A reconfigurable optical network on chip for streaming applications

Sébastien Le Beux, Hui Li, Gabriela Nicolescu, Ian O 'Connor

A Reconfigurable Optical Network on Chip for Streaming Applications

Sébastien Le Beux[1]*, Hui Li[1], Gabriela Nicolescu[2] and Ian O'Connor[1]

[1] Lyon Institute of Nanotechnology, INL-UMR5270
Ecole Centrale de Lyon, Ecully, F-69134, France

[2] Computer and Software Engineering Dept.
Ecole Polytechnique de Montréal, Canada

* Contact author: sebastien.le-beux@ec-lyon.fr

***Abstract*—Optical on-chip interconnects enable significantly increased bandwidth and decreased latency in MPSoC. They are characterized by high static energy and low dynamic energy consumption. Achieving high energy-efficient communication thus requires a certain adaptability degree to the executed application and its communication traffic in order to reduce the static energy. In this paper, we introduce a mapping method for CHAMELEON, an optical interconnect allowing configuring dedicated communication channels according to an application and its corresponding communication traffic.**

Keywords—Reconfigurable Optical Network on Chip, WDM.

I. INTRODUCTION

Optical Network-on-Chip (ONoC) is an emerging technology that is considered as one of the key solutions for the future generation of on-chip interconnects. It relies on optical waveguides to carry optical signals, so as to replace electrical interconnect and provide the low latency and high bandwidth properties of the optical interconnect. However, the design of an ONoC is a costly investment since it relies on the integration of heterogeneous technologies. Such investment thus requires a certain degree of reconfigurability of the design in order to ensure the ONoC can be adapted to various communication traffic scenarios, i.e. to be integrated in various hardware platforms. This reconfigurability degree should come with acceptable execution performances and power consumption overhead to keep benefits of the technology.

The approaches relying on MWSR (Multiple Writers Single Reader) such as Corona [3] implies that IP cores request an access to the optical interconnect. However, this requires an arbitration to avoid writing conflict, which can lead to extra latency, i.e. lower performances. Other approaches rely on SWMR (Single Writer Multiple Readers) such as ATAC [4], which implement a permanent broadcast among IP cores. While this avoids the use of arbitration, it is an energy-inefficient solution since all the data from an IP core needs to be transmitted to all the others. SWSR approaches (Single Writer Single Reader) aim at creating dedicated point-to-point communication channels between IP cores [1]. This approach is power efficient with low latency, due to the fact that the communications between IP cores are not influenced by each other (i.e. optical crossbar implementation). The main disadvantage of the SWSR approach is the lack of scalability. This scalability issue can be partially leveraged by using reduction method [1], but it comes at the price of limited connectivity which in turn impacts the reusability feature. Table 1 summarizes the main characteristics of MWSR, SWMR and SWSR approaches.

Table 1: Pros and Cons of ONoC communication schemes

Approach	Latency	Power efficiency	Reusability
MWSR	x	xx	xxx
SWMR	xxx	x	xxx
SWSR	xxx	xxx	x
CHAMELEON (reconfigurable SWSR)	xxx	xxx	xxx

In our previous works, we introduced CHAMELEON [2], a reconfigurable channel efficient optical network on chip. CHAMELEON extends the SWSR approach with a reconfigurability feature allowing opening and closing dedicated channels between IP cores. Hence, the bandwidth is shared according to the communication requirements, thus leading to a highly reusable network. The reconfiguration of CHAMELEON can be achieved at run-time, i.e. through a communication protocol, or specified at compile time, i.e. using static mapping method. The configurations generated by the later method lead to networks characterized by a low latency (no arbitration is required), highly energy-efficient (dedicated channels are used). In this paper we present a configuration of CHAMELEON implementing communications specific to the mapping of a streaming application.

