

Run-Time management of energy-performance trade-off in Optical Network-on-Chip

Jiating Luo, Van-Dung Pham, Cédric Killian, Daniel Chillet, Ian O'Connor, Olivier Sentieys, Sébastien Le Beux

▶ **To cite this version:**

Jiating Luo, Van-Dung Pham, Cédric Killian, Daniel Chillet, Ian O'Connor, et al.. Run-Time management of energy-performance trade-off in Optical Network-on-Chip. DCIS 2018 - XXXIII Conference on Design of Circuits and Integrated Systems, Nov 2018, Lyon, France. pp.1-6, 10.1109/DCIS.2018.8681477 . hal-01937350

HAL Id: hal-01937350
https://hal.science/hal-01937350

Submitted on 28 Nov 2018

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

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

Run-Time management of energy-performance trade-off in Optical Network-on-Chip

Jiating Luo*, Van-Dung Pham*, Cédric Killian*, Daniel Chillet*,
Ian O'Connor†, Olivier Sentieys*, Sébastien Le Beux†
*Univ Rennes, Inria, CNRS, IRISA, F-22305 Lannion, France, firstname.lastname@irisa.fr
†Ecole Centrale de Lyon, INL, Ecully, F-69134, France, firstname.lastname@ec-lyon.fr

***Abstract*—Optical Network-on-Chip is a promising technology to leverage the interconnection bottleneck in manycore architectures. Indeed, this technology offers high bandwidth and low latency for communications between cores or clusters of cores. However, the high static power consumption of Optical NoCs (ONoCs) calls for reconfiguration of the interconnect in order to meet the performance requirements while minimizing the required energy. This paper addresses this challenge and proposes a method allowing to define, at design time, a set of ONoC execution modes to be loaded, at run-time, according to the applications' performance and energy requirements. The proposed methodology relies on a sequencer allowing to configure Optical Network Interfaces (ONI).**

I. Introduction

Multiprocessor System-on-Chip (MPSoC) are evolving towards the integration of hundreds of cores on a single chip. Designing an efficient interconnect for such complex architectures is challenging due to the ever growing data exchange between processors. Networks on Chip (NoCs) have been proposed to overcome these issues but they are now reaching their performance limits. Indeed, packetization and depacketization of data drastically impact NoC latency [1] and the same bandwidth for all router ports is usually assumed while processors bandwidth requirements vary significantly [2]. NoC efficiency thus needs to be further improved in order to efficiently interconnect systems with still increasing number of cores and required bandwidth density [3].

To overcome the limitations of classical NoCs, nanophotonic interconnects seem a promising technology to replace support communications in MPSoCs. Indeed, Wavelength Division Multiplexing (WDM) allows the propagation of multiple signals simultaneously on a same waveguide [4], thus leading to high aggregated bandwidth. However, WDM also leads to inter-channel crosstalk noise [5], which negatively impacts the Signal to Noise Ratio (SNR) and, therefore, the Bit Error Rate (BER). The BER to be reached depends on the application requirements and is closely related to the optical devices characteristics: it depends on the photodetector sensitivity [6], on the Microring Resonator transmission spectrum, and on the optical signal power emitted by the laser sources. The higher the number of signals propagating simultaneously, the higher the crosstalk and the higher the laser output power needed. This thus leads to the following conflicting objectives: high performance communications tend to rely on an exhaustive use of the available wavelengths while energy efficient communications involve a parsimonious use of signals which occupy distinct and separate wavelengths. Using an Optical NoC (ONoC) to support the communications of a given application is thus a tedious task, especially if performance, power and BER objectives are likely to evolve with the execution context.

Hence, for an application executed on a given architecture, several solutions allow implementing the communications, each one providing an optimal solution but with a unique trade-off between energy and performance. In the state of the art, ONoCs can be classified into two families regarding the communication allocation strategy: online or offline. However, to the best of our knowledge, there is no architecture allowing run-time adaptation of the ONoC to meet energy and performance application requirements for both types of allocations.

