Energy efficient on-chip optical broadcast with partial-absorption photodiodes

Raphael Cardoso, Lubna Arif, Clement Zrounba, Fabio Pavanello, Ian O'Connor, Laurent Vivien, Leopold Virot, Sébastien Le Beux

Energy efficient on-chip optical broadcast with partial-absorption photodiodes

Raphael Cardoso*, Lubna Arif*, Clément Zrounba*, Fabio Pavanello*, Ian O'Connor*,
Laurent Vivien†, Léopold Virot‡, Sébastien Le Beux§
*Univ. Lyon - CNRS, Ecole Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon 1,
CPE Lyon - INL, UMR5270 - Écully, F-69134, France
raphael.cardoso@ec-lyon.fr
†Centre de Nanosciences et de Nanotechnologies, Université Paris-Saclay - CNRS - Palaiseau, France
‡University Grenoble Alpes, CEA LETI - Grenoble, France
§Department of Electrical and Computer Engineering, Concordia University - Montreal, Canada

***Abstract*—As we enter in the 1000-core era, broadcast-enabled interconnects are needed to synchronize processors and maintain cache coherency. While silicon photonics is a promising technology to deliver high bandwidth communications, there are currently no photonic links enabling energy efficient broadcast. Indeed, existing solutions rely on thermally sensitive microring resonators that require continuous calibration. We propose an on-chip optical link dedicated to broadcast that relies on partial-absorption photodiodes, which allow to evenly distribute the broadcasted signal power to all receivers. Since a single wavelength is used, no ring resonators are required, thus allowing a 39% reduction in the broadcast energy consumption.**

***Index Terms*—Optical networks-on-chip, partial-absorption photodiodes, broadcast, many-core systems**

I. Introduction

The diminishing returns of transistor scalability observed over the past decade shows that disruptive technological changes must be pursued to meet ever-increasing computing performance requirements. At the architectural level, many-core systems are proven to be an effective solution to achieve High-Performance Computing (HPC) [1]. In such architectures, memories are shared by multiple processing units, which calls for cache-coherency protocols. Some of these protocols require constant information broadcast from one to all other cores [2]. However, as we are approaching a thousand cores on a chip, new interconnects are needed to maintain low latency and high bandwidth broadcast. Indeed, metallic interconnects are known to be inefficient at high speeds and distances [3], leading to the proposal of the Optical Networks-on-Chip (ONoC) [4], which have the potential of achieving Tbps communication rates with low losses. The design of ONoCs is a very active research area involving design space exploration, communication protocol definition, mapping algorithms and hardware-software co-design [5].

ONoCs enabling broadcast communication usually include micro-ring resonator (MRR) filters, which allows simultaneous transmission of several signals on a same waveguide by using Wavelength-Division-Multiplexing (WDM). However, a key drawback of MRRs is the calibration power overhead required to compensate fabrication non-uniformity and thermal variation [6]. State-of-the-art calibration requires $150\ \mu$W per MRR [7] and a minimum of $2N$ MRRs would be required for N-core architectures. Energy efficient on-chip optical communications also require to minimize the laser output power. This calls for fast high responsivity photodetectors, which can be implemented using segmented photodiodes [8] such that each part converts a portion of the optical power to an electrical current that is later accumulated. These photodetectors have so far been used as end points in a link, which does not prevent from using WDM and MRRs.

In this paper, we propose a disruptive on-chip optical link enabling efficient broadcast of data. The link is designed by placing partial-absorption photodiodes (PADs) [8] along a waveguide transporting the broadcasted data, and the electrical current of each segment is directed to a different receiver. The lengths of the PADs are defined to extract an increasing fraction of the signal power, thus allowing to transmit the same power level to all receivers. Since only a single wavelength is used, no routing MRRs are required, significantly increasing the broadcast energy efficiency. Results show that up to 39% energy is saved for a four-receivers link with an off-chip laser, which could be further improved by an on-chip laser and device optimization.

The paper is organized as follows. In Section II, we first present on-chip optical broadcast links relying on WDM and MRRs. Then, we present the proposed solution, along with an energy model and a method to define the length of the PADs. We present our results in Section III and Section IV concludes this work.

