POLITECNICO DI TORINO
Repository ISTITUZIONALE

Enabling Scalable Disintegrated Computing Systems With AWGR-Based 2.5D Interconnection Networks

Original
Enabling Scalable Disintegrated Computing Systems With AWGR-Based 2.5D Interconnection Networks / Fotouhi, P; Werner, S; Proietti, R; Xiao, X; Yoo, S. J. B.. - In: JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING. - ISSN 1943-0620. - STAMPA. - 11:7(2019), pp. 333-346. [10.1364/JOCN.11.000333]

Availability:
This version is available at: 11583/2972258 since: 2022-10-12T13:09:56Z

Publisher:
Optical Society of America

Published
DOI:10.1364/JOCN.11.000333

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Optica Publishing Group (formely OSA) postprint/Author's Accepted Manuscript

"© 2019 Optica Publishing Group. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modifications of the content of this paper are prohibited."

(Article begins on next page)

12 October 2024

Enabling Scalable Disintegrated Computing Systems with AWGR-based 2.5D Interconnection Networks

Pouya Fotouhi, Sebastian Werner, Roberto Proietti, Xian Xiao and S.J. Ben Yoo
Department of Electrical and Computer Engineering, University of California, Davis, CA
Email: {pfotouhi, swerner, rproietti, xxxiao, sbyoo}@ucdavis.edu

***Abstract*—2.5D integrated systems exploiting electronic interposers to tightly integrate multiple processor dies into the same package suffer from significant performance degradation caused by the large latency overheads of their die-to-die multi-hop electrical interconnection networks. Silicon-photonic interposers with wavelength-routed interconnects can overcome this issue by enabling directly-connected, scalable topologies while exhibiting low-energy optical communication even at large distances. This paper studies the use of an Arrayed Waveguide Grating Router (AWGR) as a scalable, low-latency silicon-photonic interconnection fabric for computing systems with up to 256 cores. Our results indicate that AWGRs could be a key enabler for large-scale interposer systems, offering an average performance speed-up of at least 1.25× with 1.32× lower power for 256 cores compared to state-of-the-art electrical networks while offering a more compact solution compared to alternative photonic interconnects.**

***Index Terms*—Silicon Photonics, Interconnection Networks, 2.5D Integration**

I. Introduction

Commonly called "2.5D" integration exploits an interposer to tightly integrate processor and memory dies side-by-side within the same package, which eliminates the large parasitic from additional packaging, greatly increases in-package memory bandwidth while largely avoiding thermal challenges associated with 3D stacking [1]. In particular, *processor disintegration* [2] represents a promising approach to decrease the overall cost of 2.5D integrated systems by leveraging the higher manufacturing yield of small many-core dies compared to larger ones: for instance, instead of implementing one 64-core processor die, four 16-core processor dies integrated aside each other on an interposer can provide similar processing power at relatively higher manufacturing yield of the smaller dies and, in turn, lower overall system cost. Several commercially-available products already benefit from 2.5D integration [3][4] and future systems can be expected to further exploit the memory bandwidth and cost benefits of 2.5D integration with disintegrated processors by integrating increasing numbers of dies into the same package.

Recent studies have shown that 2.5D integrated systems put significant strain on the network-on-chip (NoC) by exhibiting high communication traffic [5]. In addition, shrinking power budgets, large physical distances, and poor technology scaling of electrical interconnects make the design of energy-efficient high-bandwidth NoCs extremely challenging. Moreover, current commercially-available systems were shown to suffer from high communication latency overheads between processors on the interposer which significantly degrade system performance [6][7]. Low-diameter topologies can potentially reduce the latency but are prohibitively expensive due to the energy consumption of electrical interconnects for interconnecting chips over large distances. This limitation could prevent 2.5D integrated systems to scale to larger number of processing dies in the future.

Silicon photonics (SiPh)–enabling optical communication on chip–features ideal physical properties to overcome these challenges, i.e., almost distance-independent energy consumption and high bandwidth density through dense wavelength-division multiplexing (DWDM)[8]. These advantages over electrical interconnects allow designers to exploit SiPh to design NoCs with 'flatter' low-diameter topologies and capitalize on their performance metrics. Moreover, their distance-independent energy consumption allows adjusting the spacing between dies on interposers to larger/varying physical distances, which was recently found to provide significant performance improvements by overcoming the 'Dark Silicon' problem caused by thermal challenges [9].

Unfortunately, enabling global all-to-all connectivity with SiPh comes with its own challenges. State-of-the-art SiPh all-to-all fabrics proposed to date are either based on optical buses [10] or wavelength-routed photonic NoCs (WRPNoCs)[11]. While bus-based designs quickly become impractical and cost-inefficient due to either large numbers of waveguides or wavelengths, WRPNoCs are based on microring resonators (MRRs) to perform wavelength-selective routing, causing power overheads for thermo-optical control and a challenging physical layout. The ideal wavelength routing fabric would provide all-to-all connectivity without excessive need for waveguides, wavelengths, and MRR heating, while enabling a compact physical implementation.

The Arrayed Waveguide Grating Router (AWGR) enable scalable, low-loss wavelength routing between all input and output ports by utilizing N wavelengths and N input and output waveguides in support of an all-to-all $N \times N$ interconnection. Recent fabrication advances in CMOS-compatible Silicon Nitride (SiN) AWGRs enable footprints of $\sim$1mm^2 [12], which is very compact as the only fabric needed for routing (as opposed to 100s of MRRs in WRPNoCs [13]). Moreover, four recent key demonstrations make AWGRs a viable candidate as a high-bandwidth interconnect. First, recently demonstrated sub-pJ Pulse-Amplitude Modulation (PAM4) transceivers at 40Gbps data rate offer high-data-rate low-energy commu-

nication on a single wavelength [14]. Second, *bit-parallel* AWGRs capable of routing multiple wavelengths to the same output port by exploiting the AWGR's cyclic routing properties now enable low levels of DWDM inside AWGRs, too [15]. Third, recent demonstrations of AWGR fabricated on an SiPh interposer show low crosstalk at scale [16]. Fourth, multiple AWGRs can now be fabricated atop each other with negligible inter-layer crosstalk, thereby eliminating any layout and area concerns of multiple AWGRs inside a NoC [17]. Based on these advancements, AWGRs could be a major enabler for energy-efficient, high-bandwidth, and scalable all-to-all connectivity in 2.5D integrated system and therefore deserve a detailed analysis of their potentials and shortcomings.

This article is a significantly extended version of our previously published paper from IEEE/ACM NOCS 2018 [18], which was the first to analyze the suitability of AWGRs as the interconnection fabric in 2.5D integrated systems, and contributes the following major additions: *(i)* a more comprehensive discussion on designing networks with AWGRs; *(ii)* a detailed analysis of AWGR's ability to scale to larger number of nodes; *(iii)* a performance and evaluation study of AWGR-based interconnects for systems of larger scale (up to 256 cores); *(iv)* a discussion of new insights on the impact and opportunities of SiPh and AWGRs on future systems. Specifically, we make the following **contributions**:

- A scalability study of large-scale 2.5D integrated computing systems showing that AWGR-based interconnection networks are a promising and suitable solution in terms of latency, bandwidth, and energy per bit.
- An exploration of different AWGR-enabled topologies, as well as their use cases and suitability to solve the challenges of interposer-based large-scale systems with up to 256 cores.
- An extensive power and performance evaluation of AWGR-based networks and a comparison to state-of-the-art interconnects with up to 256 cores (16 processor dies).

Our results show that AWGR-based topologies can offer an average speed-up of at least 1.2× (64 cores) and 1.25× (256 cores) compared to the closest electrical competitor for a range of PARSEC3.0/SPLASH-2x workloads with at least 1.32× lower power, and more than 2× reductions in average packet latency across different synthetic workloads with up to 3× sustained bandwidth at 256 nodes. These results suggest that AWGRs could be a key enabler for scaling 2.5D integrated systems both in terms of performance and power by providing a low-latency, scalable, and lower-power interconnect, and could therefore be of high impact for future 2.5D integrated computing systems.

II. 2.5D Integrated Systems

A. 2.5D Integration: Opportunities and Challenges

Increasing interposer sizes offer many opportunities for future large-scale systems inside a single package and enable higher numbers of processor and memory dies to be tightly integrated side-by-side on an interposer. Figure 1 depicts example floorplans of 2.5D integrated systems with a large many-core processor (64 and 256 cores) disintegrated into

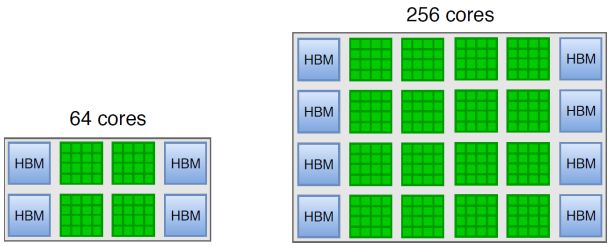


Fig. 1: Example interposer-based systems integrating 64- and 256-core processors composed of 16-core processor dies alongside 3D-stacked DRAMs (in this example, high-bandwidth memories (HBMs) [20]).

smaller 16-core dies and 3D-stacked DRAMs distributed along the two opposite edges of the chip (as commonly found in literature and commercial designs [3], [4], [2]). [1]. Such systems could enable over 1000 processor cores with hundreds of GB memory capacity tightly integrated in the same package and thereby be a key enabler for future high-performance chips operating at high energy efficiency.

