112 mW/ch for the TIA [21], respectively. To calculate the LD optical output power and thus the LD PC, the Tx-Rx optical link power budget must be taken into consideration with respect to the PD sensitivity. The insertion losses of the optical components assumed for the transmission link are: 1.5 dB for the laser-to-Si coupling, 3 dB imposed as the RM transmission penalty, 2×1.5 dB for the MUX and DEMUX, 4×0.5 dB for each one of the ADIAC couplers, 1 dB for the propagation losses in the polymer waveguide and 4 dB for the AWGR, ultimately leading to an optical power budget of 14.5 dB in total. Assuming a 10% laser wall-plug efficiency and the PD sensitivity of -12 dBm, the required LD optical power is estimated at 4.5 dBm corresponding to an estimated LD PC of 28.18 mW. Thus, the total power consumption of the active

elements in the link is estimated at 251 mW, corresponding finally to an EE of 10.04 pJ/bit and 5.02 pJ/bit for operation at 25 Gb/s and at 50 Gb/s, respectively suggesting a 38% and 69% EE improvement over the QPI's 16.2 pJ/bit [5] for operation at 25 Gb/s and 50 Gb/s, respectively.

Table 1 shows a comparison between the QPI interconnect and the presented ICT-STREAMS AWGR-based optical MSB interconnect for operation with 25 Gb/s and 50 Gb/s line-rates. The ICT-STREAMS interconnect targets to increase the limited 8-socket connectivity of QPI-based interconnects to beyond 8 sockets configurations and at the same time to minimize the number of hops to only 1 and thus the latency for inter-socket communication by exploiting the flat-topology benefits of enabled by AWGRs. By relying on 50 Gb/s components, the ICT-STREAMS interconnect can outperform QPI in terms of socket line-rate by a factor of 2.5x and 5x when operating with 25 Gb/s and 50 Gb/s line-rates, respectively. Although QPI has higher socket capacity to ICT-STREAMS interconnect at 25 Gb/s line-rates, scaling to 50 Gb/s line-rates allows the STREAMS interconnect to achieve 350 Gb/s/socket capacity that is 1.13x higher than QPI. Scaling to 50 Gb/s line-rates, allows the ICT-STREAMS interconnect to finally yield an MSB capacity of 2.8 Tb/s which is 1.14x higher than the MSB capacity of QPI. Finally, the ICT-STREAMS achieves a 38% and 69% EE improvement over the 16.2 pJ/bit efficiency of QPI [5] considering 25 Gb/s and 50 Gb/s operation, respectively.

VI. Conclusions

We presented our recent progress towards the realization of an O-band flat-topology 8-socket MSB interconnect exploiting Si-photonics Tx and routing building blocks. We reported on an 8-channel O-band Si-photonic integrated building blocks of the interconnection, i.e. an 8-channel WDM Si-Tx with 400 (8×50 Gb/s) NRZ aggregate capacity with a ~4.5 dB ER under 2.15 Vpp drive and we reviewed on our 8×8 cyclic Si-AWGR with previously validated data routing capabilities of high-speed signals. To address the IXT limitations arising from the deployment of high-XT AWGR devices in CWDM AWGR-interconnections we presented our newly proposed XT-aware wavelength allocation scheme for fully-loaded AWGR configurations by exploiting transmissions on densely-spaced wavelengths featuring also improved wavelength utilization. The XT-aware scheme was experimentally verified with 7×25 Gb/s signals and its capability to enable the full-scale 8×8 AWGR interconnection was further validated through physical layer simulations with 7×8×25 Gb/s signals achieving an aggregate data rate of 1.4 Tb/s.

References

[1] Cisco Global Cloud Index: Forecast and Methodology, 2016–2021 White Paper.[Online]:https://www.cisco.com/c/en/us/solutions/collateral /service-provider/global-cloud-index-gci/white-paper-c11-738085.html

[2] M. A. Taubenblatt, "Optical Interconnects for High-Performance Computing," *IEEE Journal of Lightwave Technology*, vol. 30, no. 4, pp. 448-457, Feb. 2012.

