

Interface Electrique/Optique pour un ONoC

Dung PHAM VAN¹, Daniel CHILLET¹, Cédric KILLIAN¹, Olivier SENTIEYS¹, Sébastien LE BEUX², Ian O'CONNOR²

¹IRISA - Université de Rennes 1

²INL - Ecole Centrale de Lyon

Résumé – La photonique sur silicium a fait d'énormes progrès depuis quelques années, permettant d'intégrer des composants optiques au sein d'architectures multi-couches. Dans ce contexte, il est donc maintenant possible de proposer l'intégration de NoC optique (ONoC) au sein d'une architecture multiprocesseurs afin de supporter des échanges extrêmement efficaces entre les cœurs de calcul. Une communication dans un ONoC s'effectue alors via des lasers, utilisable en parallèle sur des longueurs d'ondes distinctes, permettant la transmission d'information série à 10Gbit/s sur chacune d'elles. Toutefois, pour bénéficier de ces débits, les éléments clés de l'association d'un ONoC au sein d'une architecture électrique sont les interfaces électrique \leftrightarrow optique qui doivent pouvoir supporter des fréquences de conversion parallèle/série élevées. Cet article se situe dans ce contexte et propose une interface pour la gestion d'un NoC optique pour des échanges entre des tâches réparties sur différents cœurs d'une architecture multi-cœurs. L'interface proposée permet de supporter l'allocation de longueurs d'ondes et permet également une gestion de la consommation énergétique via l'utilisation de codes correcteurs d'erreurs afin de réduire la puissance d'émission des lasers. Cet article illustre le fonctionnement de cette interface et donne les résultats d'implémentation de celle-ci pour une technologie FDSOI 28nm.

Abstract – Optical Network-on-Chips are nowadays a mature technology which can be embedded in high performance multi-processors to support high bandwidth communications. This technology enables the use of several wavelengths for the same communication to increase the bandwidth and decrease the communication time. Nevertheless, supporting such high bandwidth need efficient interface between electrical and optical domains. This paper presents the interface developed to manage the wavelength allocation, but also to support the management of laser power in order to reduce the power consumption. This power consumption reduction by exploiting the tradeoff between data encoding in electrical domain and energy needed in optical domain to produce wavelengths through lasers.

1 Introduction

L'accroissement des performances dans les architectures de calcul est en grande partie tirée par l'augmentation du nombre de cœurs intégrés au sein de celles-ci. Cette évolution conduit les informaticiens à repenser leurs développements afin de profiter de la multitude de ressources de calcul pour paralléliser les applications. Celles-ci sont alors décrites par un ensemble de tâches, modélisé par un graphe de tâches exhibant des dépendances liées aux échanges de données. L'allocation de ces tâches sur les ressources de calcul est alors un processus qui doit tirer le meilleur parti des capacités de calcul de l'architecture. Toutefois, l'augmentation du parallélisme d'une application, par une découpe en un plus grand nombre de tâches, va inexorablement conduire à une augmentation du volume de données à échanger entre les ressources de calcul. Dans ce contexte, si les tâches peuvent profiter de l'augmentation de performance offerte par la multiplication des ressources de calcul, elles peuvent souffrir de ralentissement du au volume de données à échanger sur un support qui devient le goulot d'étranglement du système. Il apparaît alors crucial de doter ces architectures d'un média de communication efficace, pouvant supporter des échanges volumineux avec des temps de communication très courts. Les interconnexions de type Network-on-Chip ont largement fait leur preuve dans le contexte des architectures multiprocesseurs, malheureusement, dans un contexte où le "multi" de multi-processeur devient "massif" alors ces réseaux voient leur latence et leur consommation énergétique fortement augmenter [1]. Fort heureusement, depuis quelques années, la photonique a fait de gros progrès et des solutions d'intégration de composants