II. On-Chip Optical Broadcast Links

In this section, we first present a common optical broadcast link, then the proposed link and, finally, define a model to estimate the main performance metrics and determine the length of each PAD.

A. WDM-based link

Typical on-chip optical broadcast links, as shown in Fig. 1b, are composed of four main devices: off-chip continuous-wave (CW) lasers, electro-optic modulators based on MRRs, MRR filters and photodetectors. In such links, multiple optical

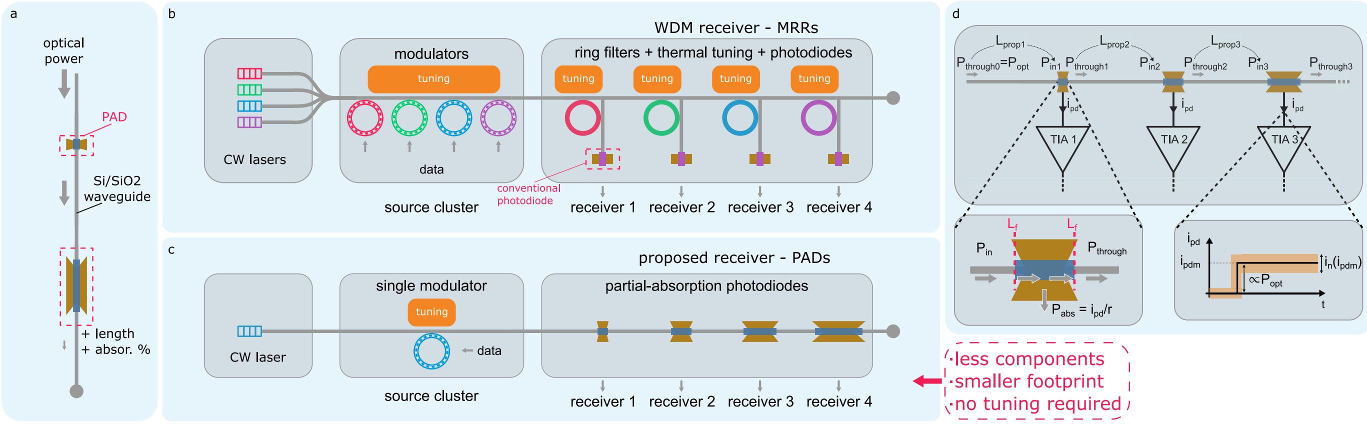

Fig. 1. (a) Partial-absorption photodiodes, (b) baseline WDM on-chip optical broadcast link for a four receiver example; the design of an M-receiver link requires M modulators and M MRR filters, (c) proposed PAD based link, which require a single modulator and does not involve any MRRs in the receiver side. (d) Receiver details for the proposed link, including necessary parameters related to the energy estimation and system-level design of PADs.

signals are modulated by the broadcasted data, and each signal is transmitted towards a broadband photodiode using WDM [9]. As illustrated in Fig. 1b for a four-receivers architecture, an M-receiver link involves M modulators, M MRR filters, and M wavelengths. The main drawbacks of such a link are i) the number of required devices that increases linearly with the number of receivers and ii) the use of WDM, which requires to calibrate the MRR filters to compensate for fabrication process non-uniformity and thermal sensitivity [6], and which can reach up to 0.12 nm/mW tuning efficiencies.

B. PAD-based Link

In order to overcome scalability and power limitations of existing broadcast links, we consider partial-absorption photodiodes (PADs), demonstrated to operate between 7 and 35 GHz with low energy dissipation [10]. Each PAD is a p-i-n photodetector with lateral silicon-germanium-silicon heterojunctions, which enable length-dependent extraction of the optical power from a waveguide, as illustrated by Fig. 1a. These key features enable to envision a link where the signal distribution is determined by photodiodes of different lengths.

The key characteristics of such a link, which is illustrated in Fig. 1c for a four-receiver example, are the following: only a single modulator is required and no MRR filters are needed. Compared to the WDM-based link, this leads to significant advantages in terms of scalability, footprint, and static power since no calibration is required. A drawback of the PADs is the loss experienced by the optical signal when propagating through each facet of the device, as illustrated by the inset in Fig. 1d. Because of these additional losses, more power is required from the laser.

