



HAL
open science

Approximate nanophotonic interconnects

Jaechul Lee, Cédric Killian, Sébastien Le Beux, Daniel Chillet

► **To cite this version:**

Jaechul Lee, Cédric Killian, Sébastien Le Beux, Daniel Chillet. Approximate nanophotonic interconnects. NOCS 2019 - 13th IEEE/ACM International Symposium on Networks-on-Chip, Oct 2019, New York, United States. pp.1-7, <10.1145/3313231.3352365>. <hal-02341667v2>

HAL Id: hal-02341667

<https://hal.science/hal-02341667v2>

Submitted on 25 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Approximate Nanophotonic Interconnects

Jaechul LEE

jaechul.lee@irisa.fr

Univ Rennes, Inria, CNRS, IRISA

F-22305 Lannion, France

Sébastien LE BEUX

slebeux@encs.concordia.ca

Dept. of Electrical and Computer Engineering

Concordia University

Montreal, Quebec, Canada

Cédric KILLIAN

cedric.killian@irisa.fr

Univ Rennes, Inria, CNRS, IRISA

F-22305 Lannion, France

Daniel CHILLET

daniel.chillet@irisa.fr

Univ Rennes, Inria, CNRS, IRISA

F-22305 Lannion, France

ABSTRACT

The energy consumption of manycore is dominated by data movement, which calls for energy-efficient and high-bandwidth interconnects. Integrated optics is promising technology to overcome the bandwidth limitations of electrical interconnects. However, it suffers from high power overhead related to low efficiency lasers, which calls for the use of approximate communications for error tolerant applications. In this context, this paper investigates the design of an Optical NoC supporting the transmission of approximate data. For this purpose, the least significant bits of floating point numbers are transmitted with low power optical signals. A transmission model allows estimating the laser power according to the targeted BER and a micro-architecture allows configuring, at run-time, the number of approximated bits and the laser output powers. Simulations results show that, compared to an interconnect involving only robust communications, approximations in the optical transmission lead to up to 42% laser power reduction for image processing application with a limited degradation at the application level.

KEYWORDS

Optical Network-on-Chip, nanophotonic interconnects, approximate communications, energy efficiency

1 INTRODUCTION

Multiprocessor System-on-Chip (MPSoC) are evolving towards the integration of hundreds of cores on a single chip. This evolution leads to significant challenges for the design of efficient interconnect due to the ever growing data exchange between processors. Electrical Network-on-Chips (NoCs) are now facing challenges related to the end of the CMOS scaling, as it already costs more energy to move data than to compute them [2].

Nanophotonic interconnects, also called Optical NoC (ONoC), are promising solutions to overcome bandwidth and power limitation. However, their implementation remains challenging due to the low efficiency of the lasers, which are key devices in such interconnects. The laser power consumption is mainly driven by the high signal power required to transmit data at the targeted Bit Error Rate (BER).

Approximate computing is emerging as an efficient method to improve energy efficiency and execution speed of embedded computing systems [18]. It relies on a accuracy reduction of the data

representations, which allows lowering design constraints and improve performances at the cost of result degradation at the application level. The paradigm has been deployed in numerous systems at both operator and memory levels. Recently, the design of a lower power Electrical NoC [1] has been investigated using approximate computing techniques. For this purpose, data are injected using low voltage supplied buffers. While the results are promising, the approach cannot be directly applied to nanophotonics interconnects, which requires control of the signal power emission directly on the laser side.

In this paper, we propose an approximate nanophotonic interconnect. It relies on the transmission of the least significant bits using low power optical signals. The resulting higher error rate allows drastic reduction in the laser power consumption. The most significant bits remain transmitted using high power signals, which allows exploring robustness and energy efficiency trade-offs for on-chip optical interconnects. The method we propose is applied on the transmission of Float Point (FP) numbers in Single Writer Multiple Reader (SWMR) optical channels. SNIPER [5] simulations are carried out for AxBench [19] benchmark applications running on shared-memory architectures. Results show that the proposed method allows reaching up to 42% laser power saving. To the best of our knowledge, this work is the first attempting to implement approximate computing concepts into nanophotonic interconnects.

The paper is organized as follow. Section 2 introduces related works and Section 3 presents the method allowing to apply approximate computing concepts in nanophotonic interconnects. Section 4 describes the considered hardware implementation and Section 5 details the simulation setup. Results are provided in Section 6 and the last concludes and gives perspective to this work.

