UNIVERSITÉ DE RENNES 1

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Informatique*

Par
Jaechul LEE

Approximate communication techniques exploration for efficient nano-photonics interconnects

Exploration de techniques de communication approximatives pour des interconnexions nano-photoniques efficaces

Thèse présentée et soutenue à Lannion, le
Unité de recherche : IRISA

Rapporteurs avant soutenance :

François VERDIER Professeur des universités, Université de Nice Côte d'Azur
Camel TANOUGAST Professeur des universités, Université de Lorraine

Composition du Jury :

Président :	Prénom NOM	Fonction et établissement d'exercice *(à préciser après la soutenance)*
Examinateur :	Frédéric PETROT	Professeur des universités, Université Grenoble Alpes
Examinatrice :	Angeliki KRITIKAKOU	Maître de conférences HDR, Université de Rennes 1
Dir. de thèse :	Daniel CHILLET	Professeur des universités, Université de Rennes 1
Co-dir. de thèse :	Cédric KILLIAN	Maître de Conférences HDR, Université de Rennes 1

Résumé

Contexte

L'évolution des architectures de calcul a progressivement conduit à des systèmes très complexes, intégrant de plus en plus de transistors soumis à des fréquences toujours plus élevées. Si cette évolution, connue sous le nom de loi de Moore, a perduré pendant de longues années, les problèmes de consommation énergétiques ont remis en cause l'accroissement continu de la fréquence d'horloge, induisant alors une possible stagnation des performances. Pour éviter cette stagnation et continuer à offrir toujours plus de performances au sein des processeurs, les concepteurs ont alors proposé de multiplier le nombre de cœurs de calcul en conservant la fréquence à une valeur élevée, mais sans qu'elle n'augmente. Par ce biais, les capacités de calcul ont continué à croître permettant alors d'implémenter des applications toujours plus complexes. Toutefois pour exploiter pleinement ces capacités de calcul, les applications ont dû être décrites sous forme de tâches ou *threads* parallèles et déployées sur l'ensemble des cœurs. Or, ce type d'implémentation conduit inévitablement à des échanges entre les cœurs de calcul ou entre les cœurs et la mémoire, nécessitant alors un support de communication efficace.

Parallèlement à cette évolution au sein des processeurs, des architectures plus spécifiques se sont développées, intégrant l'ensemble des éléments permettant de gérer un système, ces architectures sont connues sous le nom de systèmes sur puce (System-on-Chip, SoC). Ces architectures sont généralement soumises à des contraintes plus sévères que les processeurs d'usage général, en particulier elles doivent souvent être frugales en termes d'énergie pour préserver les batteries alimentant le système et limiter les recharges parfois coûteuses ou difficiles à assurer. L'évolution de ces architectures a également conduit à des systèmes fortement parallèles confrontés aux mêmes problèmes de communications précédemment évoqués.

D'une manière générale, pour une architecture disposant d'un faible nombre de cœurs de calcul, un support de communication simple, de type bus, peut être suffisant pour

assurer les échanges de données. Par contre, ce type d'interconnexion est connu pour être peu flexible et difficilement capable de suivre une augmentation importante du nombre de cœurs. Pour permettre d'augmenter le nombre de cœurs de calcul tout en assurant une bonne efficacité du système, la notion de réseau sur puce est apparue. Ces réseaux tirent leurs caractéristiques des réseaux d'ordinateurs et miment leur fonctionnement, en particulier, ils sont capables d'acheminer des messages d'un point à un autre à partir d'éléments de routage et d'informations contenues dans les entêtes des messages identifiant, entre autres, la destination à atteindre.

Pour des architectures dont le nombre de cœurs augmente très rapidement, i.e. plusieurs dizaines, centaines voire même plusieurs milliers de cœurs, les réseaux sur puce prennent une place prépondérante et peuvent devenir le goulot d'étranglement si les communications subissent des conflits et des congestions. Le parallélisme matériel disponible dans les architectures n'ayant alors plus d'intérêt puisque les temps de transferts de données deviennent majoritaires devant les temps de calcul. Dans ce contexte, des solutions innovantes doivent donc être définies pour que l'augmentation du nombre de cœurs au sein de ces architectures puisse se poursuivre tout en offrant un gain en performance.

Parallèlement à cette évolution des architectures électroniques, le domaine de la photonique s'est développé et l'intégration de composants optiques sur silicium a fait d'énormes progrès. En effet, depuis plusieurs années déjà, des composants optiques peuvent être intégrés sur du silicium sans remettre en cause tout le processus de fabrication. Bien que de dimensions plus importantes que les transistors, des lasers, des micro-résonateurs, des photo-détecteurs et des guides d'ondes sont désormais disponibles pour supporter des fonctionnalités au sein des architectures numériques classiques. Ainsi, sur la base de ces composants élémentaires, des premières solutions de réseaux sur puce optiques ont vu le jour. Ces réseaux ont la particularité d'offrir des bandes passantes très importantes ainsi que des latences de communications très courtes. Ces caractéristiques les rendent donc très attrayants pour remplacer, ou compléter, les réseaux sur puce électriques classiques qui souffrent de délais assez longs dès lors qu'un message doit atteindre un destinataire éloigné.

Travaux développés et contributions

Les avantages apportés par ces réseaux optiques (bande passante importante et faible latence) s'accompagnent toutefois d'au moins un inconvénient majeur qui est lié à leur consommation énergétique. En effet, la production d'un signal optique par une source laser est coûteuse, cela est en particulier du à la faible efficacité des lasers, de l'ordre de 15 à 20%. Par ailleurs, le signal optique subit des pertes optiques lorsqu'il transite au sein du guide d'ondes et également à chaque passage par un composant optique, même lorsque celui-ci est inactif. Face à ces constats, l'intégration de photonique sur silicium pour supporter un réseau optique doit donc s'appuyer sur des techniques permettant de profiter des bénéfices de cette technologie tout en réduisant au maximum les inconvénients.

Les travaux présentés dans ce manuscrit adressent cette problématique de l'intégration d'un réseau optique sur puce et proposent des techniques permettant de réduire la consommation énergétique en ajustant au mieux la puissance laser au niveau requis de qualité des communications.

Nos travaux exploitent les propriétés des applications, en particulier par l'utilisation du concept de communications approximatives permettant d'adapter le niveau du signal optique pour conserver globalement de bons résultats au niveau applicatif.

Les travaux que nous avons développés présentent deux contributions qui concernent dans un premier temps i) l'adaptation de la puissance laser aux types de données à transmettre dans le guide d'onde, et dans un deuxième temps, ii) une extension de ces travaux pour prendre en compte la notion de distance entre la source et la destination d'une communication.

Nous avons tout d'abord défini un modèle de coût énergétique des communications pour un réseau sur puce optique. Ce modèle prend en compte l'ensemble des *devices* optiques supportant les communications. Ce modèle est ensuite exploité pour ajuster la puissance laser optique en vue d'augmenter l'efficacité globale de l'exécution d'une application implémentée sur une architecture multi-cœurs composées de plusieurs dizaines de cœurs de calcul. Seul le type de données est pris en compte dans un premier temps, puis la distance de communication est prise en compte pour aboutir à une solution donnant un compromis intéressant entre la qualité des communications et le coût d'implémentation de la proposition.

L'ensemble de ces travaux a été validé au travers de simulations principalement basées sur l'outil Sniper. Cet outil nous a permis d'extraire des traces de communication transitant sur le réseau lors de l'exécution de *benchmarks* sur les architectures cibles. À partir de ces traces, nous avons développé des scripts en Python qui permettent d'analyser le trafic de données et d'appliquer, sur l'ensemble des communications, le modèle de coûts énergétique et temporel précédemment défini. Les résultats montrent qu'il est possible de réduire de façon importante la consommation des lasers tout en maîtrisant le coût matériel d'implémentation de la gestion du système. Notre solution est principalement comparée aux travaux Lorax qui exploitent un concept proche, mais dont le coût d'implémentation est déraisonnable pour des architectures de taille importante. En effet, l'approche développée par les auteurs de la solution Lorax proposent une gestion de la distance de communication très fine, induisant autant de niveaux de puissance laser qu'il existe de cœurs destinataires dans l'architecture. Bien que très intéressante, la solution proposée par ses auteurs n'est pas évaluée en termes d'implémentation matérielle.

Organisation du manuscrit

Le manuscrit est composé de cinq chapitres. À la suite de l'introduction, un premier chapitre dresse un état de l'art des réseaux sur puce optiques et présente les composants optiques élémentaires nécessaires pour les implémenter. Ce chapitre présente également la technique du calcul approximatif et aborde le concept des interconnexions approximatives qui sera exploité par la suite. Ce chapitre introduit la notion de qualité de communication qui est principalement basée sur un calcul d'erreur de transmission. Les approximations sur les données flottantes sont également abordées au regard de l'erreur générée dans les applications.

Le chapitre 2 de ce manuscrit aborde la conception d'un lien optique approximatif. Le principe se base sur un lien optique classique sur lequel la gestion des puissances lasers optiques est revisitée pour augmenter l'efficacité énergétique. Le chapitre décrit le lien proposé et explique comment il est possible d'adapter les niveaux de puissance optiques en fonction du type de données à transmettre. Un focus est fait sur les données flottantes dont on sait qu'elles peuvent généralement supporter des approximations sur les bits de poids faibles sans que cela n'entraine de fortes dégradations sur les résultats de certaines

applications, en particulier pour des applications de type traitement d'images, IA, etc.

Les approximations proposées sont de deux types, soit il est proposé de réduire la puissance laser (approximation), soit il est proposé d'éteindre le laser (troncation). L'extinction des laser n'étant finalement que le cas extrême de l'approximation. Nous définissons alors un mixte d'approximation et de troncation des bits de poids faibles pour atteindre une bonne efficacité énergétique. Le positionnement des seuils d'approximation et de troncation est étudié afin de juger des gains énergétiques potentiels. Afin de réduire la complexité de notre proposition, les approximations et troncations sont réalisées par ensemble de bits, et non par bit à bit.

Pour démontrer l'intérêt de notre proposition, nous avons produit un certain nombre de résultats à partir des *benchmarks* (*ApproxBench*). Sur cet ensemble d'applications, nous analysons les flux réseaux et nous calculons les proportions de données dites approximables (données flottantes) et de données dites sensibles (données entières ou instructions d'un programme). Nous montrons que les applications ne sont pas toutes adaptées au concept de communications approximatives, puisqu'il existe une grande disparité de ces proportions (variation entre 4% et 50% de données approximables).

Nous montrons ensuite que, suivant le niveau de qualité de communication requis, exprimé par le taux d'erreur binaire (Bit Error Rate, BER), la puissance laser épargnée peut être plus ou moins importante, variant de 75% à 15% pour des BER respectivement de 10^{-1} à 10^{-9}. Cette étude montre que le gain qui peut être obtenu sera très dépendant de l'application et de son niveau de résilience.

Pour compléter cette étude, l'interface réseau est alors adaptée de façon à pouvoir identifier les types des données à envoyer dans le réseau. L'identification du type est basée sur l'hypothèse d'un stockage différentié des données approximables et critiques dans des zones mémoires différentes, permettant de se baser sur une comparaison de l'adresse pour appliquer ou non le concept d'approximation. Cette solution à l'avantage d'être peu coûteuse à mettre en œuvre.

Le chapitre 3 adresse quant à lui la gestion de la distance de communication entre la source et la destination d'un message. Nous proposons une gestion plus fine de la puissance des lasers en tenant compte des pertes qui seront subits par le signal optique dans le guide d'ondes. Connaissant la puissance nécessaire à al destination pour que les photodé-

tecteurs soient en mesure de produire le courant électrique représentatif du signal optique, nous montrons qu'il est possible de remonter jusqu'à la source laser en identifiant toutes les pertes. À partir de ce modèle paramétrique, nous montrons qu'il existe un compromis dans la définition du nombre de niveaux de puissance des lasers, et nous démontrons que si l'on s'intéresse au compromis efficacité énergétique - complexité du *driver* du laser, deux niveaux lasers sont suffisants.

Pour illustrer l'intérêt de cette proposition, nous utilisons les mêmes *benchmarks* que précédemment et nous analysons dans un premier temps la répartition des distances de communications. Cette analyse permet d'exhiber l'intérêt d'une gestion de la distance puisque les résultats montrent qu'il existe une proportion assez homogène des communications courtes et longues distances. Nous montrons ensuite les gains obtenus avec notre proposition et nous les comparons à la solution Lorax.

Ce chapitre traite également du *driver* laser qui permet de piloter le courant traversant la diode émettrice du signal optique. Ce *driver* se base sur une mise en parallèle de transistors qui, lorsqu'ils sont activés, permettent d'augmenter le courant dans la diode laser et ainsi de produire un signal optique plus puissant. Le contrôleur du *driver* est également abordé afin d'avoir une évaluation complète de l'ensemble des éléments constituant notre proposition. Des synthèses matérielles, de notre solution et de la solution Lorax, sont réalisées à des fins de comparaisons complètes incluant performance et complexité d'implémentation.

Le dernier chapitre présente les conclusions du document en reprécisant les contributions et les résultats obtenus, puis propose quelques perpectives à ce travail.

Table of Contents

INTRODUCTION

Multi core era and System-on-chip

In computer history, the evolution has been swift: indeed, for several decades, the number of transistors in an Integrated Circuit (IC) is continuously doubled according to Moore's law [1]. Figure 1 presents this evolution over the years since 1970. In this graph, two interesting factors that indicate the evolution trend can be highlighted. First, we can see a steady increase in transistor density in the chip, which led to a successful results in single thread performance. It was possible thanks to engineers' efforts that increased the efficiency with nanotechnology systems whereby transistors sizes went from being millimeters to nanometers. This trend and effort allowed computer systems and chip engineering to continuously develop and, as a result, advance several technologies domains such as mobile, smart and innovative technologies, cloud computing, and robust computing. However, as the chip size is getting smaller and denser, it reacts to a high frequency more sensitively. This excessive frequency in the chip results in more power consumption, which leads to lowering the performance [2]. This significant power consumption in microprocessors becomes a real challenge to increase performance. Therefore, engineers had to find other solutions to keep following Moore's law further, in the end, to advance computer performance. For this reason, the strategy for improving performance changed, and designers started to increase the number of cores in the chips, as we can observe in Figure 1 since 2005.

This evolution is possible due to the continous grow of transistor counts in the chips, allowing to increase the core number without modification of the frequency, leading to improve the global chips performance. Ultimately, this new trend led to Multiprocessor System-on-Chip (MPSoC) era to improve Single thread performance.

The recent studies show that SoC includes a large number of Intellectual Properties (IPs) such as cores, memories and hardware accelerators. However, the increase in number of IPs provoked a large amount of data transfers due to pipeline and multi parallelization. As a result, conventional communication links, such as buses and point to-point links, cannot ensure efficient communications. To address this problem, Network-on-Chip (NoC)

was suggested as a scalable solution to manage multi core communications.

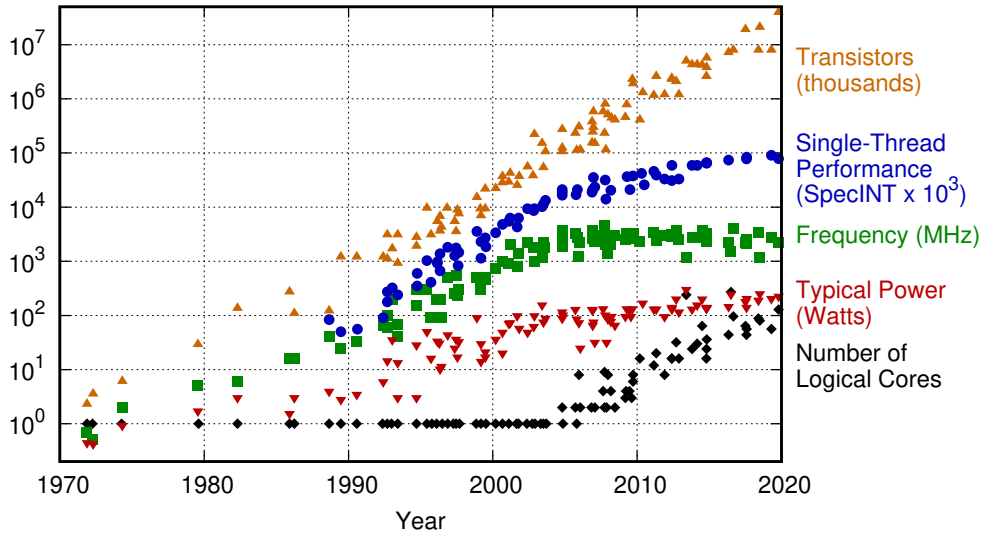


Figure 1 – 48 years of Microprocessor Trend data based on [3].

Network on chip

Nowadays, MPSoCs are evolving towards the integration of hundreds of cores on a single chip [4]. More cores in the chip enables to the parallel execution of applications but increases the needs of communications between Intellectual Property (IPs). The early MPSoCs interconnects were designed in a bus or crossbar-based communication architecture [5]. The traditional bus architecture was built with dedicated point-to-point connections, with one wire dedicated to each Intellectual Property (IPs). The bus architecture could be employed, at first, due to its straightforward architecture allowing cost-effective and simple communication system. However, bus architecture performs effectively only when the number of cores in an MPSoC is low [6]. With an increase in SoC complexity dealing with the increasing number of cores, bus communication becomes a performance bottleneck, leading to problems like scalability, increased power consumption, wire delay, and increased verification cost. These constraints motivated engineers to define alternative solutions that can provide a scalable architecture.

New concepts for interconnects came from existing networking solutions, specifically the internet. In the internet, vast number of computers are connects and communicate to share information. Instead of connecting IPs in a link of bus or point-to-point connection,

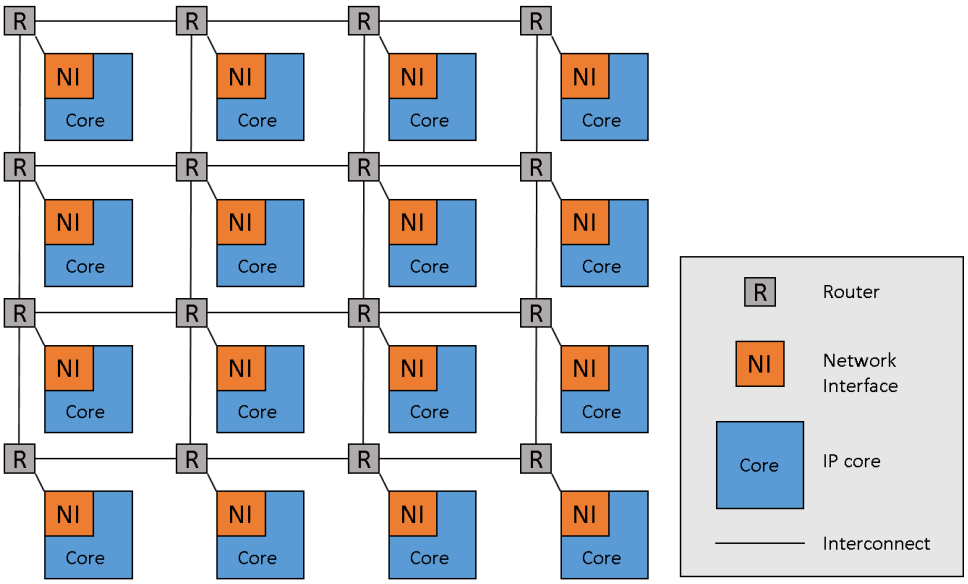


Figure 2 – NoC architecture.

IPs are connected and communicate throught network. This new paradigm is defined as networks on-chip (NoC) architecture. This solution can be seen as a miniature internet, and several existing network features and technology has been employed, such as routers, packet switching, and topologies. Therefore, more specific network techniques have migrated to NoC architecture, such as routing protocols, flow control, switching, arbitration and buffering [7, 8, 9, 10, 11].

Figure 2 illustrates very simple NoC architecture with its 4 main components: i) Routers, ii) Network Interface (NI), iii) Intellectual Property (IP), iv) interconnections. First, to handle the messages (packets) in the network, the router acts in the main role and supports network protocols and flow control. The routers are connected to each other with interconnects. The IPs in NoC can be various units such as memories, cores, accelerators, etc. Moreover, these IPs are connected through Network Interface (NI) to reach out to the network and send their messages/packets to other IPs.

Thanks to this simple implementation and management of several IPs in the network, NoC could be rapidly studied and could cope with increasing architecture size: scalability, resource reuse, performance, reduced costs, and design risk. Therefore, NoC became one of the praised solutions for the complex and large-scale SoCs that required a scalable interconnection architecture and communication system.

Importance of interconnect

As we could see in the previous section, current research and technology trends show that future MPSoC will embed hundreds or even thousands of processing elements. Moreover, as well as hardware system progresses, nowadays, applications like High-performance computing (HPC), Artificial intelligence (AI), scientific computing, video/image processing, data mining, and Big data applications are becoming more complex and sophisticated. This results in new workloads, such as pattern recognition, speech, video and text processing, speech and facial recognition, deep learning, and machine learning. These large-scale applications require high performance on a hardware system with a large amount of computation cpability and data exchange communications. To handle this a large amount of data and computation, communication models in multi-core architectures have been evolved by task parallelization and pipelining. These led to tremendous data packets movement in NoC by data aggregation, communication-intensive operations, and distributed computing. These communication overheads lead to overall system design concern for high-performance interconnects for high bandwidth communication.

This evolution leads to significant challenges for the design of efficient interconnect due to the ever-growing data exchange between processors [12]. The application requirement on hardware performance can only be achieved when architecture and technology developments provide sufficient chip-to-chip and on-chip communication performance to these massively parallel algorithms. In this context, the importance of on-chip interconnect design has become very important in terms of energy efficiency and link bandwidth [13]. However, conventional NoCs utilizing electrical interconnects 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 [13]. As the number of core increases in ENoC, the challenges of power consumption trend shift from computation bounding to data transfer bounding. This significant increase in the power dissipation of electrical interconnect becomes critical and leads to a power-hungry communication link. This power overhead requires alternate future interconnect solutions that can grant high bandwidth, low latency, scalability, and energy efficiency [14, 15].

The recent studies promote Optical interconnect is an attractive solution to address scalability issues in on-chip communication. The optical interconnect can offer low-power consumption, low latency, and high data rate and bandwidth compare to conventional electrical interconnect.

Optical interconnect NoC

Nanophotonic interconnects, also called Optical NoC (ONoC), are based on optical signals for communications in NoC. Optical signals can propagate near speed-of-light in waveguides, and this feature can be a promising alternative solution to replace ENoC by providing large bandwidth and reduced latency.

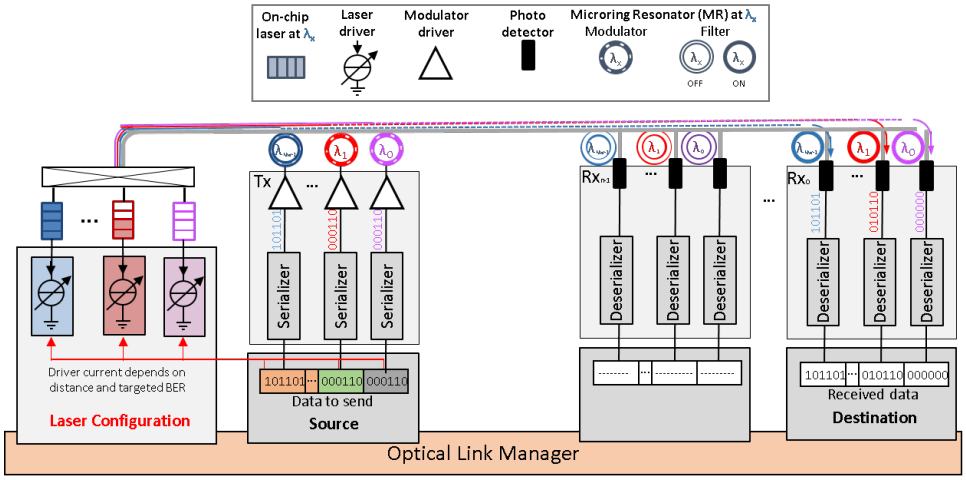


Figure 3 – Nanophotonic interconnect NoC Optical (ONoC)

Figure 3 illustrates general ONoC architecture, which implements several photonic devices such as photonic waveguides, multi-wavelength laser sources, along with microring resonators (MRRs) as modulators and photodetectors. The laser sources (either off-chip or on-chip) generate light signals at specific wavelengths at the transmission side. The optical signals are modulated through wavelength-specific Microring modulators by on-off keying (OOK) modulation to encode the information from the electrical source node. The electrical signal is generally provided as parallel bits which are serialized as a sequence of 1 and 0 to control the OOK modulation. The signals can be modulated into one or several wavelengths to reduce the transfer time. The modulated signals propagate in a shared transmission medium waveguide through multiple wavelengths with wavelength-division-multiplexed (WDM). The WDM technique enables the transmission of several signals in parallel. The transmission of each wavelength serves as a carrier for a data signal. Then, the optical signal propagates in the waveguide until the destination. When they reache the destination, the optical signals are filtered out via a serie of microring (MR) filters, which extract the targeted wavelengths of light and drive them toward

photodetectors. The photodetectors allow to convert optical signals intop electrical ones to forward them to the associated nodes (e.g., processing cores) on the chip. Each node in the ONoC communicates to multiple other nodes through Signal Writer and Multiple Reader (SWMR) communication model with WDM-enabled photonic waveguides.

Challenges of ONoC

Silicon photonics interconnection can offer advantages such as low transmission latency and high bandwidth. However, their implementations remain challenging due to several reasons;

- First, the lasers, which are key device in ONoC, suffer of low efficiency. Indeed, the laser power consumption is mainly due to the low optical gain, which needs to high electrical current. Therefore, the laser power dominates overall power in optical interconnect.

- The emitted optical signals suffer losses as they propagate through waveguides. To ensure that the optical signal at destination has sufficient power for error free transmission, it requires a high laser power to compensate such losses, so that the signal can be received at the destination node with sufficient power to enable error-free recovery of the transmitted data.

Therefore, in this thesis we focus on improving the energy efficiency of the photonic interconnection by adapting approximate computing in optical interconnect to improve efficiency of the laser power consumption, taking into account the distance losses and the targeted bit error rate (BER) for the communications. In addition, the trade-off between energy efficiency and performance communication is also explored.

Contributions

This work improves the optical interconnect based on realistic and scalable laser power management. First, we propose improved approximate communication, and we investigate bit approximation and bit truncation to reduce the NoC power consumption. We also propsoe to associate of these two techniques for each data communication to yet improve the

global system performance. This allows further design exploration and achieving a point closest to the optimal one in terms of power consumption. Thus our proposal provides more power saving with tunable output error. This optimization of approximate configuration makes optical communication more robust and energy-efficient while adapting the error tolerance of the application. Secondly, we propose laser power management based on a technique taking into account the communication distance. Our proposal is based on the communications classification by their distances between source and destination cores. However, unlike [16], only two distances are considered: short and long-distance ranges. This allows laser power configuration to be more suitable and realistic for various architectures in terms of design effort and energy efficiency.

Approximate communication in ONoC

We propose an approximate nanophotonics interconnect. Our proposal is applied to transmit Floating-Point (FP) numbers in Single Writer Multiple Reader (SWMR) optical channels. It relies on the transmission of the least significant bits (LSB) using low-power optical signals, at the cost of higher error rate. These approximate optical signals allow a drastic reduction in the laser power consumption. In parallel, to ensure the communication accuracy on most significant bits (MSBs), the laser power levels are remained using high power signals. Moreover, this 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. This homogeneous multi level transmission allows exploring robustness and energy efficiency trade-offs for on-chip optical interconnects. Simulations results demonstrate that, compared to an interconnect involving only robust communications, approximations in the optical transmission lead to up to 42% laser power reduction for Streamcluster application with a limited degradation at the application level.

Distance-aware in ONoC

We propose a distance aware approximate nanophotonic interconnect to reduce power consumption of the NoC supporting the on-chip communications. As fine-grain distance-aware management of communications is too costly, we propose a low overhead distance-aware technique based on 2 distance classes defined through a threshold. Instead of analyzing each source and destination losses to configure input laser power level, our proposal

considers only two distances of communication: short and long-distance ranges. The laser power level for long distance is defined with farthest destination from source to target a given BER. The short distance is defined with one of the intermediate destination. Based on this classification, laser power levels are adapted according to the distance and targeted BER that application requires for error tolerance. *short*/*long* distance configuration might not be a most optimum in terms of power consumption, however, it is realistic design for laser power level configuration and laser driver complexity. The results of evaluation for the proposal show drastic laser power reduction of 53% for example for Streamcluster application at the cost of less than 8% of errors. Finally, we show that our solution is scalable and outperforms state of the art solution as follows: 10% reduction in the total energy consumption, $35x$ reduction in the laser driver size and $10x$ reduction in the laser controller size.

Thesis outline

The rest of the manuscript is organized as follows:

- **Chapter 1: Related work** presents state of art of the Silicon photonics interconnect in NoC with detail of its components. In the following, the concept of approximate computing is also introduced to understand error tolerant applications and various techniques of approximate computing. For the last, state-of-art of approximate interconnects is presented.
- **Chapter 2: Approximate communication in Optical Interconnects** the idea of our first proposal in detailed with optical link hardware setup that we have utilized. It also contains simulation setup like benchmark applications and study of each application characteristic. For the last, the results of quality impact and power saving are stated.
- **Chapter 3: Distance-aware approximate optical interconnect** This chapter includes the study of laser driver complexity to compare with state-of-art and introduces multi laser power levels to increase the energy efficiency in NoC. Then, the results of our proposal are presented with comparison of state-of-art.
- **Chapter 4: Conclusion & perspectives** This chapter abstracts our contributions and presents a set of perspectives for future works.

Chapter 1

Related work

In this chapter, an overview of two domains is presented. First is introducing silicon photonics and its implementation as an NoC interconnect. Then, we address fundamental of optical interconnect with a simple link model and illustrate how it operates and its limitations, in particular, cross talk problem coming from the WDM technique. Secondly, we present the concept of approximate computing, and we explain how this concept has been extended into approximate communications.

1.1 Silicon Photonic interconnect NoC

Optical interconnect enabled by silicon photonics is praised as a promising solution for the communication bottleneck in NoC. This communication phenomenon in the data-centric era emphasizes the prominence of interconnects due to the rapid expansion of the data volume of communication. In this section, the background and property of optical interconnects are briefly introduced, and we identify several challenges to motivate our work and the scope of this dissertation.

1.1.1 History and introduction of optical interconnects

As Optical interconnects utilize optical signals as an signal carrier, it brings several benefits, such as higher bandwidth capacity, low propagation delay, and greater tolerance of electromagnetic interference compared to conventional electrical interconnects [17, 18]. Therefore, optical links are broadly explored as a substitution for electrical links. Over the past decades, this type of optical interconnect has evolved from long-distance communication to short-distance communication, as illustrated in Figure 1.1. It shows bandwidth and the footprint of interconnect modules. In this figure, the bandwidth and footprint of optical solutions are highlighted and express the opportunity of this technology for different application domains.

Conventionally, Optical interconnect modules were used for applications like LANs (local area networks), WANs (wide area networks), inter-rack, and on-board interconnects. As we can see in Figure 1.1, in LAN and WAN networks have $1Gbit/s/cm^2$ for long distance. When this optical communication comes to board-to-board communication, it can offer $100Gbit/s/cm^2$. These interconnect are based on discrete photonic devices made of compound semiconductors and photonic technology which has utilized large components. Therefore, the optical interconnect technologies could not directly be used for photonic integration within chip-to-chip communication.

The chip-to-chip interconnects evolve with Complementary metal-oxide-semiconductor (CMOS) circuits, and it requires very high-end technology due to its device size, design scheme, and materials. To solve these problems, silicon photonics technology was proposed through close electronic-photonic integration for high bandwidth communication and high energy efficiency. The silicon photonic system can achieve high-density photonic device integration and compatibility with CMOS electronic technology that targets $10Tbit/s/cm^2$ and even higher bandwidth.

1.1.2 Components of silicon photonic interconnect

Silicon photonic utilized the nano size technology that enables to be compatible with CMOS technology. In silicon photonic interconnect, there are various scenarii of component options that are possible. Figure 1.2 presents a very elementary and high level optical link with optical components blocks in a very general point of view.

For a communication system, the communication link contains three essential block parts: **A transmitter block** that gets information from upper layer and converts this information into physical signal, **a transmission medium block** that conveys the signal, and **a receiver block** that receives the signal and converts into information for upper layer.

Transmitter block it contains components like i) laser, which is source carrier of optical signal as light wave, and ii) modulator, which encodes the information data from electrical domain into bits ("1" or "0").