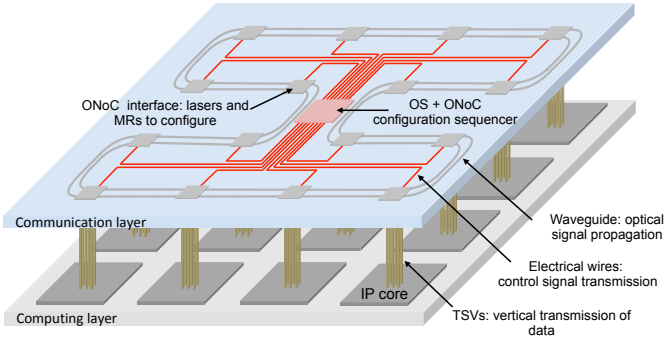


Fig. 1: Considered 3D ONoC architecture with the configuration sequencer located in the center of optical layer.

In this paper, we propose a method allowing run-time adaptation of the ONoC according to energy and performance requirements of the executed applications. The method relies on an offline framework to generate a set of ONoC configurations to be stored in ONIs. An ONoC configuration sequencer, located in the center of the optical layer (Figure 1), then selects at run-time the configuration to be loaded according to the application needs.

The paper is organized as follow: Section II presents the proposed hardware manager allowing run-time adaptation of the ONoC. Section III describes the management of the configurations of the ONIs. Finally we conclude this paper in Section IV.

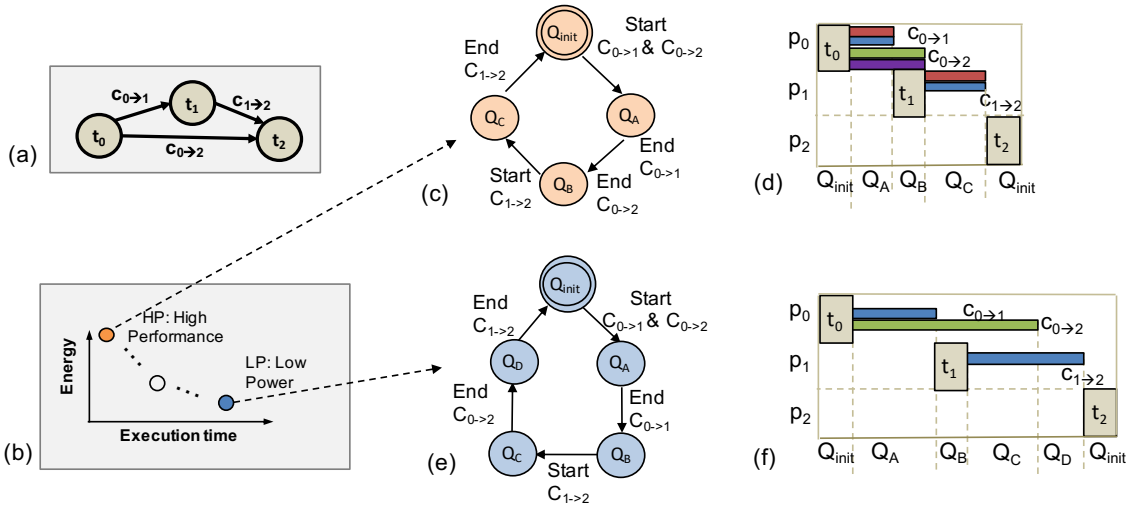


Fig. 2: Energy-performance trade-off: a) Application represented as a DAG, b) Energy versus Performance plot highlighting High Performance (HP) and Low Power (LP) modes, c-d) ONoC configuration sequence for HP mode and the resulting execution traces, e-f) ONoC configuration sequence of LP mode and the resulting traces.

II. Run-time Adaptation of ONoC Energy and Performance

A. Considered architecture and applications

We consider a 3D ONoC, i.e. a 3D integrated circuit supporting mixed-technology with electronic-controlled Optical NoC. It combines the benefits of silicon photonics and 3D stacking to overcome traditional electrical network limitations. Photonic interconnect delivers high on-chip bandwidth and low latency while 3D stacking can reduce the interconnect distance. Figure 1 shows a ring-based 3D ONoC architecture composed of two layers interconnected by TSV: i) on the bottom, an electrical layer implementing processing cores and ii) on the top, an optical layer with optical routers and waveguides.

Following the layout defined in [7], a centralized ONoC configuration sequencer driven by an Operating System allows run-time adaptation of the ONIs, as illustrated in red in Figure 1. The OS is responsible for deploying tasks among the cores and defining the strategy to adapt the ONoC to meet application requirements, such as energy consumption or execution time (called performance within this paper). The sequencer is responsible for reconfiguring the ONoC according to the mode selected by the OS.