Several studies have explored the design space of 2.5D integrated systems [5], interconnection networks extended to the interposer to increase bandwidth [1], processor disintegration to lower cost through improved overall manufacturing yield [2], and have made compelling cases for enabling exascale systems [19]. Nevertheless, Loh et al. [5] identified numerous design challenges, many of which are yet to be solved.

First, the trend towards growing numbers of high-bandwidth memory (HBM) stacks inside the same package, more channels per HBM, and wider DRAM buses to increase memory bandwidth leads to higher bandwidth demands on the NoC which will make the implementation of electrical NoCs within acceptable power envelopes extremely challenging–especially in combination with the large distances imposed by interconnecting several dies on an interposer.

Second, the NoC's clock network must deal with die-to-die-to-interposer process variations, possibly even with different technology generations of different dies or heterogeneous integration of multiple different dies. Loh et al. [5] propose to decompose the NoC into smaller, independent clock domains to have easier timing and to support dynamic voltage and frequency scaling (DVFS), indicating that topologies should ideally support clustering or be hierarchical.

Thirdly, large distances (e.g., AMD's FURY is $1011mm^2$ [4]) and routing between dies increases link latency, suggesting that disintegration comes with a performance-cost trade-off. Routing electrical signals over such distances at satisfactory speed can only be attained with power-consuming repeater circuitry, resulting in more of the power budget being dedicated to the NoC and less to the compute (assuming a system operating under a power cap) [5]. Electrical NoCs tailored to interposer-based systems were shown to be more efficient than conventional NoCs for monolithic chips [1],

[1]Although the processor dies in this example are many-core processors, heterogeneous integration of various different computing and memory chips (such as GPUs, FPGAs, non-volatile memories, etc.) have also been considered an attractive solution for future systems (and would equally benefit from the contributions of this paper) [19]

but cannot fully overcome these limitations. Especially for larger-scale systems with hundreds of cores implemented with tens of processor dies, the interconnection network represents a major obstacle to power efficiency.

B. Using Silicon Photonics To Overcome Design Challenges

Recent studies have shown the performance, power, and scalability benefits of integrated SiPh interconnects in interposer-based systems compared to their electrical counterparts, which become increasingly evident with growing number of dies and physical distances [21]. These benefits are mainly enabled by the physical properties of optical communication, which offer distance-independence in terms of energy and latency and provide high bandwidth links with supreme scalability.

The energy-efficient high-bandwidth interconnects offered by SiPh provide sufficient bisection bandwidth in the NoC to support core clustering, which, in turn, allows practical and efficient DVFS control by grouping clustered cores into separate clock domains. Besides, the discussed physical properties of SiPh allow to implement flatter topologies (and even all-to-all connectivity) in NoCs with much higher scalability than electrical interconnects. This can be leveraged to offer very low latencies even for large physical distances (like in the 256-core example in Figure 1), effectively giving the illusion of moving cores 'closer together'.

While the scientific literature is replete with proposals utilizing SiPh fabrics to construct NoCs [13][22][23], they do not study the utilization of AWGRs in interposer-based systems which provide highly scalable and energy-efficient all-to-all connectivity with just a single passive device. More importantly, we believe that the significant technological improvements of AWGRs in the last years (in terms of footprint, loss, and crosstalk) make them superior to state-of-the-art SiPh fabrics. AWGR-based NoCs eliminate the need for on-chip heating power for thermo-optical control in the switching fabric, thereby largely overcoming one of the most important concerns of optical interconnects at the chip level. In addition, in combination with an off-chip laser, we believe AWGR-based NoCs significantly reduce on-chip power consumption without performance degradation, leaving more of the power budget (constrained by thermal design point in HPC systems) to the compute and memories. The following section discusses the benefits of SiPh, AWGRs, and the topologies they enable in more detail.

III. Enabling Silicon-Photonic Technologies

A. Photonic Networks-on-chip

Figure 2 depicts a reference SiPh link with one sender and one receiver–referred to as Single-Writer-Single-Reader (SWSR) bus. An off-chip laser generates light at multiple wavelengths which is coupled into the chip and waveguide. Multiple wavelengths can either be generated by a single multi-wavelength comb laser or several single-wavelength lasers whose signals are combined into a single fiber–each with

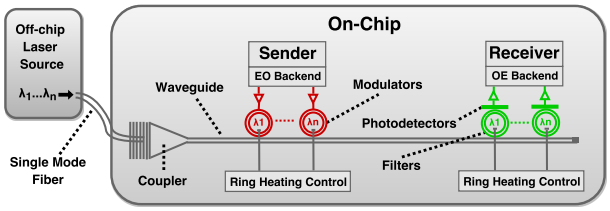


Fig. 2: An example SiPh Link [18]

different design trade-offs [2]. Modulators perform electrical-to-optical (EO) signal conversion by encoding bits onto wavelengths, which are filtered out by the receiver prior to being converted back into the electrical domain (OE) by a photodetector. Data can be transmitted on multiple wavelengths in parallel using DWDM with separate modulators and filters each tuned to one distinct wavelength. MRR heating and laser power are significant contributors to power consumption in SiPh. MRRs are used to build modulators and filters and are susceptible to temperature and process variations, thus requiring integrated/co-located heaters and control circuitry to ensure correct operation. Laser operating power depends on the number of wavelengths, optical path losses, receiver sensitivity, and laser efficiency.

There have been significant research efforts to enable on-chip lasers [24][25][26] which would reduce coupling losses. However, off-chip lasers offer higher power efficiency, easier thermal control, higher yield, and simpler heat dissipation. Also, off-chip lasers offer more practical maintenance and serviceability (e.g. testing, replacing, etc.). Therefore, our proposed architecture relies on an off-chip laser as a realistic option in the near future.

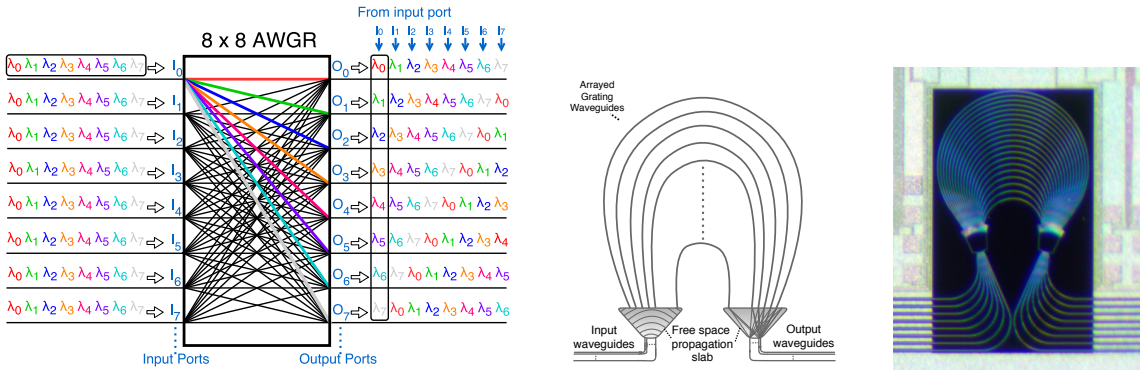
Constructing a NoC with point-to-point SWSR buses is possible but area- and layout-inefficient. Alternative approaches exploit DWDM to assign subsets of wavelengths on a single waveguide to different sources with one destination connected to the waveguide, referred to as Multiple-Writer-Single-Reader (MWSR) buses [10]. The dual to this approach is the Single-Writer-Multiple-Reader (SWMR) bus where one source can send to multiple destinations simultaneously on different wavelengths [10]. WRPNoCs reduces the number of waveguides by utilizing MRR filters to perform wavelength-selective routing [11]. These state-of-the-art approaches successfully connect all nodes in a NoC, but, as we will show in the following, do so less efficiently than AWGRs.

B. The State-Of-The-Art

Since the emergence of CMOS-compatible SiPh devices, a large number of photonic NoC architectures have been proposed (some even demonstrated [27]), typically aiming to find the most efficient way of integrating SiPh into NoCs, be it through a combination of electrical and optical interconnects in a NoC topology [28], [29], [22], WRPNoCs [11], [13], [30], [31], [32], circuit switching [33], or wavelength sharing mechanisms [34], [35], [36], [37].

SiPh have also been investigated for solving the *memory wall* problem by using optical processor-to-DRAM links,

[2]Note that a detailed description of the trade-offs of each approach is outside the scope of this paper, and we refer the interested reader to [24]



(a) Wavelength distribution inside an AWGR (b) Schematic (c) Microscope image

Fig. 3: Switching Functionality (a), Structure (b), and Optical microscope image (c) of an 8×8 AWGR [18]

which offer orders of magnitude higher bandwidth per pin than electrical interconnects [38], [39], [40], Other studies use SiPh to connect several chips on a PCB board to form a *virtual chip*, like Oracle's Macrochip [41] or Galaxy [23].