optiques sur des couches de matériaux compatibles avec le silicium deviennent réalité. Le potentiel de ce type de composants décuple les capacités de transfert au sein d'une architecture et offre alors une opportunité tout à fait intéressante pour contourner les limitations des NoC classiques. Pour être pleinement exploitable, ces composants optiques, et notamment les lasers produisant les longueurs d'ondes supportant les échanges, doivent être interfacés avec l'électronique de l'architecture multiprocesseurs. Actuellement, les lasers embarqués sur puce peuvent fournir des débits de 10Gbit/s (transmission série) et peuvent être parallélisés en émettant sur des longueurs d'ondes distinctes. Ainsi, le débit d'une communication est multiplié par le nombre de lasers alloués et permet donc de diminuer le temps de communication. Cependant, étant donné les débits élevés, les interfaces effectuant la conversion électrique vers optique, et optique vers électrique, doivent être efficaces pour que les performances des ONoC ne soient pas bridées par l'électronique associée.

Les interfaces de l'état de l'art ne considèrent pas la génération de plusieurs flux séries (cas de transmission sur plusieurs lasers), bien que cette technique soit fréquemment citée comme cas d'études [2, 3]. De plus, le lien avec le système d'exploitation est rarement présenté alors que ce dernier est le garant d'une utilisation efficace du support de communication.

Les travaux présentés dans cet article se situent dans ce contexte et illustrent l'interface que nous avons définie pour associer l'utilisation d'un NoC optique et une architecture multiprocesseurs. Cette interface est composée de trois grandes parties répondant aux trois fonctionnalités importantes permettant de prendre en

charge une donnée depuis un processeur source pour l'acheminer vers un ou plusieurs lasers en vue d'une transmission sur une ou plusieurs longueurs d'ondes. L'article présente l'objectif de ces trois fonctionnalités, et détaille le fonctionnement global de l'interface. L'article est organisé de la façon suivante. La section 2 présente un exemple d'architecture générale dans laquelle nos travaux peuvent prendre place, mais il est à signaler que la proposition présentée dans ce papier peut également être implémentée dans d'autres architectures de NoC optiques. La section 3 présente l'architecture globale ainsi que l'architecture spécifique définie pour l'interface entre les parties électronique et optique. La section 4 explique le fonctionnement de l'interface ainsi que la manière dont les données sont acheminées depuis le processeur émetteur du message vers le guide d'ondes. Enfin, la section 5 conclut cet article et présente les perspectives de ce travail.

2 Architecture générale du système

L'architecture qui sert de support pour ces travaux est l'architecture Chaméléon [3]. Elle est représentée par le schéma de la figure 1. Cette architecture est basée sur deux couches de silicium : une couche électrique contenant les processeurs et une couche contenant le média de communication optique. Ce média de communication est principalement basé sur la présence d'un guide optique permettant de guider les signaux optiques (ou longueurs d'ondes) au sein du silicium.

Sur la couche électrique, l'architecture est organisée autour d'une grille de processeurs (ou de clusters), où chaque processeur peut effectuer des requêtes de communication, se traduisant par une requête d'accès au guide d'ondes pour transférer des données à un processeur destinataire. Chaque processeur (ou cluster) dispose donc d'un ONI (Optical Network Interface) lui permettant de faire des requêtes d'allocation puis de transmettre ou de recevoir des données.

Chaque ONI est composé d'une partie émettrice, contenant des lasers et des Micro Résonateurs en anneau (MR), et d'une partie réceptrice, composée de MR et de photo-détecteurs.

Un MR permet i) d'injecter (respectivement d'extraire) un signal optique (en étant activé) dans (respectivement depuis) un guide d'ondes ; ii) de laisser passer un signal sans le modifier. Chaque MR est associé à une longueur d'onde et peut être contrôlé indépendamment (activé ou désactivé) selon les besoins de communication.