In a lossless context with a 4-receivers link, the receivers 1, 2, 3, and 4 would respectively absorb 25%, 33%, 50%, and 100% of the power at their input in order to obtain the same BER in all destinations. However, in a realistic scenario, the absorption ratio of each PAD depends not only on system-level parameters such as the target bit-error rate and number of receivers, but also on technological parameters such as the PAD capacitances and dark currents, which themselves are length-dependent.

As the PAD absorption is an important target for device-level design of its length, we evaluate it with an iterative process that considers worst-case parameters, referent to longest photodiode, associated with the last receiver in the link. The determination of the PAD absorption is closely related to the estimation of power consumption in the link, and both are presented in the following section.

C. Energy Model and Link Design Method

In the following, we present an analytical model to estimate the Energy per Broadcasted Bit (EBB). It is expressed as

$$\text{EBB} = \frac{1}{B}(P_{drive} + \frac{FSR}{2\lambda_{tot}E_{tuning}} + P_{TIA} + P_{laser}), \quad (1)$$

where B is the data rate in bits/s, P_{drive} is the modulation power, FSR is the ring Free Spectral Range, λ_{tot} is the number of wavelengths ($\lambda_{tot} = 1$ for the proposed model), and E_{tuning} is the ring modulator tuning efficiency (in nm/mW) [6]. The model also takes into account the power consumed by transconductance amplifier (P_{TIA}) and the laser (P_{laser}). The former is defined as

$$P_{TIA} = V_{DD}\left(I_B + \frac{i_{pd}}{2}\right), \quad (2)$$

such that I_B is the bias current of the integrated amplifier, designed at the transistor level using the method proposed in [4] by considering the worst-case PAD capacitance.

The laser power consumption P_{laser} is calculated given a target BER for all destinations, for which we assume the same photocurrent i_{pd} in all receivers. It is defined by considering a worst-case dark current, which is related to the last photodiode in the link, resulting in

$$i_{pd} = \text{erfcinv}^2(2\text{BER}) \times \langle {i_n}^2 \rangle^{\frac{1}{2}}, \quad (3)$$

$$\langle {i_n}^2 \rangle^{\frac{1}{2}} = \sqrt{\langle {i_{nTIA}}^2 \rangle + \langle {i_{jn}}^2 \rangle + \langle {i_{sn}}^2 \rangle}, \quad (4)$$

where $\langle {i_n}^2 \rangle^{\frac{1}{2}}$ is the RMS noise referred in the photodiode current. It is obtained from the TIA noise, photodiode junction noise and shot noise (i_{nTIA} and i_{jn} are constant, see Table I). The shot noise depends on i_{dark} (dark current) and i_{pd}, which thus require to solve the system of equations (3) and (4). The optical power to be absorbed by each photodiode is then deduced from i_{pd} and the photodiode responsivity ($P_{abs} = i_{pd}/r$).

The value of P_{abs} is necessary to define the length of the photodiodes, as well as the power that remains in the waveguide after each receiver ($P_{through}$). Since $P_{through}$ depends on numerous parameters such as the number of the photodiodes, the waveguide length and the waveguide propagation loss, we use Algorithm 1 as follows. First, we estimate the optical power to be received by the last photodetector (i.e. photodetector M) using (1), (3), (4). Since it is the last receiver, we assume that all remaining power in the waveguide is absorbed. Then, by considering the propagation losses from the previous photodetector, we obtain the optical power to remain in the waveguide at diode $M-1$ ($P_{through_{M-1}}$). We then estimate $P_{in_{M-1}}$, the optical power to be received by photodetector $M-1$, by considering the input and output facet losses L_f. In general, the input power of any diode is defined by

$$P_{in_m} = 10^{-2L_f/10} \times P_{through_m} + 10^{-L_f/10} \times P_{abs_m}, \quad (5)$$

where, for the last receiver, $P_{through} = 0$. As seen in the equation, the optical signals experience facet losses L_f when entering and leaving the detector. The lengths of the photodiodes are directly proportional to the ratio between the absorbed power and the input power attenuated by the facet loss ($abs_m = P_{abs}/(P_{in_m} \times 10^{L_f/10})$) . After iterating on all receivers in the link, we eventually obtain the required laser output power, from which P_{laser} is obtained given the lasing efficiency ($P_{laser} = P_{opt}/E_{laser}$).