— **Laser source**: The main role of the laser is to generate light signal. In order to have optimum optical link, we need to have right amount of laser power, so we do not waste the power or have bad communication quality. By injecting and adjusting the

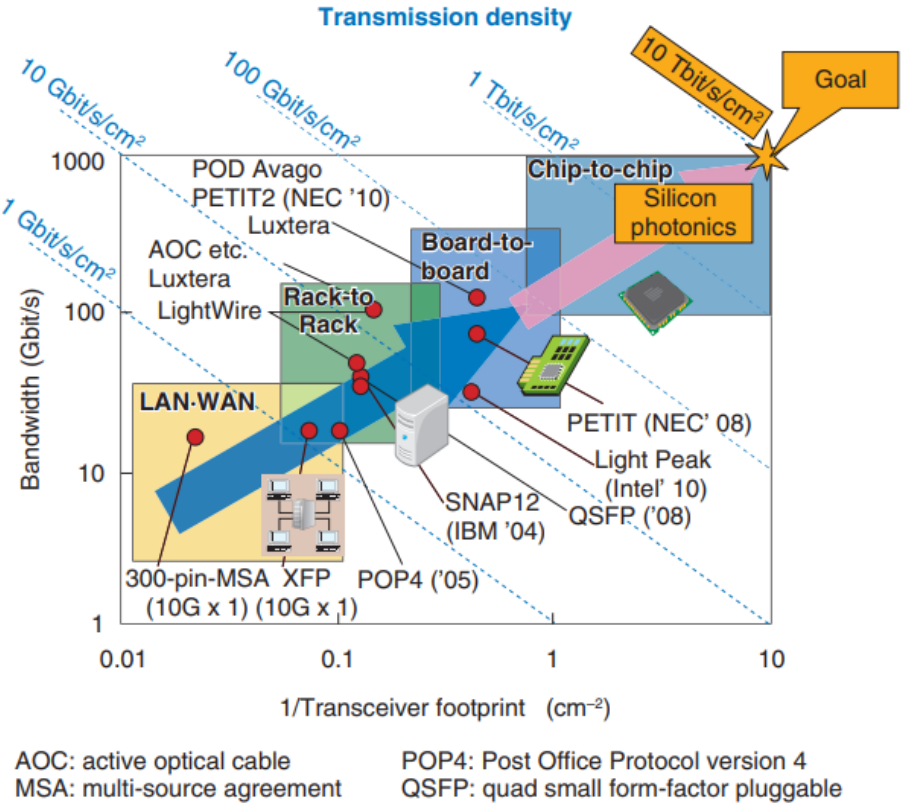


Figure 1.1 – Evolution of optical interconnect [19].

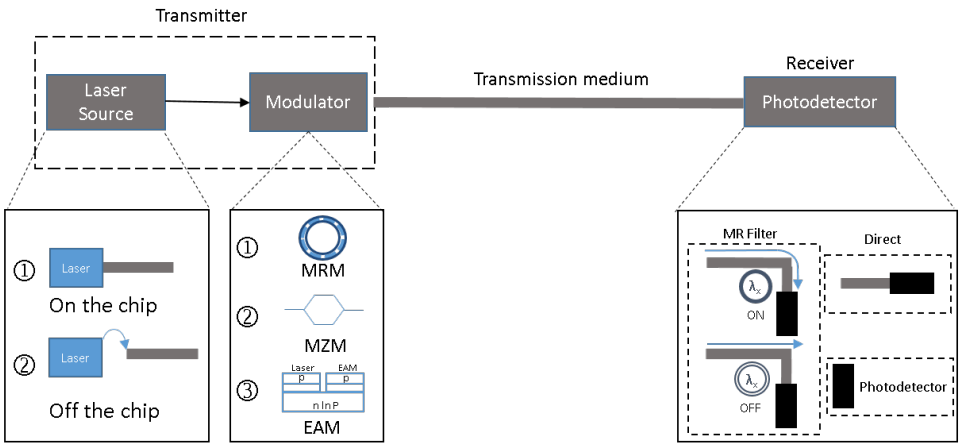


Figure 1.2 – Optical communication Block.

laser power, we can configure the power just necessary on laser. Two main locations can be found for the optical signal generation: **on chip** and **off chip** laser. Previously there are several studies that compare these two solutions with comprehensive analysis and discussed the merits and flaw of these two lasers [20, 21, 22, 23]. Compared to on-chip laser, an off-chip laser provides higher efficiency on light emitting and better temperature tolerance but it suffers from a large coupling loss, a large area and packaging cost. On the other hand, on-chip laser offers better integration in terms of area and design constrain. Therefore, when it goes to large scale NoC, on-chip laser performs better in scalability and flexibility for implementation of optical link. Therefore, considering these advantages, we consider on-chip laser implementation as an hypothesis for our works.

— **Modulator**: there are several modulation methods for the light signals: microring modulator (MRM) [24], Mach-Zehnder modulator (MZM) [25, 26] and electro-absorption modulator (EAM) [27]. Compare to other modulation model, Microring modulator provides more energy efficient due to its compact size property and easy control on optical signal. Moreover, microring modulator offers the inherent wavelength selectivity and allows wavelength division multiplexing (WDM) which can be evolve to higher dense wavelength division multiplexing (DWDM). In this thesis, we utilize MRM as the primary device for our high-speed communication link due to its energy efficiency and wavelength selectivity. The detail role of MR will be stated in section 1.1.3.

After modulation, the signal is injected into the transmission medium block to send the information.

In the block of **transmission medium**, the optical signal propagates from transmitter to receiver. Optical interconnects utilize the waveguides for signal propagation. Unlike electrical copper wire from electrical interconnect, in optical interconnect, it is possible to transmit several optical signals in a single waveguide by applying wavelength division multiplexing (WDM). WDM allows multiple signaling by simultaneously transmits signals in one transmission medium. Furthermore, several wavelength-specific lasers are allocated to support the multiple wavelengths emissions. In WDM, several wavelength-specific MRMs are placed on a shared waveguide,which allows to modulate wavelengths

individually. Moreover, the WDM signals at specific wavelengths can be extracted independently from the waveguide at receiver side. In order to avoid interference and crosstalk noise with another wavelength channel, MR should have a small ring diameter to create a large FSR, offering important wavelength spacing [28, 29]. WDM can effectively improve throughput to Gb/s and Tb/s for link bandwidth, and ease area overhead compared to single wavelength channel. However, several optical signals in one waveguide can provoke crosstalk noise that can critically distort the information in signals. The detail of crosstalk noise in WDM is detailed in 1.1.4

At the end of transmission medium, receiver block wait for the optical signal to be arrived and extract the optical signal.

At receiver block, several components and scenarii are possible, such as a photodetector and/or microring filters to convert optical into electrical one and extract the signals from waveguide. In the case of using array of mirroring filters, mirroring can extract the light signals selectively at specific wavelength. Utilizing MR filters allows to have several receivers (destination cores) in single waveguide.To reach a given destination, all the intermediate IPs set their MR in OFF-states, so optical signal continue to propagate until the right destination. When transmitted signals arrive at destination, MR filters are set to ON-state and extract the signals to drop port waveguide. At the end of drop port, photodetector is integrated to convert optical signals into electrical ones. In MR filter receiver can have multiple receiver but as MRs are arrayed at the receiver side, we also have to consider other loss which are introduced by several MRs, such as MR through loss and crosstalk noise between the wavelengths at the drop of microring filters.

In case of direct receiver which uses photodetector at the end of waveguide can have only one destination in one waveguide. This direct receiver offers simple implementation and control on network interface as it does not require MR filter controls.

For both cases, photodetector is used. Typically photodetector perform with a photodiode to sense the received optical signal and produce an input current. This current which is induced by the optical signal, is called photocurrent. In photodiode, photocurrent is converted to voltage and amplified sufficiently for data resolution. The photodetector performance is determined by its sensitivity, which is associated with the maximum data rate and amount of tolerable channel loss.

1.1.3 Physical property of optical components

In previous subsection, we saw several components of optical link. In this subsection, we detail the component's physical properties and how it operates. When a light signal propagates into the waveguide, there are several intrinsic and indirect loss from optical components. To design optimum and efficient optical link, there is a need to understand the devices.

At the beginning of optical link, the laser is located. It emits the light signal into the waveguide and acts as main role of light source. Our study is based on utilization of **on-chip laser** which provides flexibility and straightforward power level configuration accessibility. In particular, vertical Cavity Surface Emitting Laser (VCSEL) is an attractive on-chip optical laser source for short-distance optical interconnects. It offers compact size, the short rise time of VCSEL current, and high switching frequency [30, 31]. VCSEL laser emits an optical signal from its surface to the wafer. Additionally, multiple VCSEL lasers can be positioned in an array due to their direct emission into a wafer. Moreover, VCSEL is able to target several different wavelengths, which is suitable for WDM [32]. In the study of [33], devices can emit at around 850nm at large substrate with high speed-optimized VCSEL technology. In [34], the proposed VCSEL can support the variation and thermal analyses and emit optical light at BER under 10^{-9} with laser efficiency equal to 33%.

- Microring Resonator: MR can be used in several place in optical link i) as a modulator at transmission part and ii) as MR filter at receiver part. MR can offer the inherent wavelength selectivity and allows wavelength division multiplexing (WDM) which can be evolve to higher dense wavelength division multiplexing (DWDM). Therefore, in this thesis, we utilize MR as the primary device for our high-speed link due to its energy efficiency and wavelength selectivity. Figure 1.3 presents MR in NoC in very general high level point of view. Figure 1.3 a) illustrates MR with an input and a drop waveguide. There are two possible scenarii; first, when MR is activated, the light signal from the input waveguide is redirected to the drop waveguide through MR. Secondly, when MR is deactivated, the light signal continue to propagates to through port waveguide from input port. Figure 1.3 b) presents p and n doping on MR to understand electro-optic MR for on-off keying (OOK) modulation. When MR is activated the current is induced at MR to p and n doping and it extracts optical signal from the waveguide then the waveguide silicon index changes. This leads to resonance wavelength shift, achieving on-off keying

modulation. Figure 1.3 c) shows wavelength shift at targeted resonant wavelength λ. When MR is activated, there is not only wavelength shift, but also there is the phase response of MR.

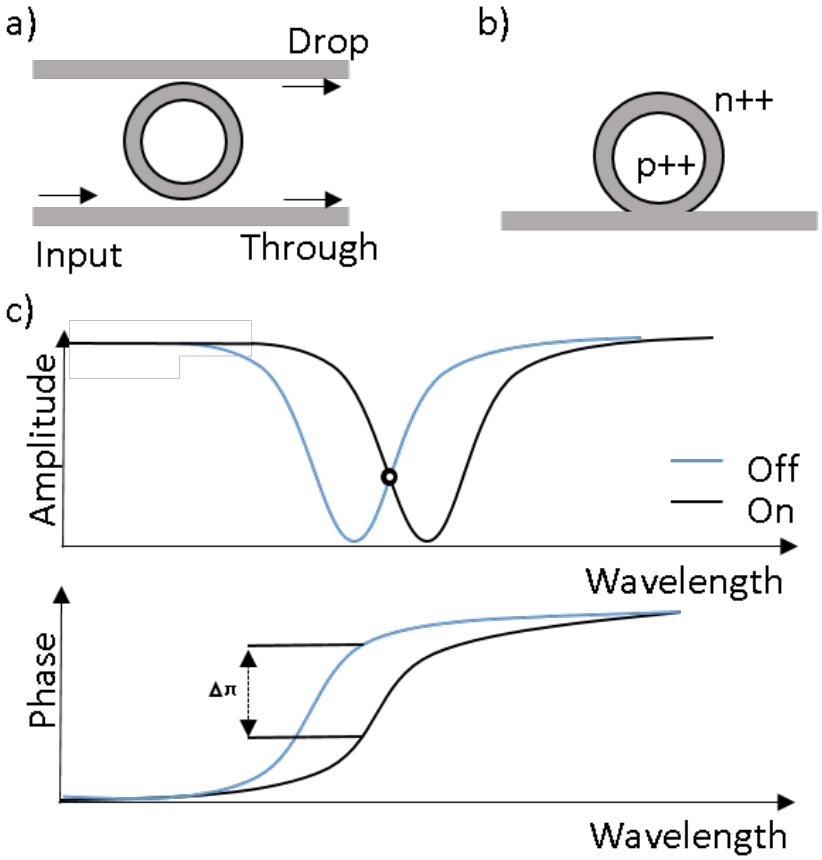


Figure 1.3 – Microring modulator a) General implementation of MRM; b) P and N doping in MRM c) states of MRM considering the drop out and pass through mode.

Now let's take look how MR actually operates.

First, **as a modulator**: in our study Microring based modulator is used for optical signal modulation [35]. Figure 1.4 presents the MR modulator that consists of a microring resonator coupled with a waveguide. By tuning the MR index properly the resonance in wavelength is shifted, which induces a strong modulation of the transmitted signal. The MR index is modulated electrically by injecting electrons and holes using a p-i-n junction embedded in the microring resonator. By turning on MR, the refractive index is adjusted by applying the bias voltage on the modulator. The equation 1.1 describes the optical

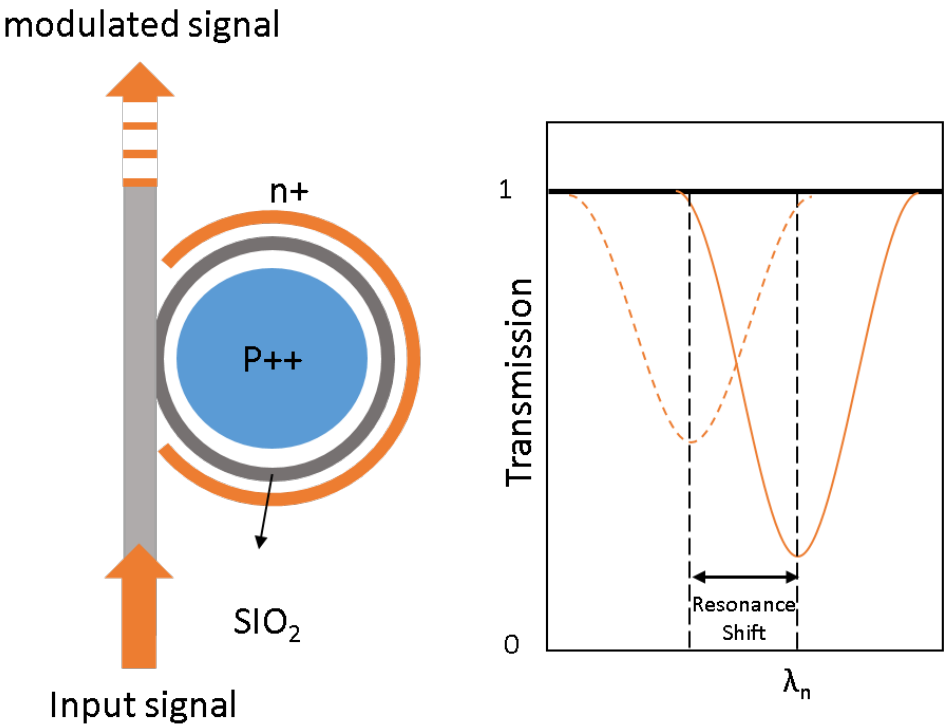


Figure 1.4 – Microring modulator.

response of a microring, where L is the optical length of the microring, α is the loss in the cavity, and t is the transmission coefficient.

$$\frac{E_{out}}{E_{in}} = \frac{-\alpha + te^{-2\pi Li/\lambda}}{-\alpha t + te^{-2\pi Li/\lambda}} \tag{1.1}$$

Second, **MR at receiver**: MR also can be used at the receiver side as a drop filter by targeting specific wavelength. Depending on the communication channel model, the bank of MRs can used to extract specific wavelengths. It extracts the light signal from waveguide at targeted wavelength and redirects the signal to drop port which leads optical signals to photodetector. Figure 1.5 a) presents microring resonator (MR) with an input and a drop waveguide. Figure 1.5 b) illustrates power transmission responses of waveguides. When MR filter is activated, input light power at through port is redirected to drop port transmission power. To understand the power rate between input power from waveguide and drop port, Figure 1.5 a) illustrates MR resonator with its coefficient. Figure 1.5 a)

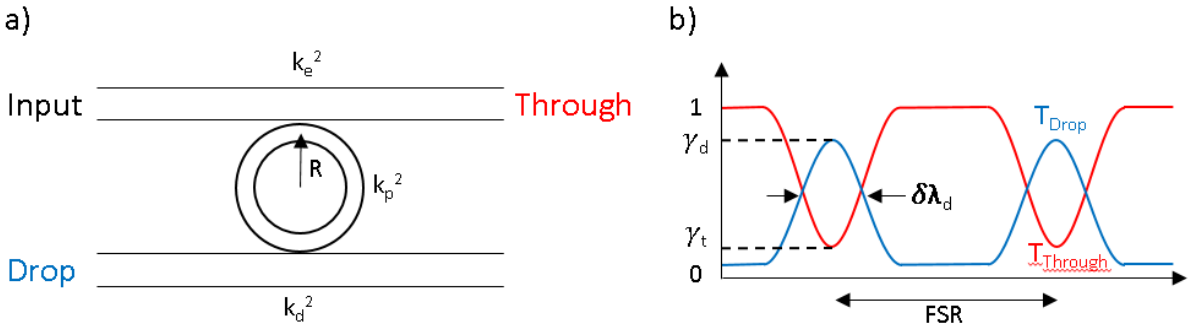


Figure 1.5 – Microring modulator.

can be presented in equation 1.2, with the coefficient of each fraction of the input optical power, fraction of light filtered in drop port and intrinsic power loss in MR respectively k_e^2, k_d^2 and k_p^2. FSR is the free spectral range (in wavelength span) in the resonator. δ is $\lambda_M R/2Q$

$$\frac{P_{drop}^{\lambda i}}{P_{in}^{\lambda i}} = \left[\frac{2k_e k_d}{k_e^2 + k_d^2 + k_p^2}\right]^2 \frac{\delta^2}{(\lambda_i - \lambda_{MR}) + \delta^2} \tag{1.2}$$

Figure 1.5 b) presents the transmission power between drop port and throughput waveguide. The minimum power transmission in the through port is γ_t at the resonance wavelength λ_0. On the other hand, at resonance wavelength λ_0, drop port waveguide transmission power reaches at high power γ_d. This leads to add-drop crosstalk $-10 \times log10(\delta\lambda_d \ / \ \delta_t)$ in dB.

- Waveguide : Like copper wire in electrical interconnect, when light signal propagates through the transmission medium, the propagation loss occurs. The loss in the waveguide can be caused by several attribute factors like light scattering, absorption, radiation loss from waveguide bending. Waveguide propagation loss is mainly due to surface scattering from the etched sidewalls. When optical signal propagates, it suffers attenuation from each reflection on the waveguide wall. The free-carrier effect can cause the absorption loss by electrons or holes in the waveguide. Furthermore, the temperature of the waveguide can increase this loss. These losses are intrinsic loss which is a physical characteristic of a waveguide. Minimizing the optical signals and the frequency that signals touches the etched interfaces can effectively reduce the waveguide propagation loss.

In [36], the authors propose to increase the waveguide width and decrease etch depth. In [36], loss shallow-ridge silicon waveguides are demonstrated with an average propagation loss of $0.274 dB/cm$.

Photodetector : the photodetector receives the optical signals which redirected from MR filters. The photodetector converts light photons into electrical current. The absorbed photons make electron–hole pairs in the depletion region. Photodetectors often contain Ge or SiGe due to their compatibility with CMOS process [37, 38]. SiGe photodetector performs with better sensitivity for shorter wavelengths than Ge photodetectors. SiGe is developed for standard CMOS processes as a strain-engineering material to improve the carrier mobility, enabling monolithic integration of silicon photonics without any requirements at CMOS process. The sensitivity of photodetector is the main factor to consider for optical link budget. The sensitivity refers to the least optical power at the destination to achieve targeted BER [39].

1.1.4 Crosstalk noise in WDM

We saw MRs that are used at both transmission and receiver side to realize modulation and multiplexing/demultiplexing of optical signal at designed optical wavelength. Moreover, the optical signals in MR can have multiple channels through out wavelength division multiplexing (WDM) which allows high bandwidth communication. However, in WDM links, channels tend to correlate and imprint negative effects on each other, such as spectral overlap of adjacent channels and polarization modes. These crosstalk effects can be found prominently where multiple MRs are employed to add or drop individual channels to or from a WDM signals (**Inter-channel crosstalk**) and when frequent switches appear between several channels of the same wavelength (**intra-channel crosstalk**). These significant signal distortion might lead to misinterpretation and affect link quality performance for high BER. Figure 1.6 presents crosstalk noise at MR in SWMR communication link using WDM. In this Figure three different wavelengths (λ_0, λ_1 and λ_2) are used in one shared waveguide medium. We assume the wavelength λ_1 as a selected communication channel for both modulation and receiving signal. Figure 1.6 a) presents crosstalk noise in MR modulator. When MR modulates at specific wavelength (λ_1 in Figure1.6) the other wavelengths λ_0 and λ_2 are also drained into the MR and these signals provoke the crosstalk noise. Figure 1.6 b) illustrates crosstalk noise in MR filters in active mode. The MR filter out the λ_1 to drop port, but small partial optical power leaks into MR and

induces crosstalk noise. If the channel space between the wavelength is large enough, in drop port, the crosstalk noise which introduced by signals in different wavelengths is very low and can be negligible.

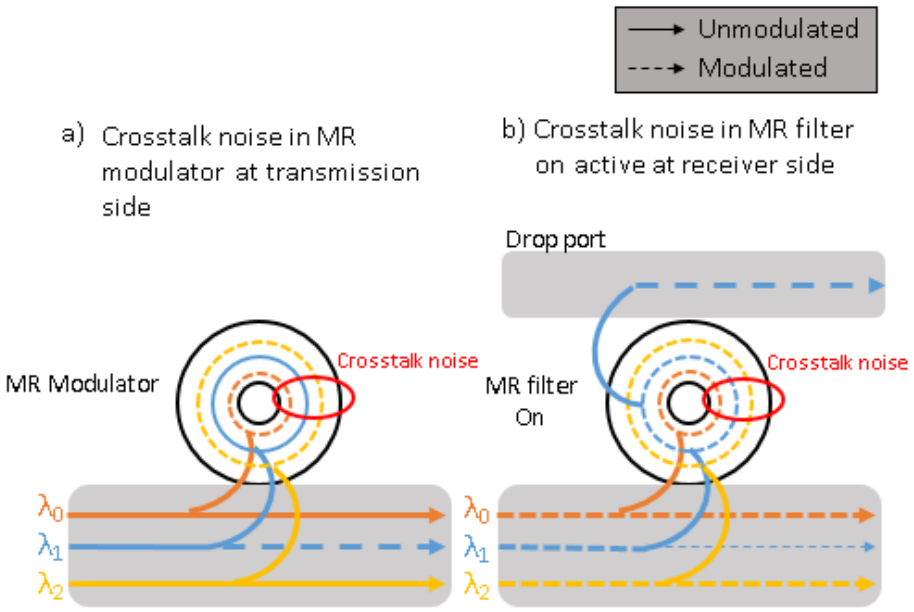


Figure 1.6 – Crosstalk noise in SWMR.

Even though MRs can selectively modulates signal using WDM, the signals that have different wavelength can be affected by MR and provoke crosstalk noise. Moreover, in WDM MR modulators are located in an array, the neighbor MRs provoke the crosstalk noise and suppress the signal power. The power of this distortion can be increased by number of different wavelength in shared medium and the channel space between them. Figure 1.7 illustrates transmission spectrum of bank of receiver resonance wavelengths of MRs with three different wavelengths (λ_0, λ_1 and λ_2). In this Figure, we can see the channel gap and modulation gap for all the wavelengths. The Figure shows two MRs which drop the optical signal at each wavelength λ_0 and λ_1. In the transmission spectrum figure we can see that λ_0 and λ_1 are neighboring wavelength which means the channel gap between them are very close. This near wavelength modulation can lead to high crosstalk noise. To reduce the crosstalk, designing larger channel spacing between the wavelengths are essential. For example, in the figure we can see that wavelength λ_2 has less impact on λ_0 which lead to less crosstalk noise. Indeed, the dense WDM can provide high bandwidth

with multiple channels. However, it can also contain large amount of crosstalk noise and might interfere the signals and distort information.

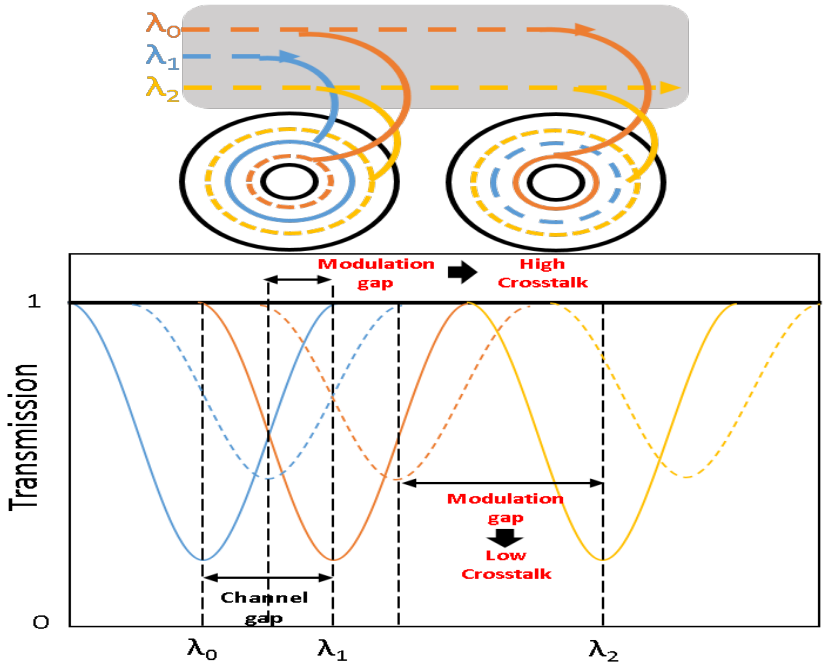


Figure 1.7 – Crosstalk noise between MR modulators.

Figure 1.8 shows crosstalk noise by the number of wavelength in same spectrum range. The first and last nodes(MR) have very low crosstalk noise due to 1 less adjacent neighbor on side. Other intermediate nodes have more MR modulator crosstalk noise. Moreover, as the number of wavelengths λ increases in the same spectrum range, the crosstalk noise increases due to shorter channel gap. Therefore, in WMD, the channel spacing is one of the main design concern to avoid high crosstalk noise. In our work, we set the FSR and channel spacing at 8nm and 1nm (120 GHz) respectively with 8 channels (wavelengths) to avoid crosstalk noise [41].

1.1.5 Optical link and ONoC architecture

As we saw from previous subsections, there are many options to define an optical link configuration. Based on our study, we choose the specific implementation of the optical link model as Figure 1.9.

Without lack of generality, as presented in Figure 1.9, conceptual diagram of optical

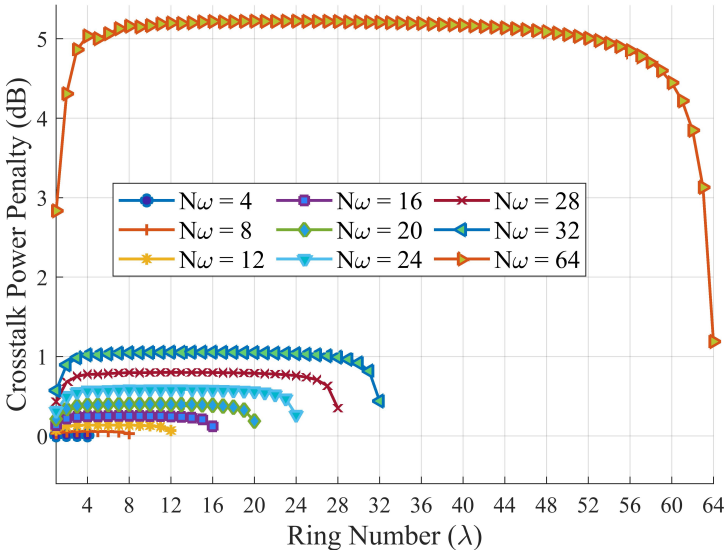


Figure 1.8 – Crosstalk noise by number of wavelengths [40].

link communication to show how information is transmitted. In Figure 1.9, three IPs connected with silicon photonic interconnect, which is consists of i) on-chip laser and MR modulator at the transmitter side, ii) SiGe Waveguide as transmission medium iii) at the receiver side, microring filter and photodetector are located. When the transmitter side receives information from IP, the laser controller turns on the laser at each wavelength λ_{N-1} by WDM communication, at moderate input power. Then, MR modulator encodes the incoming light signal with on-off-keying (OOK) at selected laser wavelength λ_{N-1}. The modulated optical signal propagates through the waveguide and there might be the another passing IPs in SWMR (Rx1 in Figure 1.9). The signals propagate in waveguide and reach at the destination. At the receiver, the bank of microring filters placed to filter out the received optical signal at a specific wavelength. The drop port of the optical filter is connected to a photodetector to convert optical signals into electrical signals. Moreover,

besides the optical component model, we also have to consider the communication link model to connect several IPs in one shared transmission medium [42]. There are several possibilities that have different pros and cons. A very simple communication model is the Single Writer and Single Reader (SWSR), which links just one transmitter (writer) and one receiver (reader). This communication model connects each IP at a point-to-point connection, just like copper wire. It performs the best in terms of latency and energy

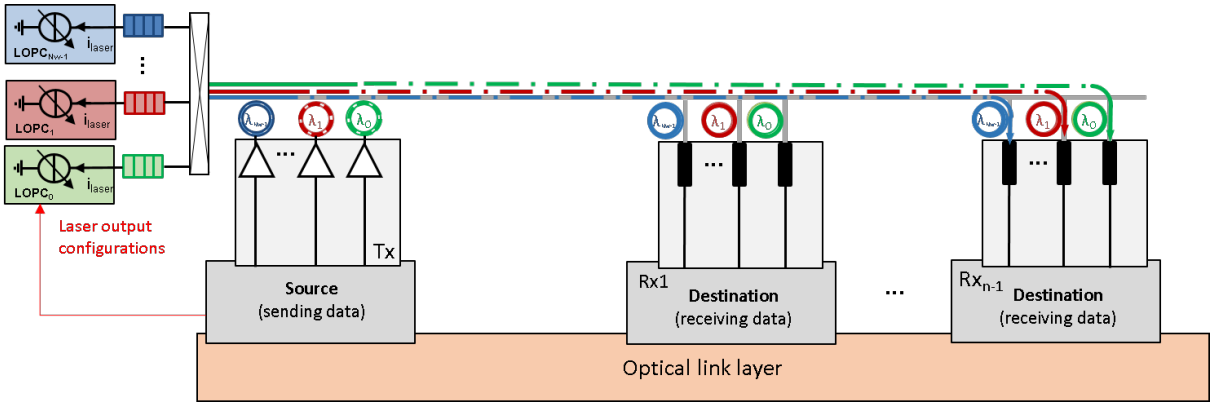


Figure 1.9 – Modeled Optical interconnect Noc (ONoC).

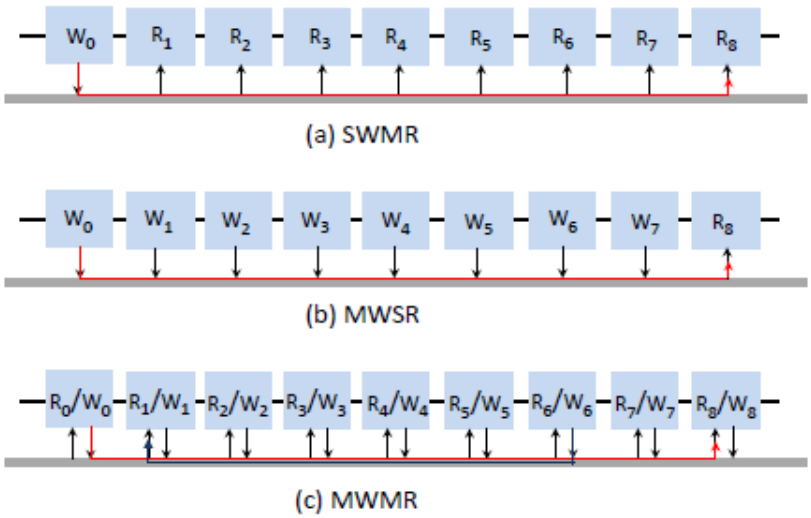


Figure 1.10 – Communication link models a) SWMR b) MWSR c) MWMR.

consumption, as it does not consider link contention and provides direct communication link. However, this communication model is not scalable due to area constrain and density.

Therefore, there are communication models that utilize multiple signaling in shared optical links to increase the efficiency of area constrain through WDM. These models are the Single Writer and Multiple Reader (SWMR), Multiple Writer and Single Reader (MWSR), and Multiple Writer and Multiple Reader (MWMR). MWSR communication model employs a local arbitration at the writer side before sending data to reserve the link to a single reader as proposed in Corona [43]. The Corona offers a high bandwidth DWDM crossbar, enabling a cache coherent design with near-uniform on-stack and memory communication latency. MWMR model optimizes the link utilization and scalability in terms of area constrain by fully sharing the channel with all the connected IPs as both writers and readers. The link contention between writers and readers is regulated by arbitration strategy [44]. However, in MWMR, the optical signal undergoes many optical devices, which leads to high optical losses. These losses from various optical devices result in high optical input laser power to meat targeted BER and suffer from power efficiency. Moreover, there are more complex model, like Dragonfly [45, 46] which combine communications models. Dragonfly architecture supports multicast and broadcast traffic in a simple and efficient way with its low diameter. The proposed architecture uses group of high-radix routers as a virtual router to increase the effective radix of the network. However, by creating groups, it relies on indirect adaptive routing and multiple global channel traversals, leading to additional complexity and increased hop count.

In this thesis, we consider SWMR as communication model as it avoids global arbitration between writers and minimizes the complexity of communication control [47, 48]. This model might not be a most energy efficient and scalable model, but it provides low-latency and high-bandwidth.

1.1.6 Energy efficient challenges in optical interconnect

Even though optical interconnect can offer several benefits like high bandwidth and low latency to NoC, it suffers from high power overhead related to low-efficiency lasers at its integration. Therefore, to overcome this problem, several attempts have been proposed to improve optical interconnect power efficiency. The energy efficiency of ONoC is addressed in several aspects. At the circuit level, in [49], the authors proposed to reduce the power loss, which comes from components such as by reducing the number of Micro Ring resonators (MR) and waveguide crossings. At the communications level, communi-

cation scheduling can improve energy efficiency thanks to less parallel communications leading to crosstalk and energy penalties [50, 51].

