Optics in Computing: from Photonic Network-on-Chip to Chip-to-Chip Interconnects and Disintegrated Architectures

T. Alexoudi, N. Terzenidis, S. Pitris, M. Moralis-Pegios, P. Maniotis, C. Vagionas, C. Mitsolidou, G. Mourgias-Alexandris, G.T. Kanellos, A. Miliou, K. Vyrsokinos and N. Pleros

***Abstract*— Following a decade of radical advances in the areas of integrated photonics and computing architectures, we discuss the use of optics in the current computing landscape attempting to re-define and refine their role based on the progress in both research fields. We present the current set of critical challenges faced by the computing industry and provide a thorough review of photonic Network-on-Chip (pNoC) architectures and experimental demonstrations, concluding to the main obstacles that still impede the materialization of these concepts. We propose the employment of optics in chip-to-chip (C2C) computing architectures rather than on-chip layouts towards reaping their benefits while avoiding technology limitations on the way to manycore set-ups. We identify multisocket boards as the most prominent application area and present recent advances in optically enabled multisocket boards, revealing successful 40Gb/s transceiver and routing capabilities via integrated photonics. These results indicate the potential to bring energy consumption down by more than 60% compared to current QuickPath Interconnect (QPI) protocol, while turning multisocket architectures into a single-hop low-latency setup for even more than 4 interconnected sockets, which form currently the electronic baseline. We go one step further and demonstrate how optically-enabled 8-socket boards can be combined via a 256x256 Hipoλaos Optical Packet Switch into a powerful 256-node disaggregated system with less than 335nsec latency, forming a highly promising solution for the latency-critical rack-scale memory disaggregation era. Finally, we discuss the perspective for disintegrated computing via optical technologies as a means to increase the number of synergized high-performance cores overcoming die area constraints, introducing also the concept of cache disintegration via the use of future off-die ultra-fast optical cache memory chiplets.**

Manuscript received Month XX, 2018; revised Month XX, 2018; accepted Month XX, 2018. Date of publication Month XX, 2018; date of current version Month XX, 2018.

T. Alexoudi, N. Terzenidis, S. Pitris, M. Moralis-Pegios, P. Maniotis, C. Vagionas, C. Mitsolidou, G. Mourgias-Alexandris, A. Miliou and N. Pleros are with the Dept. of Informatics and Center for Interdisciplinary Research & Innovation, Aristotle University of Thessaloniki, 57001, Greece (e-mail: theonial@csd.auth.gr, nterzeni@csd.auth.gr, skpitris@csd.auth.gr, mmoralis@csd.auth.gr, ppmaniot@csd.auth.gr, chvagion@csd.auth.gr, cvmitsol@auth.gr, mourgias@csd.auth.gr, amiliou@csd.auth.gr, npleros@csd.auth.gr).

K. Vyrsokinos is with the Dept. of Physics and Center for Interdisciplinary Research & Innovation, Aristotle University of Thessaloniki, 57001, Greece (email: kv@auth.gr).

G. T. Kanellos was with the Dept. of Informatics and Center for Interdisciplinary Research & Innovation, Aristotle University of Thessaloniki, 57001, Greece and is now with the Dept. of Electrical & Electronic Engineering and High-Performance Networks Research Group, University of Bristol, BS81UB, Bristol (email: gt.kanellos@bristol.ac.uk)

***Index Terms*—silicon photonics, Network-on-Chip, multisocket boards, rack-scale disaggregation, disintegrated computing, macrochip, optical memory, optical packet switch, computing architectures.**

I. Introduction

The paradigm shift experienced during the early 2000's towards dual and quad-core computing architectures [1],[2], turned communication throughput into a key factor for sustaining computational power increases. Workload parallelism and inter-core cooperation were placed among the dominant factors for increasing the number of floating-point-operations-per-second (flops), forcing computing to rely at a constantly growing degree on data movement. This obviously led to an upgraded role for the on-chip and off-chip communication infrastructure: performance advances under certain energy consumption constraints could be only accomplished via a low-power and high-bandwidth interconnect technology. This reality came almost simultaneously with the revolutionary advances triggered in the field of optical interconnects [3]-[6] and silicon photonics [7]-[10], which automatically helped to shape a highly visionary computing landscape: let data processing be done with electrons and data transport with photons, transferring the successful paradigm of long-haul optical communications even to chip-to-chip and on-chip environments [11]-[13].

In less than twenty years, optical interconnects were transformed already to a mature commercial technology for rack-to-rack [14] and even board-to-board communications [15], successfully supporting also the emerging concepts of disaggregated computing [16],[17] and leaf-spine Data Center architectures [18],[19]. The situation is somehow different when dealing with on-chip and chip-to-chip photonic technologies, where commercialization is still relatively far away despite the impressive photonic Network-on-Chip (NoC) architectural concepts [20]-[45] and experimental demonstrations [46]-[66] reported during the last 10 years. In the meantime, computing has also experienced some radical advances: it turned from simple dual- and quadcore layouts into a highly heterogeneous environment both at chip- and system-level, yielding a number of computational settings with a large variety in terms of number of cores and

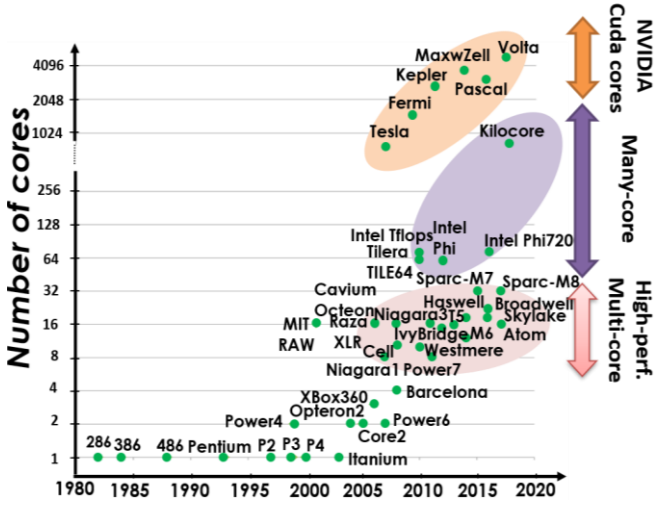

Fig. 1. Evolution from single- to many-core computing architectures

performance capabilities per core. As shown in Fig. 1, General-Purpose Graphic Processing Units (GP-GPUs) [67],[68] can host more than 4000 CUDA cores on the same die, offering, however, only a 2 Gflop per core processing power. Processing power per core increases in manycore architectures, where up to 1000 cores can be employed [69]. However, when high-performance cores are required as in the case of Chip Multiprocessor (CMP) configurations [70],[71] only a number of up to 32 cores can fit on the same die. The ideal scenario towards boosting processing power would of course imply a die that employs as many cores as a GPGPU does, but with core capabilities similar to the high-performance cores available in CMPs.

The number of high-performance cores performing as a single computational entity can scale to higher values only through multi-socket designs with 4 or maximum 8 interconnected sockets. The most recent top-class Intel Xeon 8-socket board yields a total number of up to 224 cores [72], requiring, of course, the use of high-bandwidth off-chip inter-socket interconnects. Going one step beyond the multisocket scheme, disintegration of processor dies has been coined in the recent years as a way to form macrochips that will synergize a high amount of high-performance cores, usually exploiting optical inter-die links [73]. This strongly versatile environment at chip-scale suggests a diverse set of requirements that has to be met by optics, depending on the application target. However, it creates also a new opportunity to rethink the role of optics in on- and off-chip computing, building upon the proven capabilities of optical hardware towards strengthening the compute architecture/technology co-design perspective.

In this paper, we attempt to investigate the new perspectives for optics in computing by extending our work in [74], reviewing the high-priority challenges faced currently by the computing industry and evaluating the credentials of state-of-the-art photonics to address them successfully. We provide a review of the work on photonic NoCs, highlighting the bottlenecks towards their materialization. Building on the state-of-art pNoC implementations [46]-[66] and photonics-enabled multi-socket architectures [75]-[77], we conclude to a solid case for employing integrated photonics in inter-chip multisocket and disintegrated layouts rather than in Network-on-Chip (NoC) implementations, proposing at the same time a flat-topology chip-to-chip multisocket interconnect technology. We demonstrate experimental results using integrated photonic modules towards 40Gb/s multi-socket boards (MSBs) that have the potential to scale to >8-socket designs reducing the energy consumption of conventional Quick Path Interconnect (QPI) links, significantly boosting the number of directly interconnected high-performance cores. Combined with the 256-port Hipoλaos Optical Packet Switch (OPS) that has been recently shown to support sub-μsec latencies in disaggregated computing environments [78]-[80], we evaluate via simulations, a novel optically-enabled rack-scale 256-socket disaggregated setting using a number of 32 interconnected optical 8-socket MSBs. This 256-socket setup can take advantage of traffic localization techniques towards low-latency workload execution, forming a powerful disaggregated rack-scale computing scheme with mean and p99 latencies not higher than 335nsec and 610nsec, respectively, when a 50:50 ratio between on- and off-board traffic is employed. Finally, the utilization of integrated photonics towards transferring the disaggregation concept also at chip-scale is presented, highlighting how the recent work on integrated optical RAMs [81]-[89] can presumably release completely new disintegrated architectures in the future, where precious chip real-estate can be saved by deploying ultra-fast optical cache memories that can reside off-die.

The paper is organized as follows: Section II outlines the main challenges faced today in the computing landscape, providing also a thorough overview of the research on pNoC architectures and experimental demonstrations reported over the last decade and concluding to their main limitations. Section III argues for the employment of optics in MSBs and provides preliminary experimental results on 40Gb/s flat-topology 8-node chip-to-chip (C2C) layouts using O-band integrated photonic transceiver and routing circuitry. The same section underlines the potential of optically enabled MSBs to form low-latency and powerful disaggregated computing systems when combined with the recently demonstrated 256x256 Hipoλaos OPS, presenting simulation results for a 256-node disaggregated setting. Section IV discusses the perspectives for disintegrated computing introducing also the visionary concept of cache disintegration via future off-die optical cache memory chiplets, analyzing the benefits and challenges in this visionary roadmap. Finally, Section V concludes the paper.

II. Current Challenges in Computing and The Photonic Network-on-Chip (pNoC) Escape-Way

In order to define and refine the role of optics in the current computing landscape, it is critical to identify the main challenges currently experienced by the computing industry along the complete hierarchy from on-chip through multi-socket chip-to-chip computational modules. Fig. 2 provides an illustrative overview of the main bandwidth, latency and

TABLE I: OVERVIEW OF PNOC-ENABLED COMPUTING ARCHITECTURES

Architecture	Year	Photonic technol.	Mod. & Rx energy	Data-rate (Gb/s)	Switch technol.
Photonic Torus [22]	2008	Si	0.2 pJ/bit	40	MRR
CORONA [27]	2008	3D Si	N.A.	10	MRR
Firefly [31]	2009	Si	156.25 fJ/bit	5	MRR
FONoC [30]	2009	Si	1 pJ/bit	12.5	Electr.
Si-Pho Clos [33]	2010	3D Si	N.A.	10	P2P photonic
ATAC [20]	2010	Si	300 fJ/bit	1	MRR
NSiP [23]	2010	Poly-Si & multi-layer Si3N4	N.A.	N.A.	Electr.
IRIS [52]	2011	Si	~2 pJ/bit	4	Racetrack resonators
2D-HERT [25]	2012	Si	N.A.	10	MRR
Torus O/E [35]	2012	3D Si	1.21 pj/bit	N.A.	Crux MRR
Multi-Bus NoC [28]	2013	Si	150 fJ/bit	10	Double MRR
Aurora [45]	2014	3D Si	200 fJ/bit	10	MRR
I^2CON [39]	2014	Polymer & 3D Si	102 fJ/bit	10	MRR
METEOR [21]	2014	Si	40 fJ/bit	2.3	XBar
LumiNOC [36]	2014	3D	N.A.	5	MRR
H^2ONoC [24]	2016	Si	420 fJ/bit	10	Electr.
SiS-NoC [29]	2016	Silica	50 fJ/bit	N.A.	MRR
RPNoC [40]	2016	Si	~8 pJ/bit	12.5	N.A.
IMR [37]	2016	N.A.	N.A.	1.5	MRR
TTWA [42]	2017	Si	N.A.	12.5	Broadband MRR
MRONoC [44]	2017	3D Si	N.A.	12.5	MRR
TDM-WDM ONoC [41]	2017	Si	200 fJ/bit	10	MRR
Amon [43]	2017	Si	100 fJ/bit	10	MRR

MRR: Microring Resonator

energy needs for different on-chip and off-chip interconnect layers and data transfer operations in a 20x20mm^2 processor chip fabricated by a 28nm Integrated Circuit (IC) CMOS technology. A digital processing operation performed by the core consumes only 20pJ/bit, but sending data across the chip requires 0.1pJ/bit for a 1mm long electrical link, 1pJ/bit for a 10mm link and goes up to 4pJ/bit for a link length of 40mm. When going off-chip in order to access DRAM, a high amount of 30pJ/bit is consumed, while a chip-to-chip interconnect link like QPI requires 16.2pJ/bit. Accessing L1 cache requires 0.2pJ/bit, while L2 and L3 access requires 1 and 2-4pJ/bit, respectively. Memory bandwidth reduces with increasing memory hierarchy, with L1 memory bandwidth approaching 20GB/sec and gradually decreasing when going to L2 and L3 access until an upper limit of 12.5GB/sec in the case of

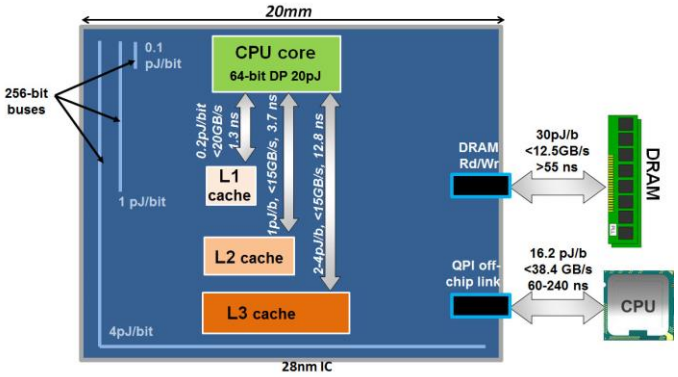

Fig. 2. Energy, bandwidth and latency requirements at different on-chip and off-chip communication needs. The size of every cache memory is bigger for larger capacity caches and their distance from the core is higher as the cache hierarchy increases.