The physical SiPh interconnects studied in this paper form the basis of all of these proposals, most of which implementing a global crossbar, which we showed is more efficient with AWGRs. The paper is the first to evaluate AWGRs for on-chip communication in large-scale interposer-based systems and to explore AWGR-based topologies most suitable to NoCs.

C. Arrayed Waveguide Grating Router

As shown in Figure 3a, all wavelengths ($\lambda_0..\lambda_7$) entering a given input port are evenly distributed across all output ports of the AWGR–one wavelength to one unique output port. One intriguing property of AWGRs is that multiple signals on the same wavelengths entering from different input ports can traverse the AWGR without interfering with each other. Therefore, multiple input waveguides can be connected to an AWGR whose wavelengths are evenly distributed to the output ports. As a result, an AWGR provides all-to-all connectivity between all input and output ports.

Figure 3b shows the schematic of an AWGR device. Wavelengths entering from the input waveguides traverse the free-space propagation region and subsequently the grating waveguides, which have a constant length increment (ΔL). Each wavelength undergoes a constant change of phase attributed to the constant length increment in the grating waveguides. Wavelengths diffracted from each waveguide of the grating interfere constructively and get refocused at the output waveguides/ports depending on the experienced array phase shift. Figure 3c illustrates a picture of a fabricated SiN 8×8 AWGR on a SiPh interposer [16]. AWGRs are a mature technology and have already been used in the telecom industry [42], allowing for years of fabrication know-how with high-yield manufacturing. The novelties that make AWGRs suitable for on-chip communication are the advancements in CMOS-compatible SiN-based AWGRs, which not only exhibit very low loss and crosstalk, but also extremely reduced footprint [12][16]. For instance, AWGR in Figure 3c) has a footprint of $< 1mm^2$.

IV. AWGR-ENABLED NETWORKS

The unique wavelength routing of AWGRs opens up many opportunities and a new design space to be explored. As we will see in this section, the structure of the AWGR, its placement of input and output ports, and all-to-all connectivity pattern are ideal for global all-to-all implementations in NoCs. In particular, bipartite graphs and all-to-all networks can be efficiently implemented with AWGRs–both of which providing flat, low-diameter topologies capable of enabling low-latency communication not attainable with electrical interconnects at high energy efficiency and compact physical implementation. This section first discusses how AWGRs can enable bipartite graphs and all-to-all topologies, followed by a discussion on enabling multi-wavelength high-bandwidth communication with AWGRs as the switching fabric and a comparison of AWGR to alternative SiPh interconnection fabrics.

A. Bipartite Graphs with AWGRs

In principle, AWGRs are *bidirectional*, i.e., light can traverse an AWGR in both directions without interference (and with the same wavelength routing pattern), effectively forming a bidirectional all-to-all switching fabric with just a single device (this logical topology is shown in Figure 4 on the left).

Two design options to implement bipartite graphs exist: 1) utilizing two AWGRs unidirectionally or 2) utilizing a single AWGR with bidirectional operation. Both enable a compact, low-loss all-to-all fabric with short and direct links between each source-destination pair and without any waveguide crossings, but have their own set of benefits and trade-offs, which will be discussed in the following for the example 4×4 bipartite graphs in Figure 4.

Constructing a bipartite graph with two separate 4×4 AWGRs–one for each direction–is easily feasible in interposer-based systems, whose size ($>1000mm^2$ is a well-established size [2]) can conveniently accommodate several AWGRs (few mm^2); however, recent demonstrations of 3D-stacked AWGRs on separate SiPh layers show that AWGRs can be integrated vertically with negligible inter-layer crosstalk and loss [17] (more details in Section IV-C2), thereby taking up the horizontal real estate of just a single AWGR. Figure 4 illustrates how two AWGRs in opposite directions stacked atop each other provide a compact implementation of a bipartite graph.

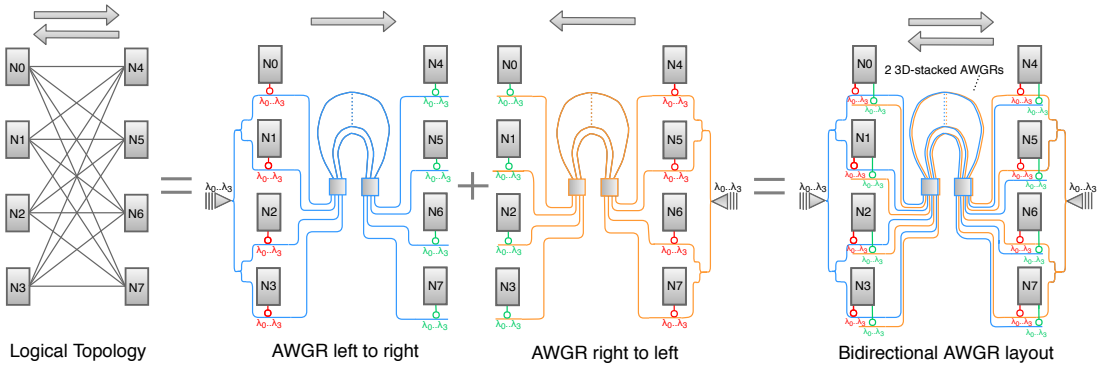


Fig. 4: Bipartite graph constructed out of two unidirectional AWGRs fabricated atop each other on separate SiPh layers. AWGR stacking enables a bipartite graph with short point-to-point links and without waveguide crossings. Note that the same can be obtained using a single $2N \times 2N$ AWGR used bidirectionally, though leading to higher crosstalk.

Utilizing a single AWGR and exploiting its symmetric, bidirectional wavelength routing operation requires an 8×8 AWGR to provide 4×4 bidirectional all-to-all connectivity (each node needs a separate input and output port to avoid filtering out its own signals). Therefore, the final layout would look exactly like the stacked AWGRs shown on the right in Figure 4, just that instead of two SiPh layers and 4×4 AWGRs, only one 8×8 AWGR on a single layer is used (in general, with such an implementation, an N×N bipartite graph needs a 2N×2N AWGR).

While both approaches enable a compact all-to-all switching fabric, each entails its own set of benefits and trade-offs, and several aspects of AWGRs should be considered when constructing all-to-all connectivty between the input/output ports. The loss inside AWGRs is mainly caused by the free-space propagation region and is relatively independent of the port count (e.g. the loss inside an 8×8 and an 16×16 AWGR is very similar [12]), meaning that a doubling of the port count to construct a bipartite graph using a single AWGR does not increase the loss inside an AWGR noticeably.

However, the footprint of an AWGR increases with the port count and utilizing two AWGRs with half the port count will result in a more compact implementation by adding an extra layer during the fabrication. In addition, a design with two separate AWGRs versus one AWGR will require smaller wavelength range (N×channel spacing of the AWGR compared to 2N×channel spacing of the AWGR).

In the following, we will stick to the 2-AWGRs case for implementing a bidirectional all-to-all fabric, as it results in reduced footprint and requires smaller wavelength range, thereby providing a more scalable fabric.

B. All-to-all Connectivity with AWGR

Although the bidirectional all-to-all fabric discussed in the previous section utilizes the AWGR in the most efficient manner in terms of footprint, layout, and loss, a true all-to-all fabric connecting all nodes directly with each other offers the ideal from a performance point-of-view (offers 1) a diameter of one which minimizes zero load latency and 2) maximum path diversity for load balancing) and could simplify the programming of many-core processors by enabling uniform memory/cache access. AWGRs provide an efficient implementation of such a fabric when connecting each input/output port to each sender/receiver, respectively.

The all-to-all utilization scenario of AWGRs, however, causes significantly higher crosstalk compared to the bidirectional use as more signals on the same wavelength are traversing the AWGR (for supporting the same number of nodes in the NoC as the bipartite graph). Moreover, as shown in Figure 3a depicting the wavelength distribution inside AWGRs, the number of wavelengths required to provide all-to-all connectivity inside the AWGR equals the number of ports (and, in turn, nodes in the NoC). A 64×64 AWGR would thus require 64 wavelengths for routing which enter the AWGR in each input port and impose crosstalk upon each other inside the AWGR. In fact, all-to-all connectivity with a single AWGR for port counts higher than 32 was shown to be challenging with SiN AWGRs (the material providing the lowest footprint and loss) due to excess crosstalk and require multiple AWGRs [43] to keep both crosstalk and laser power at feasible and practical levels. We, therefore, focus on the bipartite topology enabled by AWGRs in the remainder of this paper.