In [49, 52], an all optic interconnect topology suitable for managing optical communications is proposed. It is based on routers that reduce number of wavelengths and other optical components to reduce power consumption. In [53], the authors propose photonic on-chip network architecture composed of several sub-nets. This allows high performance and low latency by sharing optical channel between sub-nets and by dynamically schedule them without degrading throughput. Error correction code (ECC) is introduced in Optical NoC to reduce the power consumption [54]. The transmitter includes an ECC and the receiver detects faulty bits and can then correct them. Since additional faulty bits are acceptable using ECC, the optical power emitted by the laser source can be significantly reduced. In [55], the authors propose crosstalk aware wavelength scheduling. The correlation of application execution time and the energy consumption overhead is strategically analyzed to reduce the crosstalk induced by WDM (Wavelengths Division Multiplexing). In [50], the contribution focuses on dynamic bandwidth scaling based on network traffic prediction. It dynamically tunes optical channel link according to the needs of bandwidth. By turning off the unused network link, it optimizes channel utilization and reduces optical power. The technique consists on configuring the optical light intensity based on the number of paths used in a tree based optical NoC. In [56, 57, 58], based on network load of required by applications, the authors propose a network management. In [56], the authors propose to increase the laser power efficiency through the management of online activating and deactivating of L2 cache and the associated optical links. In [57], the off-chip laser power is dynamically managed at runtime to adapt the network to application characteristics, hence reducing laser power consumption. the work published in [59] proposes heterogeneous multicore architecture based on CPU-GPU cores. It reduces static optical power by predicting network load with link reservation. In [60], the authors propose to runtime power management using token based time division multiplexing. It switches ON/OFF laser source according to NoC bandwidth needs. In [61], the authors propose to balance the power consumption of lasers and the average packet latency on the ONoC. This is done by optimization of the photonic link parameters at design time, and by on-line trade-off management of power consumption and performances by limiting the cross-talk effects involved by the Microring Resonators.

1.2 Approximate computing

1.2.1 Introduction

Since the beginning of the computing system history, accuracy and reliability have been a fundamental concern for computer performance and hardware design. The computer user expects that the computer system is reliable for accurate computing and no error in communication and data management. Therefore, to prevent error in the computing system, several works are considered, such as the error detection and error correction techniques are proposed: checksum, forward error correction, and parity bit [62, 63, 64]. However, this desire for perfection might be abstract and the cost of these additional techniques is too high and sometimes not affordable.

Meanwhile, many applications do not intrinsically require perfectly accurate computation and are tolerant to errors. For example, applications like computer vision, media processing, machine learning, and sensor data analysis can cope with inaccurate information in their design [65, 66]. There are several reasons behind this tolerance with imprecision: in computer vision and image media-processing applications, sometimes the quality of output results might not be defined and vary depending on users. The output of these applications is observed by human eyes that contain perceptual limitations. There might not be an exact solution in the machine learning domain and no output result that the program is looking for. The purpose of the application is often just a suggestion to users for the best match, or it might not need to guarantee finding one unique solution. The output result can be good enough to satisfy the users. In an applications that use censored data, the input data already contains the noisy elements, which means the input data are robust to noisy and application can produce approximated output.

Nowadays, large-scale applications such as scientific computing, social media, high-performance computing, and big data applications demand modern systems with significantly high computational and storage. The capacity of hardware systems has far exceeded the available resources to fulfill the application demands. The promising solution for this dilemma is Approximate Computing (AC), which bases on exploiting the system's inherent error resiliency and tolerance to achieve energy efficiency. Instead of performing an exact computation or maintaining accurate quality, Approximate Computing allows prudent approximation or partial offenses in specification to advantage in several aspects like; efficient computation, communication power consumption, memory architecture, and better throughput by drastically reducing the computation or data accuracy quality.

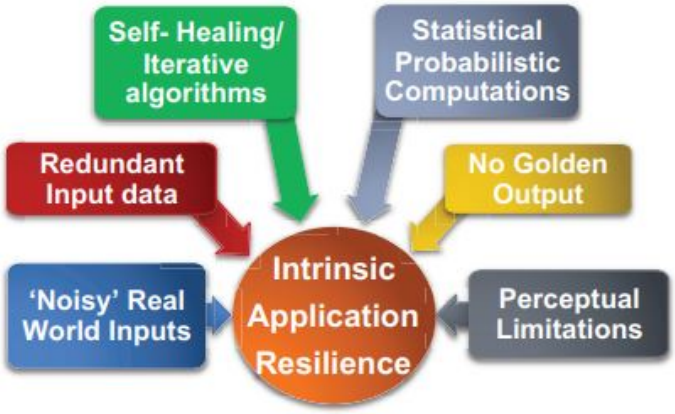


Figure 1.11 – Approximate computing applications.

Depending on where approximate computing is employed, it offers different benefits to the system. Well-known domains that adapt approximate computing are approximate arithmetic circuits:

— **Approximate adder and multipliers**
— **Energy-efficient memory**
— **Embedded sensors** for IoT applications
— **An accelerator** of Neural Network

For instance, approximation allows reducing the hardware complexity of arithmetic circuits such as adder and multiplier [67, 68, 69]. An approximate circuit contributes to the improvement of power consumption and operating speed. It also provides efficient power consumption and more effective memory management [70, 71, 72, 73]. Approximate memory architecture offers data storage with a predefined accuracy level to offer compact data storage and read/write power efficiency of memory. The work of [74, 75, 76] proposes to apply approximate computing on the Internet of Things (IoT) and sensors domains. In IoTs, sensors already contain sensing noise and inherent error-resilience in signal processing and based on redundancy in the input data. These lead in degradation of result quality and inefficient power consumption. Hence, the proposed approximate sensor leverages the approximate computing in digital sensing through low-power approximate circuits and data aggregation algorithm to provide energy efficiency. In [75], the authors propose approximate sensor which leverages the approximate computing in the real-time encoding

scheme that performs iterative threshold and approximation of wavelet coefficients for sparse encoding, thereby reducing the energy and bandwidth. Approximate computing can be applied in Neural Network by considering approximation data for both in computation and memory accesses, thereby achieving more energy savings [77, 78, 79]. The work presented in [78] proposes data pattern matching to reuse approximatly similar data in Neural Networks. It performs approximate computation reuse with relaxed precision at input patterns and designs a reconfigurable architecture to support the approximate computation reuse. It could reduce energy by 30-58% with accuracy degradation of 1-2.5%. In [80], 19-bit approximate floating-point mantissa bits are utilized with an approximate adder to achieve power saving. It not only speed up the computation, but also save the power consumption on memory and communication link.

1.2.2 Floating-point Data representation

In computer arithmetic, the representation of real numbers is one of the major concerns due to their accuracy, reliability, and stability. However, in arithmetic computing, an infinite accuracy and number representation are impossible since the numbers consist of finite bit width. Therefore, the numbers are formed in various ways: singed/unsigned integer, fixed-point and floating-point. This thesis explores the floating-point data presentation, its approximation, and the impact of number representation. To have a number representation as accurate as possible for all the numbers ranging from very small to large numbers is significantly important. To address this problem, the IEEE 754 standard defined [81] floating-point number form and it can normalized in the common base 10 (decimal) by following equation 1.3. The standard floating-point definition consists of three parts: sign, exponent, and mantissa. Moreover, the number range of floating-point can be defined in several binary formats with different bit widths of 32, 64, and 128 bits by varying the mantissa and the exponent bit width. The table 1.1 contains the bit width of various precision of floating-point number and its component bit width; mantissa, bias, exponent bias and max decimal exponent.

$$X = (-1)^S \times (1 + Mantissa) \times 2^{Exponent-Bias} \tag{1.3}$$

As we can see in Figure 1.12, the floating-point number has three components: sign $\pm$ **s**, exponent **e**, and mantissa **m**. The first bit of floating-point is sign **s**, and this indicates

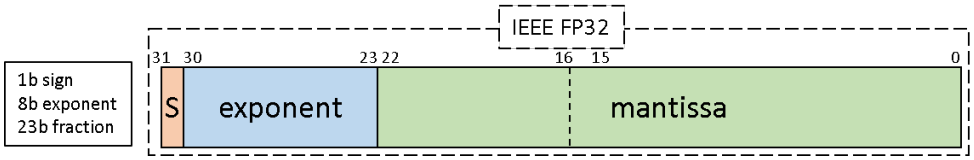


Figure 1.12 – IEEE 754 Single precision Floating-point representation.

Table 1.1 – Various precision of IEEE 754 normalized Floating-point representation

Precision	Mantissa width	Exponent width	Max decimal exponent	Exponent bias
Single precision	24	8	38.23	127
Double precision	53	11	307.95	1023
Quadruple precision	53	11	307.95	1023

the negative or positive of the value with '1' and '0' respectively. The following bits introduce exponent **e** and this represent both positive and negative exponents. These bits are designed as 2^{e-bias} : excess number of bias binary integer added. To normalized these bits, the bias is added to the actual exponent in order to obtained the stored exponent. For example, IEEE single precision floats has bias 127. Thus, to express an exponent of zero, 127 is stored in the exponent field. A stored value of 200 indicates an exponent of $(200 - 127)$, or 73. For these reasons, exponents of -127 (all 0s) and $+128$ (all 1s) are reserved for special numbers indicated in table 1.2. The rest of bits are set up for Mantissa **m**, also known as the significant bits, and it is designed for precision bits of number. It is treated as fractional fixed point binary number. Unlike fix point binary number, it is composed of an implicit bit ('1') and the fraction bits, so that a 32-bit floating-point value effectively has 24 bits of mantissa: 23 explicit fraction bits plus one implicit bit of 1.

1.2.3 Approximate techniques

From recent studies, approximate computing has been broadly employed in several domains and offered a successful solution that eases design constraints with various techniques; precision scaling, loop perforation, and voltage scaling [82, 83, 84, 85]. First, in Precision scaling, it propose different level of precision on data to gain memory and energy efficiency. For example, it proposes decreased data bit-width that scale down the data precision but gain memory density which allows save more data in memory. Secondly, loop

Sign	Exponent e	Mantissa	Value
0	00...00	00...00	+0
0	00...00	00...01 to 11...11	Positive denormalized real
0	00...01 to 11...11	XX...XX	Positive normalized real
0	11...11	00...00	$+\infty$
0	11...11	00...01 to 11...11	Signalling Not a Number(SNaN)
0	11...11	1X..XX	Quit Not a Number (QNaN)
1	00...00	00...00	-0
1	00...00	00...01 to 11...11	Negative denormalized real
1	00...01 to 11...11	XX...XX	Negative normalized real
1	11...11	00...00	$-\infty$
1	11...11	00...01 to 11...11	Signalling Not a Number(SNaN)
1	11...11	1X..XX	Quit Not a Number (QNaN)

Table 1.2 – Floating point number expression

perforation uses task skipping in iterations of a loop. By skipping some iteration in loop, it improves the computation time and power consumption. For the last, voltage scaling propose is a power management technique that the voltage in the component is adjusted, depending upon the targeted performance. It improves the power consumption in trade off with accuracy of the data in the communication. In this thesis, we explore the possibility of precision scaling and voltage scaling.

— **Precision scaling:** This technique modifies the precision by reducing the bit-width of data. It downgrades the precision of data but can improve the data density, throughput, and energy consumption. For example, instead of using a full 32-bit floating-point, shorter data bit-width is transmitted, computed, and stored. In Figure 1.13, IEEE 754 standard floating-point number is presented with various precision levels. As computing and storing 32 bit of data is sometimes expensive and heavy to compute, half-precision and bfloat-16 are proposed. bfloat has the same bit width for sign and exponent as 32-bit floating-point. With bfloat16, the system can benefit essentially almost the same precision as 32-bit floating-point, as it has the same bit width of exponent while significantly reducing power and improving throughput without a special investment of time or design cost. More various and dynamic precision scaling is possible for the different data types [86, 87]. Moreover, this technique has been employed in several domains like input/output or intermediate operands to ease the computing constrain and to improve the storage [88, 89].

— **Voltage Scaling:** This approximate technique is proposed at hardware circuit level;

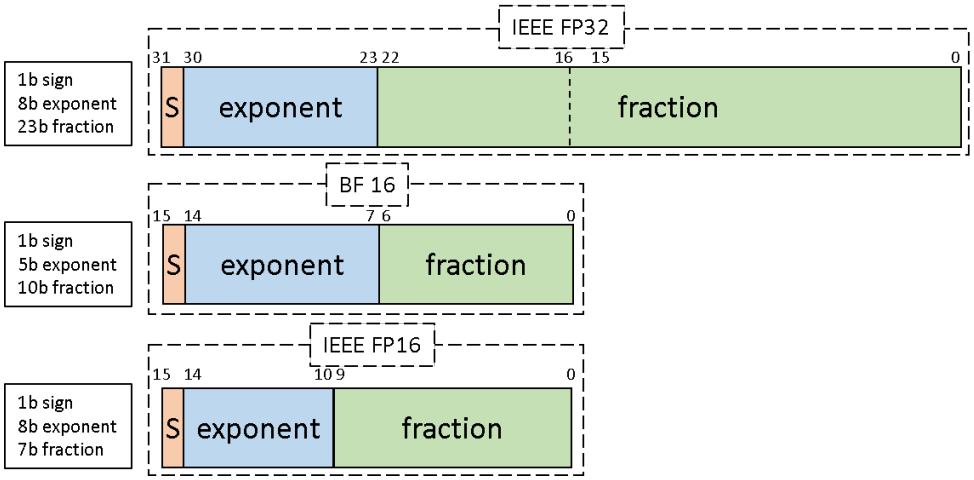


Figure 1.13 – Floating-point precision scaling.

The supplied voltage level is scaled to achieve the accuracy and power consumption the system desires [90]. [91] proposed power-aware computing through dynamic voltage scaling (DVS). It tunes the supply voltage by monitoring the error rate during operation. This voltage scaling can be implemented more dynamically; Dynamic voltage and accuracy scaling (DVAS). In [92], the authors proposed DVAS, by decreasing the processor clock frequency, the power consumption in processor is redueced which leads to improve the power efficiency. Moreover, DVAS can be implemented both in processor and chip architecture level. When implementing low voltage scaling, in data cell and interconnects, it diminishes the data bit information. It lowers and blurs the threshold of bit '0' and '1' and it impacts the hardware integrity and precision of data. When DVAS goes to architecture level, for example in interconnects in side of FPGA, it significantly decreases the power consumption by impacting the software level output results [93, 94].

1.2.4 Floating-point Data approximation

In the previous section, we saw the designed format of floating-point data to be stored and how it can be normalized. In this subsection, we first present range and limitation of floating point data and explore how floating point data suffers from an approximation techniques presented in section 1.2.3.

Figure 1.14 shows the range of a single-precision floating-point number. As we can see in this Figure, not every decimal number can be expressed precisely in the floating-point

number. The floating-point number format has its limit on representing real numbers from its conversion. The numbers that cannot be represented in floating-point format are rounded up, creating an error that is either slightly smaller or larger, depending on mantissa bits.

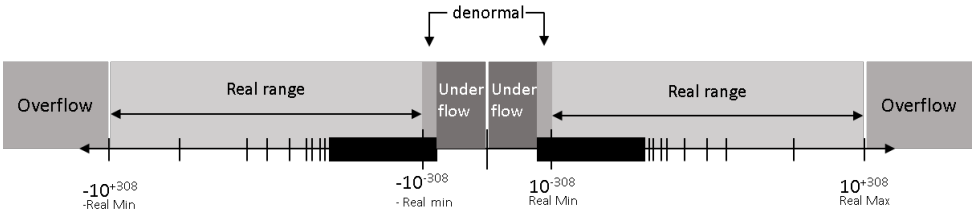


Figure 1.14 – Single precision Floating-point number range.

There are also number range that over/underflow. Overflow means that values have grown too large for the representation, in the same way that you can overflow integers. Underflow has less critical problem because it just denotes a loss of precision, which is guaranteed to be closely approximated by zero.

— **Precision scaling on Floating-point:**

Figure 1.15 presents precision scaling on single-precision floating-point data. In this precision scaling, only mantissa bits are truncated, which refers to the bit length reduction only on Least Significant Bits(LSBs). The red line 32 indicates non-approximate single-precision floating-point data, and the other 16, 12, and 10 indicate LSB bit reduction of 16, 20, and 22 with same length of sign and exponent bit length, respectively 1 and 8 bits. In the graph 1.15, we can see that when greater bit truncation is applied, lines fluctuate with higher error. As bits are truncated, the number presentation is limited, leading to a greater rounding error.

For example, in Figure 1.16, the number -27.7778 in decimal, is presented in floating point '11000001110111100011100011101111'. This number, in 22 bit reduction, is rounded up to '11000001110' which corresponds to -24, and an error rate equal to

$(-27.7778 - (-24))/-27.7778 \Longrightarrow 13.6\%$.

— **Voltage scaling on Floating-point:**

Voltage lowering techniques are applied to lower the robustness of links transmitting floating-point numbers. When Voltage scaling takes place on floating-point data, bit flip errors can occur. In Figure 1.17, the transmission of a floating point number is presented.

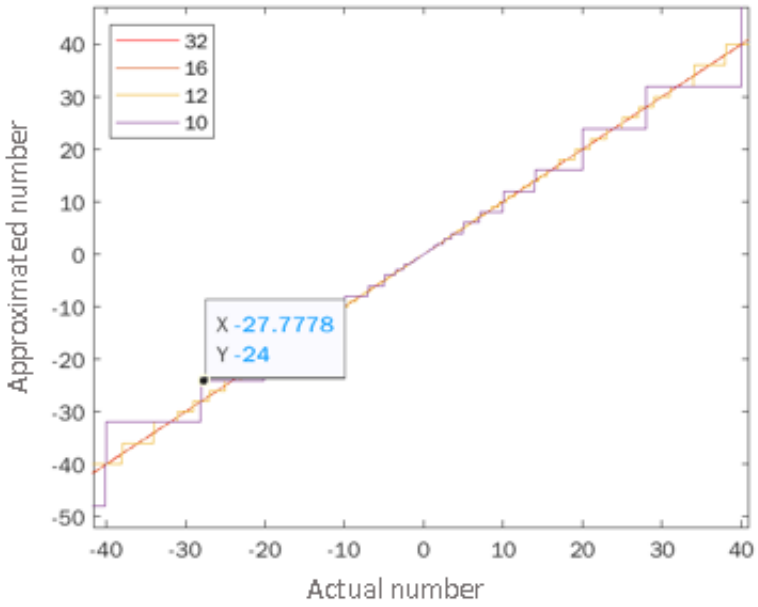


Figure 1.15 – Precision scaling of Floating-point number.

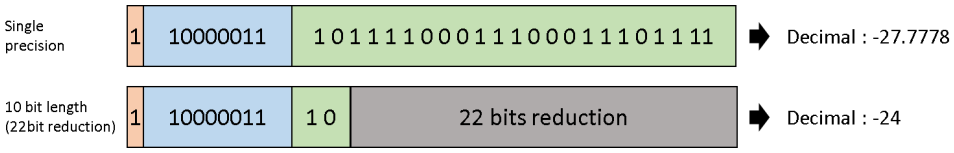


Figure 1.16 – Precision scaling on Floating-point.

Figure 1.17. a) presents accurate communication which guarantees low BER for data communication. In the case of 1.17. b), when low voltage scaling is applied, bit flip occurs, and it leads to high BER for the data communication. In voltage scaling, the output error of data can vary and be random for every communication. Though voltage scaling of output error rate is unpredictable, precise control is available compared to bit truncation as voltage can be scaled to target a specific BER.

1.2.5 Quality balance

As we saw, approximate computing is very promising but is not a panacea. As it manipulates and plays with the data's accuracy, output data can be critically damaged and

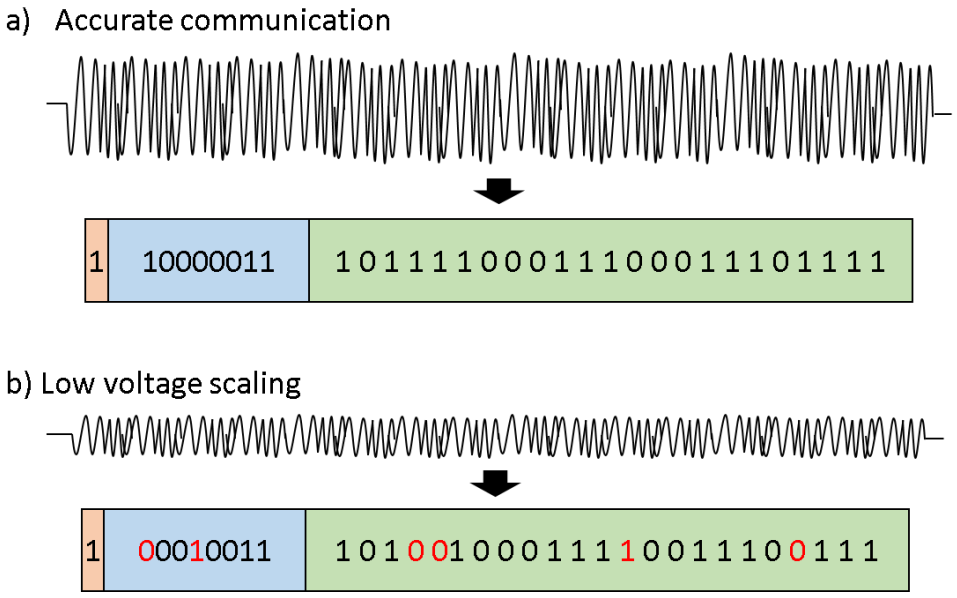


Figure 1.17 – Voltage scaling on Floating-point.

may not be acceptable on the application side. Therefore, approximate computing requires an effective and prudent approximation strategy and precision control. To overcome this challenge, [95, 96] proposed an effective and efficient quality management framework that monitors the approximation and output result at run-time to achieve controlled quality-efficiency trade-offs. In [97], the authors propose proposed outline detection and correction of approximate error. It utilizes a lightweight quality monitor to check an unacceptable and large amount of errors and fix them by rollback and re-computation. In [98], a data type qualifier is proposed to indicate approximate data and vary the precision of approximate portions in the application. In detail, the proposed type qualifier marks the variable as "approximatable" and these variables undergo through approximate qualifier, the storage, computing, and algorithm. On the other hand the variable which has to be protected, is guarded and its accuracy is guarantied. When approximate variable and accurate data are encountered, the programmer needs to explicitly specify their type to ensure careful handling of approximate data and makes the programming model safe. In [99], the method proposed ApproxBench as a standard quality metric to measure the approximation out of quantifying precision for several benchmark applications like Blackscholes, Canneal, Streamcluster, and Fluidanimate applications. This quality metric measures output error, the difference between the application output with and without approximate

communications. Each application has a specific purpose and is designed differently, so each application requires a particular error metric to measure the purpose-corresponding output error. For instance, Blackscholes is a financial pricing application that predicts the market price based on previous market values and trends. Thus, the output error can be measured with the Mean Square Error (MSE) metric, which computes the mean-error of output values as follows in Equation 1.4:

$$MSE = \sqrt{(\frac{1}{n})\sum_{i=1}^{n}(Accurate_i - Approximate_i)^2} \tag{1.4}$$

Based on this equation, approximated output values $Approximate_i$ are compared with original ones $Accurate_i$ to determine output error rate.

1.2.6 Approximate interconnect

In the previous section 1.1.6, we find out that NoC suffers from high power consumption due to a large number of data exchange communication. Indeed, as approximate computing is considered a prominent solution that provides power efficiency, several attempts implement approximate computing to optimize the power consumption in NoC. In [100], the authors propose approximate framework Approx-NoC for ENoC. The work presents a data transmission approximation that reduces the number of data transmissions by not sending the approximately similar data to improve the data locality for higher compression rate. Moreover, integer values are approximated with similar data reuse logic to minimize the area and power overheads. Figure 1.18 presents the proposed approximate controller in network interface (NI). From this work, the packets which exit to network, go through this NI and checked out to be approximated or not. When data can be approximated, it goes to a specific component VAXX which encodes/decodes data through bit compression.

The work presented in [101] proposes to regulate the packets injected in NoC to mitigate network congestion. Their proposal adoptively drops a fraction of data packets that can be approximately predicted as lost data. An approximate error correction control in NoC has been proposed for large-scale parallel systems in [102, 103]. These paper proposes configuring different lengths of error correction for the data packet. For that, it decreases the number of data bits under the error correction. They achieve latency and power effi-

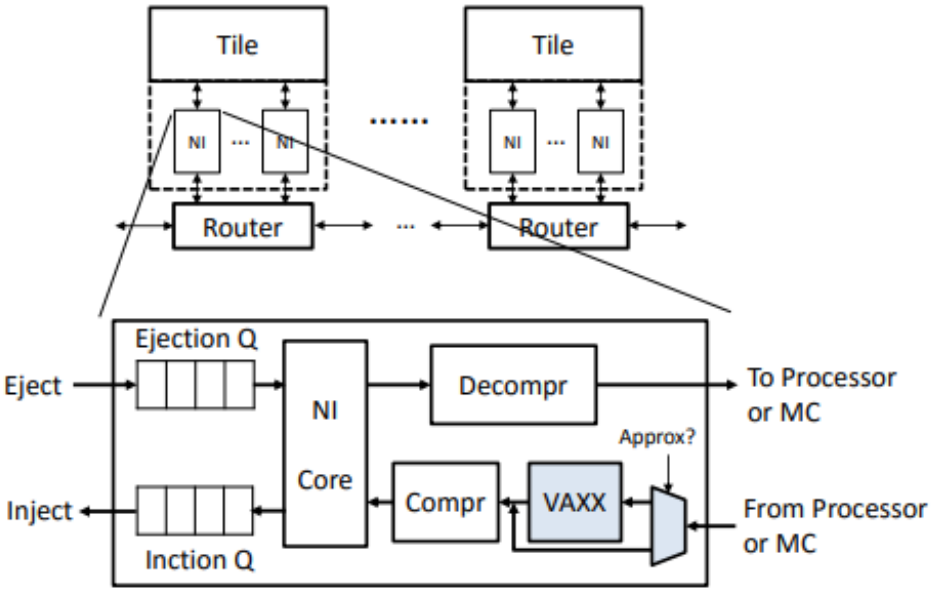


Figure 1.18 – Approx-NoC Network interface [100] .

ciencies by packet prediction and approximate error checking technique on NoC. In [104], the authors propose approximate communications on the NoC through data compression and relaxation on synchronization. Their work improves energy consumption and speeds up performance by reducing data synchronization packet that causes network bottleneck. The paper [105] proposes dual voltage communication for approximate data communication in electrical NoC approximate communication by managing the power with two voltage levels that correspond to the content of the packet. Headers and important data flits are accurately transferred at a high voltage while the approximatable flits are set to low voltage, which may provoke bit flips at the destination.

These successful study cases of approximate computing in electrical NoC have been introduced into Optical interconnect NoC as well. In [16], the authors present Lorax which truncate bits by turning off lasers, thus allowing to save power. Figure 1.19 illustrates the Lorax that adoptively controls the laser power according to the distance and apply the approximate technique, Truncation. In their system there is a specific component Gateway Interface (GWI) which connects the electrical layer of the chip to optical NoC layer. The GWI is consist of lookup table to classify the data packet to be approximated or not. In the Figure, we see SWMR optical link that sender modulate the optical signal with MRs

to destination which is ready to demodulate with MR filters. In this example, as distance to destination is short, the laser power is decreased to improve the power consumption. In case of long distance communication, Truncation is applied in the transmission. However, the truncation is only applied when the estimated power at the receiver is lower than the photodetector sensitivity, which does not allow to fully explore the accuracy and power trade-offs.

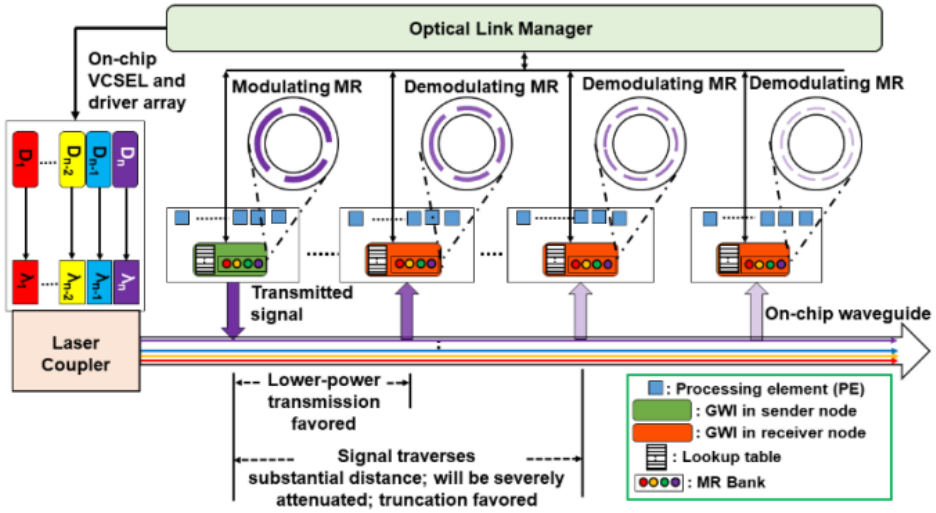


Figure 1.19 – Approximate interconnect from Lorax [16].

1.3 Summary

Nowadays, large-scale applications such as scientific computing, video/image processing, and data mining require high performance with a large amount of computation and data exchange communication. Many/multicore architectures, which integrate hundreds or thousands of cores within the same chip, fulfill the needs of hardware performance with huge parallel computation capacities. This rapid trend leads to a new communication diagram NoC, which might fulfill the great demand of efficient communication and massive parallelism between IPs in NoC. This trend of complex architectures emphasizes the on-chip interconnects as a key element to ensure communication needs, which directly affects application performance and energy consumption, and execution time. In the past decade, Electrical NoC could be an efficient solution for multicore architectures in the range of tens of cores in a chip to bypass the parallelism limitations with traditional

buses communication links. However, as the manycore era evolves to larger number cores like hundreds and thousands core in a chip, electrical interconnect NoCs faced scalability problems from latency and design area constrain and energy, power consumption due to a significant increase in the number of hops between cores.

To overcome these problems, Optical NoC (ONoC), which bases on Nanophotonic interconnects, is praised as the promising solution by providing bandwidth and latency issues, as optical signals propagate near speed-of-light in waveguides. Moreover, the optical interconnect allows very high throughput through wavelength division multiplexing (WDM). By channeling at multiple wavelengths for data communication, the optical link can potentially close the gap between the need of interconnect throughput and on-chip data processing speed. However, their implementations remain challenging due to the low efficiency of the lasers, which are vital devices in such interconnects. Indeed, the laser power consumption is mainly driven by the high signal power required to transmit data at a low BER for accurate communications.

Approximate computing is emerging as an efficient method to improve embedded computing systems' energy efficiency and execution speed. Approximate computing exploits tolerance of application quality degradation and reduces the data representations by lowering the design constraints and moderating the computation complexity. The approximate computing paradigm has been deployed in a numerous ways from devices to software and algorithms. For example, floating-point data approximation has been widely studied in several domains. Floating-point data plays an important role in number representation in science, engineering, and finance application. Approximating real numbers with floating-point numbers representation inevitably introduces rounding errors. Without careful implementation, rounding errors can be propagated and amplified, and it causes critical quality degradation at the application level. Therefore, several detailed approaches are proposed, like least significant bit (LSB) truncation and voltage scaling at communication link to improve computation performance, energy efficiency, and high memory density. In this context, this dissertation investigates the challenges of silicon photonic interconnect energy consumption and proposes approximate optical interconnect and distance-aware power laser power configuration. Moreover, we focus on techniques designing more straightforward laser power controllers than fine-grain configuration to decrease the overhead of the proposed architecture block to address more flexible scalability for NoC.

Chapter 2

Approximate communication in Optical Interconnects

2.1 Idea Overview

As we discussed in section 1.2.4, mantissa bits, and specifically LSBs, of floating-point data have less impact on number presentation. This quality feature can lead to an approximation of LSBs of floating-point data in optical communication to have meaningful power saving on communication while not affecting too much on application quality. In our work, we propose to approximate data to save power on optical communications by using two techniques: bit approximation and truncation. The bit truncation can offer high power saving as it can turn off the laser on communication. However, as we saw in section 1.2.5, truncation can generate significant error on number presentation. Bit approximation which uses low voltage scaling for communication can benefit power saving with more controllable error on data, though, it gains relatively less power saving compare to truncation.

Even-though, LSBs have a less critical impact on the number, disproportionate and excessive approximation might lead to significant quality degradation. Therefore, we need to have careful approach on approximation and it requires delicate control of data for the application quality. In our work, we define degree/intensity of approximation as *approximate level* and propose two methods for *approximate level* control. First, we define bit areas according to the approximate techniques. Then, we adjust these bit area width according to the approximate level that application requires. Whenever the application cannot tolerate much of approximation, we decrease the bit area where approximate techniques are applied. On the other hand, when application is able to tolerate large amount of error, we increase the length of bit area where truncation and bit approximation are applied. Secondly, we can control approximate level by targeting different BER on approximate bit area. Targeting high BER on a communication can profit more power saving but with

more possible errors on bits.

In the following, we present the detail of our proposals and its configuration on laser driver for various power levels. Then, we present implementation of proposed technique on optical interconnect with detail of fundamental method of power level definition.

2.2 Approximate communication in optical link

In this section, we will see the detail of proposed approximate techniques and approximate level control. It contains detail of the floating point data communication at bit level with optical laser power configuration. Then, we presents the relationship between laser power level and communication BER.

2.2.1 Proposed approximate techniques and approximate level control

In our work, we propose delicate control of approximate techniques to adopt *approximate level* that application requires. First, we defined different bit areas according to the implemented approximate technique in our work.The transmission of IEEE 754-2008 single-precision Floating-Point (FP) format regarding wavelengths (lasers) is presented in Figure 2.1. To have flexible control of *approximate level*, we divided floating-point data into three different bit areas: bit area where not approximated ($NotAx$), approximated (Ax), and truncated (A_{Trunc}) in orange, green and gray relatively. These bit areas enable different laser power configurations for the responsible lasers. As MSBs of floating-point data have a major meaning and affect data significantly, the communication of its communication has to be guaranteed. Meanwhile, LSBs are under approximate techniques: approximate bit and truncation thanks to its minor significance on number. Truncation in approximation refers to cutting off the bit and decreasing the number of bits on number representation i.e 32bits to 16 bits. However, in communication, the bit truncation refers to not containing information, which mean that all the truncated bits are set to '0's. Therefore, we can set the laser power to at off-state, called P_{off}. This can bring significant power saving for laser but it can have big distortion on number representation. Not all the LSB can handle aggressive approximations, and it might require milder approximate technique. For example, LSBs which are close to MSB bit area do not have significant meaning on number like MSBs but still have impact on number, so aggressive approximate technique

like *truncation* might not be suitable to be implemented. Therefore, we configure this bit area as approximate bit area Ax where targets higher BER for data communication. According to these defined bit areas, the laser power are configured to P_{high}, P_{low} and P_{off}. The power levels are set by targeted BER: not approximated BER (BER_{not-ax}) , approximate BER (BER_{ax}) and truncation BER (BER_{Trunc}).