La figure 1 illustre l'établissement de deux communications en parallèles. Une première communication est générée entre l'ONI_A et l'ONI_C sur une longueur d'onde illustrée en rouge. Pour ce faire, le laser "rouge" est activé, ainsi que le MR associé afin d'injecter le signal optique au sein du guide d'ondes optique. Les lasers supportent une modulation OOK et sont pilotés directement par les données binaires sérialisées venant des processeurs. C'est donc le laser qui supporte la conversion du domaine électrique vers le domaine optique. Le message est à destination de l'ONI_C. Ainsi au sein de l'ONI_B, le MR "rouge" de la partie réceptrice est désactivé, laissant passer la longueur d'onde "rouge". Au sein de l'ONI_C, le MR "rouge" est activé et permet d'extraire le signal "rouge" afin de le diriger vers le photo-dectecteur qui fait la conversion optique vers électrique.

En parallèle, l'ONI_B effectue une communication sur la longueur d'onde "verte", ce qui nécessite l'activation du MR "vert" au sein de la partie émettrice de l'ONI_B, et de celui au sein de la partie réceptrice de l'ONI_C. Comme les deux communications s'effectuent sur deux longueurs d'onde différentes, il n'y a pas de

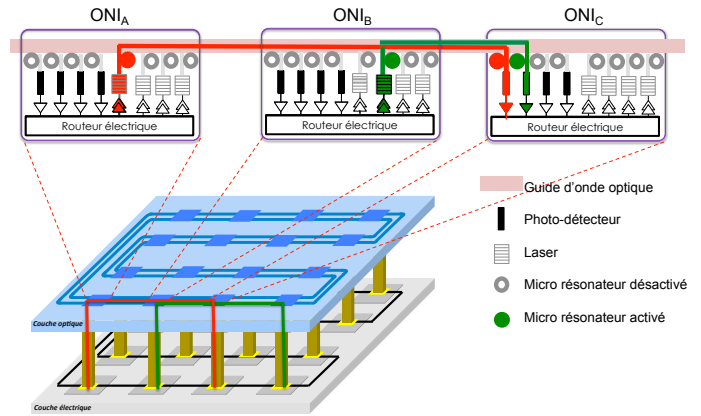


FIGURE 1 – Architecture ONoC et interfaces optiques

conflit. Dans cette chaîne de communication, les composants optiques induisent des pertes, on pourra notamment citer des pertes dues à la longueur du guide d'onde, aux courbes subies par les longueurs d'ondes au travers d'un guide d'onde non géométriquement linéaire, etc. En plus de ces pertes optiques, l'un des principaux inconvénients de la photonique sur silicium, concerne l'énergie consommée par les composants à implémenter. Ceci est particulièrement vrai pour le laser, qui dispose dans le meilleur des cas, d'un rendement de l'ordre de 5 à 10%. Pour assurer une certaine qualité de communication entre un processeur source et un processeur destination, il faut évaluer les pertes induites dans le guide d'ondes et piloter le laser avec suffisamment de puissance pour que le signal reçu soit de bonne qualité. Notons que si une certaine qualité de communication est attendue, quantifiable par un BER, alors l'objectif consiste à ne pas augmenter la puissance laser au delà du nécessaire, sachant que tout *Watt* supplémentaire dans le domaine optique nécessite 20 fois plus de puissance pour la commande électrique du laser. Dans ces conditions, il est évident que ce média de communication doit être géré de façon très précise afin de limiter autant que faire se peut les consommations inutiles.

L'étude des pertes optiques ne fait pas l'objet de ce papier, pour plus de précision sur ces points, se référer à [4]. Sur la base des études faites précédemment, nous exposons ici l'architecture de l'interface que nous avons mise en place pour piloter cet ONI.

3 Architecture de l'interface

Comme mentionné précédemment, l'interface est dotée de plusieurs éléments importants qui ont chacun un rôle bien spécifique et qui répondent à une fonctionnalité particulière. Ces fonctionnalités sont au nombre de trois et concernent :

- Fonction de codage canal des données en vue d'augmenter la protection de celles-ci et de réduire la puissance laser au strict minimum permettant de cibler un BER donné pour chaque communication ;
- Fonction de sérialisation depuis le bus processeur ayant subi un codage canal vers un ou plusieurs flux de données 1 bit ;
- Fonction d'allocation des flux de données 1 bit vers les lasers en vue de leur modulation.