2 RELATED WORKS

On-chip optical interconnects rely on the propagation of optical signals to transmit the information. The optical signals are emitted by laser sources, which account for a large part of the interconnect power consumption [15]. The laser power consumption depends on the optical power to be emitted, which is estimated according to numerous device level characteristics (e.g. photodetector sensitivity, laser efficiency, waveguide loss, and microring resonator insertion loss) and system level parameters (e.g. number of wavelengths, propagation distance). In order to adapt the laser power to communication requirements, prediction on network traffic has been

investigated in [7, 21]. In [16], the availability of the destination is validated before starting the communication. The Bit Error Rate (BER) is a key metric to evaluate the required laser output power, taking into account the device characteristics. Most ONoCs rely on Wavelength Division Multiplexing (WDM), which allows propagating multiple signals simultaneously on a same link. Since the targeted BER is the same for all transmitted signals, all the lasers emit at a same power level. Hence, while such methods lead to homogeneous energy/bit figures, which simplify the design task, they do not allow taking into account the significant of the transmitted information, which we explore in this work.

Approximate computing takes advantage of the intrinsic robustness of applications to improve their design energy efficiency, memory footprint and throughput. For instance, approximations allow reducing the hardware complexity of arithmetic circuits such as adder and multiplier [8]. It also contributes to improve power consumption of memories [12] and sensors for IoT applications [14]. In the context of electrical NoCs, approximation techniques have been deployed to reduce the number of transmitted bits [3]. In [1], voltage lowering techniques are applied to lower the robustness of links transmitting the least significant bits of FP numbers. In this work, we investigate the design of energy propositional optical links allowing to transmit the Most and Least Significant Bits at low BER and high BER respectively. To the best of our knowledge, this work is the first addressing the deployment of approximate computing for the design of a nanophotonic interconnects.

3 PROPOSED APPROXIMATE ON-CHIP OPTICAL COMMUNICATION

In this work, we aim at lowering the transmission robustness of the signals transporting the least significant bits. By tolerating approximation in the communications, constraints on the laser power are lowered, which contributes to improve the interconnect energy efficiency. In the following, the proposed generic approximate communication scheme for ONoC is first detailed. We then illustrate use case scenarios and we discuss design trade-off exploration opportunities.

3.1 Method Overview

The approach aims at maximizing the ONoC energy efficiency while minimizing the impact of the approximation at the application level. For this purpose, we discriminate bits according to their significant: most and least significant bits are transmitted under low and high BER respectively. The signals for which high error rate is tolerated will be emitted at a lower power level, thus reducing the laser power consumption. For this purpose, we assume N_{data} bit-width data to be transmitted in parallel using N_w optical wavelengths emitted by N_w lasers, as illustrated in Figure 1.

The number of consecutive bits to be sequentially transmitted at a given wavelength is defined by Bit Stream Length (BSL). The discrimination between the bits that cannot be approximated (i.e. *NotAx*) from the bits that can be approximated (i.e. *Ax*) is defined by Ax_{max} . In order to introduce flexibility in the communication system, we define Bit Protection Level (BPL). It allows adapting, for a given application, the actual number of bits to be approximated.

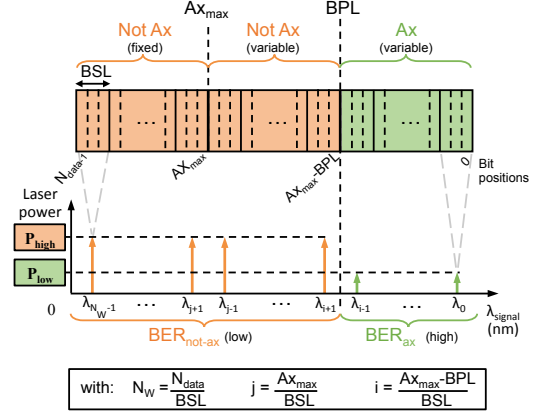


Figure 1: Overview of the proposed method to transmit approximated data in nanophotonic interconnects.

Symbol	Definition
N_{data}	Data size
Ax_{max}	Max. number of approximated bits
BPL	Bit Protection Level $0 \leq BPL \leq Ax_{max}$
N_w	Number of Lasers
BSL	Bit Stream Length
BER_{not-ax}	BER for non approximated bits
BER_{ax}	BER for approximated bits

Table 1: Parameters for approximated data

BPL value ranges from 0 to Ax_{max} . We thus obtain the following three groups of bits:

- **bits ranging from $N_{data}-1$ to Ax_{max}** correspond to the Most Significant Bits (MSB). They cannot be approximated (i.e. *Not Ax*) and they always require a robust communication. The bits are thus transmitted using high power signal level P_{high} , which is defined to allow reaching a nearly error free communication at low BER (i.e. BER_{Not-ax} , typically 10^{-12}). The transmission of the rest of the bits depends on BPL.
- **bits ranging from Ax_{max} to $Ax_{max} - BPL$** correspond to the flexible part of the bits that are transmitted without error (i.e. at BER_{Not-ax}) using P_{high} laser power.
- **bits ranging from $Ax_{max} - BPL$ to 0** are the Least significant Bits (LSB), which are approximated (*Ax*). For a given application, they are considered as not too sensitive to tolerate errors during their transmission. The laser power level of the corresponding signals is lowered to P_{low} , which leads to BER_{ax} (e.g. 10^{-5}).