C. Achieving High Node-to-Node Bandwidth in AWGRs

The efficient wavelength-distribution mechanism of AWGRs coupled with its small area footprint and low losses make it an ideal candidate for global all-to-all connectivity, especially as input/output waveguides can be directly routed to the senders/receivers, thereby minimizing path lengths and laser power. SiPh typically leverages multi-wavelength DWDM signals to increase bit-parallelism and, in turn, link bandwidth, within a single waveguide [38]. The wavelength routing attributes of AWGRs introduced in the previous section shows that AWGRs can only distribute a single wavelength between each input-/output-port pair, preventing multi-wavelength communication between nodes, and thereby limiting total port-to-port bandwidth inside an AWGR to the modulation rate (i.e., data rate per wavelength). While this would have been a serious drawback of AWGRs for the state of SiPh a few years ago, three recent technological key advances now enable high port-to-port bandwidth inside a single AWGR:

1) PAM4 Modulation: The limitation of single-wavelength communication between nodes would have been a serious concern in previous NoC studies which mostly assume On/Off keying (OOK) modulation (1 bit per symbol), a modulation rate of 10Gb/s, and DWDM levels between 16-64 (this design point was shown to provide the highest energy efficiency [44]). To satisfy bandwidth demands in NoCs without DWDM, significantly higher modulation rates than 10Gb/s would be necessary. While higher modulation rates are not significantly detrimental to the energy efficiency of the photonic components [8], clock generation/recovery and driver and SERDES circuitry consume more energy at higher data rates.

One way of increasing the data rate is using advanced modulation techniques that increase the data rate by encoding multiple bits into one symbol. Although technologically feasible, the required transceiver circuitry for such modulation techniques was shown to consume too much energy ($\sim$3pJ/bit [45]). Fortunately, Moazeni et al. [14] recently demonstrated a new PAM4 transceiver (2 bits per symbol) on a 45nm platform which only requires a 'spoked' MRR (and driver circuitry) with just 5μm in radius and 0.197pJ/bit to convert two electrical input bits into a PAM4 signal at 20Gb/s modulation rate–effectively enabling a data rate of 40Gb/s per wavelength (four times higher than 10Gb/s OOK) at extremely high energy efficiency and compact layout.

Although this PAM4 transceiver is a big step towards efficient AWGR-based NoCs, single-wavelength communication at 40Gb/s bandwidth between source-destination pairs is still significantly lower than in current electrical NoCs (e.g., the typical assumption of 128-bit wide links at 2GHz provides 256Gb/s bandwidth). Luckily, the following advances boost the AWGR bandwidth even further.

2) Spatial-division Multiplexing with AWGRs: Spatial-division multiplexing (SDM) can be used to increase bandwidth by adding links (or in our case, AWGRs) to the NoC. Although the AWGR is the only device necessary to provide all-to-all connectivity, implementing multiple AWGRs aside each other is ultimately limited by the footprint of AWGRs ($1mm^2$) and losses incurred by more complex wiring and waveguide crossings (leading to higher laser power). Although the footprint concerns are less stringent in 2.5D integrated systems where the interposer can be used for interconnection only, the more complex physical layout can be detrimental to the overall power efficiency and would require additional fabrication efforts (e.g., for tapering waveguide crossings to reduce loss [46]).

Fortunately, both the footprint and the layout concerns are overcome by recent demonstrations of AWGRs implemented on separate SiPh layers [17]. This stacked AWGR approach not only removes any area/footprint concerns, but also eliminates waveguide crossings altogether and allows to physically place the AWGR between the nodes such that path lengths, and in turn losses, are minimized, leading to reduced output power requirements at the laser source. This enables the use of SDM of AWGRs to increase bandwidth without negative impacts on laser power or area.

3) Bit-parallelism in AWGRs: One interesting feature of AWGRs is that the wavelength routing is cyclic with the period, called the Free Spectral Range (FSR), which means that an output port j can be reached by an input port i using wavelength $\lambda_{ij+k\delta}$, with δ denoting the FSR, and k an integer. This cycling behavior enables each input port to communicate with each output port using multiple wavelengths (DWDM), referred to as *bit-parallel AWGRs*. Although limited by the crosstalk inside the AWGR and the wavelength range of the laser, this bit parallelism does not need to be very high to provide sufficient bandwidth when combined with modulation rates of up to 40Gb/s (and possibly SDM). Although it has been known to be theoretically possible, only until recently, Grani et al. actually successfully demonstrated the feasibility of AWGRs with bit-parallelism by leveraging the FSR [15].

With all these recent advancements, a high-bandwidth all-to-all network can be constructed using AWGR(s). For instance, the bisection bandwidth of an 8×8 AWGR with 32Gb/s modulation rate and a bit-parallelism of 2 is $8 \times 8 \times 32 \times 2 = 4096Gb/s(4Tb/s)$, which equals the bisection bandwidth of an 8×8 2D Mesh with 128-bit wide links at 2GHz, and could even be improved further by implementing two AWGR atop each other without any impact on area footprint or layout.

D. AWGRs vs. State-of-the-art SiPh Fabrics

Figure 5 compares the physical implementation of a global all-to-all interconnect constructed with AWGRs to bus-based designs (could be either SWSR, MWSR, or SWMR, the layout of each would be the same) for 64 (a), b)) and 256 (c), d)) cores in a realistic example target system which is like the disintegrated processor design placed on an interposer discussed in Section II. We assume 16 cores per die, 8 of which are clustered at one router. The red and green lines indicate that nodes need to place MRRs to modulate and filter signals adjacent to these waveguides to enable optical communication (as introduced in Figure 2).

1) AWGRs vs. SiPh Buses: The bus-based crossbars have a U-shaped layout, which has widely been used in recent literature [22][47][34] as it allows for a crossbar implementation with a straight-forward layout and without waveguide crossings. The U-shape of the waveguides leads to longer waveguides and, in turn, path losses; however, direct links between all sender-receiver pairs would lead to a very challenging layout and introduce a large number of waveguide crossings, making the U-shaped layout the most efficient. The AWGR-based crossbar allows for direct links without imposing waveguide crossings. These benefits become more important as the system scales to a larger number of nodes (Figure 5c) and d)): while the AWGR still provides short links and a compact layout, in bus-based designs waveguides must be routed in an S-shaped fashion to be in close proximity to the nodes (otherwise, modulators and receivers would have to be driven over mm distances), not only causing a more complicated layout, but also higher waveguide losses.

Aside from these benefits of AWGRs, there are a number of additional challenges of contemporary SiPh switching fabrics that can be overcome by AWGRs. The number of waveguides in crossbars consisting of SWSRs grows quadratically with the number of nodes, which is area-inefficient and complicates

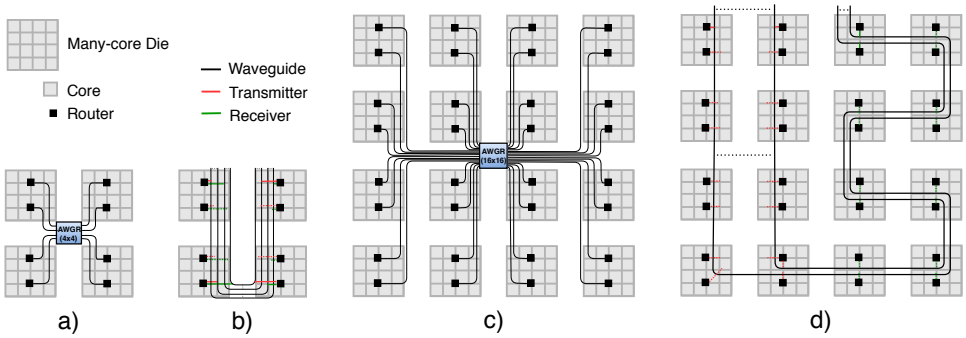


Fig. 5: Bidirectional all-to-all NoC layout with optical buses and AWGRs for 64 (a) and b)) and 256 (c) and d)) cores with a clustering of 8 cores at each router (note that, for illustration purposes, d) only shows one side (left) of the bipartite graph. The same waveguides are needed to connect the nodes on the right to those on the left).

layout. SWMRs or MWSRs overcome these issues by requiring only one waveguide per sender or receiver, respectively; however, assigning waveguides to senders/receivers complicates the physical implementation as more nodes are added to the NoC (e.g., in the SWMR case, each receiver must place MRRs at each of the senders waveguides to filter out signals). Waveguide pitches, MRR radii, and spacing between components are in the range of $\sim$5μm [34]. This results in designs in which MRRs are placed fairly far away (could be $> 100\mu$m) from the actual nodes, complicating placement of driver and heating circuitry, causing non-negligible energy consumption on the interconnect, and limiting scalability.

2) AWGRs vs. WRPNoCs: WRPNoCs (not shown in Figure 5) overcome the layout issue as each node only needs one waveguide for sending and receiving, respectively. MRR filters are strategically placed between waveguides to route wavelengths through the network to the correct destinations [11], [13], [48]. A sender merely has to modulate its data on the correct wavelengths to ensure that its data packet will arrive at the destination. WRPNoCs require fewer and shorter waveguides to create a crossbar than buses but rely on MRRs for routing which consumes heating power. Moreover, MRRs are typically distributed across the chip (depending on which layout provides the lowest losses), which complicates layout as heating circuitry must be co-located. Numerous studies dedicated just for investigating efficient WRPNoC layouts underline this issue (i.a. [11], [49]).

Besides, the number of MRRs in WRPNoCs has poor scalability as the number of nodes in a NoC increases although numerous studies with advanced topologies aimed to decrease the number of MRRs for switching (thousands of MRRs are needed for switching for NoC sizes > 32 nodes) [11], [13], [30]. This leads to significant on-chip power for thermo-optical control of MRRs, which makes them less practical than bus-based designs.