Algorithm 1 Laser optical power and PADs length estimation.

Input: $i_{pd}, r, L_f, L_{prop}, M$
Output: P_{opt}, PADs lenghts
compute P_{abs} for all receivers
$l_M \leftarrow 15\mu m$ *// required PAD length to absorb all power*
compute P_{in_M} with Eq. (5) *// power entering last PAD*
for $m =$ M-1 **to** 1 **do**
    compute $P_{through_m}$ from waveguide losses and $P_{in_{m+1}}$
    compute PAD length l_m from P_{in_m}, P_{abs} and L_f
    compute P_{in_m} with Eq. (5)
end for
compute P_{opt} from losses and P_{in_1} *// injected optical power*
return P_{opt}, PADs lengths

III. Results

In this section, we evaluate the energy efficiency of the proposed broadcast link using the model and method defined in the previous section. Comparisons with a WDM-based link are also carried out with parameters from Table I. In the first study, we compare the links in energy efficiency according to the number of receivers. In the second study, we investigate on the impact of the waveguide length in the same metric.

TABLE I
List of the parameters

	Parameter	Symbol	Value	Unit
Transmitter	Efficiency	E_{laser}	0.2	W/W
	Modulator insertion losses [7]	-	-2	dB
	Modulator driving power [11]	P_M	3	mW
PADs	Responsivity **	r	0.7	A/W
	Max. length **	l_M	15	μm
	Max. capacitance **	-	15	fF
	Max. dark current **	i_{dark}	15	nA
	Facet loss **	L_f	-1	dB
Conv. PDs	Responsivity *	r	0.7	A/W
	Capacitance *	-	2.2	fF
	Dark current *	i_{dark}	2.2	nA
Waveguide	Propagation loss per cm	-	-2	dB/cm
	Total length	-	20	mm
MRR	Through loss [12] *	-	-0.1	dB
	Drop loss [12] *	-	-0.6	dB
	Tuning efficiency [6]	E_{tuning}	0.12	nm/mW
	Full-scale range [6]	FSR	10.8	nm
System	Bit-error rate	BER	10^{-12}	1/bit
	Datarate	B	10	bits/s
	TIA noise current	$\langle {i_{nTIA}}^2 \rangle$	0.56	μA^2

* Unique for WDM links; ** Unique for PAD links.

A. EBB wrt. Number of Receivers

In this study, we evaluate the energy required to broadcast a bit (EBB) for 2, 4, 6 and 8 receivers in a link. We assume a 20 mm total waveguide length for all scenarios, which implies different distances between the uniformly spaced readers. To reach 10^{-12} BER at each receiver, an optical power of -16.6 dBm must be extracted by the photodiodes. By following the algorithm defined in the previous section, we obtain, for each scenario, the total energy consumption and the absorption ratios of all photodiodes. The energy breakdown of the results is reported in Fig. 2. It is worth mentioning that the worst-case transmitter tuning is the same for both links, since the calibration of each ring is inversely proportional to the number of wavelenghts, according to (1).

In the 4-receivers link, the photodiodes in the receivers 1, 2, 3 and 4 absorb respectively 7%, 14%, 38% and 100% of the optical power available at each point of the waveguide. The significant difference with the lossless scenario mentioned in previous section (where 25%, 33%, 50% and 100% are expected) validate the efficacy of the design method. By considering a linear relationship between the power absorption and the photodiode length and that the length required for

100% absorption is 15 μm , we obtain lengths 1 μm, 2.4 μm and 5.8 μm for receivers 1 to 3 respectively.

Regarding the energy consumption, the modulation power increases linearly with the number of receivers for WDM-based links, since each destination requires a dedicated wavelength. It accounts for 0.3 pJ/broadcasted bit for every new destination. The link we propose does not share this drawback because only a single wavelength is required for all scenarios. Regarding the TIA, no difference is observed since both approaches imply the same number of receivers.