Table 1 summarizes the parameters used in our formulations.

3.2 Quality of Result/Energy Efficiency Trade-offs

Our approach aims at improving the interconnect energy efficiency by adapting the transmission robustness to the application requirements. The exploration of the robustness/energy trade-off requires to take account of the data type and their communication rate while executing an application. We apply our approximation technique on FP numbers, which are intrinsically resilient to errors [10]. Indeed, the impact of errors introduced in the LSBs of the Mantissa is limited thanks to the normalization. To do so, we first define Ax_{max} at design-time according to the maximum approximation tolerated in the system (e.g. all the benchmark applications). Second, BPL is defined at run-time depending on the requirements specific of a given application; this allows transmitting none/part/all of the remaining bits with approximation, i.e. using signals at P_{low} instead of P_{high} . The power saving P_{saving} is formulated as follow:

$$P_{Saving} = \frac{(N_{data} - Ax_{max} + BPL) \times P_{high} + (Ax_{max} - BPL) \times P_{low}}{N_{data} \times P_{high}} \quad (1)$$

Figures 2-a and b illustrate IEEE 754-2008 double-precision and IEEE 754-2008 single-precision FP format examples, for which we assume $Ax_{max} = 48$ and $Ax_{max} = 16$ respectively. Regarding the double-precision example, BPL is set to 4: bits 44 to 47 are transmitted at P_{high} (in addition to bits 48 to 63 that are permanently protected thanks to Ax_{max}). Since we assume $BSL = 4$, the 44 unprotected bits are transmitted in parallel using 11 signals (i.e. λ_0 to λ_{10}) at power P_{low} within 4 clock cycles. For the single-precision example, only the signal at λ_0 is emitted at P_{low} since we assume $BPL = 8$ and $BSL = 8$ (i.e. each wavelength is responsible for transmitting 8 bits).

In addition to the ratio of laser emitting at P_{low} , the actual energy saving depends on i) the BERs targeted for approximated and protected bits (i.e. BER_{ax} and BER_{Not-ax}) and ii) the ratio of FP numbers transmitted on the interconnect. Indeed, not all data can be approximated. For instance, instructions and integer numbers are highly sensitive to errors and should be transferred with all signals at P_{high} , which can be achieved by setting BPL to the maximum value. The ratio of FP numbers are obtained using simulation, as detailed in Section 5.

4 HARDWARE IMPLEMENTATION AND COMMUNICATION MODEL

In the following subsections, we present the hardware architecture allowing to implement the proposed approximate communication scheme. The transmissions of data with and without approximation are detailed. Then, we introduce the communication model allowing to estimate the required laser power according to the targeted BER.

4.1 SWMR Channel Design

Without lack of generality, we illustrate in Figure 3 the proposed hardware implementation on a Single Writer Single Reader (SWSR) channel. The implementation on SWMR channels relies on the replication of the hardware blocks related to the reader. The writer (source) is connected to the reader (destination) using waveguides (one in the figure). Each waveguide allows transmitting N_w signals using WDM ($\lambda_0.. \lambda_{N_w-1}$). The signals are emitted by on-chip lasers

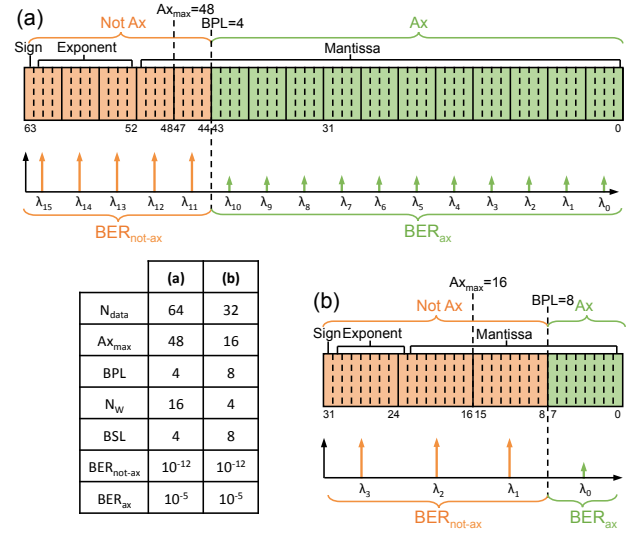


Figure 2: Example of floating point data approximation: a) IEEE 754-2008 double-precision (DP) format, b) IEEE 754-2008 single-precision (SP) format.