II. CHAMELEON ARCHITECTURE

A. Architecture Overview

Figure 1 illustrates the considered 3D architecture. It is composed of an electrical layer implementing 4x4 IP cores and an optical layer implementing CHAMELEON. It is composed of ONIs (Optical Network Interface) crossed by waveguides propagating optical signals.

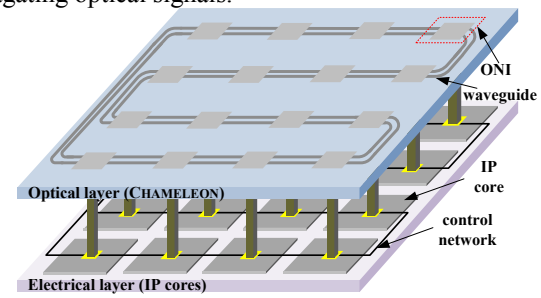

Figure 1: CHAMELEON is implemented on the optical layer, it interconnects IP cores located on the electrical layer

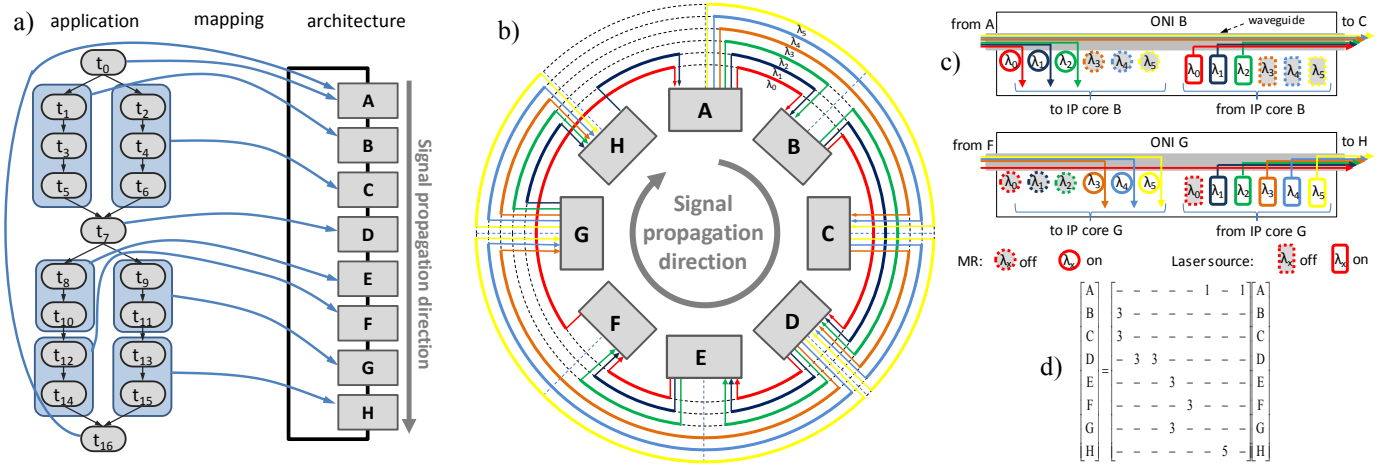

Figure 2: Configuring CHAMELEON according to communication requirements: a) task mapping on the architecture, b) configured communications channels, c) detailed configuration of ONI B and G and d) bandwidth matrix

The ONIs are composed of on-chip laser sources, Microring-Resonators (MR) and photodetectors. MRs and lasers are reconfigurable and allow opening dedicated communication channels between ONIs (i.e. connections that do not suffer from arbitration latency, similarly to SWSR architectures). In addition to **point-to-point** communication channels, it is possible to configure **multicast**, **broadcast** and **high-bandwidth** channels. If multiple waveguides propagate optical signals in both clockwise and counter-clockwise directions, **bidirectional** channels can be configured. Since CHAMELEON allows combining different channels within a single configuration at the same time, a high flexibility/reusability degree is reached.