Ces fonctionnalités sont représentées sur la figure 2 et sont rapidement détaillées dans la suite de cette section.

3.1 Encodage des données

La première fonctionnalité qui est supportée par cette interface concerne le codage des données émises dans le but de pouvoir

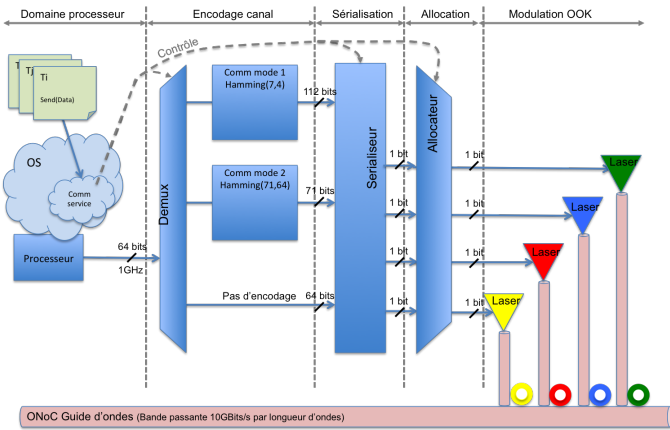


FIGURE 2 – Modèle global de l’interface pour l’émission des données dans le guide d’ondes optique.

détecter des erreurs de transmission et d’éventuellement pouvoir les corriger. On peut alors parler de codage canal de données en vue de la transmission. Cette fonctionnalité est représentée par la partie *Encodage de canal* de la figure 2. Compte tenu des débits que peut supporter l’optique intégrée (10Gbits/s), le codage envisagé doit rester simple. Nous avons opté pour un codage de Hamming ajoutant quelques bits de redondance. Ce code permet de corriger une erreur de transmission dans le mot codé. Dans cette fonctionnalité de codage de données, nous proposons l’utilisation de plusieurs modes de codage, donnant des niveaux de corrections différents. Trois modes sont alors proposés :

1) Le premier mode consiste à effectuer un codage de Hamming sur des mots de 4 bits, en ajoutant 3 bits de redondance, codage de Hamming appelé ici $H(7, 4)$. Ce codage permet une protection assez fine des données, mais il est nécessaire d’implémenter 16 encodeurs en parallèle pour un bus processeur de 64 bits, et il est coûteux en termes de temps de communication, puisque chaque message contenant N bits à transmettre va voir sa taille augmenter d’un rapport 7/4. Par ailleurs, la protection des données étant “forte”, donc le pouvoir de correction important, il est possible, pour une qualité de communication ciblée (un Bit Error Rate, BER, ciblé), de réduire la puissance d’émission du laser.

2) Le second mode consiste à effectuer un codage de Hamming sur des mots de 64 bits, en ajoutant 7 bits de redondance, codage de Hamming appelé ici $H(71, 64)$. Dans ce cas, un seul encodeur est suffisant pour un bus processeur de 64 bits, mais il conduit à une protection plus grossière des données. Il est moins coûteux que le précédent puisque chaque message va voir sa taille augmenter d’un rapport 71/64. Le surcoût temporel est donc moindre mais pour une qualité de transmission ciblée (donc pour une BER ciblé) il faut augmenter la puissance d’émission du laser.

3) Finalement, un dernier mode est proposé, il s’agit d’un mode sans protection des données, c’est-à-dire un mode dans lequel aucun codage canal des données n’est réalisé. L’avantage de ce mode de communication concerne l’absence d’impact temporel pour assurer la communication, mais l’absence de protection des données nécessite, pour une qualité de transmission définie, une puissance d’émission du laser plus importante que dans les 2 modes précédents.