and are combined into waveguides using, for instance, MultiMode Interference (MMI) couplers [6]. In case no communication occurs, the lasers are turned OFF by setting their bias current to zero. This is achieved by controlling the laser drivers through the Optical Link Manager. The design of centralized manager to configure the communication channels is out of the scope of the paper and has been already investigated in [17]. When a communication is initiated, the manager activates the lasers (the lasing effects occur after few ns typically). Then, the data to transmit are serialized and OOK modulation is carried out by MRs on the corresponding optical signals. As defined in the previous Section, the degree of serialization depends on the bit stream length (BSL), which corresponds to ratio between the data bit width (N_{data}) and the number of wavelengths (N_w). We assume that data transmission involve the use of all the signals. The design of the serializers, which has been investigated in [9, 20], is out of the scope of the paper. The modulated signals propagate along the waveguide until reaching the destination. At this stage, the signals are ejected from the waveguide using MRs in ON state. They are redirected to photodetectors from which opto-electronic conversions are carried out. The serial streams are then de-serialized back into the original data format.

The power at which the lasers emit signals depends on the approximation level of the data to transmit. We distinguish two scenario involving no approximation and approximation, as detailed below:

- **Robust transmission** is used for sensitive data such as instructions and integer numbers. It involves low BER, with BER_{not-ax} (e.g. 10^{-12}), which is guaranteed by injecting high optical power (P_{high}) from the lasers. This results in robust but power hungry communications for all bits.
- **Approximate transmission** is dedicated to data that can be approximated using the technique defined in the previous

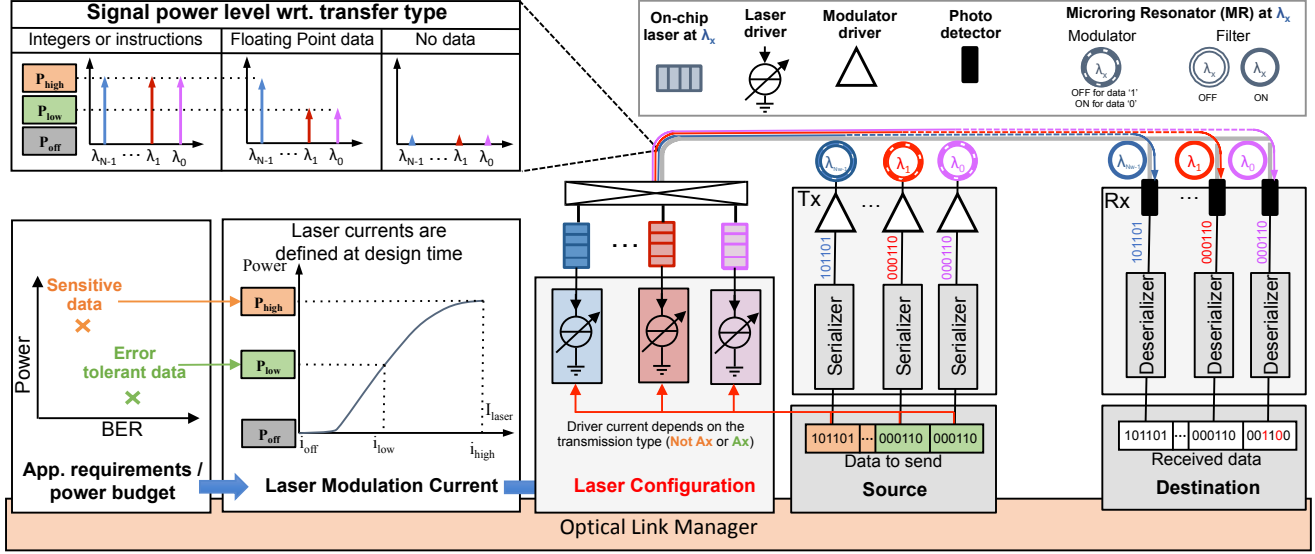


Figure 3: Optical link channel allowing the transmission of approximated data.

Section. It involves heterogeneous power level emission for the lasers: MSB and LSB are transmitted under high power P_{high} and low power P_{low} levels respectively.

The manager individually configures the laser drivers according to the selected bit protection level. The two possible power levels (P_{high} and P_{low}) are defined by evaluating the SNR on the photodetectors, as detailed in the following.