DRAM access. Latency follows the inverse path, starting from a high >55nsec value when fetching from DRAM and gradually reducing with increased memory hierarchy, with L1 access latency being around 1.3nsec. Having this overview, the main challenges today are formed around:

i) Interconnect energy consumption: A modern CPU consumes around 1.7nJ per floating-point operation, [90]-[92], being 85x higher than the 20pJ per floating point required for reaching the Exascale milestone within the gross 20MW power envelope. Current architectures rely to a large degree on data movement, with electronic interconnects forming the main energy consuming factor in both on- and off-die setups [92]. With the energy of a reasonable standard-cell-based, double-precision fused-multiply add (DFMA) being only ~20 pJ, it clearly reveals that fetching operands is much more energy-consuming than computing on them [90]-[92].

ii) Memory bandwidth at an affordable energy envelope: The turn of computing into strongly heterogeneous and parallel settings have transformed memory throughput into a key factor for increasing processing power [91]-[93], with the most efficient way for improvement still being the use of wider memory buses and hierarchical caching. However, the highest memory bandwidth per core in modern multicore processors can hardly reach 20 GB/sec [94],[95], with L1 cache latency values still being >1nsec .

iii) Die area physical constraints: The need for avoiding the latency and energy burden of DRAM access has enforced a rich on-chip L1, L2 and L3 cache hierarchy that typically occupies >40% of the chip real-estate [96]-[98], suggesting that almost half of the die area is devoted to memory and interconnects instead of processing functions.

iv) Cache coherency-induced multi- and broadcasting traffic patterns: The need for cache coherency at intra-chip multi- and manycore setups, as well as at inter-chip multisocket systems, yields communication patterns with strong multi- and broadcast characteristics, that have to be satisfied at a low-latency low-energy profile by the interconnect and network-on-chip infrastructure. Multibus ring topologies form a widely adopted multicast-enabling NoC architecture in current modern multi-core processors [99], but still the cache coherency control messages may often account for more than

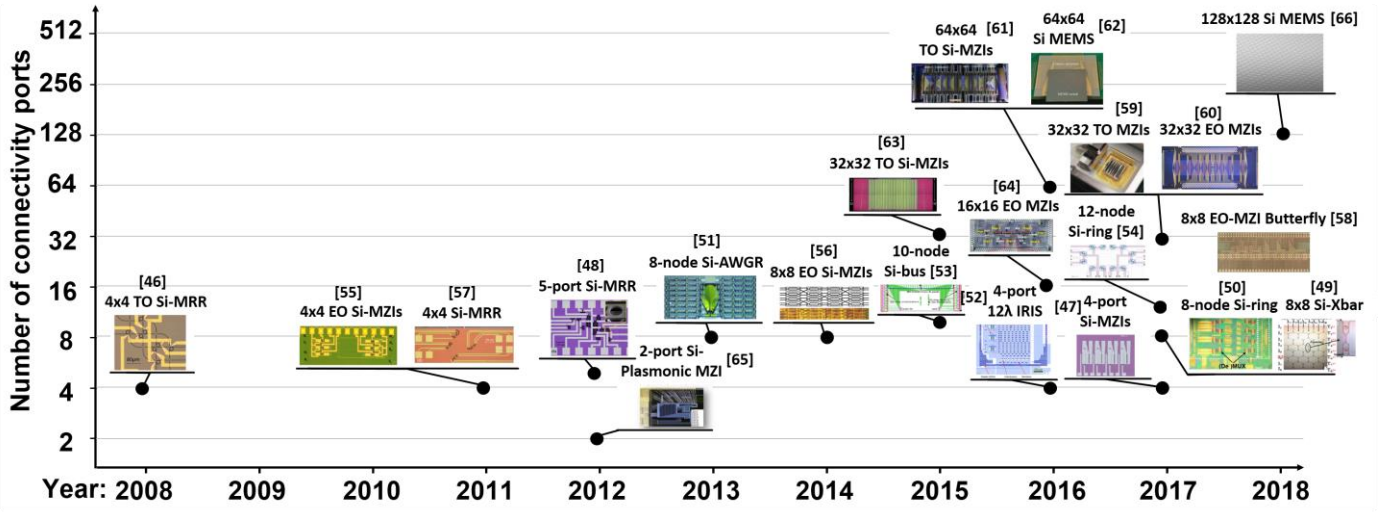

Fig. 3. Evolution of photonic Network-on-Chip and on-chip photonic switches

30% of the total available bandwidth, which may reach even 65% in multisocket settings [100]-[101].

The first and highly enthusiastic attempts to exploit photonics for overcoming the on-chip bandwidth, energy and latency bottlenecks started even a decade ago, mainly inspired by the rapidly growing field of silicon photonics. Despite the, by that time, immaturity of silicon photonic circuitry, a number of breakthrough computing architectures relying on pNoC layouts was demonstrated, proposing and utilizing novel silicon photonic transceiver and switching schemes. The pioneering work on Photonic Torus [22] and CORONA [27] architectures in 2008 was followed by important performance and energy advances in pNoC-enabled many-core designs, addressing even cache-coherency needs in a very efficient way [20]. All this work shaped a highly promising and energy efficient roadmap for many-core computing with >1000 processing cores, with the most important architectures being summarized in Table I. At the same time, it revealed the main requirements and specifications that should be met by silicon photonics towards materializing their on-chip employment in practical NoC layouts: transceiver line-rates between 1-40 Gb/s and optoelectronic conversion energies between a few tens to a few hundreds of fJ/bit were considered in the vast majority of pNoC architectures, with a more detailed breakdown of the relevant metrics per specific pNoC design being reported in Table I.

Ten years after these first efforts, photonic integration technology has reached an important maturity level and has managed indeed to achieve the performance metrics that were assumed by pNoC computing architectures: Silicon photonic modulators can now easily operate at data rates up to 56Gb/s [102]-[104] with an energy efficiency not higher than a few tens of fJ/bit [102], with recent evolution in plasmonic modulators expecting to unleash higher than 100Gbs/s operational rates with even better energy efficiency [105]. At the receiver side, SiGe has turned into a mature photodiode technology with typical operational rates up to 56Gb/s [106]. In terms of on-chip photonic connectivity, on-chip switch arrangements should guarantee connectivity among a high number of nodes in order to allow a >1000-core setting, with every node usually comprising a cluster of up to 4 cores. Fig. 3 illustrates the evolution of the most important pNoC and on-chip switch implementations reported during the last decade. Silicon switches have witnessed a remarkable progress yielding high-port connectivity arrangements with a variety of underlying physical mechanisms like the thermo-optic (TO), electro-optic (EO) and opto-mechanical effects [107] currently allowing for 32×32 EO Mach-Zehnder Interferometric (MZI)-based layouts [60], 64×64 TO MZI designs [61] and up to 128×128 Microelectromechanical switches (MEMS) [66].

All this indicates that integrated photonics can now indeed offer the line-rate, energy, footprint and connectivity credentials required by pNoC-enabled manycore computing architectures. However, the realization of a manycore machine that employs a pNoC layer seems to be still an elusive target, with the main reason being easily revealed when inspecting the non-performance-related co-integration and integration level details of a pNoC-enabled computational setting. Manycore architectures necessitate the on-die integration of a few thousands of photonic structures [21],[22],[27], residing either on 3D integration schemes [33],[39],[45] for a tighter synergy between electronics and photonics or on monolithically co-integrated electronic and photonic structures [12], with transistors and optics being almost at the same layer. Increasing the number of silicon photonic structures on the same die can currently, however, hardly scale beyond 1000 elements [38], with yield forming still a rather unknown factor at these integration scales. At the same time, 3D electro-optic integration has still not managed to fulfil the great expectations that were raised and is still struggling to overcome a number of significant challenges [108] in order to bridge photonics and electronics worlds in a 3D landscape. Last but not least, monolithic integration employing the so called “zero-change” photonics has recently accomplished some staggering achievements reporting on real workload execution over an opto-electronic die with optical core-

memory interconnection [109]-[111]. Nevertheless, this technology has still a rather long-way to go until reaching the complexity and functionality level required by a many-core pNoC design. Line-rates advances from 2.5 Gb/s [109] to the more recent 10 Gb/s [112] are focused at the transceiver modules in simple point-to-point interconnection links, still missing the functional devices that can provide on-chip routing and networking functions.

With almost the complete Photonic Integrated Circuit (PIC) technology toolkit being today available as discrete photonic chips, computing could immediately reap the benefits of optics when tailoring their use in a different architectural environment: instead of pNoC deployments and on-chip manycore processing, photonics could bring a number of advantages if employed for off-die communication in i) multisocket and ii) disintegrated layouts. Both schemes can yield a high number of directly interconnected high-performance cores, unleashing solutions that cannot be met by electronics. At the same time, this approach is fully inline with the 2.5D integration scheme that employs discrete photonic and electronic chips on the same silicon interposer and has made tremendous progress in the recent years [113]-[116]. 2.5D integration can offer tight electronic-photonic co-integration on the same interposer, significantly reducing the electronic link distances and the associated energy consumption. To this end, the employment of off-die communications via discrete photonic chips can form a viable near-term roadmap for the immediate exploitation of photons in computational settings, at least until the longer-term 3D or entirely monolithic co-integration of photonics and electronics become the steam machine of compute technology.

III. Optics for Multisocket Boards

MSB systems rely currently on electrically interconnected sockets and can be classified in two categories:

i) "glueless" configurations, where point-to-point (P2P) interconnects like Intel's QPI [117] or HT [118] can offer high-speed, low-latency, any-to-any C2C communication for a number of 4 or 8 sockets. A 4-socket setup can yield a cache-coherent layout with directly interconnected sockets and latency values that range between 60-240nsec. Scaling to 8-socket designs can only be met through dual-hop links, degrading latency performance but still comprising a very powerful cache-coherent computational setting: Intel's Xeon E7-8800 v4 was the first processor supporting 8-socket configurations and was by that time advertized as being suitable to "dominate the world" [119]. Fig. 4(a) depicts a 4-socket (4S) and 8-socket (8S) layout, respectively, along with their respective interconnects. A typical interconnect like Intel's QPI operates at a 9.6 Gb/s line-rate and consumes 16.2 pJ/bit, while the total bandwidth communicated by every socket towards all three possible directions is 38.4 GB/s, i.e. 307.2 Gb/s [120].

ii) "glued" configurations, where scaling beyond 8-socket layouts is accomplished by exploiting active switch-based setups, such as Bixby [121] and PCI-Express switches [122], in order to interconnect multiple 4- or 8-socket QPI "islands".

With latency and bandwidth comprising the main performance criteria in releasing powerful MSB configurations, "glueless" layouts offer a clear latency-advantage over the "glued" counterparts avoiding by default the use of any intermediate switch. Photonics can have a critical role in transforming "glued" into "glueless" architectures even when the number of interconnected sockets is higher than 8, enabling single-hop configurations, with Fig. 4(b) illustrating how the basic flat-topology can be accomplished for the case of an 8-Socket layout. This has been initially conceived and proposed by UC Davis in their pioneering work on Flat-Topology computing architectures [123] via Arrayed Waveguide Grating Router (AWGR) interconnects, utilizing low-latency, non-blocking and all-to-all optical connectivity credentials enabled by their cyclic-routing wavelength properties [124]. UC Davis demonstrated

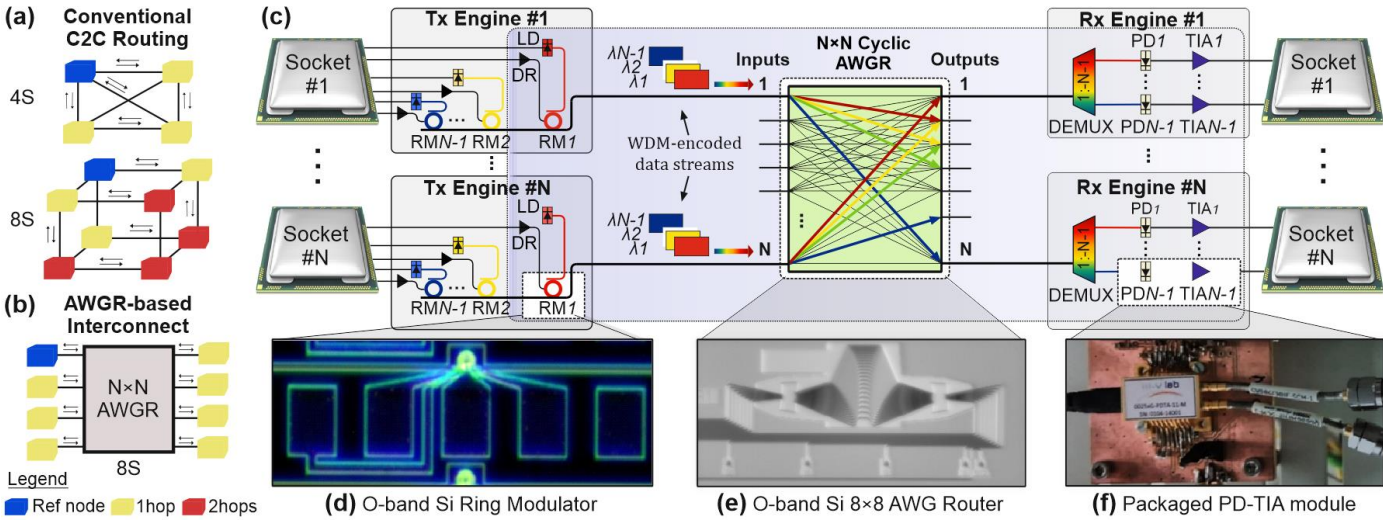

Fig. 4. (a) C2C routing in current electronic 4S and 8S MSBs, (b) Flat-topology 8S layout using AWGR-based routing, (c) proposed N×N AWGR-based optical C2C interconnect for MSB connectivity. Photonic integrated circuits employed as the basic building blocks in the 40Gb/s experimental demonstration: (d) Ring Modulator, (e) 8×8 cyclic-frequency AWGR and (f) PD-TIA module. (blue-highlighted areas: part of the architecture demonstrated experimentally, white-highlighted areas: basic building blocks used for the demonstration).

via gem5 simulations the significant execution time and energy savings accomplished over the electronic baseline [123], revealing also additional benefits when employing bit-parallel transmission and flexible bandwidth-allocation techniques [125]. Experimental demonstrations of AWGR-based interconnection for compute node architectures were, however, constrained so far in the C-band regime, limiting their compatibility with electro-optic Printed Circuit Board (PCB) technology that typically offers a low waveguide loss figure at the O-band [126]. As such, AWGR-based experimental compute node interconnect findings were reported so far only in pNoC architectural approaches, using a rather small line-rate operation of 0.3 Gb/s [127].