Using an AWGR alleviates all of the aforementioned issues. First, one input and output waveguide per node is required which allows placing all of the transceiver circuitry close to the nodes, thus simplifying layout. Second, wavelength routing does not rely on MRRs and AWGRs do not require on-chip heating (refractive index changes caused by temperature variations can either be controlled by off-chip TECs or can be avoided altogether with athermal AWGRs [50]), thus completely eliminating heating circuitry and power for routing. Also, as mentioned above, an AWGR-based crossbar does not exhibit any waveguide crossings, which lead to higher losses in WRPNoCs [11] and can only be avoided by U-shaped layouts in bus-based designs. Finally, AWGRs can be used bidirectionally or used unidirectionally stacked atop each other which allows constructing a bidirectional all-to-all fabric using just one/two passive component(s) consuming no power.

Given these benefits over state-of-the-art SiPh fabrics, AWGRs represent a promising candidate to enable low-power, low-latency, high-bandwidth, and scalable interconnection between processor dies in large-scale 2.5D integrated systems.

V. Methodology

The goal of our study is to investigate the benefits and drawbacks of AWGR-based NoC architectures and to reveal which interconnection fabric–both electrical and photonic–provides the best scalability for large-scale 2.5D integrated systems. We simulated a system based on the configuration listed in Table I, and assume a target architecture like the disintegrated processor in Figure 1 interconnected as shown in Figure 5. Each die has 16-cores, i.e., the 64- and 256-core configurations have 4 and 16 dies placed on the interposer. We assume that all interconnection fabrics are exclusively routed on the interposer. With processor dies of $\sim 74mm^2$ [2] and HBM dies of $\sim 42mm^2$ [51], with a $200\mu m$ spacing for die placement [5], the total interposer areas for 64- and 256-core configurations are $\sim 360mm^2$ and $\sim 1500mm^2$ respectively.

A. Experimental Setup

For our simulation study, we used Sniper [52] with high-performance applications from the SPLASH-2x and PAR-

TABLE I: Target System Configuration (layout as in Fig. 1)

Parameter	Description
Cores	64 and 256 cores, 16-core dies; x86 out-of-order; 2GHz
Caches	Private 32kB L1I/D and 256kB L2 per core; MSI coherence
Memory	8GB HBM2.0 per die; 1024-bit 1GHz interface
Dimensions	2mm tile width/length; 2mm spacing between dies
NoC	Routers : 128-bit at 2GHz; 5 flit deep buffers; 2 cycle traversal
	Electrical links: 128-bit at 2GHz; 1 cycle traversal
	Optical links: 64-bit at 2GHz; 1 cycle traversal
	6 virtual channels per port with virtual cut-through switching

SEC3.0 [53] benchmark suites, covering workloads of various different communication profiles. In addition, we used Garnet2.0 [54] inside gem5 [55] for performance simulation with synthetic traffic. Power and latency of the CMOS circuitry (i.e., electrical links, routers, and EO/OE backends) were modeled with DSENT [44] and a 22nm technology node. Laser power was modelled based on the formula by Li et al. [34] with 20% laser efficiency [56], -18dBm receiver sensitivity [57], 1dB coupler loss [58], 0.2dB splitter loss, 0.027dB/mm waveguide propagation loss, 0.01dB MRR through loss, 0.5dB MRR drop loss, and 0.12dB waveguide crossing loss [13][56]. We assume 1.4dB, 1.5dB, and 1.8dB loss for a 4×4, 8×8 and 16×16 SiN AWGR with -27dB, -24dB, and -20dB crosstalk, respectively [12]. Further, we assume $20\mu W/MRR$ for thermo-optical control of MRRs, and 11ps/mm signal propagation of light in silicon. Our proposal relies on off-chip static WDM lasers with 8 and 32 unique wavelengths in the 64- and 256-core cases, respectively.

B. NoCs Under Investigation

The vast majority of previously proposed NoCs make use of SiPh interconnects with optical buses, i.e., SWSR, SWMR, or MWSR (some prominent examples ATAC [28], Firefly [59], Meteor [29], Corona [37]. Therefore, we compare the bipartite graph use case ('AWGR') of AWGRs to implementations with SWSR buses. MWSR and SWMR buses assign subsets of wavelengths to each destination on one waveguide, which would require hundreds of different wavelengths for a bipartite graph supporting more than 64 cores, which would be an unrealistic design consideration. Therefore, some sort of SDM is necessary to obtain a feasible design, and since area constraints are not critical on the interposer, we opted for comparing AWGRs to SWSR buses[3]. We also added the aggressive electrical baselines 2D Mesh ('Mesh'), 2D Mesh with a clustering of 4 ('Mesh4C'), 2D Folded Torus ('FoldedTorus'), and 2D Folded Torus with a clustering of 4 ('FoldedTorus4C')–all of which utilizing XY routing–to our study to identify the benefits of SiPh in large-scale systems.

We omitted the all-to-all use case of AWGRs in our study as our analysis has shown that an optical all-to-all NoC imposes impractical laser power overheads for core counts larger than 64 and crosstalk that might render their implementation infeasible for the current state of SiPh technology. Our bidirectional AWGR NoC connects the cores of the target system as shown in Figure 5: 8 cores are clustered at each router, resulting in a 4×4 and 16×16 bipartite graph for 64 and 256 cores, respectively. The AWGRs implementing these graphs are assumed to be stacked atop each other, with one AWGR for each direction. To support a 64-bit wide link (at 2GHz), we utilize 32Gbps PAM4 signals, a bit-parallelism inside the AWGR of 2, and a SDM level of two AWGRs stacked on top each other (leading to a stacking of four AWGR in total for the entire NoC). For the SWSR implementation, we assumed four wavelengths at 32Gbps PAM4 on each waveguide.

[3]Note that a more extensive comparison between the different SiPh interconnection fabrics in terms of loss, power consumption, etc. is provided in our previous publication on this topic [18].

VI. Evaluation Results

A. Synthetic Traffic

1) Performance Results: Figure 6 shows the latency results of the NoCs under investigation for varying injection rates with uniform random, transpose, and tornado traffic to stress different corner cases of the topologies (sources compute destination nodes based on the synthetic traffic model by Dally et al. [60]). Each core in the system injects packets into the system with an increasing injection rate and packet sizes varying from 8 bytes to 72 bytes based on Garnet's pseudo cache coherence model [54]. The figures reporting latency do not show the bipartite graph implementation with SWSR as it has the same performance results as the AWGR.

Our AWGR-based topology reduces packet latency by more than $2\times$ prior to reaching network saturation compared to all alternative NoCs for both network sizes and all workloads. From a throughput point-of-view, only the folded torus topology can sustain noticeably higher throughput than the AWGR for 64 cores. For 256 cores, the AWGR-based topology dominates all other NoCs in terms of throughput, attributed to the high bisection bandwidth of the global crossbar and fewer number of hops which combined lead to less network congestion.

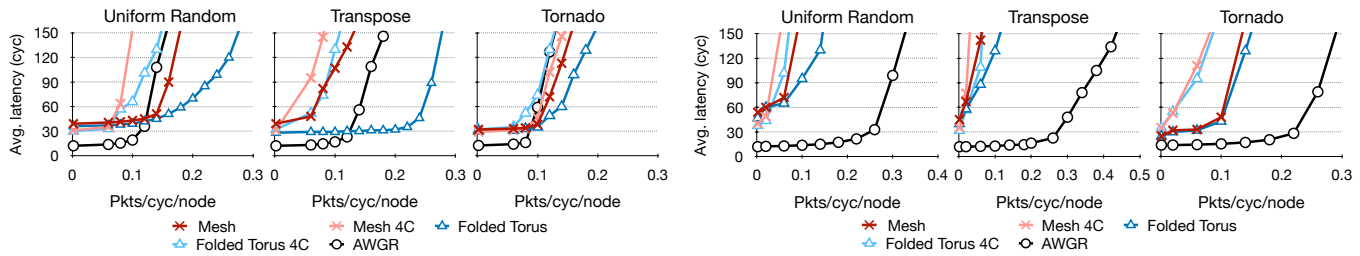
2) Power Results: Figure 7 plots the power consumption vs. injection rate, which allows to identify whether the high network loads can be sustained with satisfactory power consumption. The power results include the entire network power, i.e., leakage, dynamic, MRR heating, and off-chip laser power.

Not only does the AWGR-based topology offer much lower latency and sustains higher network loads, but also does so with less power consumption. Only the clustered versions of the electrical NoCs can compete with the AWGR, mainly due to the high leakage power overheads and high dynamic power imposed by larger number of hops in the non-clustered NoCs. Compared to the crossbar implementation with SWSR, AWGR-based topologies offer sightly less power consumption, which comes from the lower losses (and, in turn, lower laser power) in the AWGR fabric provided by shorter waveguides.