4.2 Laser Power Model

P_{high} and P_{low} depend on the targeted BER for robust and approximate communications, i.e. BER_{not-ax} and BER_{ax} respectively. To evaluate the laser power, we use the following transmission models proposed in [9]:

$$SNR = [\operatorname{erfc}^{-1}(1 - 2.BER)]^2 \quad (2)$$

$$SNR = \frac{\mathfrak{R} \times (OP_{signal} - OP_{crosstalk})}{i_n} \quad (3)$$

where SNR is the Signal to Noise Ratio calculated for an OOK modulation, OP_{signal} is the optical signal power received by a photodetector and $OP_{crosstalk}$ is the worst case crosstalk in the channel. \mathfrak{R} is the photodetector responsivity (1 A/W in this work), and i_n is the dark current ($4\mu A$). The model allows accurate evaluation of the losses experienced by the signals propagating on SWMR channels. Indeed, the losses depend on the signal wavelength and the MRs transmission. It also allows estimating the crosstalk by considering the distance between signal and MR resonant wavelengths. In our implementation, we assume 6.9dB extinction ratio for the MRs and 1dB/cm waveguide loss. Regarding the lasers, we assume CMOS compatible PCM-VCSELs [4].

5 EXPERIMENTAL SETUPS

5.1 3D ONoC architecture

We assume a 3D integrated circuit composed of an electrical layer and an optical layer implementing computing cores and the nanophotonic interconnect respectively. Figure 4-a illustrates a shared memory architecture example with 4 clusters and 4 cores per cluster. Each cluster K includes a last cache level (L2), shared among the cores in a same cluster, and each core has its own private L1 data (L1d) and instructions (L1i) caches. In this shared memory architecture, the communications on the nanophotonic interconnect are initiated by the L2 and the main memory (DRAM) through the Optical Network Interface (ONI). Each ONI is composed of a single Writer and K Readers to interface with the $K + 1$ SWMR channels.

The transmitted data on the channels are either instructions, cache coherency messages, integer numbers and FP numbers. Approximate communications are only carried out for FP numbers. The bits transmitted at BER_{ax} depend on the selected BPL. When transmitting others data, BPL is set to the maximum value by the manager.

5.2 Approximation under Cache Coherence Traffic

Cache coherency involves traffic between L2 and DRAM that is mainly initiated by cache miss or write-back access [13].

However, the traffic initiated by cache coherence protocols also involves addresses of the data to be transmitted and is thus data type agnostic. To define whether a data can be approximated or not, we use local tables to associate ranges of memory addresses to data types [10].

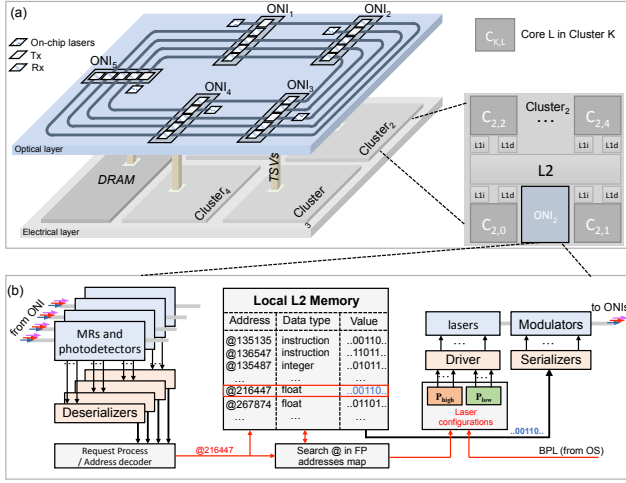


Figure 4: Architecture overview: (a) 3D ONoC and (b) SWMR channel management for approximate communications.

Table 2: Hardware parameters of the simulated architecture

Parameters	Value
Cache protocol	MSI
L1 I / D cache	32 KB
L2 cache	2048 KB
L2 cache line size	64 Bytes

The Figure 4-b illustrates the principle with a cache miss scenario. The L2 on cluster 2 receives a request containing the address of a data to be returned. The address (@216447 in the example) is compared to the address ranges of data to be approximated (FP numbers). In case of matching, the power level of lasers corresponding to the approximated bits is lowered to P_{low} . The approximation level (BPL) is application specific and can be defined at run-time by an Operating System. The manager then grants the SWMR access to the L2 that initiated the request and the data is transmitted to the serializers for signal modulation. In order to reduce the latency and hardware overhead, the whole cache line is transferred: the size of response packets (which can be approximated) is larger than packet request size (which can never be approximated). The considered memory characteristics are given in Table 2.

5.3 Simulation platform

We evaluate the proposed approximation technique on benchmark applications from Approxbench suite [11]. It provides standard quality metric to measure the approximation out of quantifying precision for blackscholes, canneal, sobel, and streamcluster applications. The benchmarks are simulated using SNIPER [5] to extract inter-core and core-to-memory communications traces.