The European H2020 project ICT-STREAMS is currently attempting to deploy the necessary silicon photonic and electro-optical PCB technology toolkit for realizing the AWGR-based MSB interconnect benefits in the O-band and at data rates up to 50Gb/s [128]. It aims to exploit wavelength division multiplexing (WDM) Silicon photonics transceiver technology at the chip edge as the socket interface and a board-pluggable O-band silicon-based AWGR as the passive routing element, as shown in a generic N-socket architecture depicted in Fig. 4(c). Each socket is electrically connected to a WDM-enabled Tx optical engine equipped with N-1 laser diodes (LD), each one operating at a different wavelength. Every LD feeds a different Ring Modulator (RM) to imprint the electrical data sent from the socket to each one of the N-1 wavelengths, so that the Tx engine comprises finally N-1 RMs along with their respective RM drivers (DR). All RMs are implemented on the same optical bus to produce the WDM-encoded data stream of each socket. The data stream generated by each socket enters the input port of the AWGR and is forwarded to the respective destination output that is dictated by the carrier wavelength and the cyclic-frequency routing properties of the AWGR [129]. In this way, every socket can forward data to any of the remaining 7 sockets by simply modulating its electrical data onto a different wavelength via the respective RM, allowing direct single-hop communication between all sockets through passive wavelength-routing. At every Rx engine, the incoming WDM-encoded data stream gets demultiplexed with a 1:(N-1) optical demultiplexer (DEMUX), so that every wavelength is received by a distinct PD. Each PD is connected to a transimpedance amplifier (TIA) that provides the socket with the respective electrical signaling.

The flat-topology AWGR-based interconnect scheme requires a higher number of transceivers compared to any intermediate switch solution, but this is exactly the feature that allows to combine WDM with AWGR's cyclic frequency characteristics towards enabling single-hop communication and retaining the lowest possible latency. Link capacity can be increased in this case by residing on channel bonding through bit-parallel schemes, as already reported in [125], by using AWGR designs for waveband instead of single wavelength routing. Utilizing an 8×8 AWGR, the optically-enabled MSB can allow single-hop all-to-all interconnection between 8 sockets, while scaling the AWGR to 16×16 layouts can yield single-hop communication even between 16 sockets, effectively turning current "glued" into "glueless" designs. The ICT-STREAMS on-board MSB aims to incorporate 50GHz single-mode O-band electro-optical PCBs [130], relying on the adiabatic coupling approach between silicon and polymer waveguides [131] for low-loss interfacing of the Silicon-Photonics (Si-Pho) transceiver and AWGR chips with the EO-PCB.

The next subsection describes the first 40Gb/s experimental demonstration of the fiber-interconnected integrated photonic building blocks when performing in the AWGR-based 8-socket MSB architecture, presenting the 40Gb/s experimental results that have been reported in [132] and extending the recently presented operation of the 8-socket architecture at 25 Gb/s [133]. The energy efficiency of the proposed 40 Gb/s chip-to-chip (C2C) photonic link is estimated at 24 pJ/bit but can dramatically go down to 5.95 pJ/bit when transferring the demonstrated fiber-pigtailed layout into an on-board assembled configuration and assuming a 10% wall-plug efficiency for the external laser. This indicates that the on-board version has the credentials to lead to 63.3% reduction in energy compared to the 16.2 pJ/bit link energy efficiency of Intel QPI [134]. Energy efficiency can be additionally improved when incorporating a broadcast-friendly transceiver layout as has been already reported in [135], which can successfully handle the broadcasted traffic typically encountered during cache coherency updates in MSBs and often comprising up to 65% of the total traffic [101]. Finally, we report on how the optically-enabled MSBs can be beneficially employed in rack-scale disaggregated systems when equipped with an additional transceiver lane for dealing with the off-board traffic and are combined with the recently demonstrated Hipoλaos high-port switch architecture [79]. By clustering the traffic exchange in on- and off-board communication ratios typically used in Data Centers, our simulation-based analysis reveals that rack-scale disaggregation among a 256-node system can be successfully accomplished for a variety of communication patterns with an ultra-low mean latency value of < 335 nsec.

A. 40 Gb/s C2C experimental setup and results

The main integrated transmitter, receiver and routing building blocks that were used for demonstrating experimentally the feasibility of the proposed C2C interconnect scheme comprise three discrete chips, i.e. a Si-based RM [136], a Si-based 8×8 AWGR routing platform [76] and a co-packaged PD-TIA [137], which are depicted in Fig. 4(d), (e) and (f), respectively, and have been already demonstrated in their operation as individual elements. The silicon O-band carrier-depletion micro-ring modulator is an all-pass ring resonator fabricated on imec's active platform with demonstrated 50 Gb/s modulation capabilities [136]. The RM can be combined with a recently developed low-power driver [138], leading to an energy efficiency of 1 pJ/bit at 40 Gb/s. For the routing platform, the demonstration relied on an O-band integrated silicon photonic 8×8 AWGR device [76] with 10 nm-channel spacing, a 5.5 nm 3-dB channel bandwidth, a maximum channel loss non-uniformity of 3.5 dB

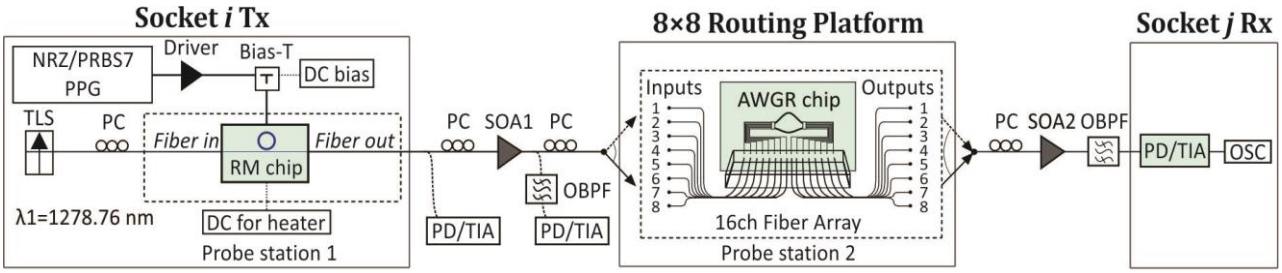

Fig. 5. Experimental setup for the 40Gb/s AWGR-based C2C demonstration

(with 2.5 dB best-case channel insertion losses) and an average channel crosstalk of 11 dB. Finally, the Rx engine employed a co-packaged uni-traveling InGaAs-InP PIN photodiode (PD) connected with a low-power TIA implemented in 0.13 μm SiGe BiCMOS [137]. The PD-TIA sensitivity for operation at 40 Gb/s is -6.4 dBm, respectively, while the energy efficiency for operation at 40 Gb/s is 3.95 pJ/bit.

The experimental setup used for the proof-of-concept demonstration at 40 Gb/s is shown in Fig. 5. A Tunable Laser Source (TLS) was used to produce a Continuous Wave (CW) signal at λ1=1278.76 nm. The RM chip was optically probed with single-mode fibers through TE-polarization Grating Couplers (GC) while an RF probe was used to access the RM electrical pads. A programmable pattern generator (PPG) was employed for producing a 40Gb/s non-return-to-zero (NRZ) pseudo-random binary sequence (PRBS7) that was amplified by a driver amplifier before being applied on the RM along with a reverse-bias DC voltage. After exiting the RM, the signal was sequentially launched into every single port out of the 8 input ports of the AWGR, utilizing a 16-channel fiber array for coupling the signal in and out of the respective AWGR input/output GCs. Depending on the AWGR input port where the incoming data signal at 1278.76 nm was launched, the data stream was routed each time to a different AWGR output port, providing in this way a total number of 8 possible C2C routing scenarios that correspond to 8 different input/output port combinations. To obtain the eye diagrams and the bit error-rate (BER) measurements of the signals, the signal was received by the PD-TIA that was connected to a digital sampling oscilloscope (OSC) and to an error detector (ED), respectively. Semiconductor optical amplifiers (SOA1 & SOA2) were incorporated in the setup after the RM and the AWGR chips, respectively, to compensate for the 9 dB input/output GC losses at the RM chip and the AWGR chip. The signal quality was also monitored directly at the output of the RM and after SOA1 using an optical band-pass filter (OBPF) with 2.5 nm 3-dB bandwidth. Polarization controllers (PC) were used to maintain proper signal polarization.

Fig. 6(a) and 6(b) show the eye diagrams of the modulated signal when connecting the PD-TIA at the RM output and at the SOA1 output, respectively, with an extinction ratio (ER) of 4.2 dB and 4.15 dB and amplitude modulation (AM) of 1 dB and 1.3 dB, respectively. Fig. 6(c)-(j) show the eye diagrams of the signal at the 8 outputs of the AWGR corresponding to the 8 routing scenarios for all possible input-output port combinations denoted as In#iOut#j, indicating clear eye openings and successful routing at 40 Gb/s with ER values of 4.38±0.31 dB and AM values of 2.3±0.3 dB, respectively. The RM was electrically driven with a peak-to-peak voltage of 2.6 V_{pp}, while the applied reverse DC bias voltage was -2.5 V. The optical power of the CW signal injected at the RM input was 8 dBm, with the modulated data signal obtained at the RM output having an average optical power level of -6.3 dBm that was amplified to 10 dBm prior entering the AWGR input. The power of the signal after being routed through the 8 different AWGR port combinations was in the range of -5 dBm to -3.1 dBm. The SOAs were both electrically driven at 175 mA during the evaluation at 40 Gb/s.

Considering the above analyzed operational conditions and a 10% wall-plug-in efficiency for the employed TLS, the energy efficiency of the entire 40Gb/s system is calculated at 24 pJ/bit. However, the 17.5pJ/bit stem from the use of SOA modules for compensating the chip I/O coupling losses, since

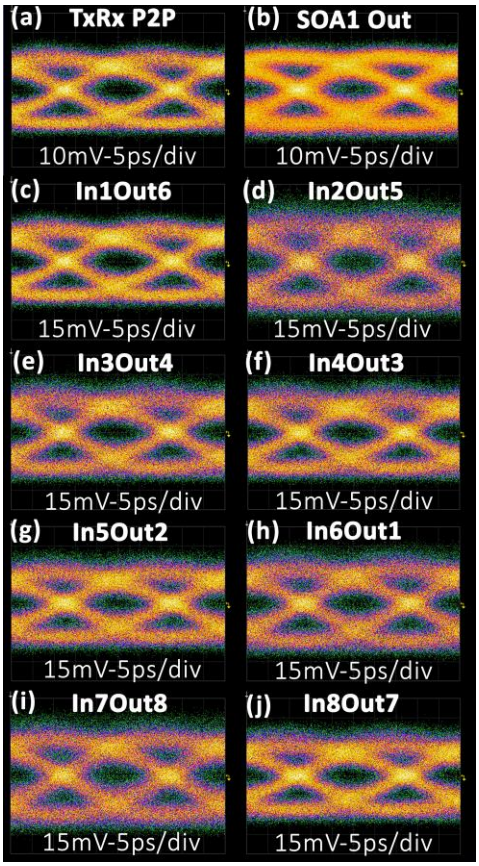

Fig. 6. Eye diagrams at 40 Gb/s: (a) at the RM output, (b) at SOA1 output, (c)-(j) after routing through the respective In#iOut#j I/O ports of the AWGR and amplified by SOA2.

every SOA consumes 350mW. By transferring this interconnect onto a polymer hosting board comprising polymer waveguides (PWG), the high losses associated with the input/output GCs of the RM and AWGR chips can be mitigated as GCs will be replaced with low-loss adiabatic coupling structures that have been shown to yield only 0.5 dB of coupling losses over the entire O-band wavelength range [131]. To this end, a potential on-board layout of the 40 Gb/s C2C interconnect will probably eliminate the need for SOAs in the transmission lines, turning C2C energy consumption into a parameter that depends solely on the power requirements of the RM and its respective electronic driver, the PD-TIA and the external LD that feeds the RM with the CW optical beam. Considering the employment of state-of-the-art RM drivers [138] and assuming an LD with 6.1 dBm output power and a 10% wall-plug efficiency, the energy efficiency of the proposed 40 Gb/s C2C photonic link was estimated at 5.95 pJ/bit that increases to 6.25 pJ/bit when incorporating also state-of-the-art SerDes [139] , assuming a LD-to-RM coupling loss of 3dB [140], a RM insertion loss of 1.5 dB, 0.5dB for every Silicon-to-polymer and polymer-to-Silicon waveguide coupling [131] and an AWGR channel insertion loss of 6dB [135]. These energy efficiency values suggest a 63.3% and 61.4% improvement, respectively, compared to the 16.2 pJ/bit link energy efficiency of Intel QPI [134].

B. Rack-Scale Disaggregated 256-node Architecture using optically-enabled 8S MSBs

The use of optics in "gluing" MSB configurations with even more than 8 sockets can yield significant performance advances also on the next Data Center (DC) hierarchy layer, i.e. at rack-scale. The transform of MSBs into single-hop flat-topology architectures can offer a low-latency and low-energy cache-coherent "island" that can be exploited for workloads with certain traffic locality characteristics. Localized traffic parts can be devoted to a single MSB, calling for MSB-to-MSB communication only for the "global" traffic exchange requests. Recent studies [141]-[143] have indicated that a heavy traffic locality can be observed within the boundaries of a Rack through a variety of emerging DC workloads, while at the same time a number of workloads span their communication capacity through the entire network hierarchy [141], requiring all-to-all connectivity. In this section, we analyze the performance of optically-enabled 8S MSBs in a rack-scale disaggregated compute environment, addressing a highly disruptive emerging computing architecture that seems to endorse the type of mixed local/global communication profile: given that compute, memory, accelerator and storage resources form a set of physically separated disaggregated resources, compute nodes are typically synergized in homogeneous pools that exhibit highly localized traffic for coherency and low-latency reasons, while at the same time require connectivity with remotely located memory or storage pools [17].