B. Application Traffic

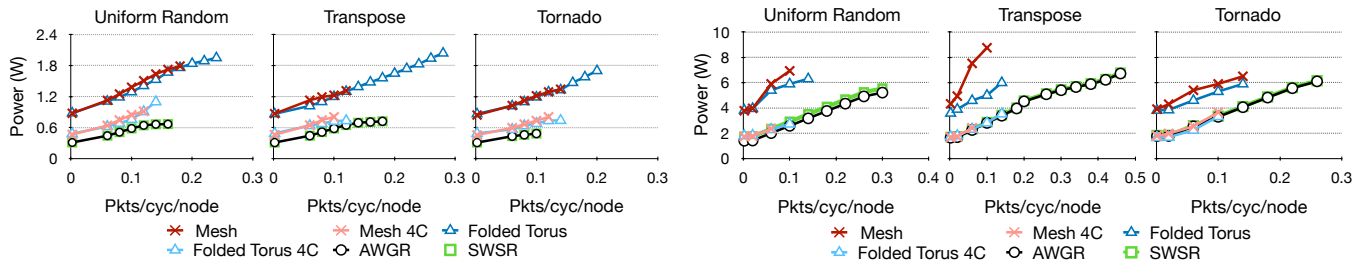
1) Performance Results: Figure 9 shows the application execution time normalized to our AWGR topology for 64 and 256 cores. For both cases, our AWGR-based topology reduces execution time of each of the simulated applications. The flat topology enabled by SiPh and the AWGR fabric offers a significantly reduced application execution time for both 64 and 256 cores. Generally, we observed that the higher the degree of data sharing in the application (and, in turn, on-chip traffic), the bigger the performance gains of the AWGR topologies, implying that applications exhibiting higher on-chip traffic profiles than those from the SPLASH2.x/PARSEC3.0 benchmark suites might benefit from AWGR-based interconnects even more.

2) Power Results: Figure 8 shows the power breakdown of all topologies for 64 and 256 cores, respectively. Breakdowns for each application are omitted for brevity, considering that we have not observed significant variations across different



(a) 64 cores (b) 256 cores

Fig. 6: Average packet latency (cyc) vs. injection rate (pkts/cyc/node) for synthetic workloads



(a) 64 cores (b) 256 cores

Fig. 7: Power consumption (W) vs. injection rate (packets/cycle/node) for synthetic workloads

workloads. The AWGR based topologies require the lowest power consumption out of all topologies for both cases, confirming the supreme scalability and energy efficiency of AWGR-based interconnects.

Leakage power is known to dominate the power budget for NoCs with buffers and virtual channels for technology nodes of 22nm and lower [61] (power gating techniques can almost halve leakage power [62], but cannot fully overcome these overheads). Deploying a high-bandwidth low-loss SiPh fabric like AWGRs allows to cluster more nodes at each router without performance drawbacks, allowing for much fewer routers in total and, in turn, less leakage power (despite the fact their routers have higher radix). Dynamic power plays an increasingly smaller role as the system size increases, which is likely due to the fairly low NoC utilization characteristics of the SPLASH-2x/PARSEC-3.0 workloads and their relatively small data sets (compared to the total size of the on-chip caches in our configuration). Multi-programmed workloads, highly virtualized systems, and applications with higher cache miss rates, data sharing, or data sets would probably benefit from the AWGR even more as it offers lower dynamic power

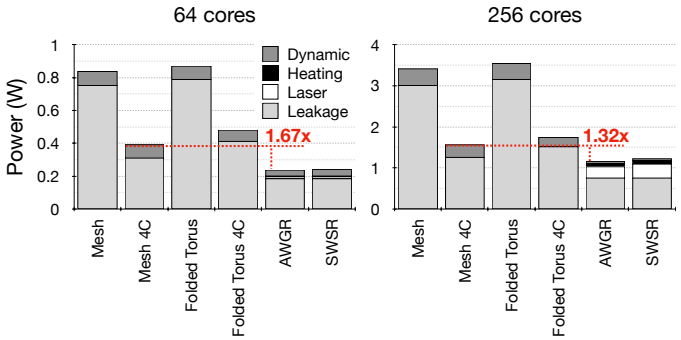


Fig. 8: Power breakdown for 64 and 256 cores

due to its low-diameter topology and distance-independent energy consumption.

The AWGR and SWSR have very similar power consumption for 64 nodes; however for 256 cores, the waveguide length of the bus based design and the number of waveguides needed (scaling quadratically with the number of nodes in SWSR crossbar) leads to significant waveguide propagation and splitter loss, and in turn to higher laser power requirements compared to the AWGR-based solution which offers short direct links between source-destination pairs.

Figure 10 plots the energy-delay-product (EDP) of the considered NoCs, workloads, and system sizes to put the performance speed-up into perspective with power consumption. In general, AWGR offer by far the most energy-efficient design. The EDP benefits compared to a SWSR bus are lower mostly because both networks provide the same performance and thus the same application execution, de-emphasizing the power reductions of the AWGR compared to the SWSR. Compared to the electrical baselines, however, AWGR improves power efficiency by at least 1.67× for both network sizes.

C. Discussion

Our results revealed that the low diameter of global bipartite graphs can have a large impact on packet latency, execution time and energy efficiency of applications in interposer-based systems. The low network diameter reduces network latency by more than 2× for low network loads, which makes them ideal for large-scale interposer-based systems executing latency-critical applications. This low latency also allows to make easier estimates on the quality of service, and makes large-scale systems easier to program as memory accesses are much less likely to have large latency differences (as it is the case in electrical NoCs).

AWGRs not only provide better performance and power metrics, but also represent a scalable and compact wavelength

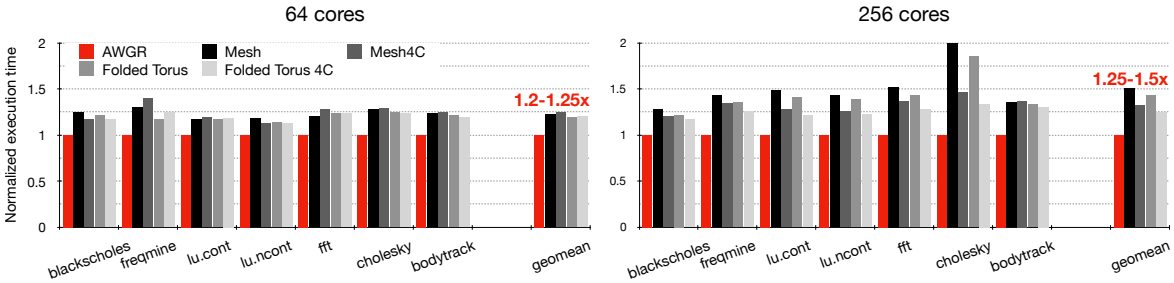


Fig. 9: Application execution time normalized to AWGR

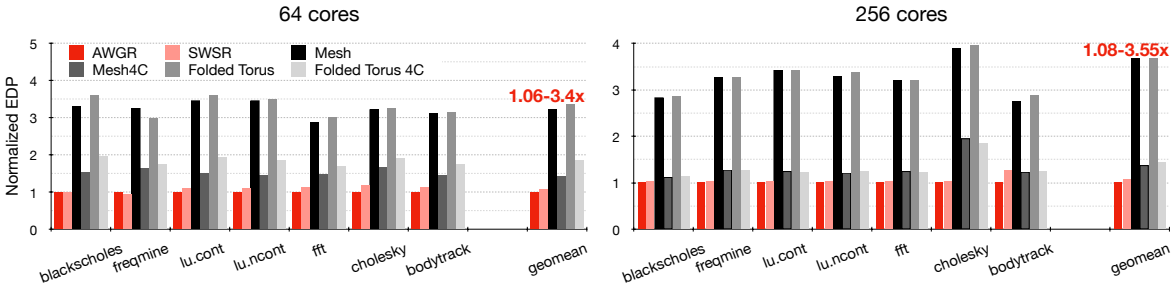


Fig. 10: Energy-delay-product normalized to AWGR

routing platform that allows for a simple, straight-forward physical layout. Rather than imposing large overheads in the number of waveguides or a complicated physical layout with MRR-based switching fabrics, the AWGR's unique wavelength routing mechanism might be a key enabler for practical future SiPh on-chip interconnects.

Our proposal requires 4×4 and 16×16 AWGRs for 64- and 256-core configurations respectively. In terms of scalability, a system with 1024 cores would require 64×64 AWGRs. Currently, there are, to the best of our knowledge, no demonstrated 64×64 SiN AWGRs to be found in literature. The loss inside AWGRs is relatively independent of the port count, and the main challenge for AWGRs with high port counts would be the crosstalk. However, there has been successful demonstrations of techniques to use multiple smaller AWGRs (in terms of port count) to provide the same functionality at lower crosstalk [43]. Also, AWGRs with much higher port counts have already been demonstrated in Si [63], albeit with considerably larger footprint ($176mm^2$ compared to $1mm^2$). This area might be negligible for a system with 512 dies (each $\sim 74mm^2$), but the interposer size/cost and crosstalk of the AWGR should be considered.

Moreover, the footprint overhead of our proposal is insignificant. Each processor die should accommodate the coupler ($2\mu m^2$ [58]), MRR ($25\mu m^2$ [14]), and backend circuitry for EO/OE conversion ($930\mu m2$ calculated using DSENT [44]) for each link. Thus, the total area occupied by optics for 64- and 256-core designs are $3828\mu m^2$ ($+0.005\%$) and $15312\mu m^2$ ($+0.0\%2$) respectively. With processor die size of $\sim 74mm^2$ [2] and $1mm^2$ for AWGRs, the aggregate overhead is 0.021% and 0.082% for 64- and 256-core configurations.