We simulate architectures with 16, 32, and 64 cores. We consider 4 cores per cluster (i.e. $L = 4$), hence the number of clusters are 4, 8, and 16. Each architecture has a shared DRAM, hence resulting in 5, 9, and 17 SWMR channels and ONIs. From the simulation results, we extract the communication traces between L2 caches

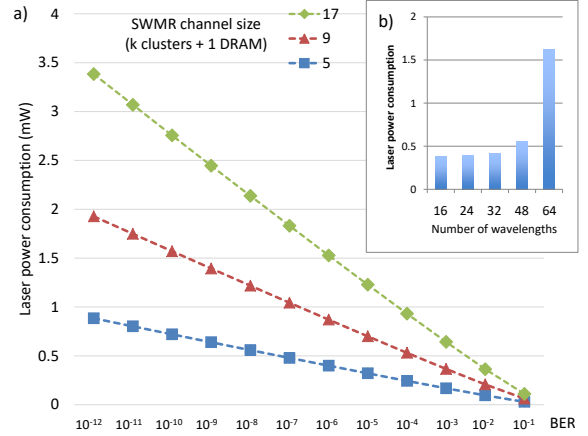


Figure 5: Laser power requirements according to a) the target BER for 16-wavelengths channels and b) the number of wavelengths for SWMR channels with 4 readers.

and the DRAM. This allows evaluating, for each application, the number and the type of data transfer on the optical interconnect. The amount of transmitted FP data gives the number of network packets to be sent using our approximation technique. The power saving is obtained by applying our model of equation 1.

6 RESULTS

6.1 laser Power Estimation

We first estimate the laser power consumption according to BER ranging from 10^{-1} to 10^{-12} . For this purpose, we model SWMR links with 4, 8, and 16 readers and the device characteristics defined in Section 4.2. We assume a 50nm Free Spectral Range (FSR) and 16 wavelengths per channel. As illustrated in Figure 5-a, the required laser power is quasi-proportional to the BER, which is explained by the relatively low crosstalk (approximately 3nm spacing between the MRs resonant wavelengths). Figure 5-b shows the impact of the crosstalk on a 4 readers link by considering 16, 32, and 64 wavelengths. We obtain quadratic growth of the power with the BER, which would thus lead to significant power saving when using high BER for transmission of approximate data.

6.2 Benchmark Characterization

Figure 6 presents, for each application, the distribution of the transmitted packet type (Instruction, Integer Data, and FP Data) according to the architecture size.

We first observe that ratio of FP numbers strongly depends on the application. For instance, Canneal and Sobel involve 4% and 80% of transmitted FP numbers respectively, which can be explain by the data type each application uses. Interestingly, the results show that the ratio of transmitted packets that can be approximated increases with the number of cores. This is due to the increasing number of shared data induced by parallel execution of the kernel. This characterization allow the designer to target on which type of data

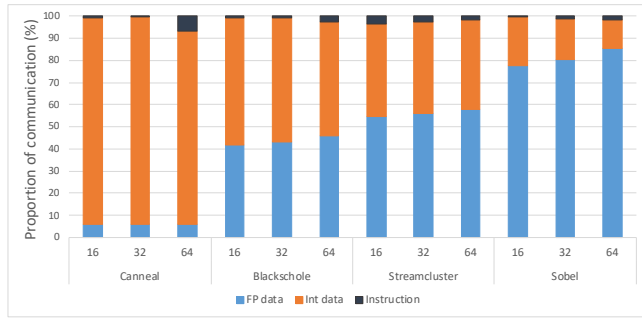


Figure 6: Benchmarks characterizations.

approximation will have more impact on power saving. We provide energy saving in the following section for Sobel application by using the proposed approach on the FP numbers which are mostly represented on the interconnect.

6.3 Power Breakdown and Quality of Result

In the following, we apply the proposed approach to approximate FP numbers in Sobel benchmark. We consider a 64 cores architecture with 16 SWMR links containing 16 wavelengths each. We use $N_{data} = 64$ and we assume $Ax_{max} = 32$. Power reduction are estimated for BER_{ax} ranging from 10^{-1} to 10^{-11} and for BPL ranging from 0 to 28. Figure 6 illustrates the power reduction with respect to a transmission of all bits at $BER_{not-ax} = 10^{-12}$ and with $BPL=32$. We observe that increasing BER_{ax} and decreasing BPL can bring more power gain respectively. Reducing both bit protection level and communication robustness allow reaching up to 42% power reduction. These results demonstrate the effectiveness of the proposed approach to drastically reduce the power consumption.