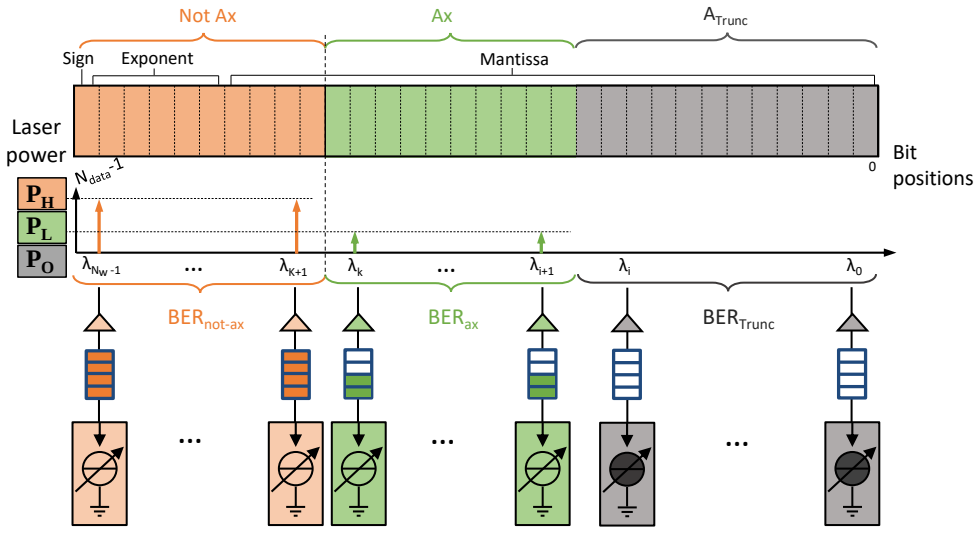


Figure 2.1 – Floating-point data scheme

These bit areas can go further for approximate level control by adjusting the bit width of approximate and truncation areas. Figure 2.2 presents various possible configuration for different approximate bit area Ax and truncation area A_{Trunc}. The graph presents three general approximate levels, which have different bit area configurations. The number of bits under no approximation, bit approximation (voltage scaling), and truncation is adjusted to have flexibility on approximate level control.

A more aggressive approximation is possible when the data is more tolerant to approximation. Thus, it can have a large bit area of truncation and approximate bit. However, when the application requires a more delicate *approximate level*, the bit area is reconfigured with a shorter bit width of truncation and approximate bit. Furthermore, these bit areas depend on application error tolerance, and sometimes, the application cannot approximate data transmission. Then we configure at No approximation $Not\ Ax$, which sets all the lasers at P_{high}.

Secondly, once the bit area is configured, the approximate level can vary by targeting different BER levels. Figure2.3 illustrates the relation between targeted BER at the

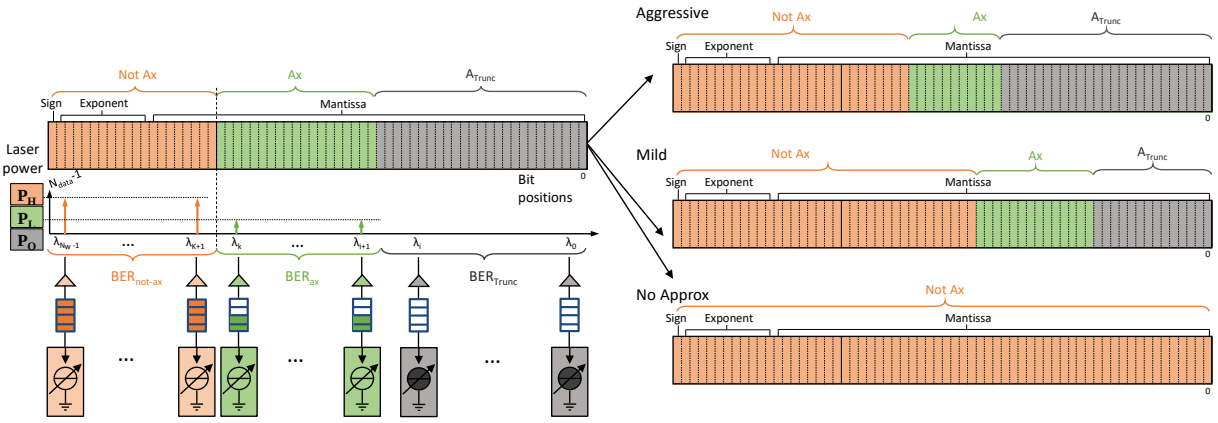


Figure 2.2 – Bit area configuration

destination and input laser source power. When the application requires high accuracy communication, it targets low BER (BER_{not-ax}) at the destination, which means high received power at the photodetector, guaranteeing an error free communication. Therefore, the input laser current has been configured in P_{high}. As it is set to high power, there is no power saving and no approximate communication. However, when targeted BER (BER_{ax}) is high, it requires low received power at photodetector, which means low input laser power P_{low}. It offers power saving by lowering the laser power level, and it directly affects the communication quality of BER. The approximation is moderately implemented in optical communication throughout this voltage scaling by lowering the input laser power. The detailed study of targeted BER and application quality is presented in section 2.5.2.

In this latter, according to these bit areas and BER for approximate level, we will explore the power saving and quality degradation effect at application side.

2.3 Proposed communication link

In previous subsection, we saw the overview of the proposed approximate optical communication link and approximate level control. In this section, we will see the detail of how the approximate techniques are implemented in optical NoC and power saving of approximate link.

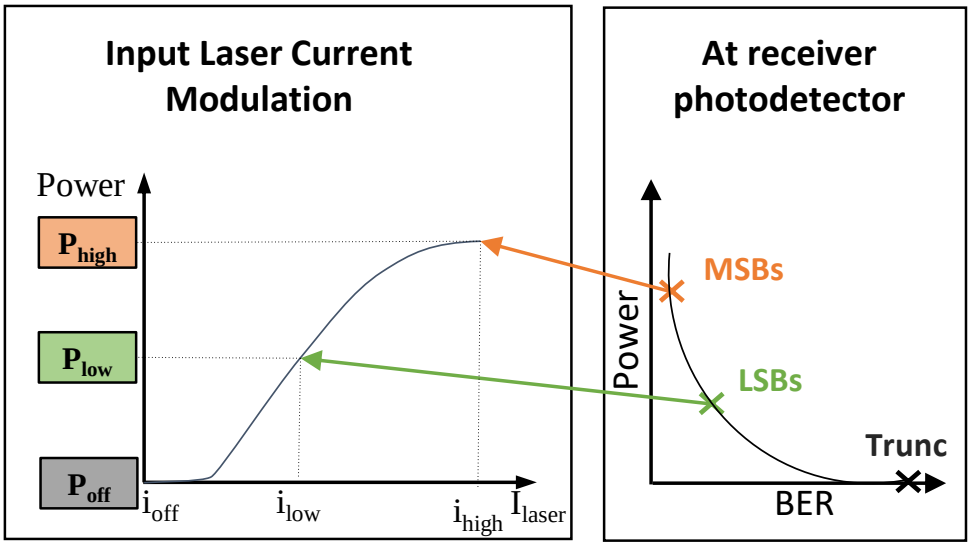


Figure 2.3 – Laser input current vs BER

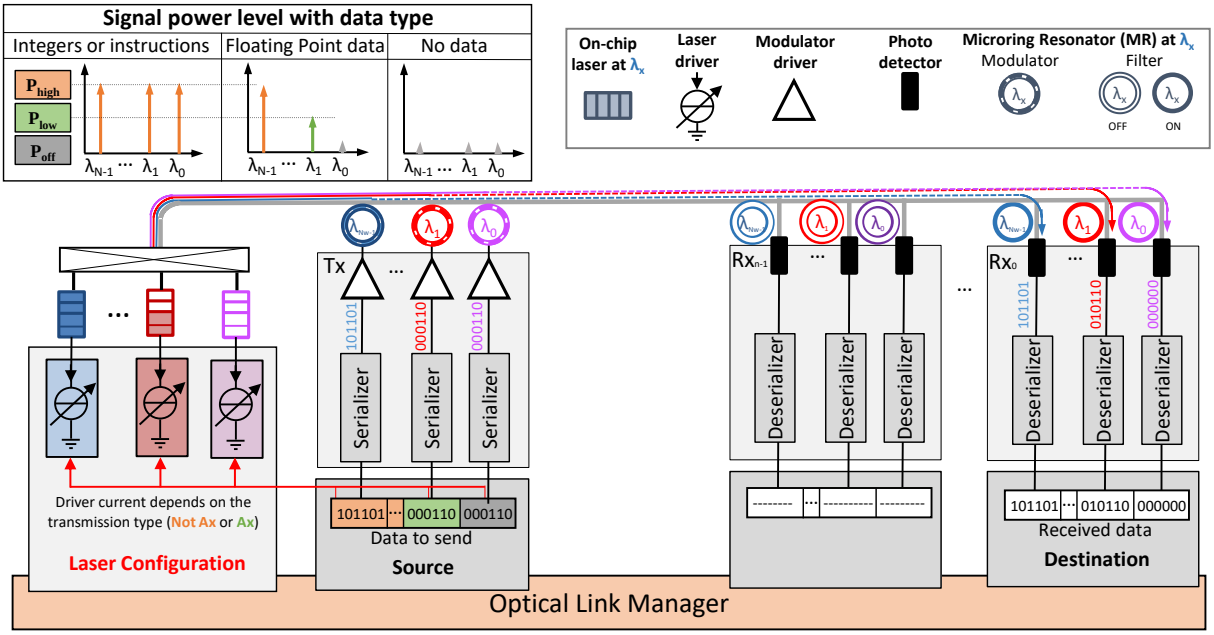


Figure 2.4 – Proposed SWMR optical link

2.3.1 SWMR Channel Design

Figure 2.4 illustrates our approach on a Single Writer Multiple Reader (SWMR) link. The transmitter (writer and optical source) is connected to several receiver (readers) using shared medium waveguide (for sake of clarity, only one SWMR link is represented in the figure). Each waveguide allows transmitting $N\lambda$ signals using WDM ($\lambda_0..\lambda_{N\lambda-1}$). The signals are emitted by on-chip lasers and are combined into waveguide using a MultiMode Interference (MMI) couplers [106]. In case no communication occurs, the lasers are turned OFF and hence do not consume energy [57]. When a communication is initiated, the writer activates the lasers (optical signals are emitted after few ns typically) and at the transmitter, the MRR modulate the optical signals using On-Off Keying (OOK) modulation according to bit data. Although, we assume that data transmissions involve the use of all the optical signals, the transmission of data are serialized according to the data bit width (N_{data}) and the number of wavelengths (N_w). 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). For instance, $N_{data} = 32$ and ($N_w = 8$ would lead to a Bit Stream Length (BSL) of 4, i.e. 4 bits are sent on each wavelength and 4 clock cycles are needed to transmit each data. The design of the serializer has been investigated in [54, 107] and is out of the scope of the thesis. The modulated signals propagate along the waveguide until reaching the destination. At this moment, all the MRs of readers are off state, except the destination. At destination, 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.

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

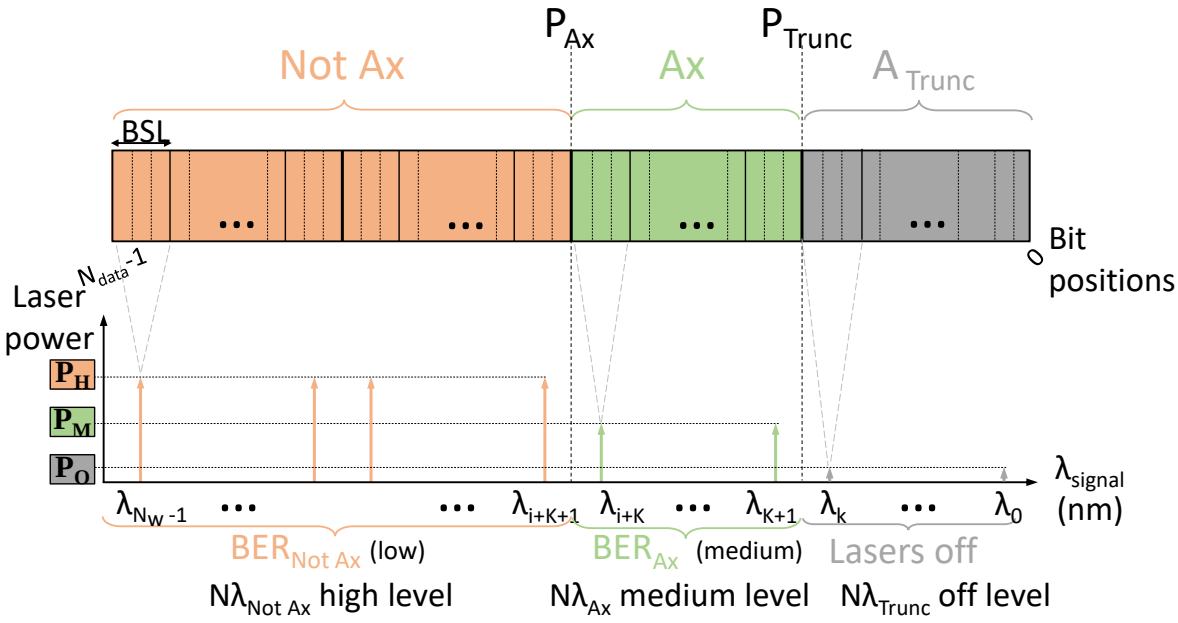
— **Robust transmission:** is used for sensitive data. For this latter, in our work we consider instructions and integer data communications. It involves low BER, noted BER_{not-ax}, (e.g. 10^{-12}), which is guaranteed by injecting a high optical power (P_{high}) from the lasers. This results in robust but power hungry communications for all bits.

— **Approximate transmission:** is dedicated to data that can be approximated, in our proposal we consider floating-point data. It involves heterogeneous power level

emission for the lasers: MSB and LSB are transmitted under high power P_{high}, low power P_{low} and truncation P_{off} levels respectively. The manager individually configures the laser drivers according to assigned bit area which meet the required communication quality.

2.3.2 Floating-point data scheme

The proposed communication method involves both approximation and truncation on data. As explained in section 2.1, to control different approximate level in NoC, we propose configurable bit areas according to the bit significance. Hence, several configurations are available to send data that can be approximated, and that can be tuned to achieve a given quality of result for an application. Therefore, we set parameters that helps the configuration of bit area for laser and wavelength.



$$\text{with: } N\lambda = \frac{N_{data}}{BSL},\ N\lambda_{Not\ Ax} = \frac{(N_{data} - P_{Ax})}{BSL},\ N\lambda_{Ax} = \frac{(P_{Ax} - P_{Trunc})}{BSL},\ N\lambda_{Trunc} = \frac{P_{Trunc}}{BSL}$$

Figure 2.5 – Approximate data scheme

Figure 2.5 illustrates how the technique can be used for the IEEE 754-2008 single-precision Floating-Point (FP) format with the three different bit areas used: $NotAx$, Ax and A_{Trunc}. These 3 groups provide flexibility to adapts the robustness of the transmitted

Parameters	**Definition**
N_{data}	Data size
P_{Ax}	Position of the higher approximated bit
P_{Trunc}	Position of the higher truncated bit
A_{NotAx}	Area of accurate bits
A_{Ax}	Area of approximated bits
A_{Trunc}	Area of truncated bits
$N\lambda$	Number of wavelengths (nb of lasers)
BSL	Bit Stream Length ($BSL = N_{data}/N\lambda$)
BER_{NotAx}	BER for non approximated bits
BER_{Ax}	BER for approximated bits

Table 2.1 – Summary of parameters used for the proposed method.

bits with respect to the bit significance. Table 2.1 summarizes the parameters used to identify and configure the different groups of bits of a data with the proposed method. While the flexibility can be controlled through the BER to reach and by selecting the bits areas to approximate or truncate (defined by A_{Ax} and A_{Trunc} on Figure 2.5), in this paper we propose to fix these parameters based on the analysis of a set of applications. Based on this bit configuration we are able to calculate the power saving from our proposal. This analysis is presented in the next subsection.

2.3.3 Power level definition

In this section, we present the method to define input laser power level P_H and P_L, according to the targeted BERs (BER_{NotAx} and BER_{Ax}). As illustrated in Figure 2.6, we plot two lines representing the received optical power with respect to the communication distance hop count for different injected optical power. The source optical power is estimated by:

$$\begin{aligned} P_s &= P_d + (Hop_{s,d} - 1) \times Loss_{mr} \times N\lambda + Hop_{s,d} \times Loss_h \\ &\quad + Loss_d + Loss_{Xtalk} \end{aligned} \tag{2.1}$$

where P_s is the optical power injected from source core (in dBm) and P_d is the optical power received at the destinations core (dBm). P_d depends on the targeted BER at the destination photodetector [41]. In this work, we defined the BER for both approximate and accurate communications. The input laser power strongly depends on the attenuation

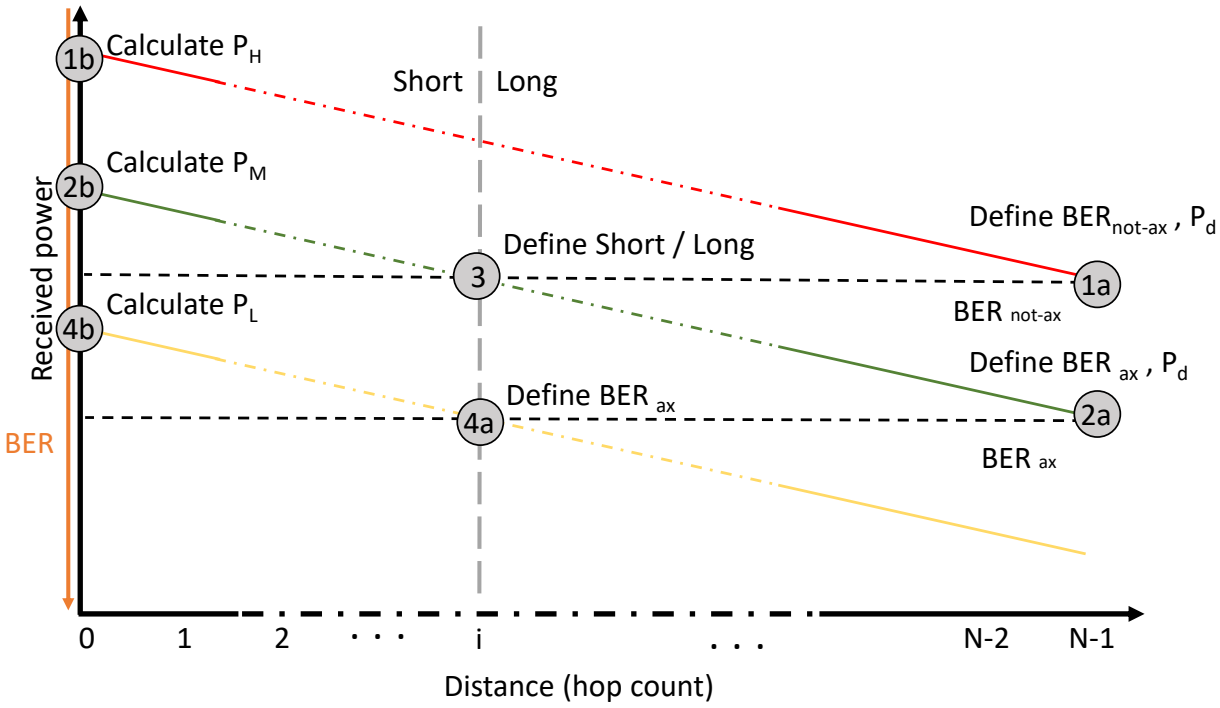


Figure 2.6 – Power level over hop distance

experienced by the propagating optical signals, which depend on: i) required power at the destination, ii) the hop count $Hop_{s,d}$, iii) the number of wavelengths (or lasers) $N\lambda$, iv) the MR through loss ($Loss_{mr}$, in dB), v) the waveguide propagation loss per Hop ($Loss_h$, in dB/hop), vi) the MR drop loss ($Loss_d$, in dB), and vii) the crosstalk loss ($Loss_{Xtalk}$, in dB). For this latter, we consider the crosstalk model from [108]. The following details the method. We first define the destination power P_d allowing to reach the targeted BER_{NotAx} at the furthest receiver (core at distance $N_c - 1$, mark (1a) in Figure 2.6). By taking into account the architecture size and the losses, we estimate the required optical power to be injected from the source P_s using equation 2.1. The corresponding signal power is illustrated by the red line. Mark (1b) corresponds to P_H, the injected optical power to inject at the source core to guarantee error-free communication for any source to destination in the architecture. Similarly, we define the power level P_d on the last receiver ($N_c - 1$) for the approximate communication obtained using BER_{Ax}, mark (2a). From equation 2.1, the required optical power at the source allows defining P_L, mark (2b).

2.3.4 Power saving model

In this subsection, we detail how power saving is achieved from our proposal bit approximation and truncation.

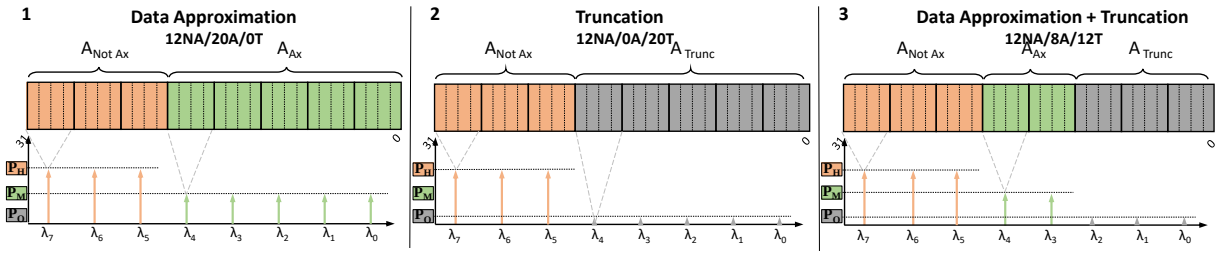


Figure 2.7 – Power saving data scheme

Figure 2.7 illustrates the 3 possible approximate configurations with our proposals. For this examples, we utilize 8 wavelengths (λ_7 to λ_0) which are allocated at 8 groups of 4 serialized bits for 32 bits long data transmission. Each scenario is noted as xNA/yA/zT with: x bits Not Approximated, y bits Approximated, z bits Truncated bits. The first scenario presents where only approximation is implemented at $A_{NotAx} = 12$ and $A_{Ax} = 20$. This scenario is noted as $12NA/20A/0T$. This leads to lasers power configuration at λ_7 to λ_5 with P_{high} and lasers at λ_4 to λ_0 with P_{low}. Second scenario, $12NA/0A/20T$, involves only truncation [16] which refers to $A_{NotAx} = 12$ and $A_{Trunc} = 20$ and it leads to laser configuration P_{high} for λ_7 to λ_5 and P_{Off} for lasers at λ_4 to λ_0. Third scenario, $12NA/8A/12T$, is a combination of approximation and truncation with $A_{NotAx} = 12$, $A_{Ax} = 8$, and $A_{Trunc} = 12$ which leads to P_{high}, P_{low} and P_{Off} for lasers at λ_7 to λ_5, λ_4 to λ_3, and λ_2 to λ_0 respectively.

To compute the power saving for these scenarii, the baseline scenario we consider is the one where all the lasers are configure with P_{high}. The power consumption of proposed optical link is presented in equation2.2a. First, MSBs of FP $[N_{data} - P_{Ax}]$ is configured with P_{high}. Secondly, approximate bit area $[P_{Ax} - P_{Trunc}]$ is configured with P_{Medium} to target approximate communication. Last, for the lasers that are in charge of Truncation bit area $[P_{Trunc} - 0]$ are set in off state P_{off}. The summation of all the power consumption from those bit areas defined the total power consumption for the proposal. The actual power saving depends of bit area configuration and targeted BER. Therefore, the power gain compare to baseline scenario can be computed with equation 2.2b

$$P = \frac{1}{BSL}[(N_{data} - P_{Ax}) \times P_{High} + (P_{Ax} - P_{Trunc}) \times P_{Low} + P_{Trunc} \times P_{Off}] \quad (2.2a)$$

$$PowerGain = \frac{P}{N\lambda \times P_H} \quad (2.2b)$$

2.4 Experimental setup

In this section, we present the system parameters of the architecture, benchmark applications and the simulation platform that are considered for the evaluations of our contribution.

2.4.1 Considered 3D architecture

For experimental setup, we consider a 3D architecture composed with two layers: electrical layer to support the computation part, and optical layer for the optical communications (as presented on Figure 2.8-a). Electrical layer contains 64 cores, which are grouped into 16 clusters of 2×2 cores. Inside of cluster, last level cache (L3) are shared with the others cores, and each core has its own private L1 data (L1d), instructions (L1i) caches and L2 cache. Furthermore, a MSI cache protocol is used to ensure coherency between the distributed caches. All these memory organization parameters are listed in Table 2.2. The total chip size considered for this section is equal to $1600mm^2$, hence the distance between two consecutive Optical Network Interfaces (ONIs) is equal to $10mm$. Moreover, every clusters are connected with ONI through TSVs to connect optical layer to reach out the network.

On top of electrical layer, optical layer is positioned for network communication. The ONI blocks are located in each clusters at optical layer and have 1 transmitting and 15 receiving waveguides which are featured with SWMR and WDM (as depicted in Figure 2.4). Each waveguide transmits 8 optical signals on different wavelengths with FSR of $8nm$, channel spacing of $1nm$ ($\tilde{1}20$ GHz) to avoid crosstalk noise between MRs, hence each having a 4 bits stream data (i.e. $BSL = 4$ and $N\lambda = 8$)

Table 2.2 – Hardware parameters of the simulated architecture

Parameters	**Values**
Nb of clusters	16 $(4 * 4)$
Nb cores by cluster	4
ONI distance	10 mm
L1 I / D caches	64 KB each private
L2 cache	512 KB private
L3 cache	4 M shared (in a cluster)
Cache protocol	MSI

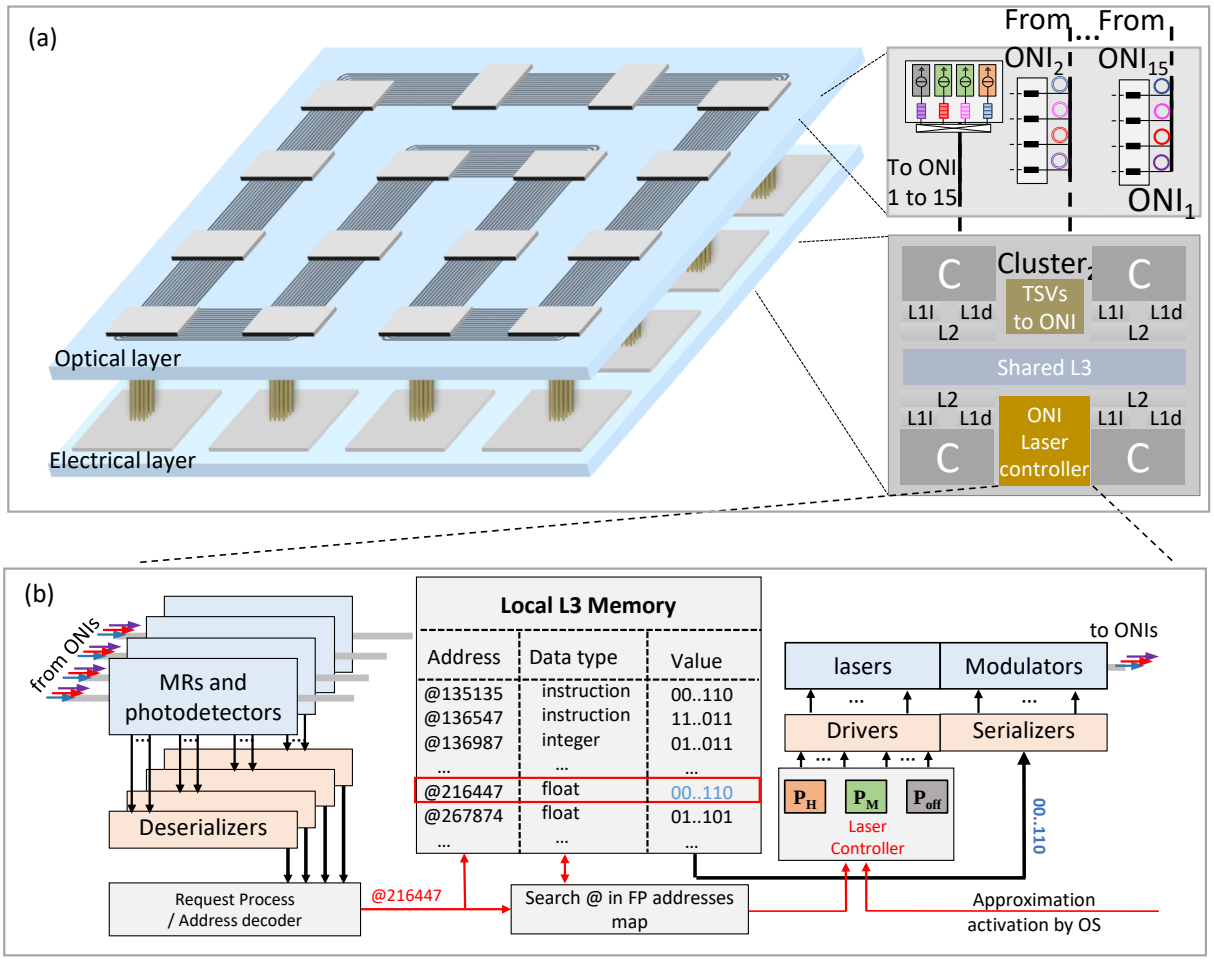


Figure 2.8 – 3D Architecture and ONI a) 3D ONoC with cluster and ONI laser controller b) Cache coherency example

2.4.2 Approximation under cache coherence traffic

Cache coherency involves traffic between cache memory and last level of cache that is mainly initiated by cache misses or write-back accesses [109]. However, the traffic initiated by cache coherence protocols also involves addresses of the data to be transmitted and is thus data type agnostic. Therefore, in LLC table, there are information of address, type of data [110, 111]. According to this information, ONIs recognize the data to approximate and configure laser power. The design of centralized optical link manager to configure the communication channels is out of the scope of the thesis and has been already investigated in [112]. In our proposal, we assume that approximation mode can be enabled by the OS based on user requirements. This is achieved by setting or resetting by a flag "Approximation activation by OS" at run-time. The details of communication between OS and ONIs is out of scope for this paper and it is studied in [113, 114]. In case approximations are disabled, all communications are carried without errors using P_{high}. Otherwise, the lasers are configured as following scenario. Figure 2.8-b) presents the principle of cache miss scenario and data transmission with approximate laser power configuration. The L3 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 cache hit, the manager then grants the SWMR access to the L3 that initiated the request and the data is transmitted to the serializers for signal modulation. Moreover, the laser controller configures the power levels of lasers corresponding to approximate level that has approximated and truncated bits area configuration. 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), as response packet contains data. The header is always transmitted with high laser power to ensure the communication quality, i.e., without approximation.

2.4.3 Simulation setup and application

We evaluate the proposed approximation technique on error tolerant benchmark applications from Approxbench suite [115]. Table 2.3 presents the four benchmark applications used in this paper: Blackscholes, Fluidanimate, Canneal and Streamcluster. The goal of each applications are financial algorithms, animation, routing and online clustering, respectively.

Table 2.3 – Used Approximate Benchmark and associated error metric [115]

Applications	**Descriptions**	**Error metrics**
Blackscholes	Financial pricing	Mean error
Canneal	VLSI routing	Routing Cost
Fluidanimate	n-body simulation	Distance in elements
Streamcluster	Online clustering	Cluster center distance

Moreover, Approxbench provides the output error metric for each application, which allows estimating the Quality of Result (QoR). QoR is a main factor to consider for approximation to monitor the quality degradation. As each application performed has specific goal, each one has a particular error metric as stated in Table 2.3. QoR can be measured by difference in output result between the application with and without approximate communications throughout the quality metric. For instance, Blackscholes is financial pricing application that predicts price of the market based on previous market values and trend. At the end of application, Blackscholes mathematical equation estimates the theoretical value of derivatives other investment instrument. For Blackscholes the output error metric uses the Mean Square Error (MSE), and is computed on mean error of output values as follow in Equation 2.3:

$$MSE = \sqrt{(\frac{1}{n})\sum_{i=1}^{n}(Accurate_i - Approximate_i)^2} \tag{2.3}$$

Based on this equation, approximated output values $Approximate_i$ are compared with original output $Accurate_i$ to determine output error. These applications are simulated in multi-core simulator Sniper [116]. From the simulations, we extracted the output results to compute the output errors and the communication traces linked to the traffic between clusters in NoC. From the communication traces, which embed information of packet type, source, destination, and sent time, we analyze communication distribution and type of data.

2.4.4 Benchmark traffic analysis

To evaluate the impact of the proposed method, we analyze the traffic distribution for the selected applications. This analysis includes several architecture size 16, 32, 64 and

128 clusters architecture. The architecture size vary with same system-level parameter defined in Table2.2 but with different number of cores.