SiPh evolves quickly and new devices enable new opportunities for NoC architectures. For instance, compelling demonstrations of on-chip lasers enable low-latency/energy adaptive laser control which can save large amounts of laser power [24][25][26]. The AWGR-based topologies proposed in this paper could, in fact, be efficiently combined with adaptive lasers to further improve power efficiency. The all-to-all style topologies enabled by AWGRs offer high path diversity, which could be exploited to perform adaptive bandwidth scaling by shutting down or turning on lasers on different paths. Although many challenges regarding stabilization mechanisms and laser turn on/off times are still needed, this could represent a great opportunity for laser power savings.

All in all, our results confirm that SiPh in general are an excellent candidate for overcoming the interconnect bottleneck in large-scale interposer-based systems, which would enable more of the power budget to be dedicated to the processor and memory dies. Using AWGRs further supplements SiPh by offering a switching fabric that allows for direct links between source-destination pairs without imposing any waveguide crossings and their associated losses and additional fabrication steps. All these attributes make AWGR a key enabling technology for future computing systems leveraging tight integration in the same package to meet performance goals at high energy efficiency.

VII. Conclusion

This paper investigated the use of AWGRs inside NoCs for interposer-based disintegrated processors to address the power, performance, and scalability drawbacks of electrical NoCs in large-scale systems, studied AWGR-based NoC topologies, and compared them to state-of-the-art SiPh interconnects and aggressive electrical baselines. Our results show that AWGRs provide significant performance speed-up, power reductions, and better scalability compared to the state of the art while enabling a practical physical implementation of low-diameter interconnection networks. AWGRs could be a key enabler of future scaling of 2.5D integrated systems with low communication latency, which could be of high impact for current and the future of computing systems that leverage tight integration.

REFERENCES

[1] N. E. Jerger, A. Kannan, Z. Li, and G. H. Loh, "Noc architectures for silicon interposer systems: Why pay for more wires when you can get them (from your interposer) for free?" in *2014 47th Annual IEEE/ACM International Symposium on Microarchitecture*. IEEE, 2014, pp. 458–470.

[2] A. Kannan, N. E. Jerger, and G. H. Loh, "Enabling interposer-based disintegration of multi-core processors," in *2015 48th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2015, pp. 546–558.

[3] NVIDIA, "NVIDIA tesla V100 GPU architecture," http://images.nvidia.com/content/volta-architecture/pdf/volta-architecture-whitepaper.pdf, 2017, [Online; accessed 03-14-2018].

[4] J. Macri, "AMD's next generation GPU and high bandwidth memory architecture: FURY," in *2015 IEEE Hot Chips 27 Symposium (HCS)*, 2015, pp. 1–26.

[5] G. H. Loh, N. E. Jerger, A. Kannan, and Y. Eckert, "Interconnect-memory challenges for multi-chip, silicon interposer systems," in *Proceedings of the 2015 international symposium on Memory Systems*. ACM, 2015, pp. 3–10.

[6] "Zen - microarchitectures - AMD," https://en.wikichip.org/wiki/amd/microarchitectures/zen#Die-die_memory_latencies, [Online; accessed 11-19-2018].

[7] A. Arunkumar, E. Bolotin, B. Cho, U. Milic, E. Ebrahimi, O. Villa, A. Jaleel, C.-J. Wu, and D. Nellans, "MCM-GPU: Multi-chip-module GPUs for continued performance scalability," *ACM SIGARCH Computer Architecture News*, vol. 45, no. 2, pp. 320–332, 2017.

[8] C. J. Nitta, M. K. Farrens, and V. Akella, "On-chip photonic interconnects: A computer architect's perspective," *Synthesis Lectures on Computer Architecture*, vol. 8, no. 5, pp. 1–111, 2013.

[9] F. Eris, A. Joshi, A. B. Kahng, Y. Ma, S. Mojumder, and T. Zhang, "Leveraging thermally-aware chiplet organization in 2.5 D systems to reclaim dark silicon," in *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2018, pp. 1441–1446.

[10] K. Bergman, L. P. Carloni, A. Biberman, J. Chan, and G. Hendry, *Photonic network-on-chip design*. Springer, 2014.

[11] L. Ramini, P. Grani, S. Bartolini, and D. Bertozzi, "Contrasting wavelength-routed optical noc topologies for power-efficient 3d-stacked multicore processors using physical-layer analysis," in *2013 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2013, pp. 1589–1594.

[12] K. Shang, S. Pathak, C. Qin, and S. B. Yoo, "Low-loss compact silicon nitride arrayed waveguide gratings for photonic integrated circuits," *IEEE Photonics Journal*, vol. 9, no. 5, pp. 1–5, 2017.

[13] P. K. Hamedani, N. E. Jerger, and S. Hessabi, "Qut: A low-power optical network-on-chip," in *2014 Eighth IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*. IEEE, 2014, pp. 80–87.

[14] S. Moazeni, S. Lin, M. Wade, L. Alloatti, R. J. Ram, M. Popović, and V. Stojanović, "A 40-Gb/s PAM-4 transmitter based on a ring-resonator optical DAC in 45-nm SOI CMOS," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 12, pp. 3503–3516, 2017.

[15] P. Grani, G. Liu, R. Proietti, and S. B. Yoo, "Bit-parallel all-to-all and flexible awgr-based optical interconnects," in *Optical Fiber Communication Conference*. Optical Society of America, 2017, pp. M3K–4.

[16] X. Xiao, Y. Zhang, R. Proietti, and S. Yoo, "Scalable awgr-based all-to-all optical interconnects with 2.5 D/3D integrated optical interposers," in *2018 IEEE Photonics Society Summer Topical Meeting Series (SUM)*. IEEE, 2018, pp. 161–162.

[17] T. Su, G. Liu, K. E. Badham, S. T. Thurman, R. L. Kendrick, A. Duncan, D. Wuchenich, C. Ogden, G. Chriqui, S. Feng, J. Chun, W. Lai, and S. J. B. Yoo, "Interferometric imaging using si3n4 photonic integrated circuits for a spider imager," *Opt. Express*, vol. 26, no. 10, pp. 12 801–12 812, May 2018.

[18] S. Werner, P. Fotouhi, R. Proietti, X. Xiao, and S. B. Yoo, "Towards energy-efficient high-throughput photonic NoCs for 2.5D integrated systems: A case for AWGRs," in *2018 Twelfth IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*. IEEE, 2018, pp. 1–8.

[19] T. Vijayaraghavan, Y. Eckert, G. H. Loh, M. J. Schulte, M. Ignatowski, B. M. Beckmann, W. C. Brantley, J. L. Greathouse, W. Huang, A. Karunanithi, O. Kayiran, M. Meswani, I. Paul, M. Poremba, S. Raasch, S. K. Reinhardt, G. Sadowski, and V. Sridharan, "Design and analysis of an APU for exascale computing," in *2017 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2017, pp. 85–96.

[20] K. Tran and J. Ahn, "HBM: Memory solution for high performance processors," *Proceedings of MemCon*, pp. 1–1, 2014.

[21] Y. Thonnart and M. Zid, "Technology assessment of silicon interposers for manycore SoCs: Active, passive, or optical?" in *2014 Eighth IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*. IEEE, 2014, pp. 168–169.

[22] S. Werner, J. Navaridas, and M. Luján, "Designing low-power, low-latency networks-on-chip by optimally combining electrical and optical links," in *2017 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2017, pp. 265–276.

[23] Y. Demir, Y. Pan, S. Song, N. Hardavellas, J. Kim, and G. Memik, "Galaxy: A high-performance energy-efficient multi-chip architecture using photonic interconnects," in *Proceedings of the 28th ACM international conference on Supercomputing*. ACM, 2014, pp. 303–312.

[24] Y. Demir and N. Hardavellas, "SLaC: Stage laser control for a flattened butterfly network," in *2016 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2016, pp. 321–332.

[25] G. Kurczveil, D. Liang, M. Fiorentino, and R. G. Beausoleil, "Robust hybrid quantum dot laser for integrated silicon photonics," *Optics express*, vol. 24, no. 14, pp. 16 167–16 174, 2016.

[26] G. Kurczveil, C. Zhang, A. Descos, D. Liang, M. Fiorentino, and R. Beausoleil, "On-chip hybrid silicon quantum dot comb laser with 14 error-free channels," in *2018 IEEE International Semiconductor Laser Conference (ISLC)*. IEEE, 2018, pp. 1–2.

[27] C. Zhang, S. Zhang, J. D. Peters, and J. E. Bowers, "8× 8× 40 gbps fully integrated silicon photonic network on chip," *Optica*, vol. 3, no. 7, pp. 785–786, 2016.

[28] G. Kurian, J. E. Miller, J. Psota, J. Eastep, J. Liu, J. Michel, L. C. Kimerling, and A. Agarwal, "ATAC: a 1000-core cache-coherent processor with on-chip optical network," in *Proceedings of the 19th international conference on Parallel architectures and compilation techniques*. ACM, 2010, pp. 477–488.