Figure 2 illustrates the dependencies between the application tasks, the allocated bandwidth and the performance. Applications are represented as a Directed Acyclic Graph (DAG) as depicted in Figure 2-a, where each vertex t_i represents a task and each directed edge ($c_{i\to j}$) represents a communication from t_i to t_j. In this example, we assume that tasks t_0, t_1, and t_2 are mapped onto different processors. Communications $c_{0\to1}$, $c_{0\to2}$, and $c_{1\to2}$ are implemented using an ONoC which is configured according to execution performance and energy requirements.

As we consider WDM communication and laser power tuning, the energy consumption and the execution time of the application are linked to the resources allocated for each communication of the DAG. Indeed, it is possible to adapt the bandwidth for each communication, by varying the number of wavelengths, and to adapt the power laser to meet a targeted communication quality through the BER (not illustrated in this figure). Figure 2-b illustrates energy versus execution time plot and shows a Pareto front of communication solutions for the architecture. This figure also highlights HP (High Performance) and LP (Low Power) modes on the Pareto front, which allows to maximize the performance and energy efficiency respectively.

Figures 2-c and 2-d are the communication graphs obtained considering i) the mapping on the processors and ii) the bandwidth allocated to each communication. While a complete communication task would consider all possible communication end-time sequence scenarios (for instance, $c_{1\to2}$ can terminate before or after $c_{1\to3}$), we consider a reduced graph for which communication times are estimated. It is worth mentioning that a reduced graph improves the scalability of our approach by simplifying the design of the controller and by reducing the memory footprint. Following the method detailed in Section III-A, each state of the communication graph is associated to a configuration Q.

Figures 2-e and 2-f show the execution traces of the HP and LP solutions. We can clearly see that the HP solution reduces the execution time of the DAG by using more bandwidth for each communication compared to the LP solution.

B. Architecture of the sequencer

The proposed approach requires two hardware blocks, as depicted in Figure 3: an ONI manager, located in each ONI, in charge of activating the Lasers and MR, and an ONoC configuration sequencer, that is shared among all the ONIs, in charge of sequencing the configurations of the ONIs.

An Optical Network Interface, ONI (illustrated on the right part of Figure 3), integrates a receiver Rx and a transmitter Tx.