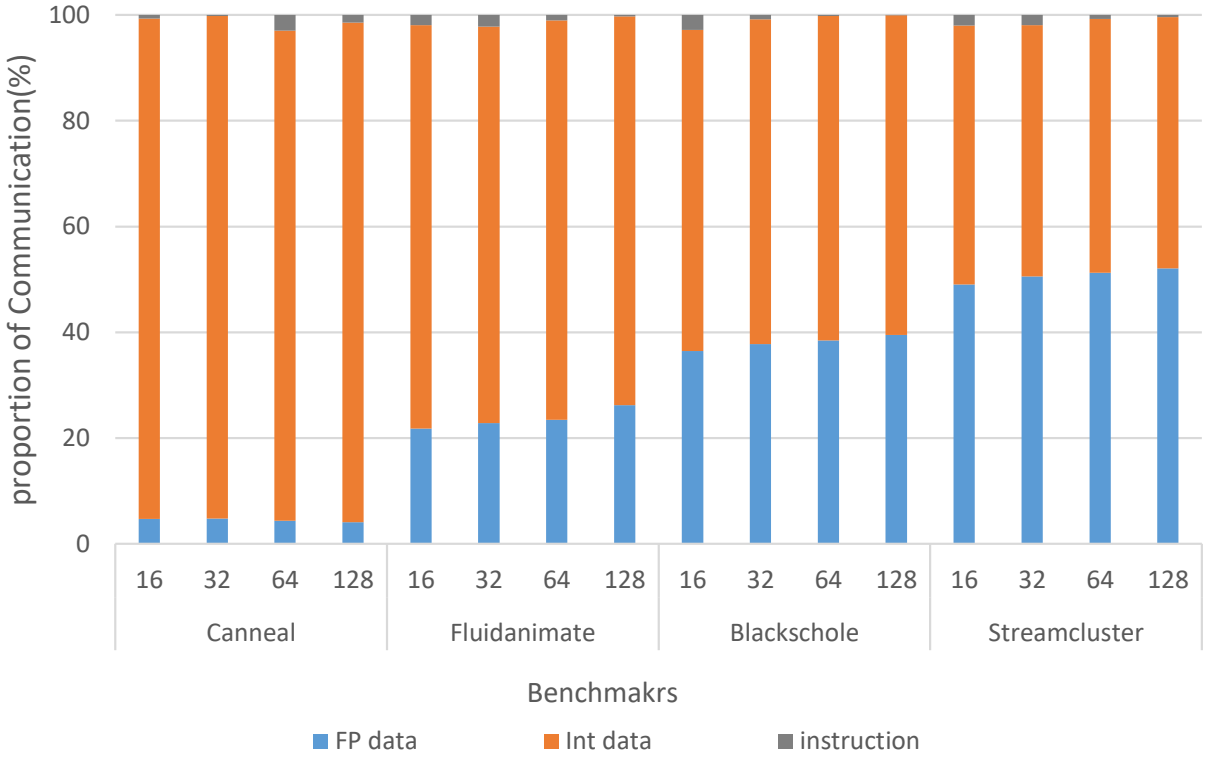


Figure 2.9 – Proportion communication over the architecture size

Figure 2.9) presents the traffic distribution by data type according to the architecture size (number of cores in the system). The communications are classified in Instruction, Integer, and Floating-Point data. We first observe that ratio of FP numbers strongly depends on the application. For instance, we can see that Canneal only have around 4% of FP data communication, while the others reached 22%, 37% and 50%, for Fluidanimate, Blackscholes and Streamcluster, respectively. This ratio of data type can be explained by how application is designed and programmed data type and it is defined at the application designing level. Moreover, 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 cores affect data parallization and induce shared data by parallel execution of the kernel. This characterization allow the designer to target on which type of application will have more efficient on power saving. From this analysis, it is evident that applying energy reduction technique on FP data for Canneal will not produce a

considerable energy reduction, unlike the other applications, and especially Streamcluster. We provide the Floating-Point (FP) ratio that refer to potential data type that may benefit from power saving by using the proposed approximate method.

2.4.5 Optical link loss parameter

In the following, we design the optical link and its loss according to system-level parameters defined in Table 2.2 and loss computation method indicated in section 2.3.3. First, according to power level definition equation 2.1, we first assume minimum optical power at the photodetector for different BER of accurate and approximate data communication respectively [41] summarized in Table 2.4.

Table 2.4 – Required power at Photodetector per BER

BER	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}	10^{-12}
Required power Photodetector dBm	-14	-13	-12	-11.2	-10.55	-10	-9.6	-9.2	-8.9	-8.6	-8.2	-8

We assume 0.02 dB MR through loss, 0.7 dB MR drop loss [117], an FSR of $8nm$ and a MR quality factor of $20,000$ [118].

Table 2.5 – ONI design parameters

Parameters	**Descriptions**	**Typical values**
	Technological parameters	
$Loss_{mr}$	MR through loss	$0.02\,dB$ [119]
$Loss_d$	MR drop loss	$0.7\,dB$ [119]
Q	MR Quality Factor	$20,000$ [118]
$Loss_h$	Waveguide loss per hop	$0.25\,dB/cm$ [117]
P_d	Optical power	$-8\,dBm$ at 10^{-12} BER
	received by photodetector	$-12\,dBm$ at 10^{-3} BER
FSR	Full Scale Range	$8nm$ [118]
$Laser_eff$	Laser efficiency	0.33 [120]
B_R	Bit-rate	10Gb/s

2.4.6 Input laser power vs BER

In proposed technique, there are various approximate BER (BER_{ax}) then there are numerous laser power level possible. To analyze power saving for proposed technique, we investigate possible input power level according to received power level at photodetector

for targeted BER. The input power is calculated with loss in waveguide and photodetector sensitivity described in section 2.3.3, equation 2.1. Figure 2.10 illustrates received power level at photodetector and input laser power by different targeted BER. First, we can see, required power at last destination of photodetector [41], in this scenario 15^{th} reader in SWMR link. Indeed, to make accurate communication without error, photodetector requires high received power. Secondly, as photodetector requires more power for low BER, laser source need to be configured with high input power. In our configuration, 10^{-12} is considered as a baseline scenario which uses accurate communication only i.e. BER_{not-ax} and no approximation for all the communication. To compare laser power consumption for different BER configurations, we analyze the laser power consumption with this baseline scenario. The other approximate laser power consume less power according to targeted BER, for instance 10^{-1} consumes 25% of laser power.

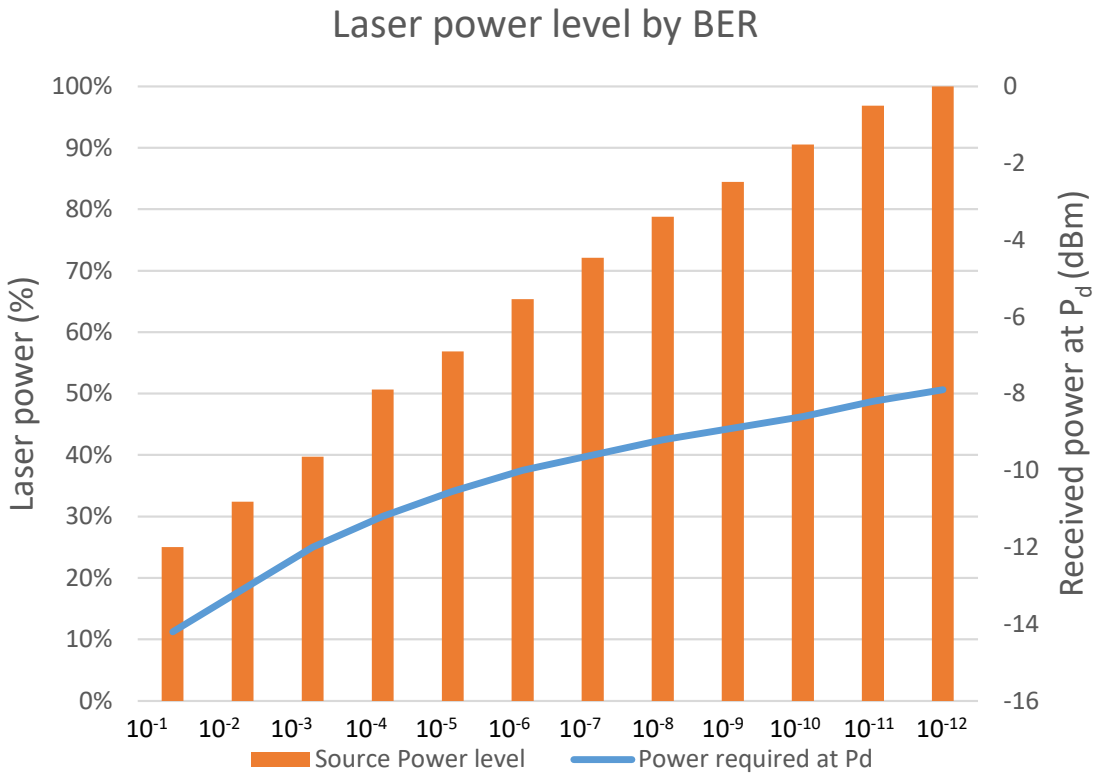


Figure 2.10 – Laser power level by BER

2.5 Quality degradation case study of Benchmark applications

In this section, we first investigate how the bit truncation and bit approximation impact on the QoR. Then, we analyze the QoR results when bit approximation and bit truncation are combined together.

2.5.1 Truncation impact on QoR

The truncation is the most straightforward technique to save energy, as certain amount of bits are not sent. However, it may also critically affect the output of application hence degrade the QoR. Therefore, we have to carefully monitor this approximate technique in order not to lose control over approximate level

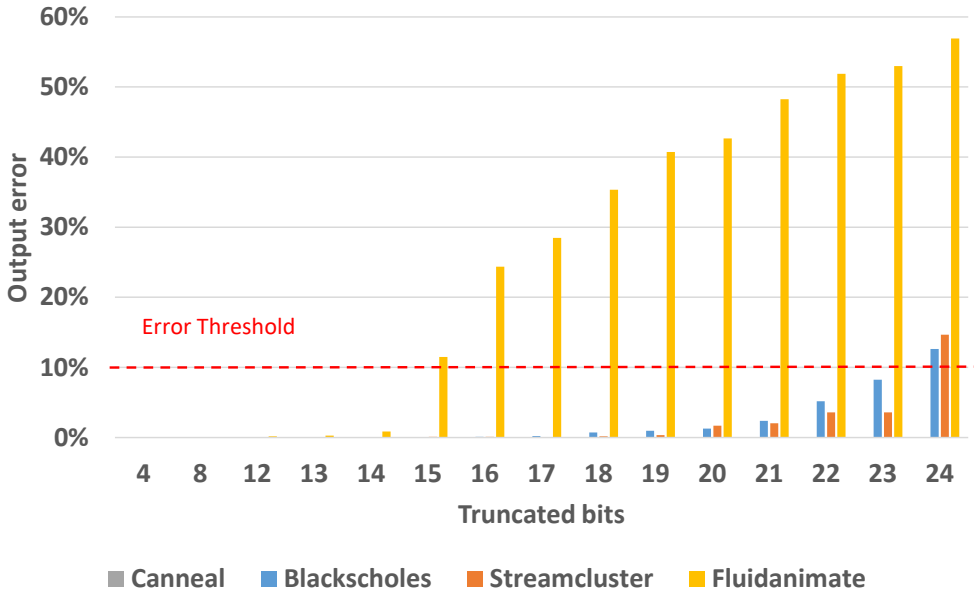


Figure 2.11 – Application Output error by Truncation bits

Figure 2.11 illustrates the output error effects when only bit truncation is applied on FP data, starting from the LSB bit. It has to be noticed that we increased the number of truncated bits four by four up to 12 bits, as below that value no effect has been sensed, then we increased this number one by one up to 24 bits, that exhibits too much error for every application. We have set an output error threshold lined at 10% which is widely used as an acceptable error range for approximate computing [121, 115, 105, 122]. From the

analysis, it is certain that applications differently react to a given approximation level. As an example, Canneal does not generate any noticeable output error with truncation. This can be explained with proportion of Floating-point data in application, as Canneal contains only 4% of FP data usage (Figure 2.9). Approximation on very small quantity of FP data does not critically impact on the whole application. Blackscholes and Streamcluster show low output error similar to Canneal until truncation of 23 bits. However, at 24 bits truncation, the produced output errors exceed the 10% threshold. Fluidanimate, meanwhile, produces a large output error, quickly reaching the 10% threshold, starting from 15 truncated bits. This application contains 22% of FP data which is quite low compared to Blackscholes and Streamcluster. However, unlike Canneal, Fluidanimate only tolerates a low level of approximation due to data sensitivity. For Blackscholes and Streamcluster, even if they contain 37% and 50% FP data, they still remain under acceptable error threshold as FP data in those applications are not sensitive and have less impact on results. This is due to data sensitivity of the data. Depends of how application is designed some data is more sensitive which mean critical to application quality. Approximating sensitive FP data critically affects output application. We can notice that all the applications suffer the 24 bits of truncation. Indeed, in FP single-precision presentation, LSB bits from 0 to 23^{th} are mantissa, and bits from 24^{th} to 31^{th} are exponent which has significant meaning for the data after the number normalization. Therefore, truncating 24^{th} and further bit truncation for all the applications will lead to a large amount of error.

However, for our experiments, we consider an ONI with 8 lasers serializing a group of four bits ($BSL = 4$) hardware design. Thus, it is possible to apply truncation by steps of 4. For instance, truncation of 20 bits or 24 bits is configurable in this architecture. Therefore, in our system, the most beneficial configuration just with truncation, under the 10% threshold, is 20 bits for A_{Trunc} (5 OFF state lasers) and 12 bits for A_{NotAx} (3 P_H lasers) for Blackscholes, Streamcluster and Canneal. For Fluidanimate, to remain under the threshold, we have to use 12 bits A_{Trunc} (3 OFF state lasers) truncation and 20 bits of A_{NotAx} (5 P_H lasers).

Even though truncation grants a large amount of power saving, it is not possible to have soft approximate level control. This limits the approximate exploration to reach further the power saving on optical lasers. To have more delicate control over the application level, our system proposes bit approximation. Instead of cutting off emitting laser (putting laser below the laser threshold), we gradually decreases the input laser power which leads to higher BER. This allows the approximation technique to adapt more progressively the

output error control.

2.5.2 Bit approximation impact on QoR

Bit approximation allows delicate control over approximate level for data approximation. As previously mentioned in section 2.1, bit approximation adjust approximate level with Bit width and targeted BER. In this subsection, we will see how bit approximate affect the QoR of each benchmark applications.

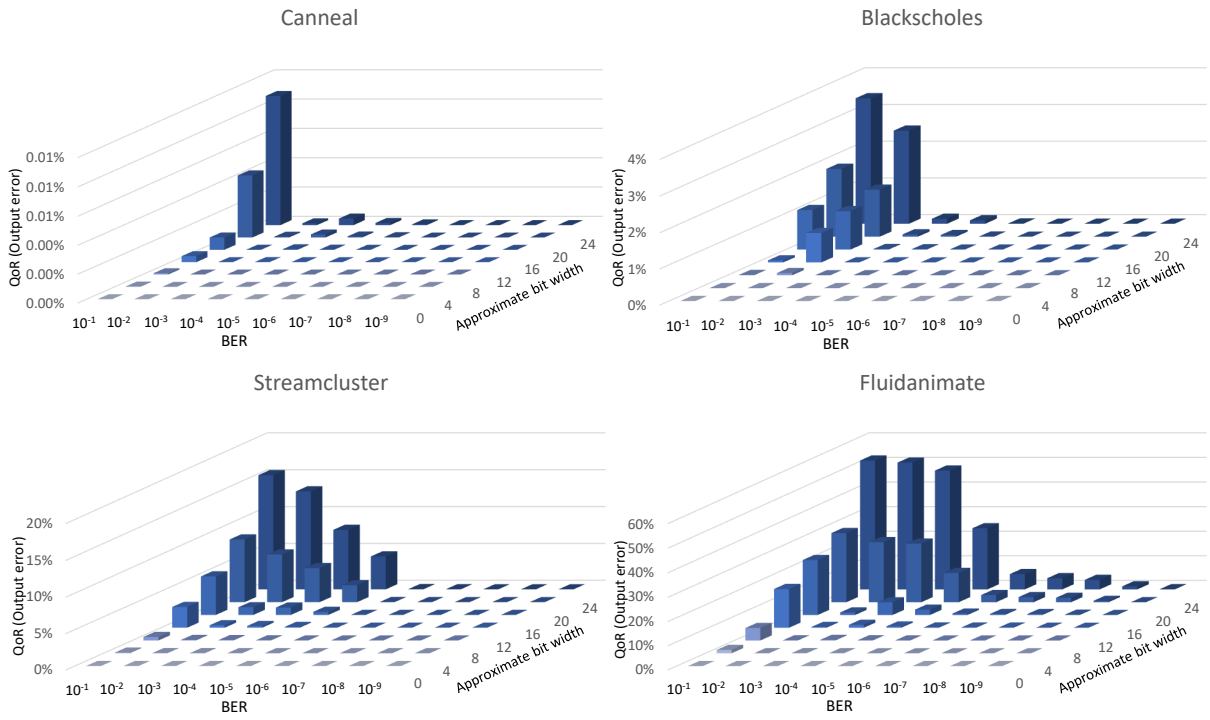


Figure 2.12 – Application Output error by BER and approximate bits

Figure 2.12, illustrates the QoR over different approximate bit width Ax on LSB from 0 to 24 bits and various approximate BER (BER_{ax}) from 10^{-1} to 10^{-9}. From this figure, we can notice that Canneal application does not much impact from approximation due to low Floating-point data in the application. However, the other applications present gradual increase in output error with both BER_{ax} and approximate bit width Ax. Interestingly, not like truncation, application such as Blackscholes and Streamcluster can have output error rate below 10% for Ax of 24 bits with BER_{ax} of 10^{-1} and 10^{-3} respectively. In case of Fluidanimate, there are many configuration that produce lower than output error

threshold (10%), such as Ax of 20 bits and BER_{ax} of 10^{-5}. Compare to truncation, bit approximation is less aggressive even on bits close to MSB for results degradation.

Moreover, there are different configurations that yield similar output error for same application. For instance, in Streamcluster, the configuration with Ax of 16 bits and BER_{ax} of 10^{-1} has output error around 4.6% similar to other configuration Ax of 20 bits and BER_{ax} of 10^{-3} and Ax of 24 bits and BER_{ax} of 10^{-5}. In order to make more profitable configuration, we have to consider power saving aspect. The detail of this comparison will be treated in the section 2.6.1

This moderate output error control allows us to go further for power consumption optimization by reaching larger bit area where under approximation or truncation. For instance, Fluidanimate does not benefit much from Truncation technique as we can turn off only 3 lasers to meet QoR. However, it does not indicates that we cannot go further approximate than 12 truncated bits. Therefore, we present the optimal mix of two approximations technique in data transmission in following subsection.

2.5.3 Mix of approximated bit and Truncation

In the previous subsection 2.5.1, we found that truncating bits close to MSBs produces large amount of errors, in particular when the truncation is done on bits coding the FP exponent. In section 2.5.2, we saw that bit approximation produces more moderate output error compare to truncation by adjusting bit area Ax and target approximate BER_{ax}. In this context, to push further the energy reduction and flexibility of the approximate computing implementation on optical interconnect, we propose mix of two approximate techniques (truncation and approximation) while adapting the application output error.

Table 2.6 – Bit area configuration

Application	Approximate Technique	
	Truncation	Bit approximation
Canneal	24 bits	24bits
Blackscholes	20 bits	24bits
Streamcluster	20 bits	24bits (BER_{ax} 10^{-3})
Fluidanimate	12 bits	20bits (BER_{ax} 10^{-5})

The table 2.6, summarize the maximum bit configuration with hardware setup from section 2.4 ($N\lambda = 16$ and BSL=4). In subsection 2.5.1, we saw that the maximum point

where we can reach with truncation was 20 bits and starting from 24 bits most of application reached out the output error threshold. Moreover, further bit truncation exploration leads to a large amount of error due to exponent truncation.

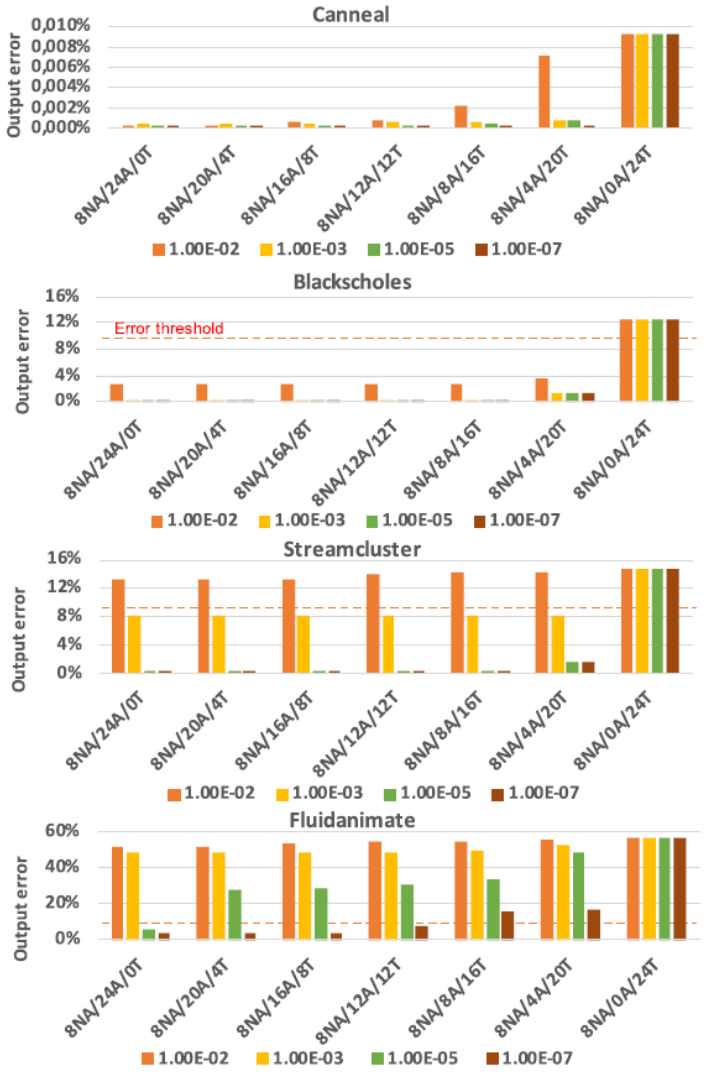


Figure 2.13 – Application Output error of mixed approximate techniques

Figure 2.13 present how combinations of both truncation and approximation techniques impact the QoR of applications. In this graph, we present the output error with respect to different bit range parameters: Not Approximated bits (NA), Approximated bits (A) and Truncated bits (T). The explored combinations range from 8 MSBs not approximated (sign and exponent of 32 bits FP data are not approximated) and the re-

maining bits approximated (noted as 8NA/24A/0T), to 8 MSBs not approximated and only truncation (noted as 8NA/0A/24T). The intermediate values are increased by steps of 4 more truncated bits instead of 4 approximated bits. The step 4 is linked to the bit stream length ($BSL = 4$ in our experiments). Regarding the approximated bit levels, we explored the aforementioned combinations with different BER_{Ax} values: 10^{-2}, 10^{-3}, 10^{-5}, and 10^{-7}.

Firstly, from the results shown in Figure 2.13, we can clearly see that the combination of mixing truncation and approximations provides a better QoR, for any BER value, than only 24 bits of truncation (represented by 8NA/0A/24T). As already explained in the previous section, applying 24 bits truncation impacts the first bit of the exponent within the Floating-Point number. Secondly, we can see that increasing BER value decreases the QoR, due to more error on bits. For instance, if we consider Fluidanimate, for 24 approximated bits (i.e. 8NA/24A/0T), the QoR goes from 2.5% to 5%, and from 48% to 50% of output error, respectively with a BER increasing from 10^{-7} to 10^{-5}, and from 10^{-3} to 10^{-2}. This is also observable for any combination of approximation with truncation. Lastly, we can notice that, for any BER value, the QoR is more degraded as the number of truncated bits increased. This is especially visible for Fluidanimate which is more sensitive to errors: for a BER value of 10^{-5}, the output error goes from 5%, to 58%, while passing by 30%, respectively with only 24 approximated bits (8NA/24A/0T), only 24 truncated bits (8NA/0A/24T), and half approximated and half truncated (8NA/12A/12T). Regarding the best combinations to keep an acceptable output result in table 2.7:

Table 2.7 – Mixed Bit area configurations

Application	Approximate Technique		
	Truncation	Bit approximation	Mix (truncation & bit approximation)
Canneal	8NA/0A/24T	8NA/24A/0T	8NA/0A/24T
Blackscholes	12NA/0A/20T	8NA/24A/0T	8NA/4A/20T (BER_{ax} 10^{-1})
Streamcluster	12NA/0A/20T	8NA/24A/0T (BER_{ax} 10^{-3})	8NA/4A/20T (BER_{ax} 10^{-3})
Fluidanimate	212 bits	12NA/20A/0T (BER_{ax} 10^{-5})	8NA/12A/12T(BER_{ax} 10^{-7})

Canneal can go for 24 truncated bits. As it is not a Floating-Point sensitive application, the combination of techniques does not provide any advantage. Blackscholes exhibits less than 10% of errors for any combination except for 24 truncated bits. Hence, it can even apply strongly degraded BER of 10^{-2}. As it is evident that truncation of 4 bits consumes less power than 4 approximated bits, we may expect that a combination with 4 approx-

imated bits and 20 truncated bits is the best solution if we target the less consuming solution with an acceptable QoR (8NA/4A/20T with a BER of 10^{-2}) for this application. A complete analysis on QoR and power savings is presented on the next section. Streamcluster has somehow the same behavior as Blackscholes, however the BER should be better than 10^{-2}. Hence, the best combination should be for 8NA/4A/20T with a BER of 10^{-3} Fluidanimate is more sensitive to data errors. For any combination, if the BER is worse than 10^{-5}, the output error is more than 40%. The combination with the higher number of truncated bits providing less than 10% of output error is for 12 approximated bits at a BER of 10^{-7} along with 12 truncated bits (8NA/12A/12T).

From these experiments, we can conclude that every application reacts differently to the approximation and truncation. However, we can see three trends:

— First, some applications are not really sensitive, like Canneal. However, this benchmark only has less than 5% of Floating-Point data in its communications. Hence, the opportunity for energy saving is not relevant.

— Secondly, some benchmarks are very sensitive to approximate, like Fluidanimate. Hence, this kind of application requires a specific and accurate combination of approximation and truncation.

— Finally, it is possible to find some applications with a non negligible amount of Floating-Point data that can be approximated using the same combination of parameters. This latter is the case for Blackscholes and Streamcluster.

From this analysis, the proposed method can be used for two way. First, application specific communication interconnect: adapting i) the number of approximated and truncated bits and ii) the BER for every application. However, this fine grain and complex configuration requires significant number of hardware control and it give overhead on ONI that will have to adapt on run time to the application needs. Secondly, generic interconnect for approximation: it is possible to determine one combination of truncated and approximated bits that suits the most of approximable applications. However, by going in that way, the ONI will not be able to approximate some applications, then the kind of application will be run using non approximated communication and not profit any power saving.

2.6 Results Power Saving

This section analyzes the power saving based on i) Floating point data approximate model, ii) the communication traffic between processors and memories by data type, and iii) the level of approximation and truncation with targeted BER for approximate and accurate communication.

2.6.1 Power saving

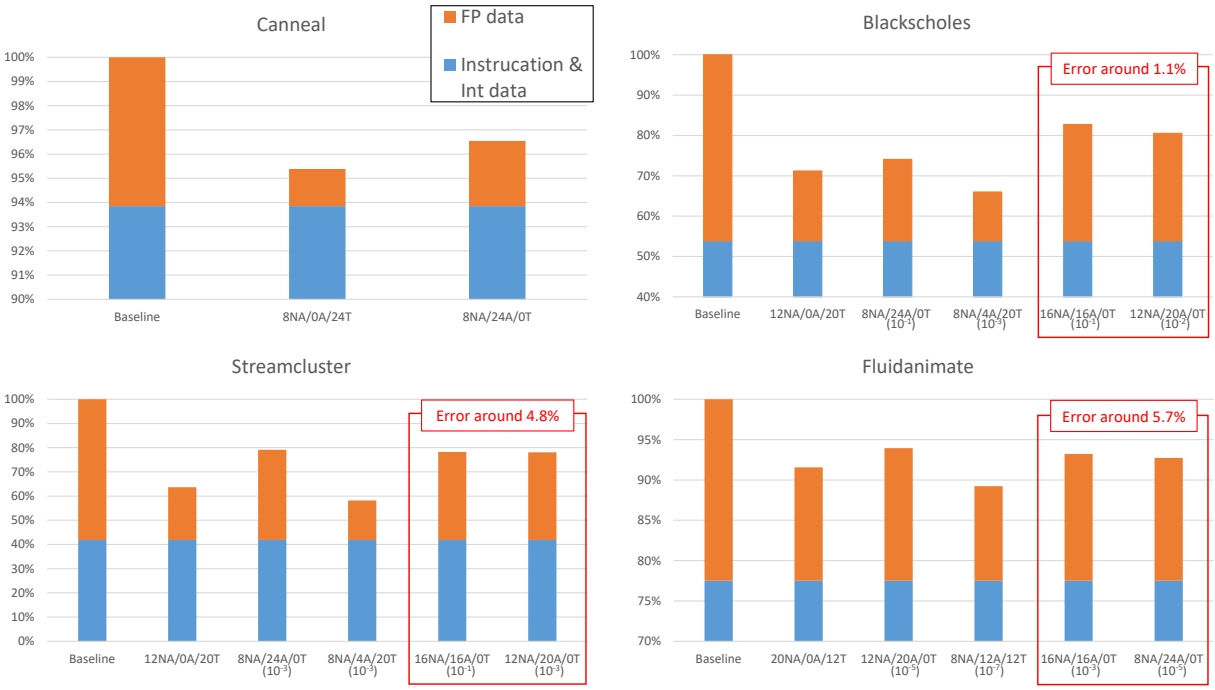


Figure 2.14 – Power saving of proposed techniques

Figure 2.14 shows power consumption for optimum configuration from previous sections 2.5.1, 2.5.2 and 2.5.3. All the power consumption of each applications are compared with their Baseline configuration which uses only highest laser power configuration P_H for accurate communication. Moreover, data type is presented in different color to clarify the approximable data and non-approximable data i.e. floating-point data are represented with orange, while the rest of communication (integer data and instructions) are represented with blue. In Figure 2.14, Canneal application achieves power saving around 4.5% and 3.4% from 6% of floating-point data communication with very low output error for

configurations $8NA/0A/24T$ and $8NA/24A/0T$ respectively. Applications that contains more floating-point data like Blackscholes, Streamcluster and Fluidanimate produce more power saving. Blackscholes can save around 29%, 26% and 32% of power for configuration of $12NA/0A/20T$, $8NA/24A/0T(BER_{ax} = 10^{-1})$ and $8NA/4A/20T(BER_{ax} = 10^{-3})$. In Streamcluster, we can observe more significant power saving: around 36%, 20% and 41% with $12NA/0A/20T$, $8NA/24A/0T(BER_{ax} = 10^{-3})$ and $8NA/4A/20T(BER_{ax} = 10^{-3})$. In case of Fluidanimate, even-though application involve a large amount of floating-point data communication, it yields trivial power saving due to low approximate level configuration. From previous sections 2.5.1, 2.5.2 and 2.5.3, we marked that Fluidanimate reacts sensitively to data approximate. Therefore, to be under 10% output error, the only possible approximate level configuration was low: $20NA/0A/12T$, $12NA/20A/0T(BER_{ax} = 10^{-5})$ and $8NA12A/12T(BER_{ax} = 10^{-7})$ and these configurations offer 8%, 6% and 11% respectively.

Moreover, in section 2.5.2, we discovered that there are several approximate level configurations that produce similar amount of output error. For instance, in Blackscholes, approximate level $16NA/16A/0T(BER_{ax} = 10^{-1})$ and $12NA/20A/0T(BER_{ax} = 10^{-2})$ produces around same output error 1.1%, but in power consumption graph 2.14, we can see $12NA/20A/0T$ $(BER_{ax} = 10^{-2})$ saves slightly more power than $16NA/16A/0T$ $(BER_{ax} = 10^{-1})$ which has less area where approximate technique is applied. This same trend can be seen in other applications.

From these experiments, we can conclude several interesting aspects that leads to power saving management in approximate optical interconnect:

1. The configuration that mixes two approximate techniques; truncation and approximation reach out further power saving while remaining under output error threshold.
2. When application contains plentiful amount of floating-point data, the power saving can be meaningful with suitable approximate level configuration.
3. Higher (more aggressive) approximate level tends to produce more compelling power saving.
4. Even with large amount of Floating-point data, power saving can be minor when number of lasers that are under approximate is low e.g. Fluidanimate application.
5. With similar output error, the number of lasers that are under approximation enacts as a more relevant aspect that might bring more power saving than targeting higher BER.

2.7 Conclusion

In this work, we have investigated the design of an Optical NoC supporting the approximate data transmission. The proposed method 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 allows configuring, at run-time, the approximate level and the laser output power. Moreover, we proposed the implementation of approximate interconnect in 3D optical NoC architecture with data type classification under cache coherence traffic. For the performance analysis, several Benchmarks which have different floating point data ratio are studied to see the approximatable traffic in the network. According to the packet trace in NoC, we explored power consumption with different laser power by several BER and bit area configuration. 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 Streamcluster application while keeping the output error threshold.