L’étude de ces différents modes de communication a été réalisée, et des résultats complets de leur efficacité sont disponibles dans [5]. À partir de cette étude montrant l’intérêt de ces différents modes, nous proposons que la décision quant au mode de communication à utiliser soit prise par le système d’exploitation au moment de la requête d’envoi de messages.

3.2 Sérialisation du flux de données

La seconde fonctionnalité à assurer dans l’interface concerne la sérialisation des données en vue d’une transmission bit à bit au travers de la modulation OOK des lasers. Cette fonctionnalité est représentée par la partie *Sérialisation* de la figure 2. Dans son principe, la sérialisation est simple, mais elle est légèrement complexifiée dans notre contexte puisque nous proposons que chaque communication puisse se faire sur plus d’une longueur d’ondes. Dans ce cas, le bloc de sérialisation doit être capable de produire plusieurs flux de données sur 1 bit à partir des données, éventuellement encodées, provenant du processeur sur 64 bits. Le bloc de sérialisation doit donc pouvoir être configuré pour pouvoir produire un ou plusieurs flux de données.

3.3 Allocation des longueurs d’ondes

Finalement, la dernière fonctionnalité que doit supporter l’interface concerne l’allocation des flux de données 1 bit sur les lasers sélectionnés pour la communication à assurer. Cette fonctionnalité est représentée par la partie *Allocation* de la figure 2. Le rôle de ce bloc consiste à *driver* les flux de données binaires produits par le sérialiseur vers les lasers ayant été retenus pour la communication. Le choix des lasers correspond aux choix de longueurs d’ondes et est réalisé par le système d’exploitation lors d’une requête de communication.

4 Fonctionnement de l’interface

Comme nous l’avons indiqué, l’interface se situe entre le système d’exploitation, qui capte les appels systèmes des tâches applicatives souhaitant transmettre des messages (ces appels systèmes sont de la forme `send(Dest, Mess)` et `rcv(Src, Mess)`) et le domaine optique supportant le transport des données (ou messages). La figure 3 illustre la circulation de paquets de données codées sur 64 bits depuis le processeur émetteur jusqu’au guide d’ondes pour le cas d’une tâche souhaitant réaliser un transfert de données et pour lequel le système d’exploitation a alloué 3 longueurs d’ondes (λ_1 “jaune”, λ_2 “rouge” et λ_4 “verte”) afin d’offrir un débit de 30 Gbits/s pour cette communication. La stratégie de sélection des longueurs d’ondes n’est pas détaillée ici, rappelons simplement qu’elle consiste à retenir les longueurs d’ondes les mieux adaptées, à l’instant de communication, pour réduire les pertes optiques dans le guide d’ondes et le *crosstalk* entre les communications en cours dans le guide d’ondes. Pour plus de détails, le lecteur pourra se référer à [4]. Cette stratégie permet alors au système d’exploitation de contrôler les différents composants de l’interface, à savoir i) le routage des données vers le bon encodage, ii) la sérialisation en un ou plusieurs flux binaires, iii) l’allocation des longueurs d’ondes pour chacun de ces flux binaire. Ce contrôle est représenté par les flèches en pointillés gris de la figure 2.

L’initiateur de la communication est toujours une tâche applicative, illustrée sur la figure 3 par la tâche $T1$ dans la partie en haut à gauche. La requête de cette tâche est captée par le système d’exploitation qui a une connaissance précise des communications en cours dans le guide d’ondes, et qui connaît donc l’état d’utilisation des longueurs d’ondes dans l’ensemble du guide. À partir de cette information globale, le système d’exploitation évalue le nombre de longueurs d’ondes qui peut être alloué pour la communication initiée par la tâche $T1$. Nous supposons ici que le système d’exploitation décide d’allouer 3 longueurs d’ondes (λ_1 , λ_2 et λ_4) pour cette communication. Sur la base d’une évaluation des pertes optiques que vont subir ces 3 longueurs d’ondes dans le guide d’ondes, le système d’exploitation est également en mesure de déterminer la puissance optique nécessaire pour assurer une