We also compare images processed with and without approximation by computing their Mean Square Error (MSE). The baseline solution leads to $MSE = 0$ and implies using only robust communication at BER_{not-ax} and maximum protection level $BPL = 32$. As it can be seen on the figure, using approximate communication with $BER_{ax} = 10^{-2}$ and $BPL = 16$ leads to acceptable degradation of the processed image ($MSE = 4.8E^{-4}$) as edge detection is still achieved. The resulting technique thus leads to 20% laser power saving without any loss in the detection accuracy. The highest power reduction (42%) is achieved for $BER_{ax} = 10^{-1}$ and $BPL = 0$. However the resulting image degradation is severe ($MSE = 4.4E^{-2}$) and could lead to wrong detection, which might not be acceptable in some applications. Interestingly, numerous intermediate solutions exist and are reachable by adapting either the laser modulation current or the protection level. The latter can be done at run-time using our approach, thus leading to true energy proportional nanophotonic interconnects.

7 CONCLUSIONS

In this work, we have investigated the design of an Optical NoC supporting the transmission of approximate data. For this purpose, the least significant bits of floating point numbers are transmitted with low power signals. A transmission model allows estimating the laser power according to the targeted BER and a micro-architecture

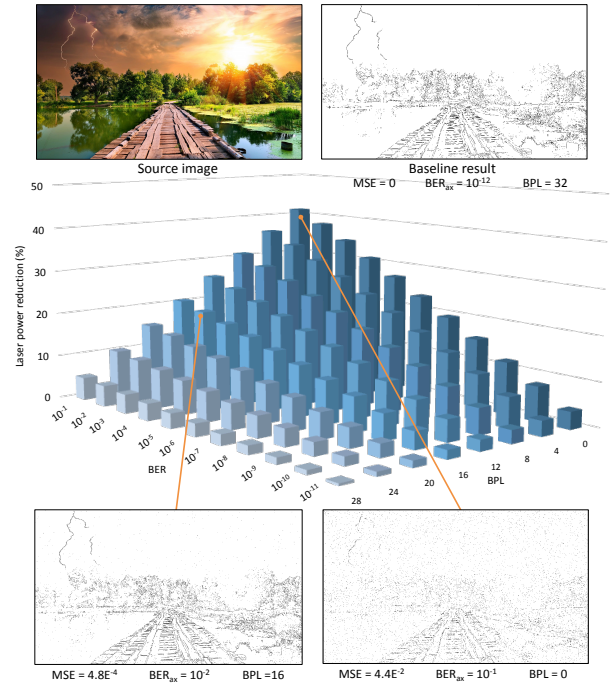


Figure 7: Laser power saving and image degradation for Sobel benchmark application. Results are given for BER_{ax} ranging from 10^{-1} to 10^{-11} and for BPL ranging from 0 to 28. We assume a 64 cores architecture with 17 clusters, $N_{data} = 64$, $Ax_{max} = 32$ and $BER_{not-ax} = 10^{-12}$.

allows configuring, at run-time, the bit protection level and the laser output power. Simulations results show that, compared to an interconnect involving only robust communication, approximations in the optical transmission lead to up to 42% laser power reduction for image processing application and shown the associated application degradation.

8 ACKNOWLEDGEMENT

This work benefited from the support of the project SHNoC ANR-18-CE25-0006 of the French National Research Agency (ANR), and a French Government support granted to the COMIN Labs excellence laboratory managed by the ANR in the *Investing for the Future* program under reference ANR-10-LABX- 07-01.

REFERENCES

- [1] A. B. Ahmed, D. Fujiki, H. Matsutani, M. Koibuchi, and H. Amano. 2018. AxNoC: Low-power Approximate Network-on-Chips using Critical-Path Isolation. In *2018 Twelfth IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*. IEEE, New York, NY, USA, 1–8. <https://doi.org/10.1109/NOCS.2018.8512158>
- [2] T. Alexoudi, N. Terzenidis, S. Pitris, M. Moralis-Pegios, P. Maniotis, C. Vagionas, C. Mitsolidou, G. Mourgias-Alexandris, G. T. Kanellos, A. Miliou, K. Vyrosokinos, and N. Pleros. 2019. Optics in Computing: From Photonic Network-on-Chip to Chip-to-Chip Interconnects and Disintegrated Architectures. *Journal of Lightwave Technology* 37, 2 (Jan 2019), 363–379. <https://doi.org/10.1109/JLT.2018.2875995>
- [3] Rahul Boyapati, Jiayi Huang, Pritam Majumder, Ki Hwan Yum, and Eun Jung Kim. 2017. APPROX-NoC: A Data Approximation Framework for Network-On-Chip Architectures. *SIGARCH Comput. Archit. News* 45, 2 (June 2017), 666–677. <https://doi.org/10.1145/3140659.3080241>