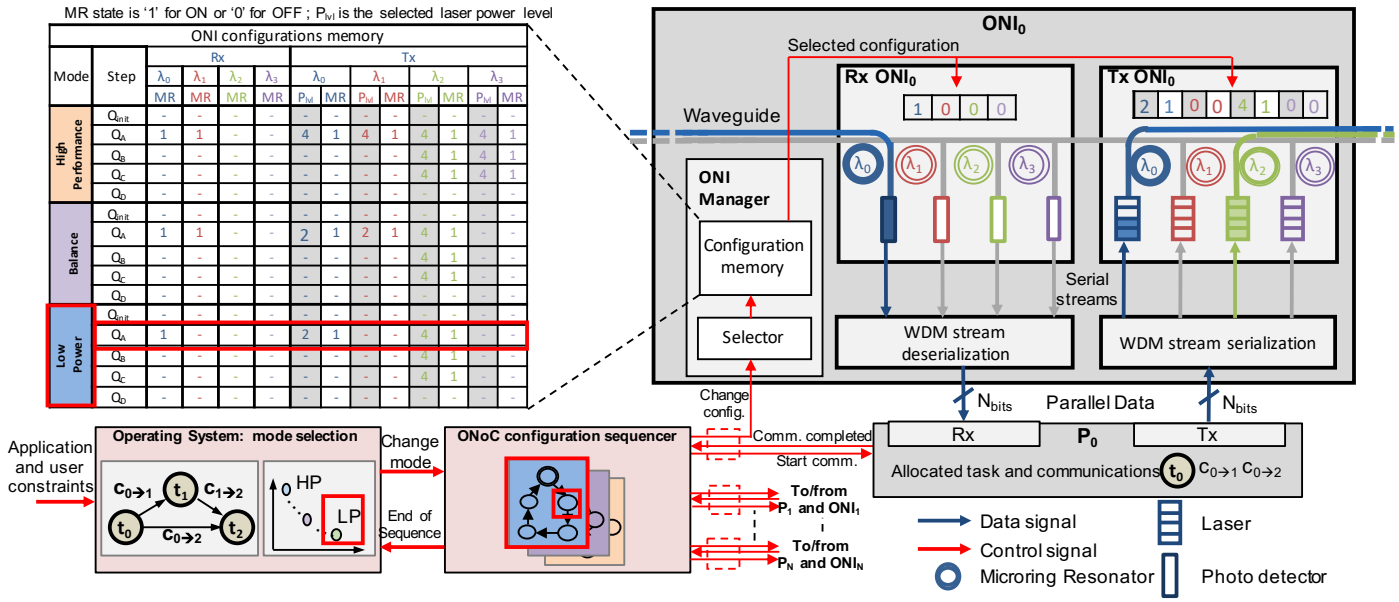


Fig. 3: Proposed Optical Network Interface architecture.

The transmitter is composed of on-chip laser sources that can emit and inject optical signal at a specific wavelength into the waveguide. The power generated by the Lasers can be configured. In this example the Lasers can be configured independently among 4 levels of power. The data are directly transmitted from these lasers through current modulation (OOK) and each laser source can also be turned OFF for energy saving. As the lasers are sending data in serial and as a core sends data in parallel on N_{bits}, a WDM stream serialization is required in order to perform the serialization and allocation of the serial data to the selected wavelengths (i.e. lasers). The receiver part includes wavelength-specific MR that can be turned ON or OFF to configure respectively drop or pass-through operations of the signals at a given wavelength. Signal dropped from a waveguide reaches a photodetector, where opto-electric data conversion generates an electrical signal suitable for the electronics part of the receiver. The considered architecture allows the reuse of wavelengths to realize multiple independent communications in a single waveguide. The receiver includes a WDM stream deserialization to perform serial to parallel conversion on the received data. Each ONI includes an ONI Manager in charge of configuring the receiver and the transmitter. It includes an ONI configurations memory which stores the configurations to apply on the MR and Lasers for different operating modes and at different steps composing a sequence of configurations.

As mentioned above, the second hardware block required is the ONoC configuration sequencer in charge of synchronizing all the ONI Managers of the ONoC. It includes a state machine explained in Section III-B as a sequence diagram. This block takes as input decision from the Operating System in order to select the configuration mode satisfying application requirements. Moreover, it is connected to the ONIs through three signals. The first is a one-bit signal indicating the ONI to start exploiting the configuration associated to the current Step and Mode selected in the ONI configuration memory. The second is also a one-bit signal generated by the ONIs indicating that a step is over. The last signal bit-width depends on the number of Modes embedded in the ONI configuration memory and allows to select the mode to apply.

C. Run-Time management of energy-performance trade-off

Figure 4 illustrates the configuration of the ONIs for each step of the configuration sequence in the Low Power mode. In this mode, p_0 sends a control signal to the configuration manager when t_0 execution ends and when data involved in $c_{0\to1}$ and $c_{0\to2}$ are ready for transfer. The manager switches from configuration Q_{init} to Q_A: Tx_0 and Rx_1 are reconfigured with the values stored in the configuration memory. Two on-chip lasers are powered on to emit optical signals for purpose of reducing energy consumption (the laser output powers are not illustrated for the sake of clarity) as shown in Figure 4-a. For low-latency purpose, the ONoC is reconfigured in parallel, which is implemented using dedicated electrical wires connecting the controller to the optical devices. Once the new configuration is loaded, the data transfer starts and, once it ends, another control signal is sent to the configuration manager to disable the communication channels. For this purpose, a new configuration (Q_B as shown in Figure 4-b) is loaded i.e. the lasers and the MR involved (blue one) in the communication are turned OFF. Figure 4-c) illustrates communication c1 $\to$ 2 carried out by the controller. The reconfiguration of the interfaces to implement a new communication has no impact on the ongoing communications: in this example, green signal on p_0 to p_2 waveguide segment is used for $c_{0\to2}$ while $c_{1\to2}$ is carried out using blue wavelength (reconfiguration $Q_B \to Q_C$). Then $c_{0\to2}$ ends and the green laser and MR are turned off as illustrated in Figure 4-d)