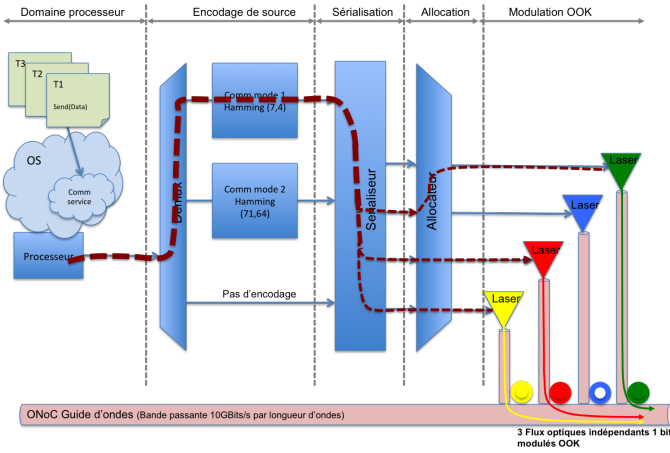


FIGURE 3 – Exemple de communication sur 3 longueurs d’ondes.

qualité de communication entre le processeur source et le processeur destination (BER ciblé, ou encore SNR ciblé). Compte tenu de la consommation des lasers, l’ajout, dans la chaîne de communication, d’un code correcteur d’erreur dans le but de réduire la puissance d’émission laser pour atteindre la même qualité de communication est évalué. Il est alors possible d’activer l’un des 2 codeurs disponibles, ou décider de ne pas coder les données afin de ne pas rallonger le temps de transfert du message complet.

Dans la figure 3, nous supposons que le choix s’est porté sur un encodage des données par un encodage de $H(7, 4)$ permettant, à partir de 16 encodeurs parallèles, de coder le bus 64 bits du processeur. Cet encodage conduit donc à l’ajout de bits de redondance, 3 bits de redondances pour 4 bits de données, conduisant alors à l’ajout de 48 bits pour encoder un bus 64 bits. La sortie de l’encodage est donc un flux de mots sur 112 bits qu’il faut alors sérialiser. Le choix d’allouer 3 longueurs d’ondes pour la communication conduit à découper le flux de mots sur 112 bits en 3 flux binaires. Finalement, les 3 flux binaires sont ensuite routés vers la modulation des lasers des longueurs d’ondes retenues. L’activation des lasers et des micro-résonateurs associés permet d’injecter les longueurs d’ondes dans le guide d’ondes.

Du côté du récepteur, une structure symétrique est mise en place, elle n’est pas présentée dans cet article faute de place. À partir des choix fait à l’émission, le système d’exploitation peut configurer l’ONI récepteur de façon à ce que celui-ci active i) les micro-résonateurs permettant d’extraire les longueurs d’ondes λ_1 , λ_2 et λ_4 , ii) la désérialisation des 3 flux binaires pour reconstituer des mots de 112 bits, iii) le décodage de $H(7, 4)$ permettant de retrouver les données transmises.

Nous avons modélisé cette interface et réalisé des synthèses de celle-ci en utilisant une technologie FDSOI 28 nm. L’interface conçu supporte la gestion de 4 longueurs d’ondes, ce qui permet de dimensionner le sérialiseur (4 sorties) et l’allocation (4 entrées vers 4 sorties). La table 1 présente les résultats de synthèse de l’interface complète et donne également une estimation de la consommation globale de cette interface pour une fréquence de fonctionnement de 1 GHz pour l’encodage, et de 10 GHz pour les autres éléments de la chaîne de communication. Pour la puissance totale consommée, plusieurs valeurs sont données, ces valeurs correspondent au choix qui est fait pour l’encodage. Lorsqu’aucun encodage n’est utilisé, les blocs $H(7, 4)$ et $H(71, 64)$ sont inactifs et leur consommation dynamique est supposée nulle, la puis-