[3] T. Alexoudi et al., "Optics in Computing: From Photonic Network-on-Chip to Chip-to-Chip Interconnects and Disintegrated Architectures," in Journal of Lightwave Technology, vol. 37, no. 2, pp. 363-379, Jan.15, 2019.

[4] D. Mulnix, "Intel Xeon Processor Family Technical Overview," July 2017

[5] Intel. An Introduction to the Intel QuickPath Interconnect. Oct. 23, 2018. [Online]: https://www.intel.com/content/www/us/en/io/quickpath-technology/quick-path-interconnect-introduction-paper.html

[6] Bull SAS, "An efficient server architecture for the virtualization of business-critical applications," White paper 2012.

[7] R. Yu *et al.*, "A scalable silicon photonic chip-scale optical switch for high performance computing systems," *OSA Optics Express.*, vol. 21, no. 26, 2013, pp. 32655-32667.

[8] P. Grani *et al.*, "Flat-topology high-throughput compute node with AWGR-based optical-Interconnects," *IEEE/OSA Journal of Lightwave Tech.*, vol. 34, no. 12, 2015, pp. 2959 – 2968.

[9] T. Lamprecht *et al.*, "EOCB-Platform for Integrated Photonic Chips Direct-on-Board Assembly within Tb/s Applications," *IEEE Electr. Compon. and Techn. Conf. (ECTC)*, San Diego, CA, 2018.

[10] M. Moralis-Pegios *et al.*, "A 160Gb/s (4x40) WDM O-band Tx subassembly using a 4-ch array of Silicon Rings co-packaged with a SiGe BiCMOS IC driver", in *Proc. 45th ECOC*, Ireland, 2019, p. W.2.B.2.

[11] E. El-Fiky *et al.*, "400 Gb/s O-band Si photonic transmitter for intra-DC optical interconnects", Op. Ex. **27**(7) 10258-68 (2019).

[12] J. B. Driscoll et al., "First 400G 8-Channel CWDM Silicon Photonic Integrated Transmitter", GFP 2018, pp. 1-2.

[13] M. Wade et al., "A Bandwidth-Dense, Low Power Electronic-Photonic Platform and Architecture for Multi-Tbps Optical I/O", ECOC 2018.

[14] S. Pitris *et al.*, "Silicon photonic 8×8 cyclic Arrayed Waveguide Grating Router for O-band on-chip communication," *Optics Express*, vol. 26, no. 5, pp. 6276-6284, 2018.

[15] H. Takahashi, K. Oda and H. Toba, "Impact of crosstalk in an arrayed-waveguide multiplexer on N× N optical interconnection," *IEEE/OSA J. of Lightw. Techn.*, vol. 14, no. 6, 1996, pp. 1097-1105.

[16] N. A. Idris and H. Tsuda, "6.4-THz-spacing, 10-channel cyclic arrayed waveguide grating for T- and O-band Coarse WDM," *IEICE Electron. Express*, vol. 13, no. 7, 2010.

[17] S. Pitris *et al.*, "Crosstalk-Aware Wavelength-Switched All-to-All Optical Interconnect Using Sub-Optimal AWGRs," in IEEE Photonics Technology Letters, vol. 31, no. 18, pp. 1507-1510, Sept.15, 2019.

[18] Ethernet Alliance roadmap. [Online]: http://www.ethernetalliance.org/

[19] S. Pitris *et al.*, "O-Band Silicon Photonic Transmitters for Datacom and Computercom Interconnects," in Journal of Lightwave Technology, vol. 37, no. 19, pp. 5140-5148, Oct.1, 2019.

[20] R. Dangel *et al.*, "Polymer Waveguides Enabling Scalable Low-Loss Adiabatic Optical Coupling for Silicon Photonics," *IEEE Journ. of Sel. Top. in Quant. Electr.*, vol. 24, no. 4, 2018, pp. 1-11.

[21] M. Moralis-Pegios *et al.*, "A 4-channel 200 Gb/s WDM O-band Silicon Photonic Transceiver sub-assembly", submitted in OSA Optics Express.

[22] Imec-ePIXfab iSiPP50G. [Online]: http://europracticeic.com