The main drawback of our approach is the laser power overhead, which is is due to the facet losses introduced by the PADs. Indeed, they account for up to 2 dB per photodiode while MRR through and drop losses in WDM-based links typically lead to less than 1 dB overhead. It is worth noticing that our model does not take into account crosstalk noise induced by the use of multiple wavelengths, which would contribute to increase the laser power required for WDM links. Most importantly, the absence of MRR tuning in the PAD link leads to significant energy saving. Overall, for a 4-receivers link, our approach leads to 7.82 pJ/broadcasted bit, which corresponds to a 39% reduction compared to the 12.89 pJ required by the WDM solution.

B. Waveguide Length

In the following, we further evaluate the scalability of the proposed link to broadcast data towards 4 receivers for a total waveguide length ranging from 20 mm to 100 mm. These lengths correspond to interposer-scale interconnects. As reported in Fig. 3, PAD links enable significant improvements in energy efficiency for very short range communication, despite a slightly larger laser output power (e.g. -1.32 dBm for PADs link wrt. -4.87 dBm for WDM link at 20 mm). However, from 96 mm, the accumulation of losses induced by long waveguide and photodiode facet losses lead to significant laser power power overhead, which does not allow to overcome the ring calibration required in WDM links.

C. Discussion

These results demonstrate the potential of the proposed PAD link to efficiently broadcast data. One of the key challenges

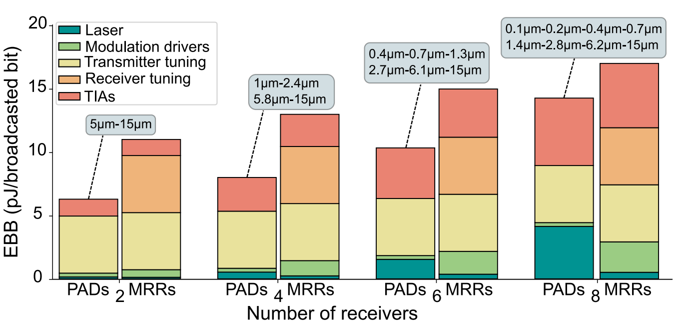

Fig. 2. Breakdown of the energy consumption per broadcasted bit for the PAD-based (proposed) and WDM-based (baseline) links. We also report the estimated length ratio for each diode in the PAD link assuming a linear relation between length and absorption.

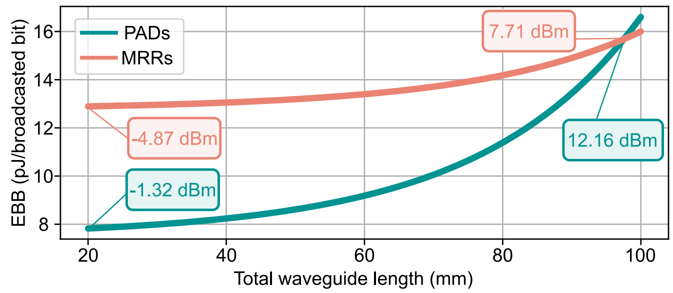

Fig. 3. EBB for 4 receivers for both PADs link (proposed) and WDM link (baseline) according to the waveguide length.

that remains is the scalability, which is mostly limited by the photodiode facet losses. This calls for device-level optimizations, which are out of the scope of the paper but will be explored in the future. Results also show that the calibration in the modulator accounts for at least 35% of the total energy consumption of the PAD link for all studied scenarios. Since our approach requires a single wavelength, we will investigate the use of on-chip lasers with direct modulation [13] to avoid the need of wavelength calibration.

IV. Conclusions

This paper presents a novel broadcast link for nanophotonic interconnects. It is mainly composed of partial-absorption photodiodes (PADs) which enable the extraction and conversion of a fraction of the power propagating in a waveguide. Unlike conventional optical broadcast links which rely on Wavelength Division Multiplexing (WDM), the proposed receiver does not require any microring resonator filter calibration. Our simulation results show that up to 39% energy saving can be obtained for a 4-receiver link, while the study involving various waveguide lengths show that the proposed link is particularly interesting for short interposer-scale communications. The perspectives include the use of on-chip lasers with direct modulation, the optimization of photodiodes to reduce facet losses and the design of multi-links for large-scale on-chip broadcast. While all the results presented in the paper focus on energy, the proposed link is also expected to bring improvements in terms of chip area, since less tuning and driver circuits are required, which will be investigated in future works. Moreover, proposed link has already been designed for a silicon photonics platform and is currently under fabrication.