Blocs Fonctionnalités	Nb blocs	Fréquence I/O GHz	Surface μm^2	Puissance totale μW
$Hamming(7, 4)$	16	1 \rightarrow 1	551	3, 14
$Hamming(71, 64)$	1	1 \rightarrow 1	491	2, 52
Sérialiseur(64 \rightarrow 4)	1	1 \rightarrow 10	1664	1, 1
Allocation(4 \rightarrow 4)	1	10 \rightarrow 10	90	0, 8
Total pour la transmission d’une data 64 bits sur 1 bit				1, 9
				4, 42
				5, 04

TABLE 1 – Résultats de synthèse de l’interface ONI.

sance consommée est alors égale à $1, 9 \mu W$. Les deux autres valeurs de consommation, $4, 42 \mu W$ et $5, 04 \mu W$, correspondent respectivement aux cas i) avec encodage $H(71, 64)$ nécessitant un seul encodeur, ii) avec encodage $H(7, 4)$ nécessitant 16 encodeurs en parallèle. Pour comparaison, la puissance optique d’un laser est de l’ordre de $800 \mu W$ pour un rendement de l’ordre de 5%, ce qui conduit à une puissance électrique pour commander le laser de l’ordre de $16 mW$. Ce dernier chiffre montre que le surcot engendré par l’ajout d’un encodage/décodage de Hamming est négligeable par rapport à la puissance laser nécessaire.

5 Conclusion

Cet article présente la structure de l’interface de communication permettant de connecter un ensemble de processeurs à un NoC optique. L’interface proposée supporte la fonction de base qui consiste à passer du domaine électrique vers le domaine optique, mais elle est également équipée de mécanismes permettant i) l’allocation de plusieurs longueurs d’ondes pour une seule et même communication dans l’objectif d’augmenter la bande passante et donc de réduire le temps de communication ; ii) l’insertion d’un encodage des données en vue de réduire la puissance nécessaire pour générer les longueurs d’ondes.

Les résultats de synthèse et de simulation que nous avons obtenus montrent que notre proposition est implémentable dans un système complet. Les perspectives de ces travaux concernent l’étude des points de fonctionnement des lasers en vue de définir plus finement les niveaux d’émission laser. En effet, dans l’étude réalisée ici, le laser est supposé commandable sur une plage continue de valeurs, alors qu’il est probable qu’un ensemble réduit de points de fonctionnement soit suffisant.

Remerciements

This work has received a French Government support granted to the COMIN Labs excellence laboratory and managed by the National Research Agency in the “Investing for the Future” program under reference ANR-10-LABX-07-01.

Références

- [1] A. Karkar, T. Mak, K. F. Tong, and A. Yakovlev, “A survey of emerging interconnects for on-chip efficient multicast and broadcast in manycores,” *IEEE Circuits and Systems Magazine*, vol. 16, no. 1, pp. 58–72, Firstquarter 2016.
- [2] X. Wang et al., “Rpnoc : a ring-based packet-switched optical network-on-chip,” *IEEE Photonics Technology Letters*, vol. 27, no. 4, pp. 423–426, 2015.
- [3] S. L. Beux, H. Li, I. O’Connor, K. Cheshmi, X. Liu, J. Trajkovic, and G. Nicolescu, “Chameleon : Channel efficient optical network-on-chip,” in *2014 Design, Automation Test in Europe Conference Exhibition (DATE)*, March 2014, pp. 1–6.
- [4] J. Luo, A. Elantably, V. Pham, C. Killian, D. Chillet, S. L. Beux, O. Sentieys, and I. O’Connor, “Performance and energy aware wavelength allocation on ring-based wdm 3d optical noc,” in *Conference on Design, Automation and Test in Europe*, Lausanne, 2017.
- [5] 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 *Design Automation Conference*, Austin, 2017.