Chapter 3

Distance-aware approximate optical interconnect

3.1 Idea Overview

In previous chapter 2, we proposed an approximate nanophotonic interconnect which improves the power efficiency in NoC. In the proposal, the source lasers are configured at heterogeneous power levels for data communication. These multi power levels allow to save power on LSBs. However, the classification of power levels is set only according to data type and it does not involve other optical signal propagation loss parameters. When optical signals propagate in waveguide, they face various losses. Moreover, these losses have different values for all the destinations due to different distances and numbers of MR crossed. In other words, in SWMR, to efficiently configure the input laser power at the source, it is necessary to consider the distance parameters to involve different loss values. Therefore, in this chapter, we explore an approximate nanophotonic interconnect to improve power consumption based on the classification of communications thought the distance between the source and destination.

Our proposal aims to manage, for a given application, power consumption regarding the Quality of Result (QoR). For this purpose, the proposed nanophotonic interconnect allows adapting the laser power level according to the communication distance and the transmitted data type. In this proposal, we used the approximate optical link presented in section 2.3.1 which consists of SWMR link with WDM waveguide channel and MR modulators. For this approximate optical link, a communication classification is done according to data type of the transiting packet. In this proposal, laser power are configured with type of data and communication distance between source and destination. When the transmission is triggered, lasers are turned ON with the appropriate power levels depending on i) the required communication quality for the specific data type and ii) the communication distance. To allow a reduction in consumption while limiting the driver

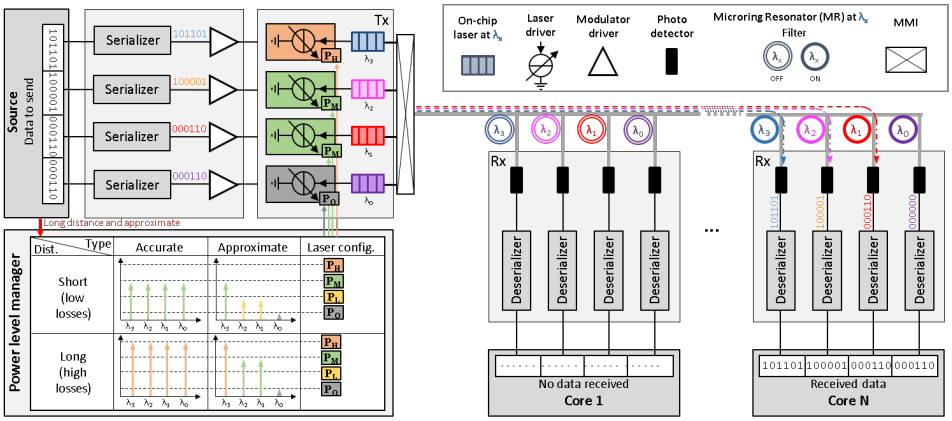


Figure 3.1 – Proposed Optical link Overview

complexity, we propose to use four laser power levels: high power (P_H), medium power (P_M) low power (P_L) and off (P_{Off}), with $P_H > P_M > P_L > P_{Off} = 0$. Figure 2.4 on page 43 presents an example assuming that the λ_3 is sent with high power, λ_2 and λ_1 are sent with medium power, while the power of λ_0 is equal to P_{Off}, meaning this wavelength is not used, which corresponds to 4 LSB bits truncation. The laser power values are calculated based on technological parameters (e.g. waveguide propagation loss and MR losses), system level parameters (e.g. number of readers) and the targeted BERs, as detailed in Section 3.3.1.

We assume that accurate data (i.e. data that cannot be approximated) and approximated data require, on the receiver side, BER_{NotAx} and BER_{Ax} respectively. Hence, assuming $BER_{NotAx} < BER_{Ax}$, approximated data is transmitted using lower laser power compared to accurate data. As previously explained, the actual laser power depends on the propagation loss and hence the communication distance. For this purpose, we define two ranges of communication distance, *Short* and *Long*, depending on the destination position on the SWMR link. This leads to the four scenarii detailed in the following:

— **Accurate & Long** scenario is used for the transmission of sensitive data (e.g. instructions and integer data) that critically impact the execution of applications. Therefore, the transmission quality is protected by targeting low BER for all the bits (e.g. $BER_{NotAx} = 10^{-12}$). To compensate the losses induced by the long-range transmission, the highest power level is used (i.e. P_H) for all the lasers.

— **Accurate & Short** scenario is also used to transmit sensitive data but at a shorter distance. Due to the lower losses, low BER is obtained for all the bits by configuring

the lasers power level to P_M.

— **Approximate & Long** scenario allows transmission of approximate data for long distance. Depending on the approximation scheme, the lasers are configured with different power levels. As detailed in Section 3.3.2, most significant bits are transmitted with low BER, intermediate bits are transmitted with higher BER and least significant bits are truncated. To achieve this, the lasers that emit signals transmitting most significant bits and approximated bits are configured to P_H and P_M respectively. Truncated bits are not transmitted and the corresponding lasers are thus set to P_{Off}.

— **Approximate & Short** scenario allows transmission of approximate data for short distance. As for the previous scenario, the configuration of the lasers depends on the bit's significance. However, as the distance between source and destination is small, it induces lower losses allowing a power reduction for the emitted signals. This leads to use the P_M and P_L power levels for most significant and intermediate bits respectively.

By adapting the laser powers according to both communication distance and level of approximation, our interconnect aims at minimizing the energy to transmit data bits. Since only three power levels are needed for all destinations with approximation levels, the design complexity of the laser driver remains comparable to solutions which require individual computation of power levels for each destination (*Distance Proportional*). The proposed method to define the laser power levels is presented in 3.3.1.

3.2 Distance aware optical interconnect

In this section, we explain the principle of distance-aware communication. Then, we present how much distance-aware power level configuration can bring the power saving.

3.2.1 Distance aware communication

In the section 2.3.3, we saw how the power levels are defined according to the loss in the waveguide and the targeted BER. The power levels can also be determined by the distance, as the optical loss varies according to the waveguide length and the number of MRs crossed. Figure 3.2 represents how the power levels are configured according to the distance aware techniques. When the power levels are configured according to each

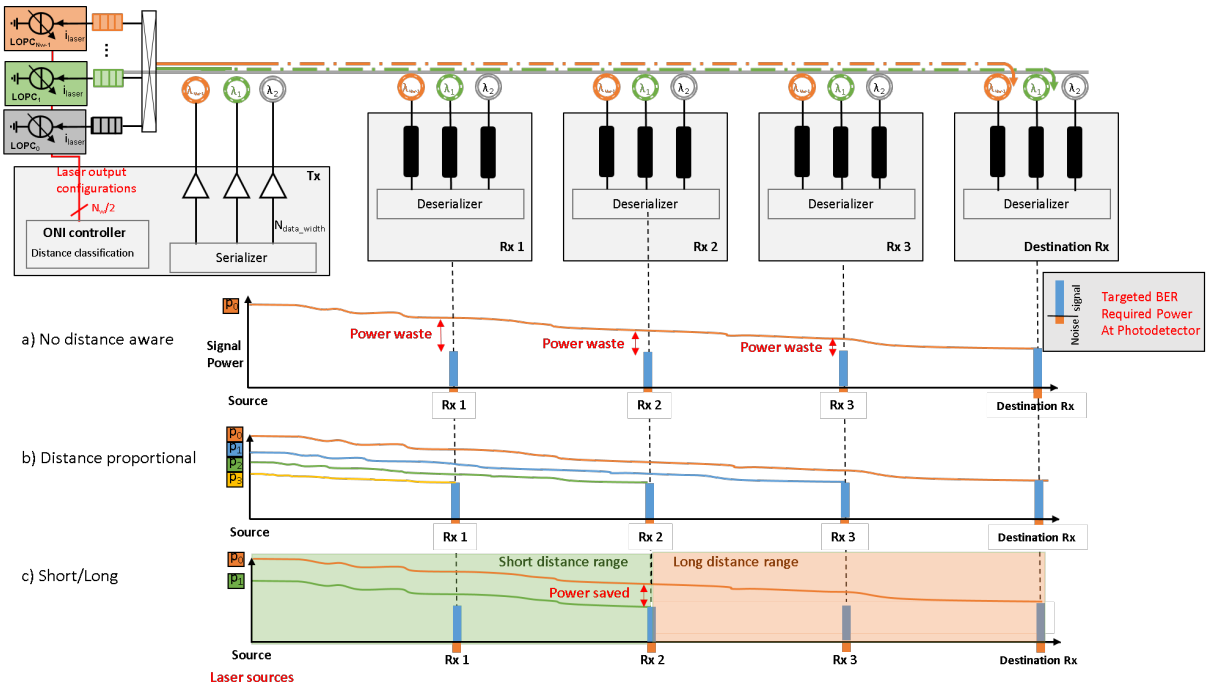


Figure 3.2 – Distance-aware techniques a) No distance aware b) Distance proportional c) Short/Long distance aware

distance to destination in Figure 3.2.a, the ONI controller used only one power level for all the destinations (*No distance aware*). It offers straightforward and very simple power level control, but there is a lot of power waste. Figure 3.2.b presents the distance proportional power level configuration. This configuration has very fine grain distance aware optical interconnect to reduce the power consumption of optical interconnect [16]. This configuration can optimize all the optical power wastes for each destination by adopting input power levels according to the different losses in the waveguide. This might be the best performance in terms of power, but it requires multiple power levels and leads to high hardware overhead for the laser controller. Figure 3.2.c illustrates the *Short/Long* distance aware, which has only two power levels that compensate optical loss due to distance from source to the last destination and to one intermediate receiver. As only two destinations exist for loss configuration, the laser control remains simple and still saves power. Even of this configuration is not the optimum, it offers interesting results when we consider the laser controller overhead and power consumption. The detail of comparison is presented in the section 3.4. In the following, we present the power consumption of each configuration to compare and check the optimum threshold for *Short/Long* distance

configuration.

3.2.2 Power consumption of Distance aware optical communication

Figure 3.3 illustrates the normalized power consumption of distance aware techniques. In this figure, no distance aware has been used for normalization to compare techniques and different threshold configurations. As we can see, *Distance proportional* from [16] can save 34% of power compare to no distance aware interconnect. The yellow bars in the histogram graph present *Short/Long* distance-aware technique with different thresholds for distance classification. For instance, threshold 8 indicates that all the destinations from source to 8^{th} destination are considered short-distances and leads to specific configuration of lasers with associate power. From 9^{th} to last destination 15^{th}, long distance is considered and laser level is set to ensure correct communication with the 15^{th} destination. Among *Short/Long* distance aware, the configuration, which sets center node 8 as a threshold, offers the optimum power-saving 25% compared to no distance-aware. Evidently, choosing center node as the threshold for *Short/Long* allows to well-balanced amount of power saved and waste.

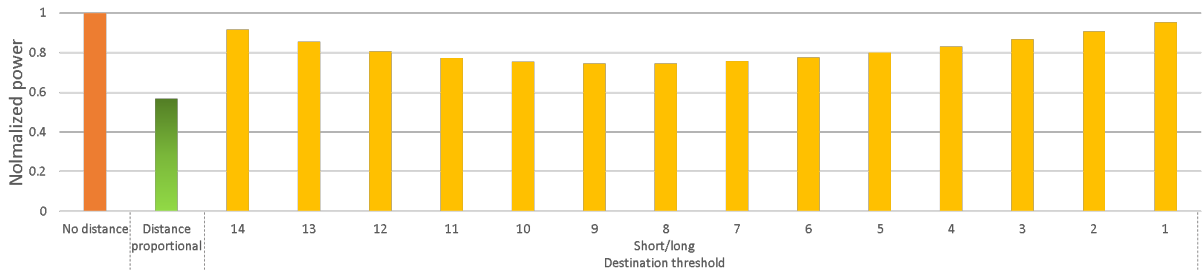


Figure 3.3 – Power consumption distance aware

3.3 Approximate distance aware optical communication

In section 3.2, we saw how distance-aware optical interconnect can save power by defining the laser power levels according to distance loss. In this section, we present distance

aware technique on top of approximate interconnect. Both distance-aware and approximate interconnect techniques allow power saving throughout a subtle laser power level control. When we integrate distance-aware on approximate interconnect, the laser power levels can be configured for both techniques. The power level can be used individually, but can be also overlapped to optimize the number of power levels. In other words, the power level, which targets approximate BER, might also adopt distance loss configuration. In the following, the method which defines the distance threshold for *Short*/*Long* classification with approximate communication is presented. Then, the proposed data schemes are illustrated with their power gains.

3.3.1 Distance aware threshold definition on approximate interconnect

As presented in section 2.3.3, the input laser power levels are defined with targeted BER (BER_{NotAx} and BER_{Ax}) which correspond to power levels P_H and P_M.

Similarly to method from section 2.3.3, Figure 3.4 illustrates power level definition for both approximate and distance aware techniques. In this figure, three lines represent the received optical power in hop count distance at each destination. The lines have been drawn with different input laser powers to adjust the required power for both distance and BER.

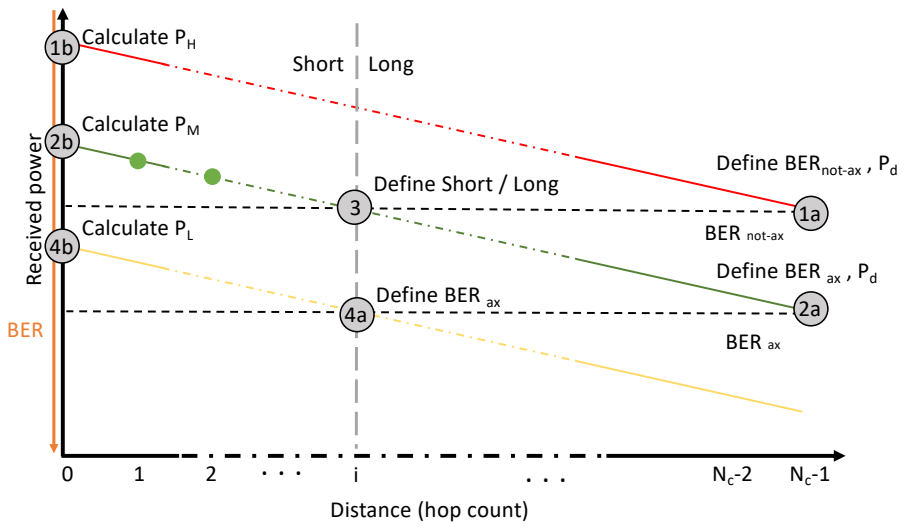


Figure 3.4 – level computation regarding the *Short*/*Long* distance.

To define the three different lines, we start with the required powers at receiver P_d for the targeted BER_{NotAx} or BER_{Ax}, depending on the data types (accurate or approximate). These points are marked (1a) and (2b) respectively in Figure 3.4). Once we know the required power at receiver and the loss on waveguide, we can calculate the input laser power P_s throughout equation 2.1. These laser powers are illustrated by the marks (1b) and (2b), and allow to define the input laser powers P_H and P_M that have different communication accuracies: accurate and approximate communications for all the destinations in the architecture.

As represented by the green line, the optical signal power for approximate communication remains high enough to reach an accurate communication for short-range communication. From this observation, we define a limit between short-range and long-range communications, mark ③. This point corresponds to the highest hop count allowing laser at P_M level to reach destination with BER_{NotAx} quality. This leads to a laser power reduction for short-range communication while maintaining robust data transmission. Finally, to efficiently transmit approximate data, we define P_d (mark (4b)) as the optical power allowing to reach BER_{Ax} for all short-range communication, mark (4a).

From this graph, we illustrate how distance aware threshold ③ is defined. Based on this threshold, the ONI can classify the destinations as short or long communications, and it can used the power level according to data type:

— P_H is used for accurate communications for long distance ;

— P_M can be used for two cases: approximate communication for long distance and accurate communication for short distance ;

— P_L is defined for approximate communication for short distance.

3.3.2 Distance aware approximate interconnect data schemes

Previous sections present how power levels are defined for accurate and approximate communications. This section presents various scenarii with multiple power levels and power saving of each configuration.

Figure 3.5 illustrates six possible communication scenarii to save energy for the transmission of 32 bits data. For this scenario, 8 wavelengths (λ_7 to λ_0) are used to simultaneously transmit 8 groups of 4 serialized bits. Each scenario is noted as $xNA/yA/zT$ with: x Not Approximated, y Approximated, z Truncated bits. Figure 3.5.a presents the long distance communication. First, Figure 3.5.a.1 illustrates a scenario where only ap-

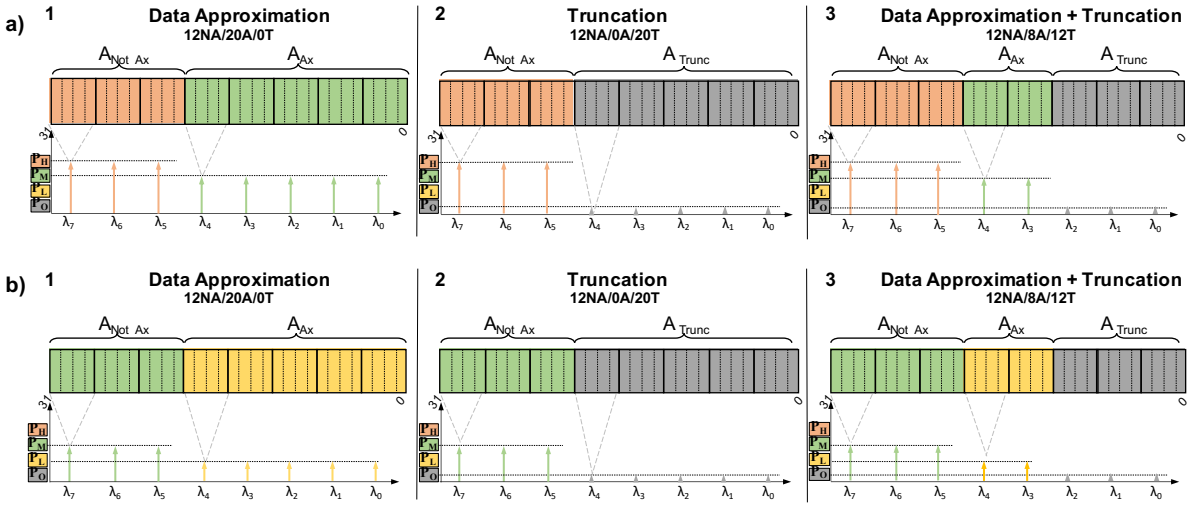


Figure 3.5 – Laser power level configurations to transmit approximated floating-point data assuming a) Long distance and b) Short distance communications. $N_{data} = 32$, $N\lambda = 8$ and $BSL = 4$. xNA/yA/zT with x Not Approximated, y Approximated, z Truncated bits

proximation is used, assuming $A_{NotAx} = 12$ and $A_{Ax} = 20$. This scenario is noted as $12NA/20A/0T$. This leads to configure lasers from λ_7 to λ_5 and from λ_4 to λ_0 configured with P_H and P_M respectively.

Second scenario, Figure 3.5.a.2, $12NA/0A/20T$, involves only truncation [16] which refers to $A_{NotAx} = 12$ and $A_{Trunc} = 20$ and it leads to laser configuration P_H for λ_7 to λ_5 and P_{Off} for lasers at λ_4 to λ_0.

Third scenario, Figure 3.5.a.3, $12NA/8A/12T$, is a combination of approximation and truncation with $A_{NotAx} = 12$, $A_{Ax} = 8$, and $A_{Trunc} = 12$ which leads to P_H, P_M and P_{Off} for lasers from λ_7 to λ_5, λ_4 to λ_3 and λ_2 to λ_0 respectively.

When these 3 scenarii are applied to short distance communication Figure 3.5.b, the input laser power can be decreased accordingly ($P_H \rightarrow P_M$ and $P_M \rightarrow P_L$).

3.3.3 Distance aware approximate interconnect power saving

Compared to the scenario where all bits are always transmitted at full power, regardless of the communication distance, our approach allows saving power, as detailed in Equation 3.1. To calculate power gain, power consumption from proposals P over Baseline scenario $\lambda \times P_H$ has to be computed. For baseline laser power is configured as P_H for all the wavelengths to guarantee accurate communication for all the distances, equation 3.1a. Moreover, in equation 3.1b, to obtain power consumption from proposal, power con-

sumption on three bit areas (defined in section 3.3.2) have to be computed. First, MSBs of FP $[N_{data} - P_{Ax}]$ is configured with P_H or P_M level according to the Short and Long distance. Secondly, approximate bit area $[P_{Ax} - P_{Trunc}]$ is configured with P_M or P_L to target approximate communication. Last, for the lasers that are in charge of Truncation bit area $[P_{Trunc} - 0]$ are set in off state P_{Off}. The summation of all the power consumption from those bit areas defined the total power consumption for the proposal. The actual power saving depends of bit area configuration, targeted BER and distance configuration.

$$PowerGain = \frac{P}{N\lambda \times P_H} \tag{3.1a}$$

$$\begin{aligned} \text{with } P &= \frac{1}{BSL}\Big[(N_{data} - P_{Ax}) \times P_1 \\ &\quad +(P_{Ax} - P_{Trunc}) \times P_2 \\ &\quad +P_{Trunc} \times P_{Off}\Big] \\ \text{and with } (P_1, P_2) &= \begin{cases} (P_H, P_M) \text{ for long distance,} \\ (P_M, P_L) \text{ for short distance.} \end{cases} \end{aligned}$$

3.4 Laser design complexity

Before getting into the results of the proposed optical link and the method to save power, we first need to understand the limitation of laser driver complexity in terms of the number of lasers, area, and power consumption for various architecture sizes and the number of channels. In this section, we propose a comprehensive study on the design complexity of the drivers for our method and Lorax. The study includes the multi level laser design complexity by comparing the work of [16], which proposed the fine grain power levels for distance, and approximate optical interconnect.

3.4.1 Laser design complexity by number of controllers

As discussed in the previous section, fine grain tuning of the optical signal power, as proposed in [16], allows reducing the laser power consumption. However, such approach

leads to significant challenges related to the design of both analog and digital electronic circuits to control the lasers. In [16], the authors consider a given approximation level for each application. This implies that each driver is capable of targeting, for each destination, any BER from 10^{-12} to 10^{-2}, so 11 power levels are necessary just to satisfy the quality of communication. Hence, for a N-cluster architecture, each driver is capable of emitting a signal at $11 \times (N-1)$ different power levels (e.g. 165 levels per laser for 16 clusters).

Since no hardware evaluation is carried out in [16], we study their driver complexity assuming only two BER values of 10^{-12} and 10^{-2} for each destination, while still assuming that the optimal BER can be used. For example, the architecture size 2×2 requires 7 power level configurations: 2 power levels targeting different BERs for each destination $2 \times (4-1)$ and one off state P_{Off}.

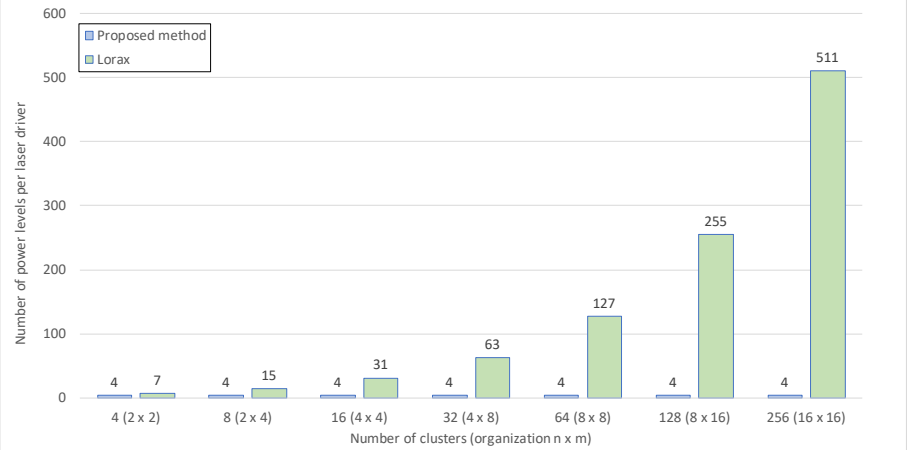


Figure 3.6 – Number of power levels per laser driver according to the network size for Lorax and our proposed method.

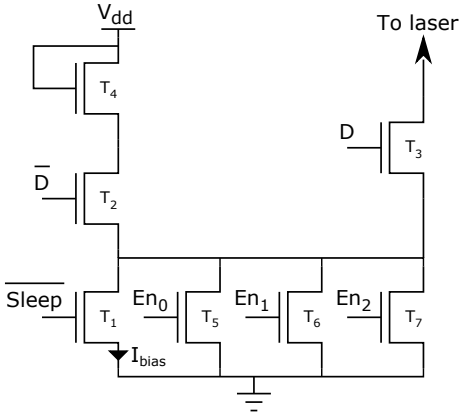
Figure 3.6 presents the number power levels for different sizes of architecture, from cluster size 2×2 to 16×16. As the architecture size expands, the number of power levels increases to match targeted BER for each destination.

The number of power levels for our proposal architecture is constant (4 levels corresponding to P_{Off}, P_H, P_M and P_L) while it linearly grows for *Lorax* despite the considered optimistic assumption. The results of *Lorax* configuration show that the number of required power levels is significant and not realistic for large architectures. For instance, a 16×16 architecture requires 511 power levels (1 level for P_{Off} and two BER levels for each of the 255 destinations), which leads to significant increase of the design complexity of the laser driver. To complete this complexity study, we explore in the next sections the

detail of laser driver in terms of power leakage and area according to the required power levels.

3.4.2 Laser design complexity according to number of transistors

In this section, we study the complexity of the CMOS drivers assuming an OOK modulation for both Lorax and our method. CMOS drivers are key components to deliver the right current to the on-chip lasers according to the selected optical power level. Figure 3.7 presents our proposed laser driver which has 4 power levels for distance aware approximate interconnect.



(a)

En_2	En_1	En_0	Current level	Power level
0	0	0	I_{bias}	P_{off}
0	0	1	I_L	P_L
0	1	1	I_M	P_M
1	1	1	I_H	P_H

(b)

Figure 3.7 – Laser drivers for 16 clusters architecture a) proposed laser driver model with 4 power levels corresponding to P_{off}, P_H, P_M and P_L) b) Combinations of inputs signals En_i to obtain a specific power level

The power levels in the driver are defined through a set of transistors. In our proposal, there are 4 power levels and 3 transistors T_5, T_6 and T_7 in the figure. These transistors are controlled by the signals En_0 to En_2, and this control enables to modify the current

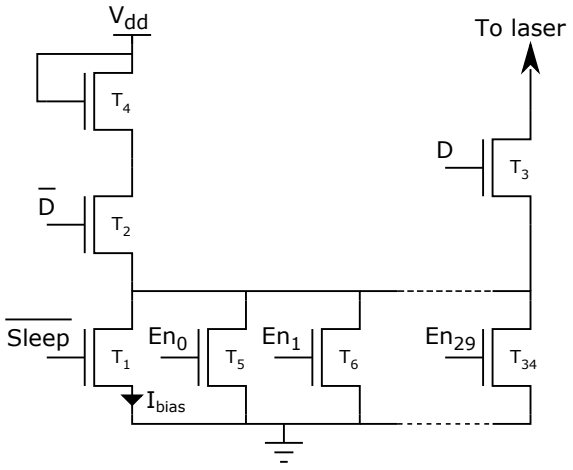
level, allowing to control the optical laser power. The drivers operate as following:

First, there is a transistor T_1 which drive a I_{bias} current to maintain the laser near the threshold. This transistor T_1 is activated with signal $/Sleep$ to distinguish the no communication and stand by state. In case a communication is required or predicted, the driver is set to stand by mode with signal $/Sleep$ setting to 1, allowing to start driving current and to prepare laser for emission. This allows reducing latency penalty occurring to switch on laser from a fully off state [123, 60]. In case no optical communication occurs or is predicted, Signal $/Sleep$ is set to 0, thus leading to energy saving. Temporal penalty and optimization of driver sleep time [124] are out of the scope of this work but would lead to the same energy reduction for both *Lorax* and our method. Secondly, when the communication must be established, which corresponds to transmission of data, the input signal D controls the laser emission through transistors T_2 and T_3. For the last, the transistors T_5 to T_7 are controlled with En_0 to En_2 to enable the multiple power levels. When the laser requires higher output power ($P_H > P_M > P_L > P_{Off}$), the signals En_i are activated to lead higher current.

Figure 3.8 presents laser driver model for 4×4 cluster architecture. The driver model is the same but to allocate more power levels, necessary to ensure fine grain distance configurations, more transistors are needed. The number of transistors increases with the number of laser output powers and the En_i are activated according to selected output power. As detailed in Figure 3.8.b), the more transistors driving current, the higher the laser output power. As we can see, the driver for *Lorax* requires 34 transistors to generate 31 laser power levels (15 destinations and 2 levels per destination, and P_{Off}). As a result, compare to our approach, which requires 7 transistors to generated the 4 power levels (i.e. P_H, P_M, P_L and P_{Off}), the *Lorax* requires 4.8× number of transistor count.

3.4.3 Laser design complexity on power leakage and footprint constrain

Table 3.1 summarizes driver estimations on leakage current and footprint for our proposed method and *Lorax*. For a $65nm$ technology node from ST Microelectronics and based on the estimation methods defined in [125, 126, 127], the leakage current for our method is $15nA$ against $125nA$ for *Lorax* (8× reduction). Regarding the footprint estimations, *Lorax* driver requires $8.89mm^2$, while our method only requires $0.25mm^2$ (35× reduction).



(a)

En_{29}	En_{28}	...	En_1	En_0	Current level	Power level
0	0	...	0	0	I_{bias}	P_{off}
0	0	...	0	1	I_1	P_{ax} Hop 1
0	0	...	1	1	I_2	P_{notax} Hop 1
...	...	...	...	...	...	...
0	1	...	1	1	I_{29}	P_{ax} Hop 15
1	1	...	1	1	I_{30}	P_{notax} Hop 15

(b)

Figure 3.8 – Laser drivers for 16 clusters architecture a) laser driver model of *Lorax*, 31 power levels corresponding to 15 destinations with 2 levels per destination and P_{off} state. b) Combinations of inputs signals En_i to obtain a specific power level.

	Number of Power levels	Transistor count	Leakage(nA)	Footprint(mm^2)
Proposed method	4	7	15.65	0.253
Lorax [16]	34	32	125.12	8.89
Lorax Increase (×)	**7.75**	**4.85**	**7.99**	**35.14**

Table 3.1 – Driver area estimations for a 16 cluster architecture.

Control bits En_i of the drivers are defined using a digital controller. The controller complexity directly depends on number of configurable laser power levels: while the 4-level driver used for our method requires a 2-input to 3-output controller, *Lorax* would require a 5-input to 30-output controller. In order to evaluate the area and power overheads, we designed the controllers at the RTL level using VHDL. The controllers were synthesized for a $65nm$ ST Microelectronics technology using Synopsis Design Compiler. We targeted a $1GHz$ frequency corresponding to core speed assumed in the previous experiment. Prime Time was used for power estimation. As seen in Table 3.2, the controller for our proposal needs only an area of $42\mu m^2$, thus leading to a 10× reduction compared to *Lorax* ($445\mu m^2$). Similar gain is observed for the power consumption: our controller consumes a total of $28.58\mu W$ (leakage $3.25\mu W$ and dynamic power $25.33\mu W$) compared to $290\mu W$ for *Lorax*.

	Power			**Area**
	Dynamic Power (μW)	Cell Leakage Power (μW)	Total power(μW)	Area (μm^2)
Proposed method	25.33	3.25	28.58	42.6
Lorax [16]	259.8	30.9	290.7	445.1
Gain	**10.2** ×	**9.5** ×	**10.1** ×	**10.4** ×

Table 3.2 – Driver controller synthesis results on $65nm$ STM technology node.

For a comprehensive comparison, we evaluate the total power consumption per laser for *Lorax* and our approach, assuming 16 clusters, BER of 10^{-3} for approximate communications and 33% laser efficiency [120].

Figure 3.9 presents the power breakdown (controller and laser) according to the communication distance. Our approach leads to power consumption of $368\mu W$ and $882\mu W$ for short-range communications (hop count 1 to 5) and long-range (hop count 6 to 15)

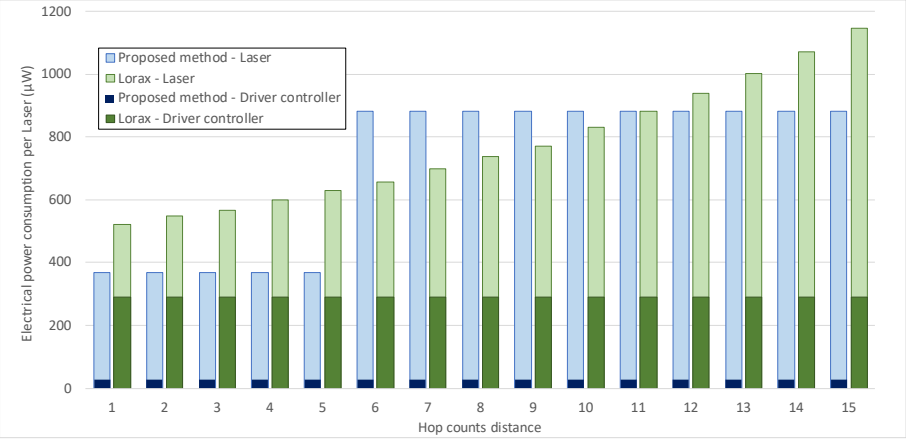


Figure 3.9 – Total power consumption per hop count for *Lorax* and the our method. Results are given for a 16 clusters architecture assuming 10^{-3} BER for approximate communications.

communications respectively. Regarding the power consumption for *Lorax*, it quadratically increases and leads to average power of i) $573\mu W$ for hop count 1 to 5 and ii) $873\mu W$ for hop count 6 to 15.. Hence, our method allows reducing power by 35% for short range communications while leading to only 2.5% power overhead for long range communications. Interestingly, our approach is more energy efficient for the longest communication range (hop count 12 to 15 in the figure). The significant gains obtained for short range communications is possible due to the lower static power consumption of the driver controller. The slight power overhead for long-range communication is acceptable considering the 10× reduction in the area footprint: while the drivers for *Lorax* would occupy a total area of $8.89mm^2$ for an architecture based on 16 clusters, our approach would take an area of $0.25mm^2$.