References

[1] Yanghui Ou, Shady Agwa, and Christopher Batten. Implementing low-diameter on-chip networks for manycore processors using a tiled physical design methodology. In *2020 14th IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*, pages 1–8. IEEE, 2020.

[2] Zainab Al-Waisi and Michael Opoku Agyeman. An overview of on-chip cache coherence protocols. In *2017 Intelligent Systems Conference (IntelliSys)*, pages 304–309. IEEE, 2017.

[3] Christos A Thraskias, Eythimios N Lallas, Niels Neumann, Laurent Schares, Bert J Offrein, Ronny Henker, Dirk Plettemeier, Frank Ellinger, Juerg Leuthold, and Ioannis Tomkos. Survey of photonic and plasmonic interconnect technologies for intra-datacenter and high-performance computing communications. *IEEE Communications Surveys & Tutorials*, 20(4):2758–2783, 2018.

[4] Ian O'Connor, Faress Tissafi-Drissi, Frédéric Gaffiot, Joni Dambre, Michiel De Wilde, Joris Van Campenhout, Dries Van Thourhout, Jan Van Campenhout, and Dirk Stroobandt. Systematic simulation-based predictive synthesis of integrated optical interconnect. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 15(8):927–940, 2007.

[5] Janibul Bashir, Eldhose Peter, and Smruti R Sarangi. A survey of on-chip optical interconnects. *ACM Computing Surveys (CSUR)*, 51(6):1–34, 2019.

[6] Aditya Narayan, Yvain Thonnart, Pascal Vivet, and Ayse K Coskun. PROWAVES: Proactive Runtime Wavelength Selection for Energy-efficient Photonic NoCs. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2020.

[7] Yvain Thonnart, Mounir Zid, José Luis Gonzalez-Jimenez, Guillaume Waltener, Robert Polster, Olivier Dubray, Florent Lepin, Stéphane Bernabé, Sylvie Menezo, Gabriel Parès, et al. A 10Gb/s Si-photonic transceiver with 150μW 120μs-lock-time digitally supervised analog microring wavelength stabilization for 1Tb/s/mm 2 Die-to-Die Optical Networks. In *2018 IEEE International Solid-State Circuits Conference-(ISSCC)*, pages 350–352. IEEE, 2018.

[8] Qianhuan Yu, Keye Sun, Qinglong Li, and Andreas Beling. Segmented waveguide photodetector with 90% quantum efficiency. *Optics express*, 26(10):12499–12505, 2018.

[9] Chen Sun, Chia-Hsin Owen Chen, George Kurian, Lan Wei, Jason Miller, Anant Agarwal, Li-Shiuan Peh, and Vladimir 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*, pages 201–210. IEEE, 2012.

[10] Daniel Benedikovic, Léopold Virot, Guy Aubin, Jean-Michel Hartmann, Farah Amar, Bertrand Szelag, Xavier Le Roux, Carlos Alonso-Ramos, Paul Crozat, Éric Cassan, et al. Comprehensive study on chip-integrated germanium pin photodetectors for energy-efficient silicon interconnects. *IEEE Journal of Quantum Electronics*, 56(1):1–9, 2019.

[11] Robert Polster, Yvain Thonnart, Guillaume Waltener, Jose-Luis Gonzalez, and Eric Cassan. Efficiency optimization of silicon photonic links in 65-nm CMOS and 28-nm FDSOI technology nodes. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 24(12):3450–3459, 2016.

[12] Johnnie Chan, Gilbert Hendry, Keren Bergman, and Luca P Carloni. Physical-layer modeling and system-level design of chip-scale photonic interconnection networks. *IEEE Transactions on computer-aided design of integrated circuits and systems*, 30(10):1507–1520, 2011.

[13] Jaechul Lee, Cédric Killian, Sébastien Le Beux, and Daniel Chillet. Approximate nanophotonic interconnects. In *Proceedings of the 13th IEEE/ACM International Symposium on Networks-on-Chip*, pages 1–7, 2019.