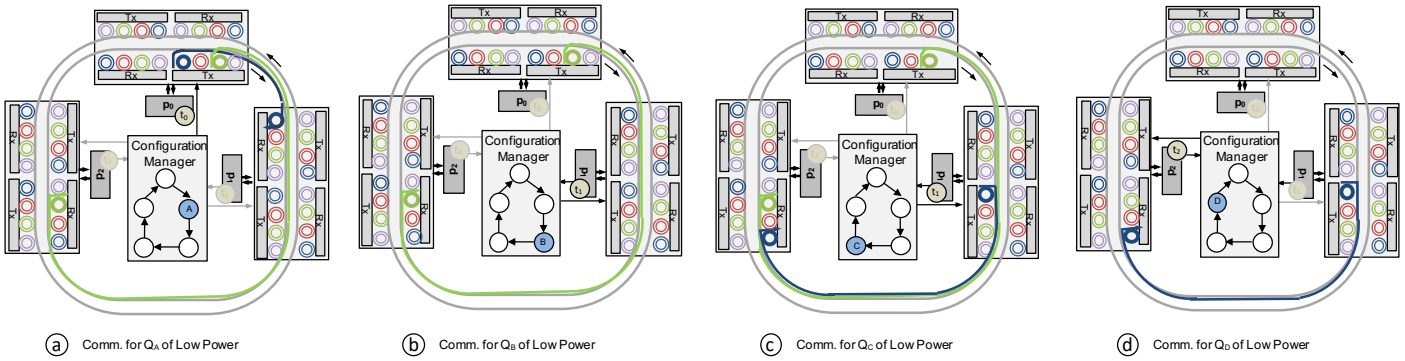


Fig. 4: illustration of the run-time management of energy-performance trade-off: a) evolution of the performance durig operation, b) ONoC ONI's configurations evolution in LP Mode.

(Q_D). Finally, the resources are released and become available for the next iteration.

III. ONoC CONFIGURATION SEQUENCER

A. Generation of configuration

We consider the use of the framework presented in [8] to generate the ONoC configuration sequences. This flow takes as inputs an application mapped onto the 3D architecture illustrated in Figure 1. The application is modeled as a task graph characterized by task execution times, amount of data transmitted between tasks, and minimum BER to be reached. The design flow relies on device parameters since they impact the performance of optical communications. Instance of parameters are photodetectors sensitivity, waveguide losses and MR model. Regarding the laser, we take the data-rate, the efficiency, the maximum output power, and the number of power levels available into account. The aim of the flow is to optimize both power consumption and application execution time. For this purpose, our framework explores both device-level and system-level parameters. Based on a set of device and system input parameters, a multi-objective optimization is carried out using a genetic algorithm due to the two contradictory objectives. In the genetic algorithm, the ONoC configuration modes are represented by chromosomes and the genes encode both wavelength allocations and laser power levels. Finally the resulting ONoC configuration modes on a Pareto front from the multi-objectives optimization are thus reported, including low-power solutions, which tends to minimize the number of used wavelengths, and high-performance solutions, for which multiple wavelengths are allocated to shorten the communication time. It has to be noticed that the framework also optimized the selection of the wavelengths in order to reduce the crosstalk. Then the obtained trade-off configuration modes could be embedded in the system and be loaded on run-time according to the execution context (e.g. high performance and low power) which is out of the scope of this paper. Refer to [8] for more details on the generation of solutions.

Figure 5 illustrates individual coding and the corresponding ONoC configuration and how we generate the communication configuration that are embedded in each ONI.

As depicted in Figure 5-a, we assume three interfaces, four wavelengths (i.e. four lasers per interface) and four laser output power levels. Tasks t_0, t_1 and t_2 are mapped on processor p_0, p_1, and p_2, respectively, which leads to optical communications between p_0 and p_1 ($c_{0\to1}$), p_0 and p_2 ($c_{0\to2}$), p_1 and p_2 ($c_{1\to2}$). The chromosome is divided into as many parts as there are communications (three in the example) [8]. Figure 5-a shows one of the solution given by the framework. It is represented as a chromosome given the bandwidth and power allocation for each communication. The first gene of each chromosome part gives the selected laser output power level; it is an integer value corresponding to the laser configuration. The following genes correspond to the wavelengths utilization: value 0 or 1 indicates that a wavelength is used or unused, respectively. In the example, both $c_{0\to1}$ and $c_{1\to2}$ use λ_0 and λ_1 while $c_{0\to2}$ is implemented using only λ_2. Obviously, the more wavelengths are allocated for a given communication, the higher the bandwidth.