B. Application Mapping

Reconfiguring CHAMELEON leads to setup of a given connectivity between IP cores that can well suit the targeted application to be executed (e.g. steaming) or the architectures communication requirements (e.g. processor to memory). The more wavelengths and waveguides there are, the higher the bandwidth is, i.e. the performance improvement wrt electrical interconnect. However, this comes at the cost of resource overhead and eventually performance penalty, meaning that design tradeoffs need to be explored. This exploration is driven by application cases, which is introduced in the following example mapping of a streaming application onto the targeted architecture. In the example, we consider 8 IP cores connected through 8 ONIs (A, B…H) with a single waveguide. We also assume that 6 wavelengths are available (λ_0, λ_1… λ_5).

Figure 2 a) represents a MP3 audio decoder application inspired from [5]. This simple application is chosen only to illustrate the approach; it is obvious that optic will be justified only for applications requiring huge data transfer. The application is represented as a Directed Acyclic Graph G = (T, E), where T is a set of tasks (i.e. t_i) and E a set of data (i.e. e_{ij} to be transferred from task i to task j). The right-hand side of Figure 2 a) represents the 8 ONIs crossed by the waveguide propagating optical signals. The arrows represent the mapping of the 17 tasks onto the IP cores connected to the ONIs, considering the constraint that source and sink are mapped on a same IP core. Such mapping solution can be obtained from methods similar to [6] and adapted to silicon photonic interconnect properties. Figure 2 b) represents the configured communication channels using a representation highlighting the use of the 6 wavelengths. Wavelengths are allocated according to the inter-ONI communications implied by the mapping (the intra-ONI communications such as $t_1 \rightarrow t_3$ are realized locally by the corresponding IP core). For instance, the data dependencies $t_0 \rightarrow t_1$ and $t_0 \rightarrow t_2$ in the application lead to the configuration of channels between A and B (with λ_0, λ_1 and λ_2), and between A and C (with λ_3, λ_4 and λ_5) respectively (i.e. a multicast is realized). In this example, 3 wavelengths are allocated to each channel, i.e. the bandwidth is equitably shared. Depending on the data dependencies between the tasks (as specified in the application model), the wavelengths can be shared differently. For instance, between G and H, 5 wavelengths and 1 wavelength are allocated to the channel implementing $t_{11} \rightarrow t_{13}$ and $t_{14} \rightarrow t_{16}$ separately. The resulting configuration leads to a contention-free execution of the application. Figure 2 c) represents the configuration of MRs and lasers in ONI B and G and Figure 2 d) gives the resulting connectivity/bandwidth matrix.

III. CONCLUSION

In this paper, we have presented a configuration of CHAMELEON defined according to the mapping of a streaming application. By avoiding contention and offering a high adaptability to the applications, we believe that CHAMELEON has the potential to be embedded in multiple platforms and reconfigured according to the user requirements.

REFERENCES

[1] I. O'Connor, et al. "Reduction Methods for Adapting Optical Network on Chip Topologies to Specific Routing Applications". In Proceedings of DCIS, 2008.

[2] S. Le Beux, et al. "Chameleon: Channel efficient Optical Network-on-Chip". In Proceedings of DATE, 2014

[3] D. Vantrease, et al. Corona: System Implications of Emerging Nanophotonic Technology. In Proceedings of the 35th Annual International Symposium on Computer Architecture (ISCA) pages 153–164, 2008.

[4] J. Psota, et al. ATAC: Improving Performance and Programmability With on-Chip Optical Networks. In Proceedings of IEEE International Symposium on Circuits and Systems, ISCAS, pages 3325–3328, 2010.

[5] Schmitz, M. T., Al-Hashimi, B. M., and Eles, P. *System-Level Design Techniques for Energy-Efficient Embedded Systems*. Kluwer Academic Publishers. 2004.

[6] C. Chou and R. Marculescu. Contention-aware Application Mapping for Network-on-Chip Communication Architectures. In Proc. *Intl. Conf. on Computer Design* (ICCD), Lake Tahoe, CA, Oct. 2008.