Rack-scale disaggregated computing has been introduced towards increasing resource utilization at a reduced energy and cost envelope [17][144][145], necessitating, however, an underlying network infrastructure that can meet a challenging set of requirements [79],[146]: low-latency performance, high-port count connectivity, as well as high data-rate operation. During the first promising demonstrations of disaggregated systems [17], optical circuit switches (OCS) have been employed to interconnect the various types of resource *bricks* due to their high-radix connectivity, scaling to hundreds of ports, along with their datarate-transparent operation. However, OCS comes at the cost of lower switching granularity values, which are not compatible with the 64-byte Last-Level Cache (LLC) word sizes and effectively limit their employment as slow reconfigurable backplanes [147]. We have recently demonstrated a high-port OPS experimental prototype called *Hipoλaos*, which relies on the combination of a set of technologies and architectures for optimizing latency performance even when the number of ports scales to 256 [78],[79] or 1024 [148]. It employs: i) a modified λ-routed Spanke architecture promoting distributed control in small input-port clusters, named as Planes, ii) small-scale *Optical Feed-Forward buffering* via optical delay lines, ensuring high-throughput while reducing latency associated with optoelectronic buffering and the respective electronic SerDes circuitry, iii) multiwavelength AWGR-based routing, utilizing the cyclic routing properties of AWGRs in order to extend the switch radix through a collision-less WDM routing mechanism. Every single Hipoλaos Plane is controlled by means of a Field Programmable Gate Array (FPGA) unit. A more detailed description of the Hipoλaos architecture can be found in [78]. The Hipoλaos architecture can enable also the realization of multicast functionality, building upon the proven efficiency of AWGR devices in multicast operations [149], while its integration roadmap has already been reported in a preliminary 9×9 switch prototype exploiting μm SOI technology for the most challenging integration part, i.e. its Optical Delay Lines [80].

The Hipoλaos switch architecture has been already demonstrated via simulations in 256- and 1024-node experiments exploiting uniform [78] and synthetic [79] traffic profiles without any localized traffic characteristics. Taking advantage of its latency-optimized character, we employ here the 256-port Hipoλaos layout towards connecting 256 nodes clustered in 8-node MSBs and evaluate the network performance when considering a mixed local/global communication profile, forming a dual-layer locality-aware Rack interconnection scheme.

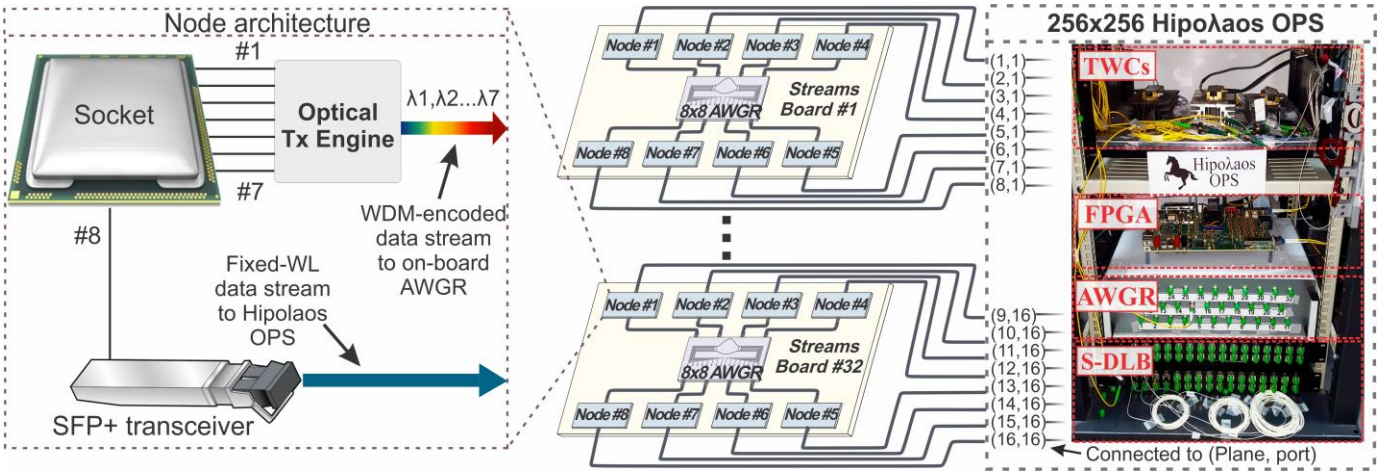

Fig. 7. Illustration of a locality-aware Rack interconnection scheme employing 32 Streams boards, with 8 nodes each, interconnected to a 256x256 Hipoλaos switch. On the left inset, the internal node architecture is presented, while on the right an actual photo of the Hipoλaos experimental prototype is presented (TWCs: Tunable Wavelength Converters, FPGA: Field Programmable Gate Arrays, S-DLB: Shared Delay Line Bank).

Fig. 7 presents a schematic illustration of a 256-node DC system comprising 32 optically-enabled 8-Socket MSBs, with every MSB incorporating 8 network nodes. Every node in the proposed dual-layer network hierarchy is connected via different optical links to an on-board 8×8 AWGR, serving as the intra-board routing infrastructure, as well as to a Hipoλaos-based 256 port switch, providing inter-board all-to-all connectivity. The node interface architecture is depicted in the left inset of Fig. 7, where a socket (CPU or memory) can communicate with any of the remaining 7 on-board nodes by utilizing links #1 to #7, following the transceiver engine layout analyzed in Section III.A. The WDM-encoded data stream, comprising seven lambdas, is forwarded to the on-board AWGR device where every wavelength channel is finally delivered to a different end node. This first layer of switching can ensure 100% throughput of on-board traffic, being in agreement with the requirement for transparent localized traffic forwarding. At the same time, the latency associated with header processing and scheduling is minimized as this is carried out at the network edge, i.e. the socket, for a single-hop collision-less architecture.

The second inter-board layer in the DC switching topology can be accessed through link #8, which forwards inter-board traffic via a fixed-wavelength optical data stream to the Hipoλaos switch. In this design, we have assumed that every socket will connect over an electrical lane to a board-pluggable SFP+ device, but this also could be an additional single-wavelength optical engine directly attached at the socket interface similar to the WDM on-board transceiver engine. The internal architecture of the 256-port Hipoλaos layout has been described in detail in [78] and comprises 16 switch Planes with every Plane aggregating traffic from 16 nodes. In the current architecture with 32 8S MSBs, the input port allocation per switch plane is performed so that Node#i, $1 \leq i \leq 8$, from the odd-numbered boards#j, j=1,3,..,31, connects to the input#k, k=(j+1)/2, of Plane#l, l=i, denoted as input (l,k) of the switch. Moreover, Nodes#i, $1 \leq i \leq 8$, from the even-numbered boards#j, j=2,4,..,32, connect to the input#k, k=j/2, of Plane#l, l=8+i. The proposed port-allocation scheme groups packets from 2 adjacent boards into the respective contention resolution stages of the switch, ensuring minimum contention between the different packets

Throughput and latency performance analysis of the 256-node system, depicted in Fig. 8, has been carried out via simulations using the Omnet++ platform, extending the OptoHPC simulator reported in [150]. A synchronous slotted network operation has been modelled, following the time-slotted operational characteristics usually employed in several high-port-count optical packet switch demonstrations reported during the last years for DataCenter applications. As such, packets are generated at predefined packet-slots, each one lasting for 57.6ns. The traffic profile was customized to distribute a certain percentage of the total traffic generated by every node, uniformly to nodes of the same board (intra-board traffic), while the rest of the traffic was uniformly distributed to nodes residing on the remaining 31 boards (inter-board traffic). In order to offer a thorough evaluation of the architecture's performance in terms of latency, both mean as well as p99 packet delay metrics were collected by the simulation.

The modelled DC system featured node-to-switch and node-to-AWGR channel datarate of 10 Gb/s, along with fixed size packet-length of 72 bytes, comprising 8-bytes for header, synchronization and guardband requirements and 64 bytes data payload, matching the size of a single cache-line transfer. The 10Gb/s line-rate has been selected so as to comply with the experimentally reported values of the first Hipoλaos prototype, despite the optically-enabled MSB has the credentials to scale at 40Gb/s as outlined in Section III.A. However, given that the Hipoλaos switch architecture relies on individual technologies that can provably perform at 40Gb/s [151], it should be expected that a full 40Gb/s line-rate analysis could be supported by the next Hipoλaos prototype in the near future. Regarding the Hipoλaos processing latency, it was assumed to be equal to 456ns in accordance with the experimental results [78], while the propagation latency for the various optical components of the switch (fibers, amplifiers, AWGRs), excluding the optical delay lines, was

modelled to be 35ns.

In order to perform a versatile evaluation of the proposed architecture under different traffic locality patterns, we have considered in our analysis two different cases for the percentage of the intra-/inter-board traffic; 50/50 and 75/25. Performance has been evaluated as a function of the available packet-buffers in the Hipoλaos switch, with the number of buffers ranging for 0 to 4 and corresponding to the maximum number of buffers experimentally demonstrated in [79].

Fig. 8(a) to (c) present the simulation results for the case of 50/50 intra-/inter-board traffic distribution. Fig. 8(a) presents the respective throughput versus the offered load results, concerning the total network traffic (both intra- and inter-board) for different numbers of buffers per Hipoλaos Delay-Line-Bank (DLB). As expected, throughput increases almost linearly with increasing buffer size, reaching 100% for 100% offered load when employing more than 2 packet-size buffers. This can be easily explained when taking into account that 50% of the traffic remains on-board and experiences collisionless routing through the 8x8 AWGR interconnect, while every Hipolaos switch Plane aggregates the remaining the 50% intra-board traffic from nodes uniformly distributed in the different boards of the system. Fig. 8(b) presents the mean packet delay versus offered load, showing that latency ranges between 297ns and 335ns for a buffer size between 0 and 4 packet slots and for loads until 100%. Fig. 8(c) presents the p99 delay results vs. the offered load, revealing a minimum p99 value of 553ns, mainly attributed to the traffic forwarded through the Hipoλaos switch, while reaching 606ns for maximum load and 4 packet-slot buffers per Hipoλaos switch plane [78],[79]. As can be observed, the p99 delay metrics perform an almost step-wise "jump" as contention starts to occur, due to the fact that packets are forwarded to longer delay-line buffers that introduce delays in packet duration granularity. For the 0-buffer case latency remains constant as no retransmission mechanism is used for dropped packets, but even when reaching 100% throughput with more than 2-packet-size buffers the p99 latency doesn't exceed 606nsec. It is important to note that the only point of congestion in the architecture was identified at the Hipoλaos switch, since the intra-board AWGR switching scheme is able to offer 100% throughput with latency values owing solely to the inter-socket data propagation delay, which was assumed to be 2nsec.

Fig. 8(d) to (f) present the simulation results for the case of 75/25 intra-/inter-board traffic distribution. As expected, throughput is slightly increased in all cases, due to the fact that a lower percentage of traffic is headed towards the Hipoλaos switch, where congestion occurs. At the same time mean packet latency is decreased, reaching its maximum value of 215ns with more than 2 packet-buffers. Finally, the p99 latency values remain constant at 553ns, as no extra delay line is utilized in the Hipoλaos DLB blocks.

With sub-μsec latency considered as the main performance target for current memory disaggregated systems [17], the mean and p99 latency values of this novel Hipoλaos-based architecture with clustered optically-enabled 8-Socket MSBs reveals an excellent potential for a practical interconnect solution that can bring latency down to just a few 100's of nanosecond. Allowing on-board nodes to cluster in single-hop configurations over AWGR-based interconnects can yield minimized latency when combined with proper workload allocation for strengthening board-level traffic localization, while off-board traffic benefits from the latency-optimized dynamic switch characteristics of the Hipoλaos design. Latency and energy consumption metrics of this novel disaggregated compute architecture are expected to improve drastically when scaling Hipoλaos operational data-rates to 40Gb/s, making this compatible with the 40Gb/s silicon photonic transmitter engines reported in Section III.A. Finally, this layout could in principle form the basis for replacing the massive QPI "island" interconnection supported by a number of switch technologies like Bixby's [121] and PCI express [122], yielding a powerful network of cache-coherent islands at a maximum p99 latency value just above 600nsec even when a balanced 50/50 traffic locality pattern is followed. This

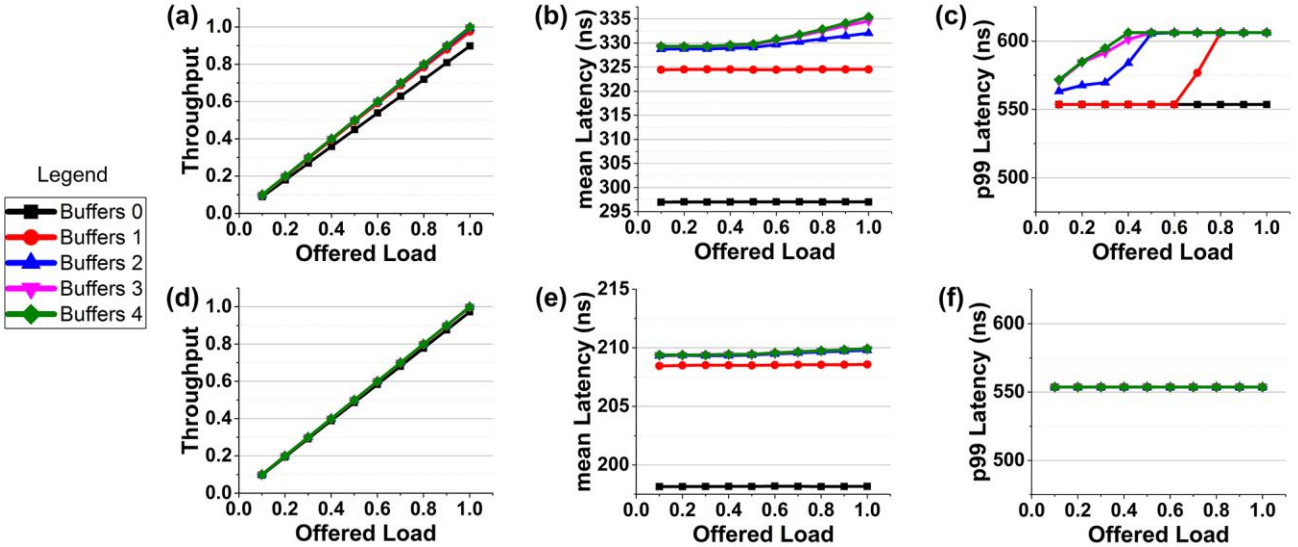

Fig. 8. Simulation results for different number of buffers per DLB (a) Throughput – 50/50, (b) Mean latency – 50/50, (c) P99 latency – 50/50, (d) Throughput – 75/25, (e) Mean latency – 75/25, (f) P99 latency – 75/25

implies a mean latency for a 256-node system that is slightly higher of the maximum average 240nsec latency experienced by an electronic QPI-based 4S MSB, i.e. a 64x bigger computational setting with slightly higher latency compared to current 4-Socket systems. In a real environment, where probably a packet retransmission mechanism has to be incorporated to ensure packet loss avoidance, latency would probably be slightly higher, while the Hipoλaos switch should accommodate some additional mechanism for informing the source node about a dropped packet.