- [4] C. Sciancalepore *et al.* 2012. Thermal, Modal, and Polarization Features of Double Photonic Crystal Vertical-Cavity Surface-Emitting Lasers. *IEEE Photonics Journal* 4, 2 (April 2012), 399–410. <https://doi.org/10.1109/JPHOT.2012.2189378>
- [5] Trevor E. Carlson, Wim Heirman, Stijn Eyerman, Ibrahim Hur, and Lieven Eeckhout. 2014. An Evaluation of High-Level Mechanistic Core Models. *ACM Transactions on Architecture and Code Optimization (TACO)*, Article 5 (2014), 23 pages. <https://doi.org/10.1145/2629677>
- [6] F. Mandorlo *et al.* 2012. Controlled Multi-Wavelength Emission in Full CMOS Compatible Micro-Lasers for on Chip Interconnections. *Journal of Lightwave Technology* 30, 19 (Oct 2012), 3073–3080. <https://doi.org/10.1109/JLT.2012.2209863>
- [7] P. K. Hamedani, N. E. Jerger, and S. Hessabi. 2014. QuT: A low-power optical Network-on-Chip. In *2014 Eighth IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*. 80–87. <https://doi.org/10.1109/NOCS.2014.7008765>
- [8] Jie Han and Michael Orshansky. 2013. Approximate computing: An emerging paradigm for energy-efficient design. *2013 18th IEEE European Test Symposium (ETS)* (2013), 1–6.
- [9] C. Killian, D. Chillet, S. Le Beux, Van-Dung Pham, O. Sentieys, and I. O'Connor. 2017. Energy and performance trade-off in nanophotonic interconnects using coding techniques. In *2017 54th ACM/EDAC/IEEE Design Automation Conference (DAC)*. IEEE, New York, NY, USA, 1–6. <https://doi.org/10.1145/3061639.3062237>
- [10] Adrian Sampson, Werner Dietl, Emily Fortuna, Danushen Gnanaprasam, Luis Ceze, and Dan Grossman. 2011. EnerJ: Approximate Data Types for Safe and General Low-power Computation. *SIGPLAN Not.* 46, 6 (June 2011), 164–174. <https://doi.org/10.1145/1993316.1993518>
- [11] Adrian Sampson, Werner Dietl, Emily Fortuna, Danushen Gnanaprasam, Luis Ceze, and Dan Grossman. 2015. ACCEPT: a programmer-Guided compiler Framework for Practical Approximate Computing. 1, 6 (June 2015), 11. <https://doi.org/10.1145/1993316.1993518>
- [12] A. Sampson, J. Nelson, K. Strauss, and L. Ceze. 2013. Approximate storage in solid-state memories. In *2013 46th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. 25–36.
- [13] Yan Solihin. 2015. *Fundamentals of Parallel Multicore Architecture*.
- [14] Phillip Stanley-Marbell and Martin C. Rinard. 2018. A Hardware Platform for Efficient Multi-Modal Sensing with Adaptive Approximation. *CoRR* abs/1804.09241 (2018).
- [15] C. Sun, C. O. Chen, G. Kurian, L. Wei, J. Miller, A. Agarwal, L. Peh, and V. Stojanovic. 2012. 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*. 201–210. <https://doi.org/10.1109/NOCS.2012.31>
- [16] S. Werner, J. Navaridas, and M. Luján. 2017. Efficient sharing of optical resources in low-power optical networks-on-chip. *IEEE/OSA Journal of Optical Communications and Networking* 9, 5 (May 2017), 364–374. <https://doi.org/10.1364/JOCN.9.000364>
- [17] Xiaowen Wu, Jiang Xu, Yaoyao Ye, Zhehui Wang, Mahdi Nikdast, and Xuan Wang. 2014. SUOR: Sectioned Unidirectional Optical Ring for Chip Multiprocessor. *J. Emerg. Technol. Comput. Syst.* 10, 4, Article 29 (June 2014), 25 pages. <https://doi.org/10.1145/2600072>
- [18] Q. Xu, T. Mytkowicz, and N. S. Kim. 2016. Approximate Computing: A Survey. *IEEE Design Test* 33, 1 (Feb 2016), 8–22. <https://doi.org/10.1109/MDAT.2015.2505723>
- [19] Amir Yazdanbakhsh, Divya Mahajan, Pejman Lotfi-Kamran, and Hadi Esmaeilzadeh. 2016. AxBench: A Benchmark Suite for Approximate Computing Across the System Stack.
- [20] Z. Wang *et al.* 2016. A Holistic Modeling and Analysis of Optical-Electrical Interfaces for Inter/Intra-chip Interconnects. *IEEE TVLSI* 24, 7 (July 2016), 2462–2474. <https://doi.org/10.1109/TVLSI.2015.2511065>
- [21] L. Zhou and A. K. Kodi. 2013. PROBE: Prediction-based optical bandwidth scaling for energy-efficient NoCs. In *2013 Seventh IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*. 1–8. <https://doi.org/10.1109/NoCS.2013.6558413>