[29] S. Bahirat and S. Pasricha, "Meteor: Hybrid photonic ring-mesh network-on-chip for multicore architectures," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 13, no. 3s, p. 116, 2014.

[30] S. Werner, J. Navaridas, and M. Luján, "Amon: An advanced mesh-like optical noc," in *2015 IEEE 23rd Annual Symposium on High-Performance Interconnects*. IEEE, 2015, pp. 52–59.

[31] M. Bahadori, S. Rumley, H. Jayatilleka, K. Murray, N. A. Jaeger, L. Chrostowski, S. Shekhar, and K. Bergman, "Crosstalk penalty in microring-based silicon photonic interconnect systems," *Journal of Lightwave Technology*, vol. 34, no. 17, pp. 4043–4052, 2016.

[32] D. Nikolova, D. M. Calhoun, Y. Liu, S. Rumley, A. Novack, T. Baehr-Jones, M. Hochberg, and K. Bergman, "Modular architecture for fully non-blocking silicon photonic switch fabric," *Microsystems & Nanoengineering*, vol. 3, p. 16071, 2017.

[33] A. Shacham, K. Bergman, and L. P. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," *IEEE Transactions on Computers*, vol. 57, no. 9, pp. 1246–1260, 2008.

[34] C. Li, M. Browning, P. V. 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.

[35] Y. Pan, J. Kim, and G. Memik, "Featherweight: low-cost optical arbitration with QoS support," in *2011 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2011, pp. 105–116.

[36] A. Zulfiqar, P. Koka, H. Schwetman, M. Lipasti, X. Zheng, and A. Krishnamoorthy, "Wavelength stealing: an opportunistic approach to channel sharing in multi-chip photonic interconnects," in *2013 46th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2013, pp. 222–233.

[37] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. G. Beausoleil, and J. H. Ahn, "Corona: System implications of emerging nanophotonic technology," in *ACM SIGARCH Computer Architecture News*, vol. 36, no. 3. IEEE Computer Society, 2008, pp. 153–164.

[38] P. Grani, R. Proietti, V. Akella, and S. B. Yoo, "Design and evaluation of AWGR-based photonic NoC architectures for 2.5 D integrated high performance computing systems," in *2017 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2017, pp. 289–300.

[39] Z. Wang, Z. Pang, P. Yang, J. Xu, X. Chen, R. K. Maeda, Z. Wang, L. H. Duong, H. Li, and Z. Wang, "MOCA: An inter/intra-chip optical network for memory," in *Proceedings of the 54th Annual Design Automation Conference 2017*. ACM, 2017, p. 86.

[40] C. Batten, A. Joshi, J. Orcutt, A. Khilo, B. Moss, C. W. Holzwarth, M. A. Popovic, H. Li, H. I. Smith, J. L. Hoyt *et al.*, "Building many-core processor-to-DRAM networks with monolithic CMOS silicon photonics," *IEEE Micro*, vol. 29, no. 4, pp. 8–21, 2009.

[41] P. Koka, M. O. McCracken, H. Schwetman, X. Zheng, R. Ho, and A. V. Krishnamoorthy, "Silicon-photonic network architectures for scalable, power-efficient multi-chip systems," in *ACM SIGARCH Computer Architecture News*, vol. 38, no. 3. ACM, 2010, pp. 117–128.

[42] S. Kamei, M. Ishii, M. Itoh, T. Shibata, Y. Inoue, and T. Kitagawa, "64× 64-channel uniform-loss and cyclic-frequency arrayed-waveguide grating router module," *Electronics Letters*, vol. 39, no. 1, pp. 83–84, 2003.

[43] R. Proietti, X. Xiao, K. Zhang, G. Liu, H. Lu, P. Fotouhi, J. Messig, and S. Yoo, "Experimental demonstration of a 64-port wavelength routing thin-CLOS system for data center switching architectures," *Journal of Optical Communications and Networking*, vol. 10, no. 7, pp. B49–B57, 2018.

[44] C. Sun, C.-H. O. Chen, G. Kurian, L. Wei, J. Miller, A. Agarwal, L.-S. Peh, and V. Stojanovic, "DSENT-a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *2012 IEEE/ACM Sixth International Symposium on Networks-on-Chip*. IEEE, 2012, pp. 201–210.

[45] I. G. Thakkar, S. V. R. Chittamuru, and S. Pasricha, "Improving the reliability and energy-efficiency of high-bandwidth photonic NoC architectures with multilevel signaling," in *2017 Eleventh IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*, 2017, pp. 1–8.

[46] Y. Zhang, A. Hosseini, X. Xu, D. Kwong, and R. T. Chen, "Ultralow-loss silicon waveguide crossing using bloch modes in index-engineered cascaded multimode-interference couplers," *Optics letters*, vol. 38, no. 18, pp. 3608–3611, 2013.

[47] A. Joshi, C. Batten, Y.-J. Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic, "Silicon-photonic clos networks for global on-chip communication," in *2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*. IEEE, 2009, pp. 124–133.

[48] I. O'Connor, M. Briere, E. Drouard, A. Kazmierczak, F. Tissafi-Drissi, D. Navarro, F. Mieyeville, J. Dambre, D. Stroobandt, J.-M. Fedeli *et al.*, "Towards reconfigurable optical networks on chip." vol. 5, 2005, pp. 121–128.

[49] M. Ortín-Obón, M. Tala, L. Ramini, V. Viñals-Yufera, and D. Bertozzi, "Contrasting laser power requirements of wavelength-routed optical noc topologies subject to the floorplanning, placement, and routing constraints of a 3-d-stacked system," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 7, pp. 2081–2094, July 2017.

[50] T. Huang, X. Liu, H. Zhang, S. Gong, and A. Zhang, "Athermal arrayed waveguide grating wavelength division multiplexer," 2016, uS Patent 9,519,103.

[51] L. Li, P. Chia, P. Ton, M. Nagar, S. Patil, J. Xue, J. Delacruz, M. Voicu, J. Hellings, B. Isaacson *et al.*, "3D SiP with organic interposer for ASIC and memory integration," in *2016 IEEE 66th Electronic Components and Technology Conference (ECTC)*. IEEE, 2016, pp. 1445–1450.

[52] W. Heirman, T. Carlson, and L. Eeckhout, "Sniper: Scalable and accurate parallel multi-core simulation," in *8th International Summer School on Advanced Computer Architecture and Compilation for High-Performance and Embedded Systems (ACACES-2012)*, 2012, pp. 91–94.

[53] X. Zhan, Y. Bao, C. Bienia, and K. Li, "PARSEC3.0: A multicore benchmark suite with network stacks and SPLASH-2X," *ACM SIGARCH Computer Architecture News*, vol. 44, no. 5, pp. 1–16, 2017.

[54] N. Agarwal, T. Krishna, L.-S. Peh, and N. K. Jha, "GARNET: A detailed on-chip network model inside a full-system simulator," in *2009 IEEE international symposium on performance analysis of systems and software*. IEEE, 2009, pp. 33–42.

[55] N. Binkert, B. Beckmann, G. Black, S. K. Reinhardt, A. Saidi, A. Basu, J. Hestness, D. R. Hower, T. Krishna, S. Sardashti *et al.*, *ACM SIGARCH Computer Architecture News*, vol. 39, no. 2, pp. 1–7, 2011.

[56] M. Bahadori, S. Rumley, D. Nikolova, and K. Bergman, "Comprehensive design space exploration of silicon photonic interconnects," *Journal of Lightwave Technology*, vol. 34, no. 12, pp. 2975–2987, 2016.

[57] M. Nada, S. Kanazawa, H. Yamazaki, Y. Nakanishi, W. Kobayashi, Y. Doi, T. Ohyama, T. Ohno, K. Takahata, T. Hashimoto *et al.*, "High-linearity avalanche photodiode for 40-km transmission with 28-gbaud PAM4," in *Optical Fiber Communication Conference*. Optical Society of America, 2015, pp. M3C–2.

[58] Y. Zhang, K. Shang, Y.-C. Ling, and S. B. Yoo, in *2018 Conference on Lasers and Electro-Optics (CLEO)*. IEEE, 2018, pp. 1–2.

[59] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, and A. Choudhary, "Firefly: Illuminating future network-on-chip with nanophotonics," in *ACM SIGARCH Computer Architecture News*, vol. 37, no. 3. ACM, 2009, pp. 429–440.

[60] W. J. Dally and B. P. Towles, *Principles and practices of interconnection networks*. Elsevier, 2004.

[61] R. Parikh, R. Das, and V. Bertacco, "Power-aware nocs through routing and topology reconfiguration," in *Proceedings of the 51st Annual Design Automation Conference*. ACM, 2014, pp. 1–6.

[62] R. Das, S. Narayanasamy, S. Satpathy, and R. G. Dreslinski, "Catnap: energy proportional multiple network-on-chip." in *ISCA*, 2013, pp. 320–331.

[63] S. Cheung, T. Su, K. Okamoto, and S. Yoo, "Ultra-compact silicon photonic 512× 512 25 ghz arrayed waveguide grating router," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 4, pp. 310–316, 2014.