IV. From C2C and Rack-Scale Disaggregation to Disintegrated Computing: Chiplet interconnects with Off-Die Optical Caching

Although MSBs can yield directly interconnected multicore sockets reaching unprecedented performance metrics, they still don't cope with some of the major bottlenecks faced by the computing industry and analyzed in Section II: Memory bandwidth, die area and cache coherency-induced traffic overhead continue to comply with the limitations outlined in Section II. Optically-enabled MSBs hold the potential to yield higher memory bandwidths at a lower off-die interconnect energy envelope, but they still comply to the architectural rule of connecting several dies together without intervening at on-die level. At the same time, the Hipoλaos optical packet switch architecture can yield a high number of interconnected MSBs in a low-latency disaggregated environment, but can obviously be applied only at the next level of compute hierarchy, i.e. C2C or Board-to-Board. As such, on-die computing architectures remain intact and every single die continues to follow the typical design rules for on- and off-chip connectivity: a) a rigid computational setting with pre-defined and rich on-die caching and b) a number of cores that scales inversely with single-core performance and has limited scale-out potential.

To cope with die area constraints allowing for a high number of high-performance cores to communicate within a single computational setting, the pioneering and visionary work of [152] and [93] introduced the concepts of disintegrating computing and macro-chips. Disintegrated computing departs from the conventional monolithic chip layouts and proposes the aggregation of several discrete smaller dies, termed chiplets, into a so called macro-chip, instead of having a single large die. This scheme can overcome area and yield limitations allowing the total silicon area to scale even beyond reticle size limits, with optical switch infrastructures connecting between the multiple physically separated chiplets, as shown in Fig. 9(a).

Disintegrated computing continues, however, to consider electronic cache memories as an indispensable part of every single die, so that the die area is still shared between processing and memory functions. This is obviously enforced by the requirement to have data as close as possible to the core, so that they can be fetched within even a single processor core-cycle in order to yield stall-free execution at least in the cases of Level-1 (L1) cache hits. Any attempt to

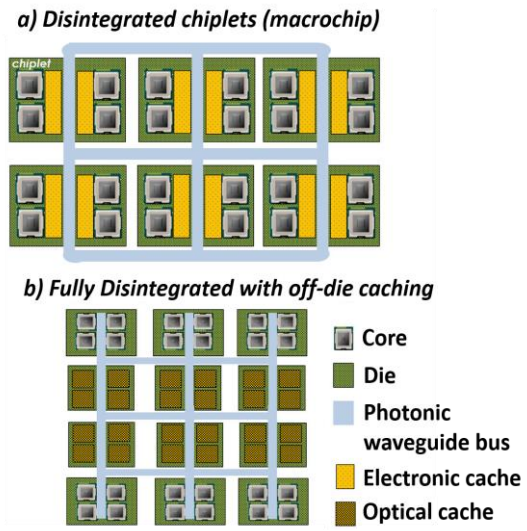

Fig. 9: a) A disintegrated architecture forming a macrochip that comprises six smaller-die chiplets, b) Fully disintegrated setting using also off-die caches as discrete chiplets. Multiple rings are shown for the photonic waveguide bus networking topology but different topologies can be applied as well.

bring cache memories off-die would necessitate an ultra-fast cache and core-to-cache interconnect technology that could operate at a multiple of the core frequency, so that the cache bus and the cache memory could release the data within a few cache clock cycles that will have in total still a duration lower than a single core cycle. With electronic Static RAM technology frequencies not exceeding a few GHz [153], any intervention on the traditional on-die core-cache architectural paradigm will most probably fail even at its conception in case electronics continue to comprise the steam machine of caching functions.

However, the recent advances on the still new technology of Optical Static RAMs and the first designs of optical cache memories [154]-[168] might allow an alternative visionary route towards an expanded disintegrated compute architecture with off-die shared optical caching [74]. This is illustrated in Fig. 9(b), where two types of chiplets are now connected over an optical network infrastructure: processing chiplets including only cores and being devoted to processing functions, and optical cache chiplets that can be accessed by any processing chiplet. Although this is still a highly visionary path with a plethora of challenges to be addressed prior being considered as a viable solution, it is certainly of interest to investigate the unique benefits that may arise by such a platform, reviewing also some of the first recent results obtained when restricting the analysis on a single processing and single optical caching chiplet.

Fig. 10(a) presents a typical example of a modern Chip Multiprocessor (CMP) with multi-level electronic caches and an indicative number of eight processing cores. Specifically, the standard approach is to put dedicated L1d and L1i caches at each core that run at the same speed with the core in order to maintain stall-free core operation assuming cache hits. Down the memory hierarchy, a second unified Level-2 (L2) cache stores both instructions and data, and, depending on the number of cores and the target application, Level-3 (L3)

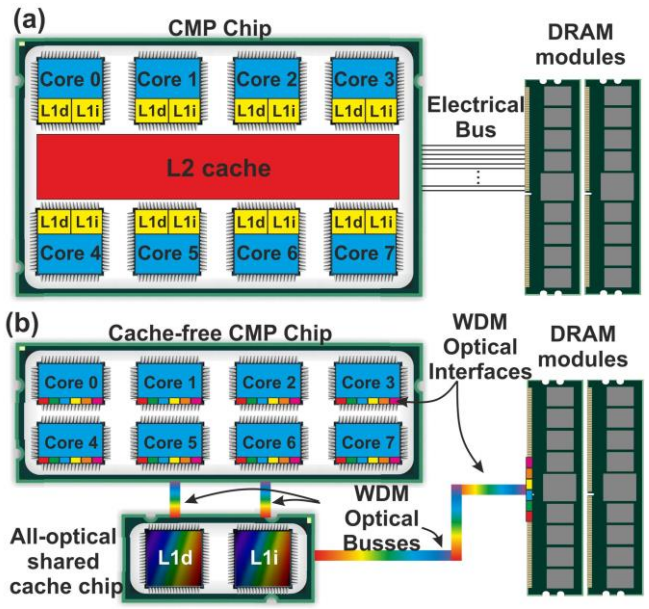

Fig. 10: (a) Conventional CMP architecture with on-chip Cache Memories and Electrical Bus for CPU-MM communication (b) The proposed CMP architecture with off-chip optical Cache Memories between CPU-MM and Optical Busses between them

caches may be eventually also employed and shared among the cores. Last, the Main Memory (MM) connects to the CPU chip with a spatially multiplexed electrical bus. Although L2 and L3 are slower than L1, they are much faster to access than MM and, typically much larger in size than L1, diminishing thus the penalty of an L1 miss.

Releasing the CMP from its electrical caches would save a significant fraction of more than 40% of the die area, yielding a cache-free CMP. This is illustrated in Fig. 10(b), where caching has been disintegrated from processing by using the optical cache memory technology presented in [164]. In the proposed CMP architecture of Fig. 10(b), the shared L1 cache is an optical cache memory technology, connected to CPU and MM via optical waveguides. The direct sharing of the cache among the cores does not necessarily stall the core operation as the optical cache operates at significant higher speeds than the electronic cores, serving concurrently multiple requests from many cores during each electronic core cycle [167]. As can be seen in Fig. 10(b), the proposed optical-bus-based CMP architecture comprises three discrete subsystems: (i) the cache-free CMP chiplet (8 cores are shown as in Fig. 10(a)), (ii) the optical cache chiplet with separate L1i and L1d caches lying next to the CMP chiplet, and (iii) the MM module. The interconnection system between the three subsystems consists of three optical buses with proper WDM optical interfaces at the edge of the CPU cores and the MM. Note that optical to electronic conversion is not required at the cache-memory connection as the optical cache memory operates completely in the optical domain. More details about the optical interface technologies that are being considered in the proposed scheme can be found in [167]. The short access time of the optical cache memory layer can theoretically sidestep any bottleneck phenomena arising from the aggregation of the multiple memory requests from the different cores to the single cache. At the same time, the shared buffering approach eliminates the coherency issues faced by multiple discrete caches in conventional CMP configurations, as data is cached uniquely in the proposed system.

Assuming, for example, an optical CMP-to-cache bus speed and optical cache operational speed of 16GHz, as has been modelled in [164], with a reasonable processing core clock speed of 2GHz, the cache access system performs 8x faster than the processing cores. This indicates that the optical cache can serve all 8 processing cores within a single 2GHz cycle. Regarding latency, every core has 8 cache clock cycles available to complete its request within a single core clock cycle, including of course optoelectronic conversion at the CMP interface, propagation in the optical bus and cache accessing. Assuming a bus length of 1cm, which can be considered as a reasonable value within a macrochip System-in-Package, the time-of-flight is just 50psec for a waveguide-based bus refractive index of 1.5. With optoelectronic conversion taking place at the bus clock speed and at the Memory Address and Memory Buffer Register (MAR and MBR, respectively) interfaces, ultra-fast cache access latency can be obviously easily retained. For detailed timing diagrams that present the optical cache circuitry operation at various stages for both Read and Write operations and the TDM-based access scheme followed in the proposed system of Fig. 10 (b), please refer to [164] and [167], respectively.

This has been extensively analyzed in [167], where also the performance of the system depicted in Fig.10 was thoroughly investigated via detailed simulations using the gem5 simulation engine and the PARSEC benchmark suite. The main findings when comparing the system of Fig.10(a) with the system of Fig.10(b) for the same amount of total cache capacity can be summarized as follows [167]:

- The use of a shared L1 cache yields an important reduction in the cache miss rate of more than 75%, especially when executing parallel programs with high data sharing and exchange needs among their threads; the high volumes of data exchange increase the traffic and consequently the miss rate among the dedicated L1d caches in typical architectures with dedicated L1 caching.
- The shared L1 cache negates the need for cache coherency updates and cache coherency protocols, simplifying the program execution and contributing significantly in cache miss ratio reduction by cancelling all cache coherency misses.
- Cache miss ratio reduction and concurrent multiple core service translate to important execution time speed-up factors that were shown to range between 10% and 20% for computational settings that employed cache capacities equal to the Sparc T5 processor [169] and IBM's Power7 processor [170], respectively.

Extending this concept into a macrochip layout with multiple core and optical cache chiplets can bring additional benefits, since caching will be rather utilized as a pool of resources that will facilitate time and energy savings. Moreover, it can transform computing from a rigid into a versatile and flexible environment, where caching and processing resources can be exploited on demand depending

on the workload requests, allowing eventually also for cache and processing power upgrades similar to the way that DRAM upgrades are currently being performed. These challenging steps simply project the trajectory of mimicking the currently attempted rack-scale disaggregation concept in the chip-scale domain: disintegrate processing, interconnects and memory, introducing at the same time caching as a new type of disintegrated resource.

Building, however, an ultra-fast optical cache memory at the capacity and energy consumption metrics required for this type of applications is a highly challenging task and has still a long way to go. Witnessing, however, the limitations in electronic Static RAM (SRAM) technology that tends to trade-off between access times and energy efficiency [171],[172], optics might have a chance to penetrate even into the traditional stronghold of electronics, i.e. caching. Electronic SRAMs have opted for an increased access time from 150psec to 300psec in order to break the energy efficiency limit of 1fJ/bit as they moved from 45nm to 16nm technology [172]. At the same time, optical SRAM cell architectures have been demonstrated via a variety of SOA-based layouts at Read/Write speeds up to 10Gb/s [156]-[158] with theoretical predictions going up to 40Gb/s [173], [174] and have recently managed to migrate into the low-energy and small-footprint InP-on-Silicon photonic crystal platform that revealed 50psec access times with just 13fJ/bit energy requirements [158]. With optics offering a natural platform for higher operational speeds within the same power envelope, these advances create a unique opportunity for moving from single optical RAM cell to complete optical cache memory module demonstrations in order to counteract the access time-energy efficiency trade-off of electronic SRAMs.

V. Conclusion

We have reviewed the main challenges in current computing related to interconnect energy consumption, memory bandwidth, die are and cache coherency-associated traffic characteristics, overviewing the research attempts over the last decade to resolve these issues via pNoC-enabled manycore architectures. After analyzing the co-integration aspects as the main limiting factors towards the realization of pNoC-based computing, we have defined a new role for photonics in the landscape of computing related to off-die communication infrastructure. In this respect, we discuss how optics can yield single-hop low-latency multisocket boards for even more than 4 interconnected sockets, demonstrating experimental results for 40Gb/s C2C interconnection in a 8-node setup via integrated photonic transmitter and routing circuits. Combining 8-socket optical boards with a Hipoλaos optical packet switch, photonics can yield a powerful 256-node compute disaggregated system with latency values that go well below the sub-usec threshold considered for memory disaggregation environments. Finally, the perspectives and opportunities for scaling disaggregation down to chip-level and enabling disintegrated macrochip architectures are discussed, bringing a new visionary approach for functional disintegration via off-die ultra-fast optical cache memories. Building upon the recent developments of optical Static RAM cell technologies and optical cache memory designs, this article discusses how processing, caching and networking can form a pool of resources within a disintegrated system, eventually allowing for migrating from the rigid computational settings of today to a versatile and scalable macrochip environment in the future.

Acknowledgments

This work has been partially supported by the European H2020 projects ICT-STREAMS (Contract No. 688172) and L3MATRIX (Contract No. 688544). For the work on off-die optical caches, Dr. Alexoudi acknowledges support from the IKY scholarships program that is co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the action entitled “Reinforcement of Postdoctoral Researchers”, in the framework of the Operational Programme “Human Resources Development Program, Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) 2014 – 2020.

References

[1] R. Kalla, Balaram Sinharoy and J. M. Tendler, "IBM Power5 chip: a dual-core multithreaded processor," in *IEEE Micro*, vol. 24, no. 2, pp. 40-47, Mar-Apr 2004.

[2] J. Dorsey et al., "An Integrated Quad-Core Opteron Processor," 2007 IEEE International Solid-State Circuits Conference. Digest of Technical Papers, San Francisco, CA, 2007, pp. 102-103.

[3] K. Bergman, "Photonic Networks for Intra-Chip, Inter-Chip, and Box-to-Box Interconnects in High Performance Computing," Eur. Conf. on Optical Comm. (ECOC) 2006 Tu1.2.1, Cannes, France, Sep. 2006

[4] D. A. B. Miller, "Rationale and challenges for optical interconnects to electronic chips," in Proceedings of the IEEE, vol. 88, no. 6, pp. 728-749, June 2000.