3.5 Experimental setup

In this section, the experimental setup for distance aware approximate interconnect is presented. It includes the hardware implementation of both distance and approximate interconnect for 64 cores 4×4 cluster architecture. Then, we explore the ONI design according to the system-level parameters used for our simulation. In the following, the error tolerant applications are analyzed for FP data communication based on distance hope counts.

3.5.1 Hardware implementation and communication model

For the experiment, same 3D hardware implementation from section 2.4.1 and cache coherent traffic from 2.4.2 are used (Figure 3.10.a. Based on this architecture, the proposed ONI configures the communication mode based on data type (approximate or accurate) and on communication distance.

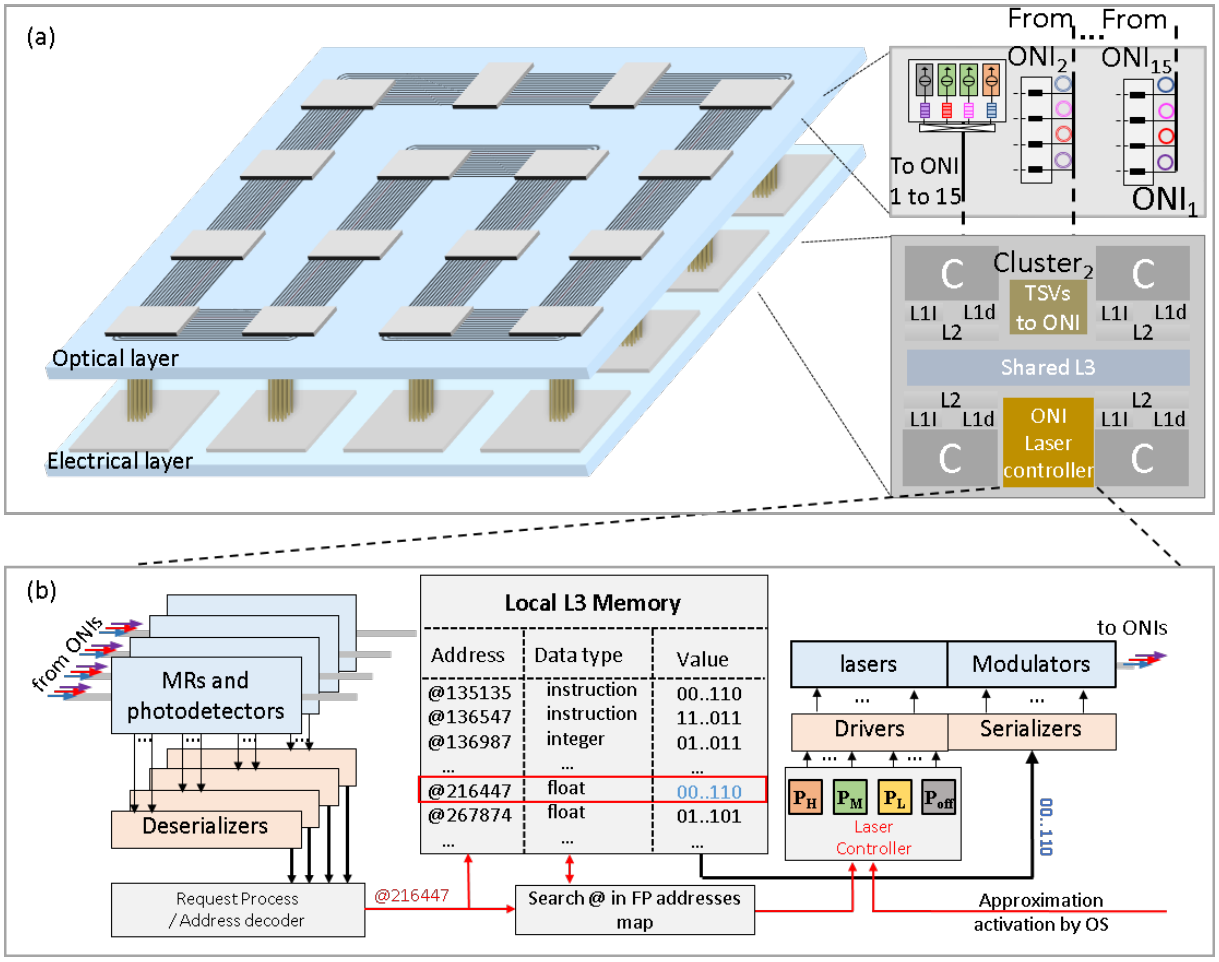


Figure 3.10 – Proposed 3D architecture and cache coherent system.

The architecture model has control flag coming from operating system to ”activate approximate mode”. In case approximations are disabled, all communications are carried without errors using P_H and P_M for long-range and short-range communications respectively. When approximate mode is activated, the lasers are configured as follows. First, we assume that i) accurate data and approximate data are located in different address

spaces of the L3 memory and ii) ranges of addresses associated to a given data type are stored in local tables. Then, from the packet destination ID, the ONI evaluates whether the communication distance is short or long. Both informations (accurate/approximate and Short/Long) allow the ONI manager to configure the laser power level from the laser configuration table. The details of communication between OS and ONIs is out of scope for this paper and it is studied in [113]. By considering all these parameters, Laser controller configures laser powers at the different levels (P_H, P_M, P_L and, P_{Off}). Thus, cores have two ways to communicate, one without error and one with approximation mode.

The Figure 3.10.b illustrates the same example of a cache miss scenario from section 2.4.2. The L3 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 levels of lasers corresponding to the approximated and truncated bits are configured.

To limit the hardware complexity, the proposed architecture only integrates one approximation and truncation configuration (see laser complexity section 3.4 for more details) and can be activated at run-time by an Operating System with a flag signal [113, 114]. The manager then grants the SWMR access to the L3 that initiated the request and the data is transmitted to the serializers for signal modulation. We assume the use of a packet header integrating the IDs of source and destination cores, data address, data type, and flag indicating if the transmitted data are approximated. The header is always transmitted with high laser power, i.e. without approximation.

3.5.2 ONI specification

In the following, we design the ONIs according to system level parameters defined in Table 3.3 and by considering $BER_{NotAx} = 10^{-12}$ and $BER_{Ax} = 10^{-3}$. Our goal is twofold: first we aim to define power levels P_H, P_M and P_L. Second, we aim to define the hop count from which long-range communications start.

We first assume minimum optical power of -8 dBm and -12 dBm for accurate and approximate data reception respectively [41]. We assume 0.02 dB MR through loss, 0.7 dB MR drop loss, [117] an FSR of $8nm$ and a MR quality factor of $20,000$ [118]. We consider distance between ONIs ranging from 0.8 cm to 1.3 cm and consider waveguide losses of 1 dB/cm [128] and 0.25 dB/cm [129]. While 0.25 dB/cm corresponds to ridge waveguides not compatible for side-coupling with microrings [130], our aim is to demonstrate the ability of the method to also consider prospective propagation loss and

Table 3.3 – ONI and hardware design parameters

Parameters	Descriptions	Typical values
	Technological parameters	
$Loss_{mr}$	MR through loss	$0.02\,dB$ [119]
$Loss_d$	MR drop loss	$0.7\,dB$ [119]
Q	MR Quality Factor	$20,000$ [118]
$Loss_h$	Waveguide loss per hop	$0.25\,dB/cm$ [117]
P_d	Optical power	$-8\,dBm$ at 10^{-12} BER
	received by photodetector	$-12\,dBm$ at 10^{-3} BER
FSR	Full Scale Range	$8nm$ [118]
$Laser_eff$	Laser efficiency	0.33 [120]
B_R	Bit-rate	10Gb/s
	ONI Parameters	
BER_{NotAx}	BER without approximation	10^{-12}
BER_{Ax}	BER with approximation	10^{-3}
P_H	High power	$707\,\mu W$
P_M	Medium power	$281\,\mu W$
P_L	Low power	$112\,\mu W$
$N\lambda$	Number of wavelengths	8
Short distance	Nb of hops from	1 to 5 hops
Long distance	source to destination	6 to 15 hops

to carry our comparison with *Lorax*.

Figure 3.11.a reports the exponential growth of the laser powers for increasing distance between ONI, assuming the two considered propagation losses. P_H, P_M and P_L are estimated based on the method defined in Section 3.3.1 The losses breakdown to communicate with the last receiver are defined in Figure 3.11.b for a number of wavelengths ranging from 8 to 24 and for a $0.25dB/cm$ waveguide loss. The waveguide loss, which dominates the total losses for a small number of wavelengths, is independent from the number of lasers since it only depends on the communication distance. On the contrary, the MR through losses linearly increase with the number of wavelengths due to the higher number of crossed MRs. Regarding the crosstalk losses, it only represents 1.3% for 8 wavelengths. Indeed, for an FSR of $8nm$, the channel spacing is $1nm$ ($\sim 120GHz$), which results in negligible crosstalk losses as demonstrated in [118]. For a same FSR, increasing the number of wavelengths to 24 leads to 7.9% of losses provided by the crosstalk, which is acceptable and can be lowered by increasing the channel spacing [108]. In the rest of the study, we will consider 8 wavelengths.

P_H, P_M and P_L calculations are detailed for a distance between ONIs equal to 0.25 cm and to 1 dB/cm for Figure 3.11.c and Figure 3.11.d respectively. As we can see in Figure 3.11.c (respectively Figure 3.11.d), to reach BER_{NotAx} at ONI 15, the higher transmission losses are compensated by an increase of P_H up to $707\mu W$ for $0.25dB/cm$ (respectively up to $9549\mu W$ for 1 dB/cm). Similarly, to reach BER_{Ax} at ONI 15, P_M of $281\mu W$ is needed for 0.25 dB/cm, which leads to BER_{NotAx} at ONI 5, while P_M must be fixed at $3801\mu W$ for $1dB/cm$, which leads to BER_{NotAx} at ONI 11. Hence, short range communications for 1 dB/cm include ONIs 1 to 11, against ONIs 1 to 5 for 0.25 dB/cm. In other words, the total losses experienced by the optical signal for hops 11 to 15 under 1 dB/cm are similar to the losses for hops 5 to 15 under 0.25 dB/cm. For comparison purpose with [16], we consider power values obtained for 0.25 dB/cm in the rest of this section ($P_H = 707\mu W$, $P_M = 281\mu W$ and $P_L = 112\mu W$), as summarized in Table 3.3.

3.5.3 Distance aware communication

To evaluate the impact of the proposed method, we analyze the traffic distribution for the selected applications. This analysis includes both hop distance and data type, within the considered 16 clusters architecture.

Figure 3.12.a presents ratio of communication by hop count. We can note that the communication traffics over the network are well balanced for all the hop distances and for any application. From this balanced distributed communication, we can say that packets are uniformly exchanged among the cores. Secondly, Figure 3.12.b presents the traffic distribution by data type. The communications are classified in Instruction, Integer, and Floating-Point data. The ratio of data type is linked to the application and it is defined at the application designing level. Only Floating-Point (FP) ratio refers to potential data type that may benefit from power saving by using the proposed approximate method. We can see that Canneal only have around 4% of FP data communication, while the others reached 22%, 44% and 60%, for Fluidanimate, Blacksholes and Streamcluster, respectively. From this analysis, it is evident that applying energy reduction technique on FP data for Canneal will not produce a considerable energy reduction, unlike the other applications, and especially Streamcluster.

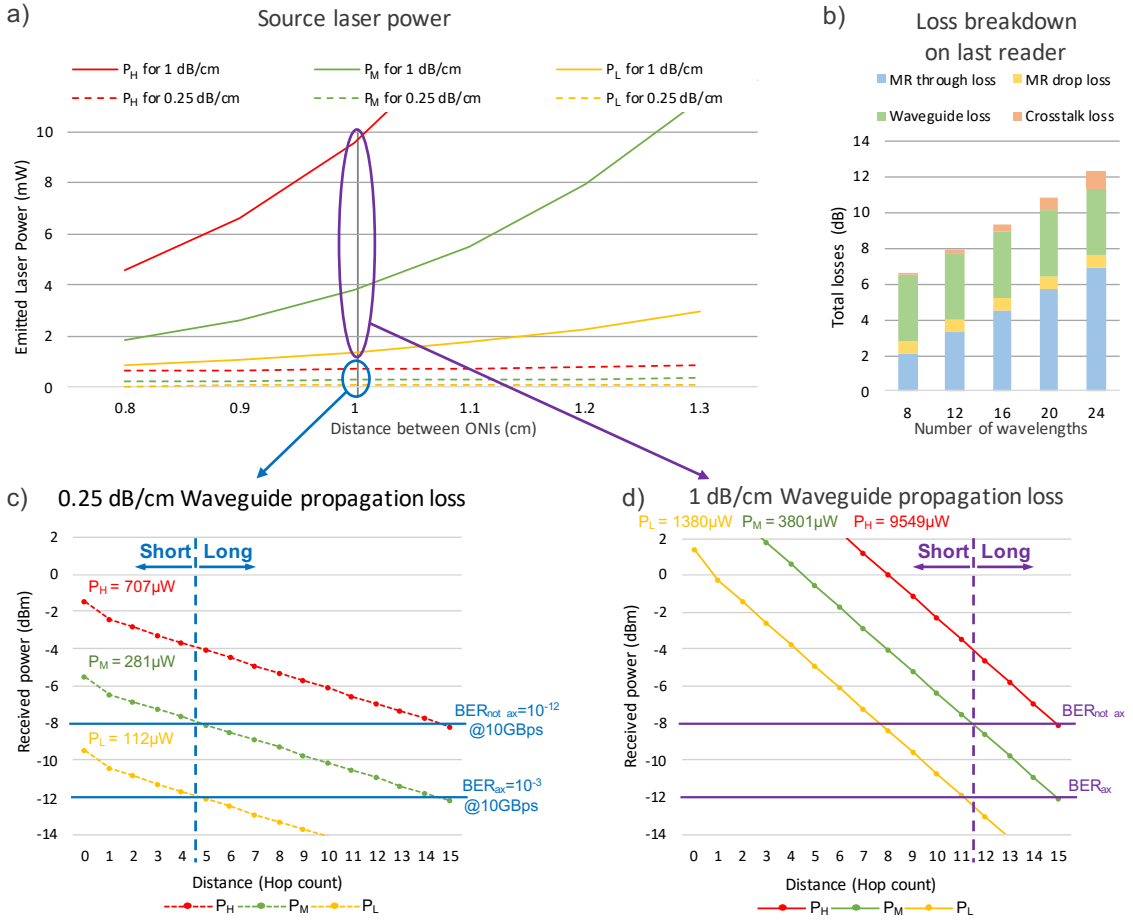


Figure 3.11 – a) Required laser output power according to the distance between ONIs and for $BER_{ax} = 10^{-3}$ and $BER_{notax} = 10^{-12}$. b) Communication losses overview to reach the last receiver with $0.25dB/cm$ waveguide propagation loss for 8 to 24 wavelengths. Details of the method to define Short/Long-range communications and to calculate P_H, P_M and P_L for a ONI distance of $10mm$ are given for c) $0.25dB/cm$ and d) $1dB/cm$ waveguide propagation losses.

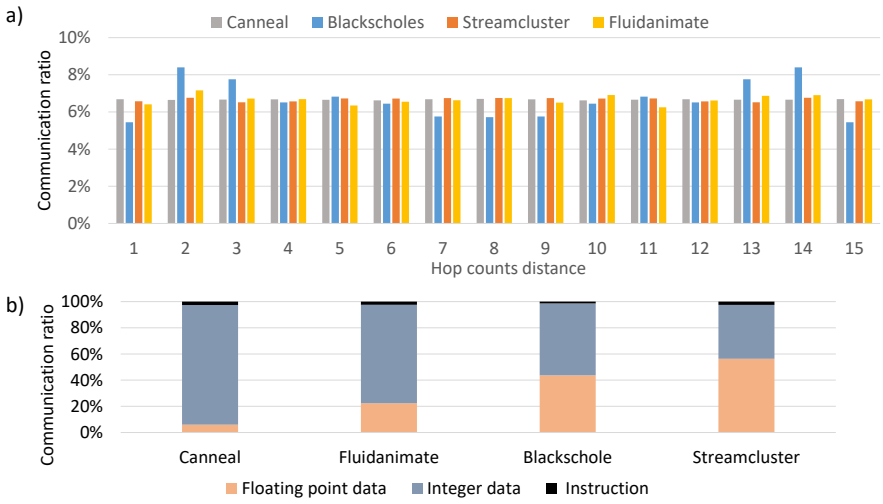


Figure 3.12 – Traffic distributions on 16 clusters architecture: a) by hop count b) by data type.

3.6 Power Consumption Vs QoR

This section presents the power saving of our distance aware approximate optical interconnect considering the following cases: i) approximation and truncation with the targeted BER for approximate and accurate communication from section 2.6 and ii) distance aware approximate data communication model from 3.2.

The exploration of QoR–power-saving trade-off is provided and the solutions are compared to the ideal ones. As discussed in the previous section, not all the Approxbench applications lead to significant power saving under approximate communications schemes. Nevertheless, to reduce the design complexity, we decided to use the same approximation scheme, i.e. $8NA/4A/20T$ configuration with $BER_{Ax} = 10^{-3}$, for all the applications. In the following, we mainly focus our studies on the StreamCluster application.

3.6.1 Power saving analysis

Figure 3.13 shows the power saving of several methods with different distance aware techniques (*No distance aware*, *Distance Proportional*, and *Short/Long*) and different approximate configuration (one approximate technique either *approximatedbit* or *truncation* and mixed approximate techniques).

First, we explore the distance aware techniques on Streamcluster in terms of power saving. All the results are compared with the optical power of the *Baseline* configuration,

which only uses the highest optical power level ($P_H = 707\mu W$), without regards on data type and distance of communication. The floating-point data are represented with orange colors, while the rest of communication (integer data and instructions) are represented with blue colors. The light orange, respectively light blue, represents the contribution of short distance communications, i.e. hop count less than 6, respectively long distance communications, i.e. hop count greater than 5. As previously defined, the threshold between long and short distance communications is equal to 5. Considering this distance threshold, the *Baseline* shows that 33% of transfers are short communications, while 66% are long communications. Regarding the data type, up to 58% can be transmitted using approximation (see FP ratio in Figure 3.12.b for Streamcluster application).

The second stacked vertical bars, named *Short/Long* on Figure 3.13, show the power contribution when the lasers power levels are adapted according to *Short* or *Long* distances. In this case, the long distance communications are sent with the highest power P_H, while the power is reduced to P_M for short distance communications and for any data type. This results in 20% of power saving compared to the *Baseline*.

The last method that only benefits from distance consideration for power saving is *Distance proportional* illustrated by the third stacked bars. It assumes to use the exact required power for any distance within the architecture [16], hence it requires as many power levels as distance hops available. This solution is the best in terms of power reduction for accurate communications. It allows to save 21% and 41% more power compared to *Short/Long* and *Baseline* methods, respectively. This solution avoids to waste any extra power for the transmission. However, as we saw in section 3.4, the laser design complexity for fine grain multi power levels is too complex and requires significant effort.

Second, we explore the different approximate configurations by applying different number of bit truncation and approximation. The scenario $12NA/20A/0T$ involves approximation for 20 LSB of FP data, with a BER of 10^{-3}, hence by setting laser power to $P_M = 281\,\mu W$ and without distance management [131]. The system benefits from 21.8% power saving compared to *Baseline*, which is due to a 62% reduction to transmit FP data. Obviously, no power saving is obtained for non-FP data. By assuming *Short/Long* distance management from the previous method, approximated bits will be sent using $P_L = 112\,\mu W$ and $P_M = 281\,\mu W$ for short and long distances, respectively. An overall power reduction from 78.1% to 63.6% is achieved.

Power consumption for communications that only use truncation, without distance management, corresponds to configuration $12NA/0A/20T$. In this case, 20 LSB are trun-

cated on FP data communication by turning off lasers. The architecture benefits from 36% power saving compared to *Baseline* throughout all communications. Moreover, if *Short/Long* distance management is associated to this solution, hence leading to $12NA/0A/20T+$ *Short/Long* result, it leads to an extra 13% power saving, then reaching a total power consumption of only 50.9% compared to the *Baseline*.

By combining truncation and approximation without considering the distance, scenario $8NA/4A/20T$ shows 40% of power reduction. Associated to *Short/Long* distance management, $(8NA/4A/20T + Short/Long)$, the power required for the communication drops to 47%, thus leading to a significant improvement. All these results demonstrate the needs to combine approximation techniques based on both optical link and data representation.

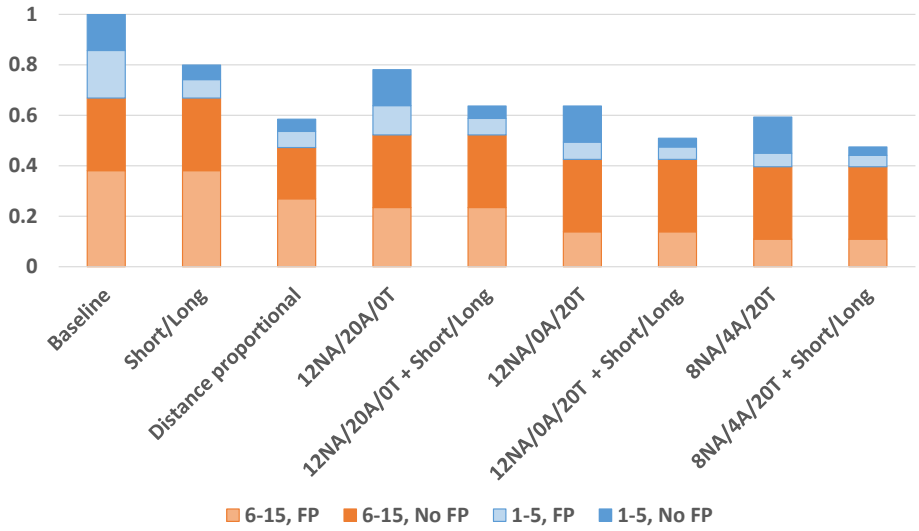


Figure 3.13 – Normalized power consumption for the considered communication schemes.

3.6.2 Power saving vs. Quality of Results Trade-offs

To fully evaluate the different techniques, we have to consider both power consumption and application output error. Figure 3.14 highlights the design space with all the possible techniques to handle approximate communications in ONoC (including the following different BER_{Ax}: 10^{-2}, 10^{-3}, 10^{-5} and 10^{-7}). The bottom of Figure shows zooms of valid solutions from the top of Figure. All optical power consumption have been normalized to the *Baseline* which is represented by the red cross. This baseline configuration uses P_H power level for all lasers, hence no distance management and no approximation are

considered. For more clarity, we grouped the results by different communication distance managements: *No Distance aware*, *Short/Long* and *Distance proportional*, in red, blue, and green, respectively.

For each group, the solutions without approximation nor truncation are represented with a cross. Moreover, in each group, some points are located on Pareto front lines which indicate the most effective results in terms of power consumption and output error. The solutions on Pareto front lines are represented with triangles and the configurations are highlighted, e.g. $4A/20T$ 10^{-3} stands for 4 approximated bits, 20 truncated bits, and $BER_{ax} = 10^{-3}$. For more clarity in Figure 3.14, the number of non-approximated bits is not shown, but can be easily deduct as the total number of bits of FP data is equal to $32bits$. On the right gray part, we also represent the solutions providing more than 10% of output error.

First of all, this figure illustrates the huge exploration space with high differences in terms of power versus QoR. Unsurprisingly, the *distance proportional* solutions are the less power consuming, but with a high complexity, as already mentioned before. However, unlike what we have observed in Figure 3.13, there is less power consumption difference in Pareto front lines between the distance aware techniques. For example, in Fig 3.13, there are approximately 20% power consumption difference between the baseline .and *Short/Long* solution. However, in Figure 3.14, there is less difference in between Pareto front lines, around 13% and 12% between *No Distance aware* and *Short/Long*, *Short/Long* and *Distance proportional* respectively.

Regarding the best solutions in terms of trade-offs between power consumption and output errors, the same trends on the different color groups can be observed. The first main cloud of solutions, zoomed in Fig 3.14.b, is dominated with three points on the Pareto fronts in each group. All these solutions provide less than 0.15% of output errors. The solutions $16NA/0/16T$ with a maximum of 16 truncated bits provide a very small degradation of the output error, but allows to save a huge amount of power compared to the baseline, and for any group of distance managements. Indeed, more than 20% of power reduction is obtained for these solutions. We can find find some points are located in higher position than different distance communication technique. It means some of approximated configurations have less power consumption than higher distance communication control. For instance, upper the red line(Pareto front lines of *No Distance aware*), there are several points from *Short/Long* distance aware. Moreover, we can find same trend in *distance proportional* configuration in green dots. Some configurations in *distance proportional*

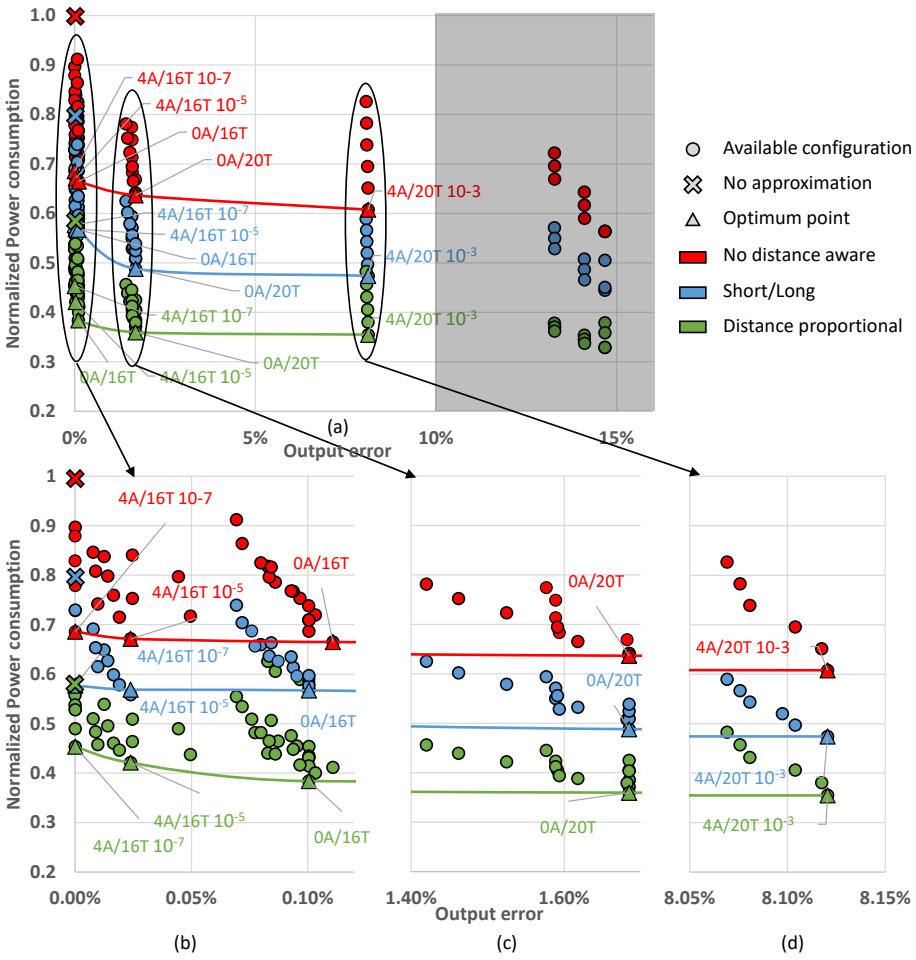


Figure 3.14 – Power consumption and output error trade-off: (a) design space, (b), (c), and (d) are zooms.

can consume more than the best solution with only two distances considered.

The second main trend of points, zoomed in Figure 3.14.c, is dominated by the solutions $12NA/0A/20T$. These solutions do not exceed 1.7% output errors while pushing forward the power consumption reduction. The last trend of solutions, that do not exceed the tolerable output error of 10%, is zoomed in Fig 3.14.d. These solutions are dominated by the $8NA/4A/20T$ with $BER_{ax} = 10^{-3}$. These solutions provide less than 8.15% of output error while providing the more important power consumption reduction.

To conclude this exploration, first we can state that the design space is huge, and designing an ONI able to configure all configurations is unrealistic. Secondly, we highlighted that the different solutions can be grouped in four trends, almost no error (Fig 3.14.b), few errors (Fig 3.14.c), reasonable errors (Fig 3.14.d), and unacceptable errors. All trends are dominated by one or few optimal points, hence limiting the number of interesting solutions to embed in the ONIs. As already mentioned, we target to limit the ONI complexity and its management by only considering one low power communication scheme. Thus, the most relevant solution to provide the highest power savings is $8NA/4A/20T$ with $BER_{ax} = 10^{-3}$ which provides a reasonable output error. Finally, regarding the distance of communications, the proposed solution (which considers two group of distance, i.e. $Short/Long$) provides efficient result compared to no distance management technique. Compared to the optimal solution induced by fully distance proportional links, our solution consumes only 12% additional power. However, it is important to notice that our ONI design is realistic as it requires only four laser powers, compared to [16] for which the complexity linearly increases with both number of cores and number of BER levels.

3.7 Conclusion

In this chapter, we explored the distance aware approximate optical interconnect to improve the power consumption in NoC. First, we presented different distance aware techniques to adapt the power levels to optical losses in waveguides. Then, we observed how much power can be saved with different distance thresholds to find optimum point.

Second, we illustrated the approximate optical interconnect which is based on a distance classification and on the level of approximation acceptable for the communications. This proposition allows the distance classification into two sets *Short* and *Long* and computes the laser power required at source to ensure a sufficient optical power at destination for approximate communication. This allows more efficient power level control for the

ONI and reduces required number of power levels. Moreover, we present a comprehensive study on the design complexity of the drivers for our method and *Lorax* [16]. The fine grain laser power level tuning from *Lorax* requires significant effort on laser design and has large overhead on both number of controllers and transistors in terms of power leakage and area constrain.

Chapter 4

Conclusion and Perspectives

4.1 Overview

As applications evolve in complexity, conventional electric interconnect (NoC) suffer from limitation on bandwidth in on-chip communication. Furthermore, the energy consumption of thus NoCs increases as data transfers also increase.

Therefore, photonics on silicon appears has real interesting opportunity to overcome these two challenges. Indeed, optical NoC, based on optical devices, offers large bandwidth, low link area link and low latency.

However, optical interconnect has several cons in spite of its prominent pros. It suffers from high power consumption overhead related to low-efficiency lasers and from crosstalk noise (electro-magnetic effect between the wavelength channels). These two negative effect must be taken into account and managed to limit their impacts on NoC in terms of power efficiency.

At the same time, approximate computing has been developed as a promising computation paradigm for improving energy and computation efficiency. This computation paradigm is based on exploiting the error resiliency and tolerance in computation and brings large amount of efficiency with neglecting output errors. Approximate computing uses data approximation, rather than computing accurate data. Based on this concept, computation can be simplified and this leads to power consumption efficiency, better throughput and memory area.

Therefore, in this dissertation, we proposed approximate optical NoC and distance-aware optical interconnect to overcome limit in Optical NoC in terms of energy efficiency.

First, we propose an approximate nanophotonic interconnect in NoC. For that, we take advantage of approximate computing paradigm, and we proposed to transmit the least significant bits of floating point data with low power optical signals. This leads to higher error rate at application results, but it can drastically reduce the laser power consumption.

The data approximation is done by keeping most significant bits to in high power

leading to guarantee low BER. The case study of various possible configurations allows exploring robustness and energy efficiency trade-offs for on-chip optical interconnects. The proposed techniques are implemented in Optical NoC with single writer multiple reader and wavelength division multiplexing optical channels. Several Benchmark applications are tested in designed ONoC to carry out the results. For example in Streamcluster application, the results demonstrate a power saving up to 42% compared to classical method. To the best of our knowledge, this work was the first attempt that implements approximate computing concepts into nanophotonic interconnects.

Second, we proposed distance-aware approximate optical interconnect to improve power consumption of on-chip communication. The proposed distance aware technique classifies the communication packets by its distance from source to destination cores, and configures the transmission laser power. In single writer and multiple reader, the link is shared with between several readers for one writer. Due to waveguide losses, the amount of required power to achieve targeted BER is different for each destinations. To maximize the power efficiency a dynamic fine-grained laser power control is necessary but it can lead to a complex controller if the number of power levels become high. Therefore, in our proposal, we grouped the destinations into two groups; short and long communications. The results demonstrate that exploiting of these communication distances (defined as short/long configurations), 20% of power can be saved compared to the Baseline solution using high power for all lasers. Finally, we illustrated that our two techniques can be implemented together to yet increase the power saving. For example, for Streamcluster application, mixing the two techniques could save 47% of overall application power consumption.

4.2 Perspectives

4.2.1 Optical network interface

In our study, multi laser power levels are used to configure approximate data communication and to adapt different losses induced by distance from source and destination cores. We assumed that the configuration for these complex laser power managements are done in Optical Network Interface (ONI). The ONI can perform dynamic laser power configuration at runtim based on each data to transmit and destination of communication. However, we did not carried out the detail of this ONI. To fairly compare trade-off

between power consumption efficiency and multi power level laser power configurations, the dynamic ONI runtime management must be defined in micro architecture.

4.2.2 Smart ONI

Once the dynamic ONI defined, we can propose to implement several techniques to make ONI more smart. For this proposal, a set of several parameters can be define like; approximate level, communication distance, number of channel (WDM and DWDM) and link priority. Based on these parameters, a smart ONI can have several approximate levels that can adapt to different applications and user needs to control output errors. As we saw in our study, for the same approximation scheme, the sensibilities of applications are different. In this case, the threshold between short and Long communications could be to be adapted to obtain a better energy efficiency. Moreover, a smart ONI could have control over number of wavelength used at runtime. For example, we could imagine more dense WDM for short distance communications to increase the bandwidth, which could improve application performances. Lastly, in parallel computing some data packets are need to be prioritized to reduce waiting times during applications execution. By dynamically booking our apprximate and distance aware link for certain packets, power consumption efficiency and memory efficiency could be increased. .