The configuration of each interface is obtained as follows, and is illustrated in Figure 5-b.

MR state and Laser power levels are extracted from the chromosome for each Tx and Rx of the ONIs. As we can see in this example, the Tx of p_0 will handle two communications that use different wavelengths and power levels.

Then, the configurations are merged and the results configuration of each Tx and Rx are obtained as shown in 5-c.

Finally the configuration can be applied to the ONIs as illustrated in 5-d. First, optical channels are open by switching ON the MRs involved in communications (i.e. MRs localized in the transmitter Tx of the source processor and the receiver Rx of the destination). Then, the power of the optical signals propagating through the channels is defined according to the selected lasers power level. In the example, at the transmitter of p_0 interface, three MRs are turned ON to implement communications $c_{0\to1}$ (λ_0 and λ_1) and $c_{0\to2}$ (λ_2). On the receiver side, the MRs corresponding to λ_0 and λ_1 are set to the ON state in p_1 interface, while the MR corresponding to λ_1 remains OFF to let the signal reaching p_2 interface were it will be dropped. In the chromosome part dedicated to $c_{0\to1}$, the laser output power level is set to 2: in p_0 interface, the lasers emitting at

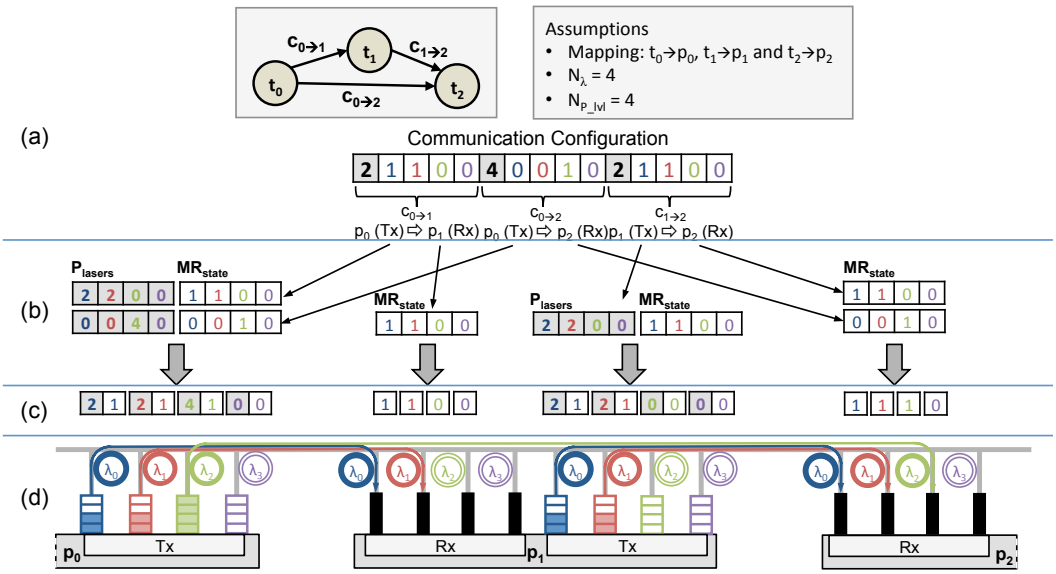


Fig. 5: Generation of the configurations: a) DAG description, architectural assumptions and the results of the offline optimization framework, b) Extraction of each ONI configuration, c) Result of the ONI configurations, d) Configuration of each ONI with respect to the extraction configurations.

wavelength λ_0 and λ_1 are set to 50% of the maximum power. In the same interface, laser at λ_2 is set to 100% to match value 4 in corresponding gene for $c_{0\to2}$.

B. ONoC sequencing illustration

Figure 6 illustrates the relation between the Operating System (OS), the ONoC configuration Sequence, and the ONIs.