[5] F. Benner, M. Ignatowski, J. A. Kash, D. M. Kuchta and M. B. Ritter, "Exploitation of optical interconnects in future server architectures," in IBM Journal of Research and Development, vol. 49, no. 4.5, pp. 755-775, July 2005.

[6] L. Schares et al., "Terabus: Terabit/Second-Class Card-Level Optical Interconnect Technologies," in IEEE Journal of Selected Topics in Quantum Electronics, vol. 12, no. 5, pp. 1032-1044, Sept.-Oct. 2006.

[7] M. Lipson, "Guiding, modulating, and emitting light on Silicon-challenges and opportunities," in Journal of Lightwave Technology, vol. 23, no. 12, pp. 4222-4238, Dec. 2005.

[8] B. Jalali, M. Paniccia and G. Reed, "Silicon photonics," in IEEE Microwave Magazine, vol. 7, no. 3, pp. 58-68, June 2006.

[9] C. Gunn, "CMOS Photonics for High-Speed Interconnects," in IEEE Micro, vol. 26, no. 2, pp. 58-66, March-April 2006.

[10] B. Analui, D. Guckenberger, D. Kucharski and A. Narasimha, "A Fully Integrated 20-Gb/s Optoelectronic Transceiver Implemented in a Standard 0.13-mu CMOS SOI Technology," in IEEE Journal of Solid-State Circuits, vol. 41, no. 12, pp. 2945-2955, Dec. 2006.

[11] N. Kirman et al., "Leveraging Optical Technology in Future Bus-based Chip Multiprocessors," 2006 39th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'06), Orlando, FL, 2006, pp. 492-503.

[12] C. Baten, A. Joshi, J. Orcutt, A. Khilo, B. Moss, C.W. Holzwarth, M.A. Popovic, H. Li, H.I. Smith, J.L. Hoyt, F.X. Kartner, R.J. Ram, V. Stojanovic, K. Asanovic, Building many-core processor-to-DRAM networks with monolithic CMOS silicon photonics, IEEE Micro 29 (4) (2009) 8–21.

[13] Hewlett Packard Enterprise The machine. [Online]. Available: https://www.labs.hpe.com/the-machine

[14] Intel® Silicon Photonics 100G PSM4 Optical Transceiver. [Online]. Available: https://www.intel.com/content/www/us/en/architecture-and-technology/silicon-photonics/optical-transceiver-100g-psm4-qsfp28-brief.html

[15] Luxtera 2x100G-PSM4 OptoPHY Product Family [Online]. Available: http://www.luxtera.com/embedded-optics/

[16] G. Zervas, H. Yuan, A. Saljoghei, Q. Chen and V. Mishra, "Optically disaggregated data centers with minimal remote memory latency: Technologies, architectures, and resource allocation [Invited]," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. A270-A285, Feb. 2018.

[17] M. Bielski I. Syrirgos, K. Katrinis, D. Syrivelis, A. Reale, D. Theodoropoulos, N. Alachiotis, D. Pnevmatikatos, E. Pap, G. Zervas, V. Mishra, A. Saljoghei, A. Rigo, J.Fernando Zazo, S. Lopez-Buedo, F.Zyulkyarov, M. Enrico and O. Gonzalez de Dios, "dReDBox: Materializing a Full-stack Rack-scale System Prototype of a Next-Generation Disaggregated Datacenter," *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*.

[18] Lenovo Introduction to Spine-Leaf Networking Designs [Online]. Available: https://lenovopress.com/lp0573.pdf

[19] Cisco Data Center Spine-and-Leaf Architecture: Design Overview White Paper. [Online]. Available: https://www.cisco.com/c/en/us/products/collateral/switches/nexus-7000-series-switches/white-paper-c11-737022.html

[20] G. Kurian, J. Miller, J. Psota, J. Eastep, J. Liu, J. Michel, L. Kimerling and A. Agarwal, "ATAC", *Proceedings of the 19th international conference on Parallel architectures and compilation techniques - PACT '10*, 2010.

[21] S. Bahirat and S. Pasricha, "METEOR", *ACM Transactions on Embedded Computing Systems*, vol. 13, no. 3, pp. 1-33, 2014. A. Shacham, K. Bergman and L. P. Carloni, "Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors," in *IEEE Transactions on Computers*, vol. 57, no. 9, pp. 1246-1260, Sept. 2008.

[22] A. Shacham, K. Bergman and L. Carloni, "Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors", *IEEE Transactions on Computers*, vol. 57, no. 9, pp. 1246-1260, 2008.

[23] M. Cianchetti, N. Sherwood-Droz, and C. Batten. Implementing System-in-Package with Nanophotonic Interconnect. Workshop on the Interaction between Nanophotonic Devices and Systems (in conj. with MICRO-43), December 2010.

[24] E. Fusella and A. Cilardo, "H^2ONoC: A Hybrid Optical–Electronic NoC Based on Hybrid Topology," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 1, pp. 330-343, Jan. 2017.

[25] S. Koohi and S. Hessabi, "All-Optical Wavelength-Routed Architecture for a Power-Efficient Network on Chip," in *IEEE Transactions on Computers*, vol. 63, no. 3, pp. 777-792, March 2014.

[26] J. H. Ahn, N. Binkert, A. Davis, M. McLaren and R. S. Schreiber, "HyperX: topology, routing, and packaging of efficient large-scale networks," *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, Portland, OR, 2009, pp. 1-11.

[27] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. Beausoleil and J. Ahn, "Corona: System Implications of Emerging Nanophotonic Technology", *2008 International Symposium on Computer Architecture*, 2008.

[28] C. Chen and A. Joshi, "Runtime Management of Laser Power in Silicon-Photonic Multibus NoC Architecture," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 2, pp. 3700713-3700713, March-April 2013.

[29] E. Kakoulli, V. Soteriou, C. Koutsides and K. Kalli, "Silica-Embedded Silicon Nanophotonic On-Chip Networks," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 6, pp. 978-991, June 2017.

[30] H. Gu, J. Xu and W. Zhang, "A low-power fat tree-based optical Network-On-Chip for multiprocessor system-on-chip", *2009 Design, Automation & Test in Europe Conference & Exhibition*, 2009.

[31] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang and A. Choudhary, "Firefly", *ACM SIGARCH Computer Architecture News*, vol. 37, no. 3, p. 429, 2009.

[32] Mark J. Cianchetti, Joseph C. Kerekes, and David H. Albonesi. 2009. Phastlane: a rapid transit optical routing network. *SIGARCH Comput. Archit. News* 37, 3 (June 2009), 441-450.

[33] A. Joshi, C. Batten, Y. Kwon, S. Beamer, I. Shamim, K. Asanovic and V. Stojanovic, "Silicon-photonic clos networks for global on-chip communication", *2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*, 2009.

[34] S. Le Beux, J. Trajkovic, I. O'Connor, G. Nicolescu, G. Bois and P. Paulin, "Optical Ring Network-on-Chip (ORNoC): Architecture and design methodology," *2011 Design, Automation & Test in Europe*, Grenoble, 2011, pp. 1-6.

[35] Y. Ye, J. Xu, X. Wu, W. Zhang, W. Liu and M. Nikdast, "A Torus-Based Hierarchical Optical-Electronic Network-on-Chip for Multiprocessor System-on-Chip", *ACM Journal on Emerging Technologies in Computing Systems*, vol. 8, no. 1, pp. 1-26, 2012.

[36] C. Li, M. Browning, P. Gratz and S. Palermo, "LumiNOC: A Power-Efficient, High-Performance, Photonic Network-on-Chip", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 33, no. 6, pp. 826-838, 2014.

[37] S. Liu, T. Chen, L. Li, X. Feng, Z. Xu, H. Chen, F. Chong and Y. Chen, "IMR: High-Performance Low-Cost Multi-Ring NoCs", *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 6, pp. 1700-1712, 2016.

[38] Z. Li, M. Mohamed, X. Chen, H. Zhou, A. Mickelson, L. Shang and M. Vachharajani, "Iris", *ACM Journal on Emerging Technologies in Computing Systems*, vol. 7, no. 2, pp. 1-22, 2011.

[39] X. Wu, J. Xu, Y. Ye, X. Wang, M. Nikdast, Z. Wang and Z. Wang, "An Inter/Intra-Chip Optical Network for Manycore Processors", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 4, pp. 678-691, 2015.

[40] X. Wang, H. Gu, Y. Yang, K. Wang and Q. Hao, "RPNoC: A Ring-Based Packet-Switched Optical Network-on-Chip," in *IEEE Photonics Technology Letters*, vol. 27, no. 4, pp. 423-426, Feb.15, 15 2015.

[41] H. Gu, Z. Wang, B. Zhang, Y. Yang and K. Wang, "Time-division-multiplexing–wavelength- division-multiplexing-based architecture for ONoC," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 5, pp. 351-363, May 2017.

[42] B. Zhang, H. Gu, K. Wang, Y. Yang and W. Tan, "Low polling time TDM ONOC with direction-based wavelength assignment," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 6, pp. 479-488, June 2017.

[43] S. Werner, J. Navaridas and M. Luján, "Efficient sharing of optical resources in low-power optical networks-on-chip," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 5, pp. 364-374, May 2017.

[44] H. Gu, K. Chen, Y. Yang, Z. Chen and B. Zhang, "MRONoC: A Low Latency and Energy Efficient on Chip Optical Interconnect Architecture," in *IEEE Photonics Journal*, vol. 9, no. 1, pp. 1-12, Feb. 2017.

[45] Z. Li, A. Qouneh, M. Joshi, W. Zhang, X. Fu and T. Li, "Aurora: A Cross-Layer Solution for Thermally Resilient Photonic Network-on-Chip," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 1, pp. 170-183, Jan. 2015.

[46] N. Sherwood-Droz, H. Wang, L. Chen, B. Lee, A. Biberman, K. Bergman and M. Lipson, "Optical 4x4 hitless slicon router for optical networks-on-chip (NoC)", *Optics Express*, vol. 16, no. 20, p. 15915, 2008.

[47] H. Jia, Y. Xia, L. Zhang, J. Ding, X. Fu, and L. Yang, "Four-port Optical Switch for Fat-tree Photonic Network-on-Chip," Journal of Lightwave Technology, vol. 35, no. 15, pp. 3237–3241, 2017.

[48] Lin Yang, Ruiqiang Ji, Lei Zhang, Jianfeng Ding, Yonghui Tian, Ping Zhou, Yangyang Lu, Weiwei Zhu, "Optical routers with ultra-low power consumption for photonic networks-on-chip," 2012 Conference on Lasers and Electro-Optics (CLEO), San Jose, CA, 2012, pp. 1-2.

[49] G. Fan, R. Orobtchouk, B. Han, Y. Li and H. Li, "8 x 8 wavelength router of optical network on chip", Optics Express, vol. 25, no. 20, p. 23677, 2017.

[50] C. Zhang, S. Zhang, J. Peters and J. Bowers, "8 × 8 × 40 Gbps fully integrated silicon photonic network on chip", *Optica*, vol. 3, no. 7, p. 785, 2016.

[51] R. Yu, S. Cheung, Y. Li, K. Okamoto, R. Proietti, Y. Yin and S. J. B. Yoo, "A scalable silicon photonic chip-scale optical switch for high performance computing systems", Optics Express, vol. 21, no. 26, p. 32655, 2013.

[52] F. Testa, C. Oton, C. Kopp, J. Lee, R. Ortuno, R. Enne, S. Tondini, G. Chiaretti, A. Bianchi, P. Pintus, M. Kim, D. Fowler, J. Ayucar, M. Hofbauer, M. Mancinelli, M. Fournier, G. Preve, N. Zecevic, C. Manganelli, C. Castellan, G. Pares, O. Lemonnier, F. Gambini, P. Labeye, M. Romagnoli, L. Pavesi, H. Zimmermann, F. Di Pasquale and S. Stracca, "Design and Implementation of an Integrated Reconfigurable Silicon Photonics Switch Matrix in IRIS Project", *IEEE Journal of*

Selected Topics in Quantum Electronics, vol. 22, no. 6, pp. 155-168, 2016.

[53] P. Dong, Y. Chen, T. Gu, L. Buhl, D. Neilson and J. Sinsky, "Reconfigurable 100 Gb/s Silicon Photonic Network-on-Chip", *Optical Fiber Communication Conference*, 2014.

[54] F. Gambini, P. Pintus, S. Faralli, M. Chiesa, G. Preve, I. Cerutti and N. Andriolli, "Experimental demonstration of a 24-port packaged multi-microring network-on-chip in silicon photonic platform", *Optics Express*, vol. 25, no. 18, p. 22004, 2017.

[55] M. Yang, W. Green, S. Assefa, J. Van Campenhout, B. Lee, C. Jahnes, F. Doany, C. Schow, J. Kash and Y. Vlasov, "Non-Blocking 4x4 Electro-Optic Silicon Switch for On-Chip Photonic Networks", *Optics Express*, vol. 19, no. 1, p. 47, 2010.

[56] B. Lee, A. Rylyakov, W. Green, S. Assefa, C. Baks, R. Rimolo-Donadio, D. Kuchta, M. Khater, T. Barwicz, C. Reinholm, E. Kiewra, S. Shank, C. Schow and Y. Vlasov, "Monolithic Silicon Integration of Scaled Photonic Switch Fabrics, CMOS Logic, and Device Driver Circuits", *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 743-751, 2014.

[57] T. Hu, H. Qiu, P. Yu, C. Qiu, W. Wang, X. Jiang, M. Yang and J. Yang, "Wavelength-selective 4×4 nonblocking silicon optical router for networks-on-chip", *Optics Letters*, vol. 36, no. 23, p. 4710, 2011.

[58] N. Dupuis, A. Rylyakov, C. Schow, D. Kuchta, C. Baks, J. Orcutt, D. Gill, W. Green and B. Lee, "Nanosecond-scale Mach-Zehnder-based CMOS Photonic Switch Fabrics", *Journal of Lightwave Technology*, pp. 1-1, 2016.

[59] P. Dumais, D. Goodwill, D. Celo, J. Jiang, C. Zhang, F. Zhao, X. Tu, C. Zhang, S. Yan, J. He, M. Li, W. Liu, Y. Wei, D. Geng, H. Mehrvar and E. Bernier, "Silicon Photonic Switch Subsystem With 900 Monolithically Integrated Calibration Photodiodes and 64-Fiber Package", *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 233-238, 2018.

[60] L. Qiao, W. Tang and T. Chu, "32 × 32 silicon electro-optic switch with built-in monitors and balanced-status units", *Scientific Reports*, vol. 7, no. 1, 2017.