4.2.3 Truncation and data insertion

In our study, we saw that quite a lot of bits could be truncated in floating point data without real impact on results quality for several applications. When bits are truncated the lasers are set to off or stand by state to improve the power efficiency.

As some lasers are not used to send data D_t at time t, we can imagine to allocate them at time t for several bits of data D_{t+1} which wil have been sent at time $t+1$. This bit reallocation could lead to reduce the transmission time of a complete message with a reduce power consumption overhead coming for the increase of controller complexity.

BIBLIOGRAPHY

[1] G. E. Moore, "Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp.114 ff.," *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 33–35, 1965.

[2] M. Bohr, "A 30 year retrospective on dennard's mosfet scaling paper," *IEEE Solid-State Circuits Society Newsletter*, vol. 12, no. 1, pp. 11–13, 2007.

[3] C. Kozyrakis, A. Kansal, S. Sankar, and K. Vaid, "Server engineering insights for large-scale online services," *Micro, IEEE*, vol. 30, pp. 8 – 19, 09 2010.

[4] S. Z. Ahmed, G. Sassatelli, L. Torres, and L. Rougé, "Survey of new trends in industry for programmable hardware: Fpgas, mppas, mpsocs, structured asics, efpgas and new wave of innovation in fpgas," in *2010 International Conference on Field Programmable Logic and Applications*, pp. 291–297, 2010.

[5] S. Charles and P. Mishra, "A survey of network-on-chip security attacks and countermeasures," *ACM Comput. Surv.*, vol. 54, May 2021.

[6] "Amba specification," *https://developer.arm.com/products/architecture/amba-protocol.*, 1999.

[7] C. Wissem, B. Attia, A. Noureddine, A. Zitouni, and R. Tourki, "A quality of service network on chip based on a new priority arbitration mechanism," in *ICM 2011 Proceeding*, pp. 1–6, 2011.

[8] P. Pande, C. Grecu, A. Ivanov, and R. Saleh, "Design of a switch for network on chip applications," in *Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03.*, vol. 5, pp. V–V, 2003.

[9] V. Rantala, T. Lehtonen, and J. Plosila, "Network on chip routing algorithms," 12 2008.

[10] U. Ogras and R. Marculescu, "Prediction-based flow control for network-on-chip traffic," in *2006 43rd ACM/IEEE Design Automation Conference*, pp. 839–844, 2006.

[11] Y. Yin, S. Chen, and X. Hu, "Input buffer planning for network-on-chip router design," in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, vol. 13, pp. V13–201–V13–204, 2010.

[12] T. Alexoudi, N. Terzenidis, S. Pitris, M. Moralis-Pegios, P. Maniotis, C. Vagionas, C. Mitsolidou, G. Mourgias-Alexandris, G. T. Kanellos, A. Miliou, K. Vyrsokinos, and N. Pleros, "Optics in computing: From photonic network-on-chip to chip-to-chip interconnects and disintegrated architectures," vol. 37, pp. 363–379, 2019.

[13] A. Asad, A. Dorostkar, and F. A. Mohammadi, "A novel power model for future heterogeneous 3d chip-multiprocessors in the dark silicon age," *EURASIP Journal on Embedded Systems*, vol. 2018, pp. 1–16, 2018.

[14] A. Karkar, T. Mak, K. Tong, and A. Yakovlev, "A survey of emerging interconnects for on-chip efficient multicast and broadcast in many-cores," *IEEE Circuits and Systems Magazine*, vol. 16, no. 1, pp. 58–72, 2016.

[15] W. Wolf, A. A. Jerraya, and G. Martin, "Multiprocessor system-on-chip (mpsoc) technology," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 10, pp. 1701–1713, 2008.

[16] F. Sunny, A. Mirza, I. G. Thakkar, S. Pasricha, and M. Nikdast, "LORAX: loss-aware approximations for energy-efficient silicon photonic networks-on-chip," *CoRR*, vol. abs/2002.11289, 2020.

[17] R. G. Beausoleil, M. McLaren, and N. P. Jouppi, "Photonic architectures for high-performance data centers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 2, pp. 3700109–3700109, 2013.

[18] M. J. R. Heck, H.-W. Chen, A. W. Fang, B. R. Koch, D. Liang, H. Park, M. N. Sysak, and J. E. Bowers, "Hybrid silicon photonics for optical interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 2, pp. 333–346, 2011.

[19] K. Yamada, Y. Urino, T. Nakamura, and Y. Arakawa, "Integrated silicon-based optical interconnect for fast, compact, energy-efficient electronic circuit systems," vol. 11, 02 2013.

[20] M. J. R. Heck and J. E. Bowers, "Energy efficient and energy proportional optical interconnects for multi-core processors: Driving the need for on-chip sources," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 4, pp. 332–343, 2014.

[21] Z. Zhou, B. Yin, and J. Michel, "On-chip light sources for silicon photonics," *Light-Science Applications*, vol. 4, no. 11, 2015.

[22] A. H. Perrott, L. Caro, M. Dernaika, and F. H. Peters, "A comparison between off and on-chip injection locking in a photonic integrated circuit," *Photonics*, vol. 6, no. 4, 2019.

[23] D. C. Byrne, J. P. Engelstaedter, W.-H. Guo, Q. Y. Lu, B. Corbett, B. Roycroft, J. O'Callaghan, F. H. Peters, and J. F. Donegan, "Discretely tunable semiconductor lasers suitable for photonic integration," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 15, no. 3, pp. 482–487, 2009.

[24] R. Li, D. Patel, A. Samani, E. El-Fiky, Y. Wang, Z. Xing, L. Xu, and D. V. Plant, "Analysis and experimental study of a silicon photonic single mrm-assisted mzi pam-4 modulator," *IEEE Photonics Journal*, vol. 9, no. 6, pp. 1–7, 2017.

[25] L. Cheng, S. Aditya, and A. Nirmalathas, "An exact analytical model for dispersive transmission in microwave fiber-optic links using mach-zehnder external modulator," *IEEE Photonics Technology Letters*, vol. 17, no. 7, pp. 1525–1527, 2005.

[26] B. Enning, "Signal processing in a mach-zehnder intensity modulator for coherent and incoherent light using copropagating and counterpropagating electrical signals," *IEEE Photonics Technology Letters*, vol. 17, no. 11, pp. 2424–2426, 2005.

[27] J.-Y. Lee, S.-H. Lee, and J.-I. Song, "All-optical frequency up-conversion technique utilizing a frequency-doubling optoelectronic oscillator based on electro-absorption modulators," in *Microwave Photonics (MWP) and the 2014 9th Asia-Pacific Microwave Photonics Conference (APMP) 2014 International Topical Meeting on*, pp. 226–228, 2014.

[28] M. Bahadori, S. Rumley, H. Jayatilleka, K. Murray, N. A. F. 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.

[29] 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.

[30] K. Iga, "Vertical-cavity surface-emitting laser (vcsel)," *Proceedings of the IEEE*, vol. 101, no. 10, pp. 2229–2233, 2013.

[31] S. Arslan, S. A. A. Shah, and H. Kim, "Power efficient current driver based on negative boosting for high-speed lasers," *Electronics*, vol. 8, no. 11, 2019.

[32] C. S. Wang, J. Geske, G. Berdin, F. Talantov, T. Cardellino, H. Garrett, D. Millenheft, V. Kumsomboone, and D. Renner, "Recent advances in high-power vcsel arrays," in *2012 IEEE Photonics Society Summer Topical Meeting Series*, pp. 3–4, 2012.

[33] A. Demir, G. Zhao, S. Freisem, X. Liu, and D. G. Deppe, "Scaling properties of lithographic VCSELs," in *Vertical-Cavity Surface-Emitting Lasers XV* (J. K. Guenter and C. Lei, eds.), vol. 7952, pp. 171 – 176, International Society for Optics and Photonics, SPIE, 2011.

[34] X. Chen, J. Feng, J. Xu, J. Zhang, and S. Chen, "Simultaneously tolerate thermal and process variations through indirect feedback tuning for silicon photonic networks," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 7, pp. 1409–1422, 2021.

[35] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature*, vol. 435, pp. 325–327, May 2005.

[36] P. Dong, W. Qian, S. Liao, H. Liang, C.-C. Kung, N.-N. Feng, R. Shafiiha, J. Fong, D. Feng, A. Krishnamoorthy, and M. Asghari Khiavi, "Low loss silicon waveguides for application of optical interconnects," pp. 191 – 192, 08 2010.

[37] L. Vivien, D. Marris-Morini, J. Mangeney, P. Crozat, E. Cassan, S. Laval, J.-M. Fedeli, J. Damlencourt, and Y. Lecunff, "42 ghz waveguide germanium-on-silicon vertical pin photodetector," in *2008 5th IEEE International Conference on Group IV Photonics*, pp. 185–187, 2008.

[38] E. Timurdogan, Z. Su, K. Settaluri, S. Lin, S. Moazeni, C. Sun, G. Leake, D. D. Coolbaugh, B. R. Moss, M. Moresco, V. Stojanović, and M. R. Watts, "An ultra low power 3d integrated intra-chip silicon electronic-photonic link," in *2015 Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, 2015.

[39] X. Zeng, Z. Huang, B. Wang, D. Liang, M. Fiorentino, and R. G. Beausoleil, "Silicon-germanium avalanche photodiodes with direct control of electric field in charge multiplication region," *Optica*, vol. 6, pp. 772–777, Jun 2019.

[40] J. O. Sosa, C. Killian, H. B. Ammar, and D. Chillet, "Min/max time limits and energy penalty of communication scheduling in ring-based ONoC," in *NoCArc 2020 - 13th IEEE/ACM International Workshop on Network on Chip Architectures*, no. 1-6, (On-line, France), Oct. 2020.

[41] B. Radi, M. S. Nezami, M. Taherzadeh-Sani, F. Nabki, M. Ménard, and O. Liboiron-Ladouceur, "A 22-gb/s time-interleaved low-power optical receiver with a two-bit integrating front end," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 1, pp. 310–323, 2021.

[42] N. Michit, P. L. Maître, B. Charbonnier, R. Orobtchouk, A. Michard, and J. F. Carpentier, "Scalable highly flexible wdm switch for onoc architectures," in *OPTO*, 2018.

[43] 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 *2008 International Symposium on Computer Architecture*, pp. 153–164, 2008.

[44] P. Grani and S. Bartolini, "Design options for optical ring interconnect in future client devices," vol. 10, jun 2014.

[45] J. Kim, W. J. Dally, S. Scott, and D. Abts, "Technology-driven, highly-scalable dragonfly topology," in *Proceedings of the 35th Annual International Symposium on Computer Architecture*, ISCA '08, (USA), p. 77–88, IEEE Computer Society, 2008.

[46] N. Akbari and M. Modarressi, "A high-performance network-on-chip topology for neuromorphic architectures," in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, vol. 2, pp. 9–16, 2017.

[47] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, and A. N. Choudhary, "Firefly: illuminating future network-on-chip with nanophotonics," in *ISCA '09*, 2009.

[48] N. Kirman, M. Kirman, R. K. Dokania, J. F. Martinez, A. B. Apsel, M. A. Watkins, and D. H. Albonesi, "Leveraging optical technology in future bus-based chip multiprocessors," in *2006 39th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'06)*, pp. 492–503, 2006.

[49] 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)*, pp. 80–87, 2014.

[50] L. Zhou and A. K. Kodi, "Probe: Prediction-based optical bandwidth scaling for energy-efficient nocs," in *2013 Seventh IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*, pp. 1–8, April 2013.

[51] J. Postman, T. Krishna, C. Edmonds, L. Peh, and P. Chiang, "Swift: A low-power network-on-chip implementing the token flow control router architecture with swing-reduced interconnects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 8, pp. 1432–1446, 2013.

[52] Y.-H. Kao and H. J. Chao, "Design of a bufferless photonic clos network-on-chip architecture," *IEEE Transactions on Computers*, vol. 63, no. 3, pp. 764–776, 2014.

[53] 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.

[54] C. Killian, D. Chillet, S. L. Beux, V. Pham, O. Sentieys, and I. O'Connor, "Energy and performance trade-off in nanophotonic interconnects using coding techniques," in *Proceedings of the 54th Annual Design Automation Conference, DAC 2017, Austin, TX, USA, June 18-22, 2017*, pp. 85:1–85:6, ACM, 2017.

[55] J. O. Sosa, C. Killian, H. B. Ammar, and D. Chillet, "Min/max time limits and energy penalty of communication scheduling in ring-based ONoC," in *NoCArc 2020*, (On-line, France), Oct. 2020.

[56] C. Chen, J. L. Abellán, and A. Joshi, "Managing laser power in silicon-photonic noc through cache and noc reconfiguration," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 6, pp. 972–985, 2015.

[57] B. Neel, M. Kennedy, and A. Kodi, "Dynamic power reduction techniques in on-chip photonic interconnects," in *Proceedings of the 25th Edition on Great Lakes Symposium on VLSI*, GLSVLSI '15, (New York, NY, USA), p. 249–252, Association for Computing Machinery, 2015.

[58] J. Bashir and S. R. Sarangi, "Predict, share, and recycle your way to low-power nanophotonic networks," *J. Emerg. Technol. Comput. Syst.*, vol. 16, Oct. 2019.

[59] S. Van Winkle, A. K. Kodi, R. Bunescu, and A. Louri, "Extending the power-efficiency and performance of photonic interconnects for heterogeneous multicores with machine learning," in *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pp. 480–491, 2018.

[60] C. Chen and A. Joshi, "Runtime management of laser power in silicon-photonic multibus noc architecture," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 2, pp. 3700713–3700713, 2013.

[61] S. S. Vatsavai, V. S. P. Karempudi, and I. G. Thakkar, "PROTEUS: rule-based self-adaptation in photonic nocs for loss-aware co-management of laser power and performance," in *14th IEEE/ACM International Symposium on Networks-on-Chip, NOCS 2020, Hamburg, Germany, September 24-25, 2020*, pp. 1–8, IEEE, 2020.

[62] E. Khan, S. Lehmann, H. Gunji, and M. Ghanbari, "Iterative error detection and correction of h.263 coded video for wireless networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 12, pp. 1294–1307, 2004.

[63] Y. Zhang and Q. Yuan, "A multiple bits error correction method based on cyclic redundancy check codes," in *2008 9th International Conference on Signal Processing*, pp. 1808–1810, 2008.

[64] R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, 1962.

[65] S. Venkataramani, S. T. Chakradhar, K. Roy, and A. Raghunathan, "Approximate computing and the quest for computing efficiency," DAC '15, (New York, NY, USA), Association for Computing Machinery, 2015.

[66] K. Roy and A. Raghunathan, "Approximate computing: An energy-efficient computing technique for error resilient applications," in *2015 IEEE Computer Society Annual Symposium on VLSI*, pp. 473–475, 2015.

[67] V. Camus, M. Cacciotti, J. Schlachter, and C. Enz, "Design of approximate circuits by fabrication of false timing paths: The carry cut-back adder," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 4, pp. 746–757, 2018.

[68] A. Tsukahara and A. Kanasugi, "Design of approximate arithmetic circuits within tolerance," in *2018 10th Computer Science and Electronic Engineering (CEEC)*, pp. 79–82, 2018.

[69] C. Labrado, H. Thapliyal, and F. Lombardi, "Design of majority logic based approximate arithmetic circuits," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4, 2017.

[70] A. Ranjan, A. Raha, V. Raghunathan, and A. Raghunathan, "Approximate memory compression," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 28, no. 4, pp. 980–991, 2020.

[71] S. Amanollahi, M. Kamal, A. Afzali-Kusha, and M. Pedram, "Circuit-level techniques for logic and memory blocks in approximate computing systemsx," *Proceedings of the IEEE*, vol. 108, no. 12, pp. 2150–2177, 2020.

[72] J. Xu, S. Kim, B. Nikolic, and Y. S. Shao, "Memory-efficient hardware performance counters with approximate-counting algorithms," in *2021 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, pp. 226–228, 2021.

[73] D. T. Nguyen, H. Kim, H.-J. Lee, and I.-J. Chang, "An approximate memory architecture for a reduction of refresh power consumption in deep learning applications," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, 2018.

[74] A. Siddique, O. Hasan, F. Khalid, and M. Shafique, "Approxcs: Near-sensor approximate compressed sensing for iot-healthcare systems," 2018.

[75] A. Ghosh, A. Raha, and A. Mukherjee, "Energy-efficient iot-health monitoring system using approximate computing," *Internet of Things*, vol. 9, p. 100166, 03 2020.

[76] J. Li, M. Siddula, X. Cheng, W. Cheng, Z. Tian, and Y. Li, "Approximate data aggregation in sensor equipped iot networks," *Tsinghua Science and Technology*, vol. 25, no. 1, pp. 44–55, 2020.

[77] S. Venkataramani, A. Ranjan, K. Roy, and A. Raghunathan, "Axnn: Energy-efficient neuromorphic systems using approximate computing," in *2014 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pp. 27–32, 2014.

[78] D. Ma, X. Yin, M. Niemier, X. S. Hu, and X. Jiao, "Axr-nn: Approximate computation reuse for energy-efficient convolutional neural networks," in *Proceedings of the 2020 on Great Lakes Symposium on VLSI*, GLSVLSI '20, (New York, NY, USA), p. 363–368, Association for Computing Machinery, 2020.

[79] Q. Zhang, T. Wang, Y. Tian, F. Yuan, and Q. Xu, "Approxann: An approximate computing framework for artificial neural network," vol. 2015, pp. 701–706, 04 2015.

[80] M. Kwak, J. Lee, H. Seo, M. Sung, and Y. Kim, "Training and inference using approximate floating-point arithmetic for energy efficient spiking neural network processors," in *2021 International Conference on Electronics, Information, and Communication (ICEIC)*, pp. 1–2, 2021.

[81] "Ieee standard for floating-point arithmetic," *IEEE Std 754-2008*, pp. 1–70, 2008.

[82] A. Lashgar, E. Atoofian, and A. Baniasadi, "Loop perforation in openacc," in *2018 IEEE Intl Conf on Parallel Distributed Processing with Applications, Ubiquitous Computing Communications, Big Data Cloud Computing, Social Computing Networking, Sustainable Computing Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)*, pp. 163–170, 2018.

[83] H. Omar, M. Ahmad, and O. Khan, "Graphtuner: An input dependence aware loop perforation scheme for efficient execution of approximated graph algorithms," in *2017 IEEE International Conference on Computer Design (ICCD)*, pp. 201–208, 2017.

[84] S. P. Crago and D. Yeung, "Reducing data movement with approximate computing techniques," in *2016 IEEE International Conference on Rebooting Computing (ICRC)*, pp. 1–4, 2016.

[85] A. Schöll, C. Braun, and H.-J. Wunderlich, "Energy-efficient and error-resilient iterative solvers for approximate computing," in *2017 IEEE 23rd International Symposium on On-Line Testing and Robust System Design (IOLTS)*, pp. 237–239, 2017.

[86] G. Michelogiannakis, X. S. Li, D. H. Bailey, and J. Shalf, "Extending summation precision for network reduction operations," in *2013 25th International Symposium on Computer Architecture and High Performance Computing*, pp. 41–48, 2013.

[87] X. Cheng, M. Zhou, X. Song, M. Gu, and J. Sun, "Tolerating c integer error via precision elevation," *IEEE Transactions on Computers*, vol. 68, no. 2, pp. 270–286, 2019.

[88] A. Abdelfattah, S. Tomov, and J. Dongarra, "Towards half-precision computation for complex matrices: A case study for mixed precision solvers on gpus," in *2019 IEEE/ACM 10th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA)*, pp. 17–24, 2019.

[89] V. Alieksieiev, "One approach of approximation for incoming data stream in iot based monitoring system," in *2018 IEEE Second International Conference on Data Stream Mining Processing (DSMP)*, pp. 94–97, 2018.

[90] H. Esmaeilzadeh, A. Sampson, L. Ceze, and D. Burger, "Architecture support for disciplined approximate programming," *SIGPLAN Not.*, vol. 47, p. 301–312, mar 2012.

[91] D. Ernst, N. Kim, S. Das, S. Pant, R. Rao, T. Pham, C. Ziesler, D. Blaauw, T. Austin, K. Flautner, and T. Mudge, "Razor: A low-power pipeline based on circuit-level timing speculation," pp. 7– 18, 01 2004.

[92] B. Moons and M. Verhelst, "Dvas: Dynamic voltage accuracy scaling for increased energy-efficiency in approximate computing," in *2015 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pp. 237–242, 2015.

[93] P. Mantovani, E. G. Cota, K. Tien, C. Pilato, G. Di Guglielmo, K. Shepard, and L. P. Carlon, "An fpga-based infrastructure for fine-grained dvfs analysis in high-performance embedded systems," in *2016 53nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 1–6, 2016.

[94] D. Garg and R. Sharma, "Low power multiplier using dynamic voltage and frequency scaling (dvfs)," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 560–564, 2014.

[95] T. Wang, Q. Zhang, N. S. Kim, and Q. Xu, "On effective and efficient quality management for approximate computing," in *Proceedings of the 2016 International Symposium on Low Power Electronics and Design*, ISLPED '16, (New York, NY, USA), p. 156–161, Association for Computing Machinery, 2016.

[96] B. Grigorian, N. Farahpour, and G. Reinman, "Brainiac: Bringing reliable accuracy into neurally-implemented approximate computing," in *2015 IEEE 21st International Symposium on High Performance Computer Architecture (HPCA)*, pp. 615–626, 2015.

[97] D. S. Khudia, B. Zamirai, M. Samadi, and S. Mahlke, "Rumba: An online quality management system for approximate computing," in *2015 ACM/IEEE 42nd Annual International Symposium on Computer Architecture (ISCA)*, pp. 554–566, 2015.

[98] A. Sampson, W. Dietl, E. Fortuna, D. Gnanapragasam, L. Ceze, and D. Grossman, "Enerj: Approximate data types for safe and general low-power computation," *SIGPLAN Not.*, vol. 46, p. 164–174, June 2011.

[99] A. Sampson, W. Dietl, E. Fortuna, D. Gnanapragasam, L. Ceze, and D. Grossman, "Accept: A programmer-guided compiler framework for practical approximate computing," 2015.

[100] R. Boyapati, J. Huang, P. Majumder, K. H. Yum, and E. J. Kim, "Approx-noc: A data approximation framework for network-on-chip architectures," in

2017 ACM/IEEE 44th Annual International Symposium on Computer Architecture (ISCA), pp. 666–677, 2017.

[101] L. Wang, X. Wang, and Y. Wang, "Abdtr: Approximation-based dynamic traffic regulation for networks-on-chip systems," in *2017 IEEE International Conference on Computer Design (ICCD)*, pp. 153–160, 2017.

[102] Y. Chen, M. F. Reza, and A. Louri, "Dec-noc: An approximate framework based on dynamic error control with applications to energy-efficient nocs," in *2018 IEEE 36th International Conference on Computer Design (ICCD)*, pp. 480–487, 2018.

[103] M. F. Reza and P. Ampadu, "Approximate communication strategies for energy-efficient and high performance noc: Opportunities and challenges," in *Proceedings of the 2019 on Great Lakes Symposium on VLSI*, (New York, NY, USA), Association for Computing Machinery, 2019.

[104] F. Betzel, K. Khatamifard, H. Suresh, D. J. Lilja, J. Sartori, and U. Karpuzcu, "Approximate communication: Techniques for reducing communication bottlenecks in large-scale parallel systems," *ACM Comput. Surv.*, vol. 51, Jan. 2018.

[105] A. B. Ahmed, D. Fujiki, H. Matsutani, M. Koibuchi, and H. Amano, "Axnoc: Low-power approximate network-on-chips using critical-path isolation," in *2018 Twelfth IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*, (New York, NY, USA), pp. 1–8, IEEE, Oct 2018.

[106] F. Mandorlo *et al.*, "Controlled multi-wavelength emission in full cmos compatible micro-lasers for on chip interconnections," *Journal of Lightwave Technology*, vol. 30, pp. 3073–3080, Oct 2012.

[107] Z. Wang a *et al.*, "A holistic modeling and analysis of optical-electrical interfaces for inter/intra-chip interconnects," *IEEE TVLSI*, vol. 24, pp. 2462–2474, July 2016.

[108] M. Bahadori, S. Rumley, H. Jayatilleka, K. Murray, N. A. F. Jaeger, L. Chrostowski, S. Shekhar, and K. Bergman, "Crosstalk penalty in microring-based silicon photonic interconnect systems," *J. Lightwave Technol.*, vol. 34, pp. 4043–4052, Sep 2016.

[109] Y. Solihin, *Fundamentals of Parallel Multicore Architecture*. Chapman & Hall/CRC, 1st ed., 11 2015.

[110] S. Wuytack, J. da Silva, F. Catthoor, G. de Jong, and C. Ykman-Couvreur, "Memory management for embedded network applications," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 18, no. 5, pp. 533–544, 1999.

[111] C. Ykman-Couvreur, J. Lambrecht, A. van der Togt, and F. Catthoor, "Multi-objective abstract data type refinement for mapping tables in telecom network applications," vol. 38, no. 2 supplement, 2002.

[112] X. Wu, J. Xu, Y. Ye, Z. Wang, M. Nikdast, and X. Wang, "Suor: Sectioned undirectional optical ring for chip multiprocessor," *J. Emerg. Technol. Comput. Syst.*, vol. 10, June 2014.

[113] J. Luo, V.-D. Pham, C. Killian, D. Chillet, I. O'Connor, O. Sentieys, and S. Le Beux, "Run-time management of energy-performance trade-off in optical network-on-chip," in *2018 Conference on Design of Circuits and Integrated Systems (DCIS)*, pp. 1–6, 2018.

[114] T.-J. Kao and A. Louri, "Optical multilevel signaling for high bandwidth and power-efficient on-chip interconnects," *IEEE Photonics Technology Letters*, vol. 27, no. 19, pp. 2051–2054, 2015.

[115] A. Sampson, W. Dietl, E. Fortuna, D. Gnanapragasam, L. Ceze, and D. Grossman, "Accept: a programmer-guided compiler framework for practical approximate computing," *University of Washington Technical Report UW-CSE-15-01*, vol. 1, June 2015.

[116] T. E. Carlson, W. Heirman, S. Eyerman, I. Hur, and L. Eeckhout, "An evaluation of high-level mechanistic core models," *ACM Transactions on Architecture and Code Optimization (TACO)*, 2014.

[117] P. Dong, W. Qian, S. Liao, H. Liang, C. Kung, N. Feng, R. Shafiiha, J. Fong, D. Feng, A. V. Krishnamoorthy, and M. Asghari, "Low loss silicon waveguides for application of optical interconnects," in *IEEE Photonics Society Summer Topicals 2010*, pp. 191–192, 2010.

[118] 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.

[119] F. Xia, M. Rooks, L. Sekaric, and Y. Vlasov, "Ultra-compact high order ring resonator filters using submicron silicon photonic wires for on-chip optical interconnects," *Opt. Express*, vol. 15, pp. 11934–11941, Sep 2007.

[120] X. Chen, J. Feng, J. Xu, J. Zhang, and S. Chen, "Simultaneously tolerate thermal and process variations through indirect feedback tuning for silicon photonic net-

works," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 7, pp. 1409–1422, 2021.

[121] A. Sampson, W. Dietl, E. Fortuna, D. Gnanapragasam, L. Ceze, and D. Grossman, "Enerj: Approximate data types for safe and general low-power computation," *SIGPLAN Not.*, vol. 46, pp. 164–174, June 2011.

[122] D. Mahajan, K. Ramkrishnan, R. Jariwala, A. Yazdanbakhsh, J. Park, B. Thwaites, A. Nagendrakumar, A. Rahimi, H. Esmaeilzadeh, and K. Bazargan, "Axilog: Abstractions for approximate hardware design and reuse," *IEEE Micro*, vol. 35, no. 5, pp. 16–30, 2015.

[123] M. T. Hosseinabadi and N. Ansari, "Multi-power-level energy saving management for passive optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 6, no. 11, pp. 965–973, 2014.

[124] F. Lan, R. Wu, C. Zhang, Y. Pan, and K. T. Cheng, "Dlps: Dynamic laser power scaling for optical network-on-chip," in *2017 22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 726–731, 2017.

[125] Y. Agrawal, R. Chandel, and R. Dhiman, "Design and analysis of efficient multilevel receiver for current mode interconnect system," in *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*, pp. 1–6, 2014.

[126] M. Dave, M. Jain, M. Shojaei Baghini, and D. Sharma, "A variation tolerant current-mode signaling scheme for on-chip interconnects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 2, pp. 342–353, 2013.

[127] Y. Agrawal, R. Chandel, and R. Dhiman, "High-performance current mode receiver design for on-chip vlsi interconnects," *Advances in Intelligent Systems and Computing*, vol. 343, pp. 527–536, 01 2015.

[128] N. M. Fahrenkopf, C. McDonough, G. L. Leake, Z. Su, E. Timurdogan, and D. D. Coolbaugh, "The aim photonics mpw: A highly accessible cutting edge technology for rapid prototyping of photonic integrated circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–6, 2019.

[129] P. Dong, W. Qian, S. Liao, H. Liang, C.-C. Kung, N.-N. Feng, R. Shafiiha, J. Fong, D. Feng, A. V. Krishnamoorthy, and M. Asghari, "Low loss shallow-ridge silicon waveguides," *Opt. Express*, vol. 18, pp. 14474–14479, Jul 2010.

[130] Q. Wilmart, H. El Dirani, N. Tyler, D. Fowler, S. Malhouitre, S. Garcia, M. Casale, S. Kerdiles, K. Hassan, C. Monat, X. Letartre, A. Kamel, M. Pu, K. Yvind, L. K.

Oxenløwe, W. Rabaud, C. Sciancalepore, B. Szelag, and S. Olivier, "A versatile silicon-silicon nitride photonics platform for enhanced functionalities and applications," *Applied Sciences*, vol. 9, no. 2, 2019.

[131] J. Lee, C. Killian, S. L. Beux, and D. Chillet, "Approximate nanophotonic interconnects," in *Proceedings of the 13th IEEE/ACM International Symposium on Networks-on-Chip*, NOCS '19, (New York, NY, USA), Association for Computing Machinery, 2019.

List of Figures

List of Tables

UNIVERSITÉ DE RENNES 1

Titre : Exploration de techniques de communication approximatives pour des interconnexions nano-photoniques efficaces

Mot clés : Réseaux Optique sur puce, Système sur Puce

Résumé : Depuis leur apparition, les systèmes sur puce (SoC) ont évolué en intégrant progessivement un nombre croissant de processeurs contenant des milliards de transistors. Avec cette évolution des SoC, de nouveaux sujets de recherche ont surgi, en particulier concernant les interconnexions entre les cœurs. Le concept de réseau sur puce (NoC) a été proposé comme solution pour supporter des communications plus dynamiques pour ce type d'architectures. Néanmoins, ces NoCs électriques présentent des inconvénients en termes de latence et de consommation. Pour surmonter ces problèmes, les réseaux optiques (ONoC) sont apparus comme une réelle opportunité de par leur faible latence de communication et une grande bande passante. Toutefois, ces réseaux souffrent d'une faible efficacité énergétique. Dans cette thèse, nous proposons des techniques pour gérer la consommation d'énergie des ONoCs. Pour cela, nous exploitons le concept d'approximation et nous l'appliquons à ces réseaux pour proposer des communications précise ou approximables. Notre proposition s'applique aux nombres à virgule flottante (FP) en utilisant des signaux optiques de faible puissance pour les LSB, au prix d'un taux d'erreur plus élevé. Ces signaux optiques approximatifs permettent une réduction drastique de la consommation d'énergie du laser. En parallèle, pour assurer la précision de la communication sur les MSB, les niveaux de puissance laser sont maintenus en utilisant des signaux de forte puissance. Les résultats des simulations montrent qu'une réduction de 42% de la puissance laser peut être obtenue pour l'application Streamcluster avec une dégradation limitée au niveau de l'application. De plus, nous proposons de gérer les communications en fonction de la distance entre les cœurs source et destination. Pour limiter la complexité, nous proposons une technique de gestion de la distance avec un faible coût basée sur deux classes de distances : courtes ou longues. Les résultats de notre évaluation montrent une réduction drastique de la puissance laser de 20% pour l'application Streamcluster.

Title: Approximate communication techniques exploration for efficient nano-photonics interconnects

Keywords: Optical network on chip, System-on-chip

Abstract: Over the years, System-on-Chip (SoC) has evolved from a single processor in a chip to multi/many processors in chips containing billions of transistors. With the evolution of SoC, new research topics have risen on interconnect between processors in a chip. **Network-on-Chip(NoC)** has been proposed as a solution for more dynamic communication links to connect large number of Intellectual Property (IP)s. Secondly, to overcome drawbacks of electrical NoC, **Optical communication link** has been proposed as a promissing solution. This type of NoC provides low latency and high bandwith, but it suffers from low power efficiency. In this thesis, we address this topic and we aim to develop techniques to manage laser power consumption. To address this challenge, we exploit the approximation concept and we apply it to the ONoC to propose two types of communications: Accurate and **Approximate communication**. Our proposal is applied to Floating-Point (FP) numbers by using low-power optical signals for LSB, at the cost of higher error rate. These approximate optical signals allow a drastic reduction in the laser power consumption. In parallel, to ensure the communication accuracy on MSBs, the laser power levels are remained using high power signals. Simulations results demonstrate that a reduction of to 42% of laser power can be obtained for Streamcluster application with a limited degradation at the application level. Furthermore, we propose to manage the communications accordong to the distance between source and destination cores. However, this fine-grain distance-aware management could be too costly, so we propose a low overhead distance-aware technique based on only two distances **Short/Long** classes. The results of our evaluation show drastic laser power reduction of 20%for example for Streamcluster application.