Regarding the communication through the ONoC, the OS is only in charge of defining the mode (energy-performance trade-off) of the communication. It is possible to change the mode at the beginning of a new sequence. Regarding the sequencing of steps within a sequence, the ONoC configuration sequencer sends the order to apply the ONI configuration for Tx Rx by using a one-bit signal to the ONI manager. The ONI Manager reads the associated memory containing the Tx and Rx configurations, hence the communication can start.

When the communications are over for a processor, or when no communication are necessary in a step, the ONI manager associated to a core sends back a one-bit signal to the ONoC configuration sequencer. When all the communications are completed, the ONoC configuration sequencer indicates to all the ONI to go through the next step. Hence, the aforementioned sequencing is repeated.

IV. Result

In this section we show the energy versus performance trade-offs available in an ONoC. We consider the architecture in Fig. 1 in which each ONI is connected to a cluster of four electrically connected cores. We assume two waveguides, eight wavelengths per waveguide, five configurable levels of electrical power for the Lasers: [2, 4, 6, 8, 10] mW. The Table I summarizes the technological parameters used to generate the ONoC configurations with the framework proposed in [8]. Regarding the applications, we use a random task graph generator that provides applications including from 52 to 107 tasks and from 80 to 158 communications. The task execution time values are randomly selected between [100, 1000] cc and the communication volumes are randomly selected within [100, 1000] bytes range. The targeted BER is 10^{-9} and each task is randomly mapped on a dedicated core. As we assume shared memory within a same cluster, no latency is assumed for intra-cluster communications.

Table II summarize the characteristics of the task graphs and the total energy and execution time in clock cycles (cc) required. Since the generation of the ONoC configuration leads to a Pareto front with several solutions, we only show the solutions with i) the lowest energy consumption (denoted Low Power in the table) and i) the lowest execution time (High Perf.). As shown in Table II, the solutions offer, on average, 44% energy variation and 71% execution time variation trade-offs. From these solutions, the designers can implement in the configuration memory of the ONI manager several solutions that provide intermediate trade-offs. These results demonstrate the efficiency of flexible wavelengths allocation and laser output power tuning to adapt the nanophotonic interconnects according to application requirements.

Parameter	Value	Ref
Waveguide propagation loss	-0.274 dB/cm	[9]
Photodetector sensitivity	-20 dBm	[10]
Laser efficiency	15%	[10]
$\Delta\lambda$	0.4 nm	[11]
FSR	8 nm	[11]
$-3dB$ MR bandwidth	0.26 nm	[11]

TABLE I: Technological parameters.

V. Conclusion

In this paper we presented a method to offer dynamicity to ONoC in terms of energy versus performance trade-offs.