[61] L. Qiao, W. Tang and T. Chu, "Ultra-large-scale silicon optical switches," *2016 IEEE 13th International Conference on Group IV Photonics (GFP)*, Shanghai, 2016, pp. 1-2.

[62] T. J. Seok, N. Quack, S. Han, W. Zhang, R. S. Muller and M. C. Wu, "64×64 Low-loss and broadband digital silicon photonic MEMS switches," *2015 European Conference on Optical Communication (ECOC)*, Valencia, 2015, pp. 1-3.

[63] K. Tanizawa, K. Suzuki, M. Toyama, M. Ohtsuka, N. Yokoyama, K. Matsumaro, M. Seki, K. Koshino, T. Sugaya, S. Suda, G. Cong, T. Kimura, K. Ikeda, S. Namiki and H. Kawashima, "Ultra-compact 32 × 32 strictly-non-blocking Si-wire optical switch with fan-out LGA interposer", *Optics Express*, vol. 23, no. 13, p. 17599, 2015.

[64] L. Lu, S. Zhao, L. Zhou, D. Li, Z. Li, M. Wang, X. Li and J. Chen, "16 × 16 non-blocking silicon optical switch based on electro-optic Mach-Zehnder interferometers", *Optics Express*, vol. 24, no. 9, p. 9295, 2016.

[65] S. Papaioannou, D. Kalavrouziotis, K. Vyrsokinos, J. Weeber, K. Hassan, L. Markey, A. Dereux, A. Kumar, S. Bozhevolnyi, M. Baus, T. Tekin, D. Apostolopoulos, H. Avramopoulos and N. Pleros, "Active plasmonics in WDM traffic switching applications", *Scientific Reports*, vol. 2, no. 1, 2012.

[66] K. Kwon, T. Seok, J. Henriksson, J. Luo, L. Ochikubo, J. Jacobs, R. Muller and M. Wu, "128×128 Silicon Photonic MEMS Switch with Scalable Row/Column Addressing", *Conference on Lasers and Electro-Optics*, 2018.

[67] J. Kider, NVIDIA Fermi architecture [Online]. Available: http://www.seas.upenn.edu/~cis565/Lectures2011/Lecture16_Fermi.pdf

[68] D.B. Kirk, W.W. Hwu, NVIDIA G80 architecture and CUDA programming [Online]. Available: http://tjwallas.weebly.com/uploads/3/5/1/9/3519640/nvidia_g80_architecture_and_cuda_programming.pdf

[69] B. Bohnenstiehl, A. Stillmaker, J. Pimentel, T. Andreas, B. Liu, A. Tran, E. Adeagbo and B. Baas, "KiloCore: A 32-nm 1000-Processor Computational Array", *IEEE Journal of Solid-State Circuits*, vol. 52, no. 4, pp. 891-902, 2017.

[70] Intel® Xeon® Platinum 8180 Processor. [Online]. Available: https://ark.intel.com/products/120496

[71] Oracle SPARC M8 Processor. [Online] Available: http://www.oracle.com/us/products/servers-storage/sparc-m8-processor-ds-3864282.pdf

[72] Supermicro SuperServer 7089P-TR4T [Online]. Available: http://www.supermicro.com/products/system/7U/7089/SYS-7089P-TR4T.cfm

[73] K. Raj, J. Cunningham, R. Ho, X. Zheng, H. Schwetman, P. Koka, M. McCracken, J. Lexau, G. Li, H. Thacker, I. Shubin, Y. Luo, J. Yao, M. Asghari, T. Pinguet, J. Mitchell and A. Krishnamoorthy, ""Macrochip" computer systems enabled by silicon photonic interconnects", *Optoelectronic Interconnects and Component Integration IX*, 2010.

[74] N. Pleros, "Silicon Photonics and plasmonics towards Network-on-Chip functionalities for disaggregated computing", Optical Fiber Comm. Conf. (OFC) 2018, Tu3F.4, San Diego, USA, Mar. 2018

[75] P. Grani, R. Proietti, S. Cheung and S. J. B. Yoo, "Flat-Topology High-Throughput Compute Node With AWGR-Based Optical-Interconnects", *Journal of Lightwave Technology*, vol. 34, no. 12, pp. 2959-2968, 2016.

[76] S. Pitris, G. Dabos, C. Mitsolidou, T. Alexoudi, P. De Heyn, J. Van Campenhout, R. Broeke, G. T. Kanellos and N. Pleros, "Silicon photonic 8 × 8 cyclic Arrayed Waveguide Grating Router for O-band on-chip communication", Opt. Express vol. 26, issue 5, pp. 6276-6284 (2018).

[77] G. Dabos, S. Pitris, C. Mitsolidou, T. Alexoudi, D. Fitsios, M. Cherchi , M. Harjanne, T. Aalto, G. T. Kanellos, N. Pleros, "Thick-SOI Echelle grating for any-to-any wavelength-routing interconnection in multisocket computing environments", SPIE OPTO Photonics West 2017, February 2017

[78] N. Terzenidis, M. Moralis-Pegios, G. Mourgias-Alexandris, K. Vyrsokinos, N. Pleros "High-port low-latency optical switch architecture with optical feed-forward buffering for 256-node disaggregated data centers," *Opt. Express*, vol. 26, pp. 8756-8766, 2018.

[79] N. Terzenidis, M. Moralis-Pegios, G. Mourgias-Alexandris, T. Alexoudi, K. Vyrsokinos and N. Pleros, "High-Port and Low-Latency Optical Switches for Disaggregated Data Centers: The Hipoλaos Switch Architecture [Invited]", *Journal of Optical Communications and Networking*, vol. 10, no. 7, p. B102, 2018.

[80] M. Moralis-Pegios, N. Terzenidis, G. Mourgias-Alexandris, M. Cherchi, M. Harjanne, T. Aalto, A. Miliou, K. Vyrsokinos and N. Pleros, "Multicast-Enabling Optical Switch Design Employing Si Buffering and Routing Elements," *IEEE Photon. Technol. Lett*, vol. 30, pp. 712-715, 2018.

[81] N. Pleros, D. Apostolopoulos, D. Petrantonakis, C. Stamatiadis and H. Avramopoulos, "Optical Static RAM Cell", IEEE Photon. Tech. Lett., 21, 2, 73-75, Jan. 2009

[82] D. Fitsios, T. Alexoudi, A. Bazin, P. Monnier, R. Raj, A. Miliou, G.T. Kanellos, N. Pleros and F. Raineri, "An ultra-compact III–V-on-Si Photonic Crystal memory for Flip-Flop operation at 5 Gb/s", OSA Opt. Expr., 24, 4, 4270-4277, Feb. 2016

[83] T. Alexoudi et al., "III–V-on-Si Photonic Crystal Nanocavity Laser Technology for Optical Static Random Access Memories", IEEE J. of Sel. Top. in Quant. Electron., 22, 6, Nov.-Dec. 2016

[84] T. Alexoudi et al., "III-V/SOI Photonic Crystal nanolaser for high-speed wavelength conversion and memory operation", in Proc. Optical Fiber Comm. Conf. (OFC 2016), Tu2K.1, Los Angeles, CA, USA, March 2016

[85] S. Pitris et al., "WDM-Enabled Optical RAM at 5 Gb/s Using a Monolithic InP Flip-Flop Chip" IEEE Photon. Journal, 8, 2, Apr. 2016

[86] S. Pitris et al., "An Optical Content Addressable Memory (CAM) Cell for Address Look-up at 10Gb/s", IEEE Photon. Techn. Lett., 28, 16, 1790 – 1793, Aug. 2016

[87] C. Vagionas, P. Maniotis, S. Pitris, A. Miliou, N. Pleros, "Integrated Optical Content Addressable Memories (CAM) and Optical Random Access Memories (RAM) for Ultra-Fast Address Look-Up Operations" Applied Sciences, vol. 7, no 7, 700, Jul. 2017

[88] L. Liu, et al. "An ultra-small, low-power, all-optical flip-flop memory on a silicon chip," Nat., 4, 182-187, Jan. 2010.

[89] K. Nozaki, et al., "Ultralow-power all optical RAM based on nanocavities," Nat. Photon., 6, 248–252, Feb. 2012.

[90] S. Keckler et al, "GPUs and the future of parallel computing", IEEE Micro, 31, 5, 7-17, Oct. 2011

[91] S. Parker, "The Evolution of GPU Accelerated Computing",Extreme Scale Computing, IL, USA, July 29, 2013

[92] B. Dally, "Challenges for Future Computing Systems", HiPEAC 2015, Amsterdam, NL, 2015

[93] Y. Demir et al, "Galaxy: A High-Performance Energy-Efficient Multi-Chip Architecture Using Photonic Interconnects", ACM Intern. Conf. on Supercomputing (ICS), pp. 303–312, Munich, Germany, June 2014

[94] S. Saini et al, "Performance Evaluation of the Intel Sandy Bridge Based NASA Pleiades Using Scientific and Engineering Applications", NAS Technical Report: NAS-2015-05

[95] S. Saini et al, "Performance Evaluation of an Intel Haswell- and Ivy Bridge-Based Supercomputer Using Scientific and Engineering Applications", NASA Technical Report: NAS-2016-12