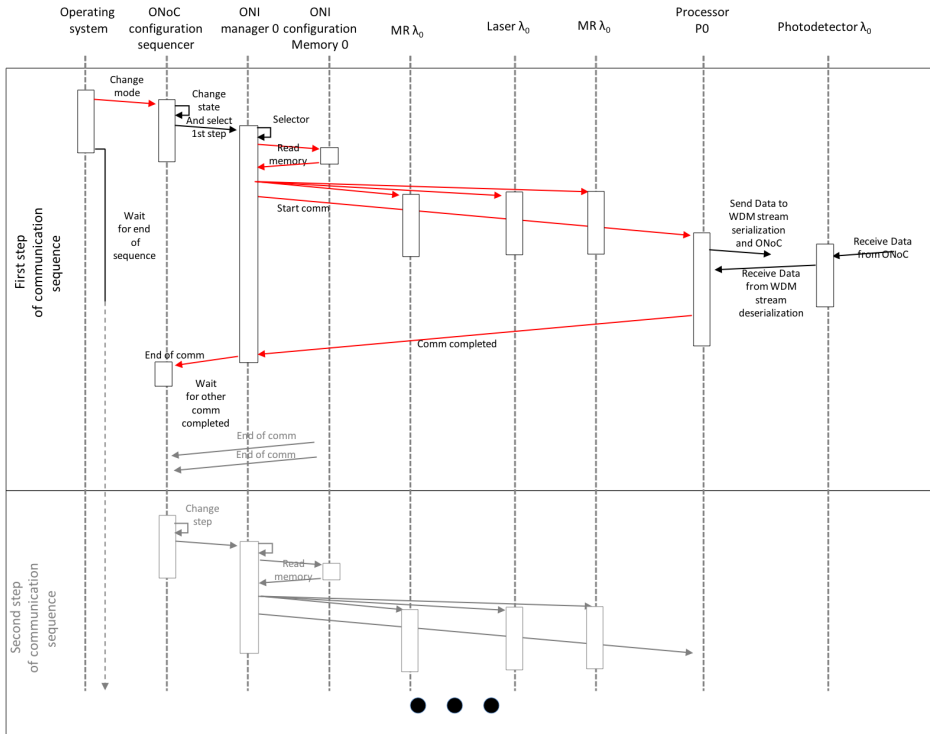


Fig. 6: Sequence diagram detailing the interactions between the Operating Software, the ONoC configuration Sequencer, and ONIs.

Graph ID	Number of Tasks/Comms	Energy (nj)			Execution Time (kcc)		
		Low Power	High Perf.	Variation	Low Power	High Perf.	Variation
TG 1	55/80	141	187	1.33	41.2	23.8	1.73
TG 2	52/78	118	194	1.64	37.9	23.6	1.62
TG 3	57/82	102	120	1.18	39.9	24.6	1.84
TG 4	60/92	141	195	1.38	37.3	22.8	1.62
TG 5	63/93	128	193	1.51	41.2	24.6	1.89
TG 6	62/92	147	236	1.61	43.2	24.4	1.75
TG 7	56/87	103	148	1.44	31.7	19.9	1.57
TG 8	63/91	112	156	1.39	37.5	23.3	1.60
Average		**124**	**179**	**1.44**	**38.7**	**22.7**	**1.71**

TABLE II: Energy-performance trade-off variation possibilities.

The hardware blocks in charge of the run-time management of the ONoC have been presented and the method generates the allocation solution through an offline optimizing framework. The ONoC dynamicity is exploited by the Operating System regarding the application requirements. An off line analyze of the application mapped on the architecture allows to extract all the possible configurations to support the communications Based on this analyze, a sequence of communication configurations is then controlled step by step by the Operating System to ensure the current requirements.

ACKNOWLEDGEMENT

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. J. Luo is supported by China Scholarship Council (CSC).

REFERENCES

[1] J. D. Owens, W. J. Dally, R. Ho, D. N. Jayasimha, S. W. Keckler, and L. S. Peh, "Research challenges for on-chip interconnection networks," *IEEE Micro*, vol. 27, no. 5, pp. 96–108, Sept 2007.

[2] A. Cilardo and E. Fusella, "Design automation for application-specific on-chip interconnects: A survey," *Integration, the VLSI Journal*, vol. 52, pp. 102–121, 2016.

[3] C. Batten, A. Joshi, V. Stojanovic, and K. Asanovic, "Designing chip-level nanophotonic interconnection networks," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 2, pp. 137–153, June 2012.

[4] A. K. Dutta *et al.*, *WDM technologies: optical networks*. Academic Press, 2004.

[5] L. H. K. Duong *et al.*, "Coherent and incoherent crosstalk noise analyses in interchip/intrachip optical interconnection networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 2475–2487, July 2016.

[6] E. Fusella and A. Cilardo, "Lighting up on-chip communications with photonics: Design tradeoffs for optical noc architectures," *IEEE Circuits and Systems Magazine*, vol. 16, no. 3, pp. 4–14, thirdquarter 2016.

[7] X. Wu, J. Xu, Y. Ye, Z. Wang, M. Nikdast, and X. Wang, "Suor: Sectioned undirectional optical ring for chip multiprocessor," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 10, no. 4, p. 29, 2014.

[8] J. Luo, C. Killian, S. Le Beux, D. Chillet, O. Sentieys, and I. O'Connor, "Offline optimization of wavelength allocation and laser power in nanophotonic interconnects," *J. Emerg. Technol. Comput. Syst.*, vol. 14, no. 2, pp. 24:1–24:19, Jul. 2018.

[9] 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 silicon waveguides for application of optical interconnects," in *IEEE Photonics Society Summer Topicals 2010*, July 2010, pp. 191–192.

[10] M. Kennedy and A. K. Kodi, "Laser pooling: Static and dynamic laser power allocation for on-chip optical interconnects," *Journal of Lightwave Technology*, vol. 35, no. 15, pp. 3159–3167, Aug 2017.

[11] 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, June 2016.