[96]S. Borkar, A.A. Chien, "The future of microprocessors," Commun. ACM 54 (5) (2011) 67–77.
[97]L. Zhao, R. Iyer, S. Makineni, J. Moses, R. Illikkal, D. Newell, "Performance, area and bandwidth implications on large-scale CMP cache design," in: Proceedings of the Workshop on Chip Multiprocessor Memory Systems and Interconnects, Phoenix, AZ, USA, 2007.
[98]P. Kongetira, K. Aingaran, K. Olukotun, "Niagara: a 32-way multithreaded sparc processor," IEEE Micro 25 (2) (2005) 21–29.
[99]M. Kumashikar, S. Bendi, S. Nimmagadda, A. Deka and A. Agarwal, "14nm Broadwell Xeon® processor family: Design methodologies and optimizations", *2017 IEEE Asian Solid-State Circuits Conference (A-SSCC)*, 2017.
[100] H. Zhao, A. Shriraman, S. Kumar and S. Dwarkadas, "Protozoa", *ACM SIGARCH Computer Architecture News*, vol. 41, no. 3, p. 547, 2013.
[101] Bull SAS. An efficient server architecture for the virtualization of business-critical applications. white paper 2012. [Online]. Available: https://docuri.com/download/bullion-efficient-server-architecture-for-virtualization_59c1dc51f581710b28689168_pdf
[102] M. Pantouvaki, S. Srinivasan, Y. Ban, P. De Heyn, P. Verheyen, G. Lepage, H. Chen, J. De Coster, N. Golshani, S. Balakrishnan, P. Absil and J. Van Campenhout, "Active Components for 50 Gb/s NRZ-OOK Optical Interconnects in a Silicon Photonics Platform", *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 631-638, 2017.
[103] M. Streshinsky, R. Ding, Y. Liu, A. Novack, Y. Yang, Y. Ma, X. Tu, E. Chee, A. Lim, P. Lo, T. Baehr-Jones and M. Hochberg, "Low power 50 Gb/s silicon traveling wave Mach-Zehnder modulator near 1300 nm", *Optics Express*, vol. 21, no. 25, p. 30350, 2013.
[104] D. Thomson, F. Gardes, J. Fedeli, S. Zlatanovic, Y. Hu, B. Kuo, E. Myslivets, N. Alic, S. Radic, G. Mashanovich and G. Reed, "50-Gb/s Silicon Optical Modulator", *IEEE Photonics Technology Letters*, vol. 24, no. 4, pp. 234-236, 2012.
[105] C. Hoessbacher, A. Josten, B. Baeuerle, Y. Fedoryshyn, H. Hettrich, Y. Salamin, W. Heni, C. Haffner, C. Kaiser, R. Schmid, D. Elder, D. Hillerkuss, M. Möller, L. Dalton and J. Leuthold, "Plasmonic modulator with >170 GHz bandwidth demonstrated at 100 GBd NRZ", *Optics Express*, vol. 25, no. 3, p. 1762, 2017.
[106] J. Van Campenhout et al., "Silicon Photonics for 56G NRZ Optical Interconnects," 2018 Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, 2018, pp. 1-3.
[107] B. Lee, "Silicon Photonic Switching: Technology and Architecture", *2017 European Conference on Optical Communication (ECOC)*, 2017.
[108] S. J. B. Yoo, B. Guan and R. Scott, "Heterogeneous 2D/3D photonic integrated microsystems", *Microsystems & Nanoengineering*, vol. 2, no. 1, 2016
[109] C. Sun, M. Wade, Y. Lee, J. Orcutt, L. Alloatti, M. Georgas, A. Waterman, J. Shainline, R. Avizienis, S. Lin, B. Moss, R. Kumar, F. Pavanello, A. Atabaki, H. Cook, A. Ou, J. Leu, Y. Chen, K. Asanović, R. Ram, M. Popović and V. Stojanović, "Single-chip microprocessor that communicates directly using light", *Nature*, vol. 528, no. 7583, pp. 534-538, 2015.
[110] D. Brunina, A. S. Garg, H. Wang, C. P. Lai, K. Bergman, "Experimental Demonstration of Optically-Connected SDRAM," in Proc. Photonics in Switching 2010, PMC5, July 2010.
[111] Y. Yin, R. Proietti, X. Ye, S. J. B. Yoo, V. Akella, "Experimental Demonstration of optical Processor-Memory Interconnection", in Proc. AIAI2010, 213-216, Beijing, China, Oct. 2010.
[112] A. Atabaki, S. Moazeni, F. Pavanello, H. Gevorgyan, J. Notaros, L. Alloatti, M. Wade, C. Sun, S. Kruger, K. Qubaisi, I. Wang, B. Zhang, A. Khilo, C. Baiocco, M. Popović, V. Stojanović and R. Ram, "Monolithic Optical Transceivers in 65 nm Bulk CMOS", *Optical Fiber Communication Conference*, 2018.
[113] C. Li, T. Li, G. Guelbenzu, B. Smalbrugge, R. Stabile and O. Raz, "Chip Scale 12-Channel 10 Gb/s Optical Transmitter and Receiver Subassemblies Based on Wet Etched Silicon Interposer," in J. of Lightwave Technol., vol. 35, no. 15, pp. 3229-3236, 1 Aug.1, 2017.
[114] C. Li, R. Stabile, F. Kraemer, T. Li and O. Raz, "400 Gbps 2-Dimensional Optical Receiver Assembled on Wet Etched Silicon Interposer," 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), San Diego, CA, 2018, pp. 848-853.
[115] D. Kim, K. Y. Au, H. Y. L. X. Luo, Y. L. Ye, S. Bhattacharya and G. Q. Lo, "2.5D Silicon optical interposer for 400 Gbps electronic-photonic integrated circuit platform packaging," 2017 IEEE 19th Electronics Packaging Technology Conference (EPTC), Singapore, 2017, pp. 1-4.
[116] Xiaowu Zhang et al, "Heterogeneous 2.5D integration on through silicon interposer", Applied Physics Reviews, Vol. 2, No.2, 021308, Mar. 2015
[117] Intel. An Introduction to the Intel QuickPath Interconnect [Online]. Available: https://www.intel.com/content/www/us/en/io/quickpath-technology/quick-path-interconnect-introduction-paper.html
[118] J. Duato, "HyperTransport; technology tutorial," *2009 IEEE Hot Chips 21 Symposium (HCS)*, Stanford, CA, 2009, pp. 1-53.
[119] Intel, "Intel® Xeon® Processor E7-8800/4800/2800 Families", Intel, 2011, [Online]. Available: https://www.intel.com/content/www/us/en/processors/xeon/xeon-e7-8800-4800-2800-families-vol-2-datasheet.html.
[120] R. Maddox, G. Singh and R. Safranek, *Weaving high performance multiprocessor fabric*. Hillsboro, Intel Press, 2009.
[121] T. Wicki and J. Schulz, "Bixby: The scalability and coherence directory ASIC in Oracle's highly scalable enterprise systems," *2013 IEEE Hot Chips 25 Symposium (HCS)*, Stanford, CA, 2013, pp. 1-34.
[122] J. Ajanovic, "PCI express 3.0 overview," *2009 IEEE Hot Chips 21 Symposium (HCS)*, Stanford, CA, 2009, pp. 1-61.
[123] P. Grani, R. Proietti, S. Cheung and S. J. B. Yoo, "Flat-Topology High-Throughput Compute Node With AWGR-Based Optical-Interconnects," in *Journal of Lightwave Technology*, vol. 34, no. 12, pp. 2959-2968, June 15 2016.
[124] R. Proietti, Z. Cao, C. Nitta, Y. Li and S. J. B. Yoo, "A Scalable, Low-Latency, High-Throughput, Optical Interconnect Architecture Based on Arrayed Waveguide Grating Routers", Journal of Lightwave Technology, vol. 33, no. 4, pp. 911-920, 2015.
[125] P. Grani, G. Liu, R. Proietti and S. J. B. Yoo, "Bit-parallel all-to-all and flexible AWGR-based optical interconnects," *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, Los Angeles, CA, 2017, pp. 1-3.
[126] A. Sugama, K. Kawaguchi, M. Nishizawa, H. Muranaka and Y. Arakawa, "Development of high-density single-mode polymer waveguides with low crosstalk for chip-to-chip optical interconnection", Optics Express, vol. 21, no. 20, p. 24231, 2013.
[127] R. Yu, S. Cheung, Y. Li, K. Okamoto, R. Proietti, Y. Yin and S. J. B. Yoo, "A scalable silicon photonic chip-scale optical switch for high performance computing systems", Optics Express, vol. 21, no. 26, p. 32655, 2013.
[128] G. T. Kanellos and N. Pleros, "WDM mid-board optics for chip-to-chip wavelength routing interconnects in the H2020 ICT-STREAMS," Proc. SPIE 10109, Optical Interconnects XVII, 101090D, 20 February 2017.
[129] X. Leijtens, B. Kuhlow and M. Smit, "Arrayed Waveguide Gratings", Springer Series in Optical Sciences, pp. 125-187.
[130] T. Lamprecht, A. Brudener, J. Lambrecht, H. Ramon, R. Premerlani, X. Yin and F. Betschon, "EOCB-Platform for Integrated Photonic Chips Direct-on-Board Assembly within Tb/s Applications", presented at IEEE 68th Electronic Components and Technology Conference (ECTE), pp. 854–858 (2018).
[131] R. Dangel *et al*., "Polymer Waveguides Enabling Scalable Low-Loss Adiabatic Optical Coupling for Silicon Photonics," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 4, pp. 1-11, July-Aug. 2018.
[132] S. Pitris et al., "A 40 Gb/s chip-to-chip interconnect for 8-socket direct connectivity using integrated photonics," submitted at IEEE Photon. J, July 2018.
[133] M. Moralis-Pegios et al, "Chip-to-Chip Interconnect for 8-socket direct connectivity using 25Gb/s O-band integrated transceiver and routing circuits.", to be presented at 48th European Conference on Optical Communication (ECOC), Rome, Italy, Sept. 2018.
[134] C. Gough, I. Steiner and W. Saunders, *Energy Efficient Servers*, Apress, 2015.
[135] S. Pitris *et al*., "O-band Energy-efficient Broadcast-friendly Interconnection Scheme with SiPho Mach-Zehnder Modulator (MZM) & Arrayed Waveguide Grating Router (AWGR)," *2018 Optical Fiber Communications Conference and Exposition (OFC)*, San Diego, CA, 2018, pp. 1-3.
[136] M. Pantouvaki *et al*., "Active Components for 50 Gb/s NRZ-OOK Optical Interconnects in a Silicon Photonics Platform," in *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 631-638, Feb.15, 15 2017.
[137] B. Moeneclaey *et al*., "A 40-Gb/s Transimpedance Amplifier for Optical Links," in *IEEE Photonics Technology Letters*, vol. 27, no. 13, pp. 1375-1378, July1, 1 2015.
[138] H. Ramon *et al*., "Low-Power 56Gb/s NRZ Microring Modulator Driver in 28nm FDSOI CMOS," in *IEEE Photonics Technology Letters*, vol. 30, no. 5, pp. 467-470, March1, 1 2018.

[139] V. Stojanović et al., "Monolithic silicon-photonic platforms in state-of-the-art CMOS SOI processes [Invited]," Optics Express vol. 26, p. 13106-13121 (2018)

[140] M. Seifried *et al*., "Monolithically Integrated CMOS-Compatible III–V on Silicon Lasers," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 6, pp. 1-9, Nov.-Dec. 2018

[141] Arjun Roy, Hongyi Zeng, Jasmeet Bagga, George Porter, Alex C. Snoeren, "Inside the Social Network's (Datacenter) Network", ACM SIGCOMM Computer Communication Review, vol. 45, pp. 123, 2015.

[142] C. Delimitrou, S. Sankar, A. Kansal and C. Kozyrakis, "ECHO: Recreating network traffic maps for datacenters with tens of thousands of servers," 2012 IEEE International Symposium on Workload Characterization (IISWC), La Jolla, CA, 2012, pp. 14-24.

[143] Srikanth Kandula, Sudipta Sengupta, Albert Greenberg, Parveen Patel, and Ronnie Chaiken. 2009. The nature of data center traffic: measurements & analysis. In Proceedings of the 9th ACM SIGCOMM conference on Internet measurement (IMC '09). ACM, New York, NY, USA, 202-208.

[144] H. M. Mohammad Ali, T. E. H. El-Gorashi, A. Q. Lawey and J. M. H. Elmirghani, "Future Energy Efficient Data Centers with Disaggregated Servers," in *Journal of Lightwave Technology*, vol. PP, no. 99, pp. 1-1.

[145] A. D. Papaioannou, R. Nejabati and D. Simeonidou, "The Benefits of a Disaggregated Data Centre: A Resource Allocation Approach," *2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-7.

[146] P. X. Gao, A. Narayan, S. Karandikar, J. Carreira, S. Han, R. Agarwal, S. Ratnasamy, and S. Shenker. 2016. Network requirements for resource disaggregation. In *Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation* (OSDI'16). USENIX Association, Berkeley, CA, USA, 249-264.

[147] Q. Chen, V. Mishra, N. Parsons and G. Zervas, "Hardware programmable network function service chain on optical rack-scale data centers," 2017 Optical Fiber Communications Conference and Exhibition (OFC), Los Angeles, CA, 2017, pp. 1-3.

[148] N. Terzenidis, M. Moralis-Pegios, G. Mourgias-Alexandris, K. Vyrsokinos, N. Pleros, " A 1024-port sub-μsec latency Optical Packet Switch using the Hipoλaos λ-routed modified Spanke switch architecture," *2018 European Conference on Optical Communication (ECOC)*, Rome, 2018, to be published..

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

[150] P. Maniotis *et al*., "Application-Oriented On-Board Optical Technologies for HPCs," in *Journal of Lightwave Technology*, vol. 35, no. 15, pp. 3197-3213, 1 Aug.1, 2017

[151] M. Spyropoulou, N. Pleros, K. Vyrsokinos, D. Apostolopoulos, M. Bougioukos, D. Petrantonakis, A. Miliou, and H. Avramopoulos, "40Gb/s NRZ Wavelength Conversion using a Differentially-Biased SOA-MZI: Theory and Experiment", *J. Lightw. Technol*., vol. 29, no. 10, pp. 1489 – 1499, May 2011.

[152] P. Koka, M. McCracken, H. Schwetman, X. Zheng, R. Ho and A. Krishnamoorthy, "Silicon-photonic network architectures for scalable, power-efficient multi-chip systems" in ACM SIGARCH Computer Architecture News, 38(3), p.117, 2010.

[153] E. Karl et al. "A 4.6GHz 162Mb SRAM design in 22nm tri-gate CMOS technology with integrated active VMIN-enhancing assist circuitry" IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), pp. 230 – 232, 19-23 Feb., San Francisco, CA, 2012.

[154] N. Pleros, D. Apostolopoulos, D. Petrantonakis, C. Stamatiadis and H. Avramopoulos, "Optical Static RAM Cell", IEEE Photon. Tech. Lett., 21, 2, 73-75, Jan. 2009

[155] S. Pitris et al., "WDM-Enabled Optical RAM at 5 Gb/s Using a Monolithic InP Flip-Flop Chip" IEEE Photon. Journal, 8, 2, Apr. 2016

[156] C. Vagionas et al., "Optical RAM and Flip-Flops Using Bit-Input Wavelength Diversity and SOA-XGM Switches", IEEE Journal of Light. Techn., 30, 18, Sep. 15, 2012

[157] D. Fitsios et al., "Dual-Wavelength Bit Input Optical RAM with three SOA XGM switches," IEEE PTL, 24, 1142-1144, 2012.

[158] T. Alexoudi et al, "III–V-on-Si Photonic Crystal nanocavity laser technology for optical Static Random Access Memories (SRAMs)", IEEE J. of Selected Topics in Quantum Electronics, vol. 22, no. 6, pp. 1-10, Nov.-Dec. 2016

[159] T. Alexoudi et al., "Optical Cache memory Peripheral Circuitry: Row and Column Address Selectors for Optical Static RAM Banks", IEEE J. of Light. Techn., special issue on Optical Interconnect, 31, 24, 4098 – 4110, Dec. 2013.

[160] T. Alexoudi et al., "Optical RAM Row Access with WDM-Enabled All Passive Row/Column Decoders", IEEE Photonics Technology Letters, 26, 7, 671 – 674, April 2014.

[161] C. Vagionas et al., "Column Address Selection in Optical RAMs With Positive and Negative Logic Row Access," in IEEE Photonics Journal, vol. 5, no. 6, aID. 7800410, Dec. 2013

[162] C. Vagionas et al., "All-Optical Tag Comparison for Hit/Miss Decision in Optical Cache Memories" IEEE Photon. Techn. Lett., 28, 7, 713 - 716, Dec. 2015

[163] C. Mitsolidou, C. Vagionas, S. Pitris, J. Bos, P. Maniotis, D. Tsiokos, and N. Pleros, "All-Optical Tag Comparator for 10Gb/s WDM-enabled Optical Cache Memory Architectures," in Proceedings of Optical Fiber Communication Conference and Exposition (OFC), paper Th2A.53, Anaheim, CA, 20-24 March 2016

[164] P. Maniotis, D. Fitsios, G.T. Kanellos, N. Pleros, "Optical Buffering for Chip Multiprocessors: a 16GHz Optical Cache Memory Architecture", IEEE J. of Lightwave Technol., Vol. 31, No. 24, pp. 4175-4191, Dec. 2013

[165] P. Maniotis et al., "A 16GHz Optical Cache Memory Architecture for Set-Associative Mapping in Chip Multiprocessors", Optical Fiber Commun. Conf. (OFC) 2014, San Francisco, CA, USA, March 2014

[166] P. Maniotis et al., "A novel Chip-Multiprocessor Architecture with optically interconnected shared L1 Optical Cache Memory", Optical Fiber Commun. Conf. (OFC) 2014, San Francisco, CA, USA, March 2014

[167] P. Maniotis, S. Gitzenis, L. Tassiulas and N. Pleros, "An Optically-enabled Chip-Multiprocessor architecture using a single-level shared Optical Cache Memory", Optical Switching and Networking Journal, Vol. 22, pp 54–68, Nov. 2016

[168] P. Maniotis, S. Gitzenis, L. Tassiulas, N. Pleros, "High-Speed Optical Cache Memory as Single-Level Shared Cache in Chip-Multiprocessor Architectures," Workshop on Exploiting Silicon Photonics for Energy-Efficient High Performance Computing (SiPhotonics), HiPEAC conference, Amsterdam, Netherlands, 19-21 January 2015

[169] J. Feehrer et al., "The Oracle Sparc T5 16-Core Processor Scales to Eight Sockets," IEEE Micro, Vol. 33, no. 2, pp. 48-57, Apr. 2013.

[170] D. F. Wendel et al., "POWER7™ a highly parallel scalable multi-core high end server processor", IEEE J. Solid-State Circuits, vol. 46, no. 1, pp. 145-161, Nov. 2010.

[171] A. Chen, J. Hutchby, V. Zhirnov and G. Bourianoff, "Emerging Nanoelectronic Devices", Wiley, ISBN: 978-1-118-44774-1, Jan. 2015

[172] T. Alexoudi, G.T. Kanellos and N. Pleros, "Optical Flip-Flop and Optical RAM technologies", (Invited Review Article), submitted at Nature Photonics 2018 (under revision)

[173] D. Fitsios et al., "Memory speed analysis of optical RAM and optical flip-flop circuits based on coupled SOA-MZI gates," IEEE J. Sel. Topics Quantum Electron., 18, 2, 1006-1015, Mar/Apr 2012

[174] C. Vagionas, D. Fitsios, K. Vyrsokinos, G. T. Kanellos, A. Miliou and N. Pleros, "XPM- and XGM-Based Optical RAM Memories: Frequency and Time Domain Theoretical Analysis," in IEEE Journal of Quantum Electronics, vol. 50, no. 8, pp. 1-15, Aug. 2014