



**HAL**  
open science

## Dynamically Reconfigurable RF-NoC with Distance-Aware Routing Algorithm

Thomas Romera, Alexandre Brière, Julien Denoulet

► **To cite this version:**

Thomas Romera, Alexandre Brière, Julien Denoulet. Dynamically Reconfigurable RF-NoC with Distance-Aware Routing Algorithm. 14th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC 2019), Jul 2019, York, United Kingdom. 10.1109/ReCoSoC48741.2019.9034949 . hal-02176370

**HAL Id: hal-02176370**

**<https://hal.science/hal-02176370>**

Submitted on 8 Jul 2019

**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.

# Dynamically Reconfigurable RF-NoC with Distance-Aware Routing Algorithm

Thomas Romera  
Sorbonne Université, CNRS, LIP6  
Paris, France  
thomas.romera@lip6.fr

Alexandre Brière  
ESIEA, ESIEA N&S  
Paris, France  
alexandre.briere@esiea.fr

Julien Denoulet  
Sorbonne Université, CNRS, LIP6  
Paris, France  
julien.denoulet@lip6.fr

**Abstract**—With the growing number of cores per processor, on-chip communications are becoming a critical issue. Conventional wired mesh Networks-on-Chips (NoC) are reaching their scalability limit, paving the way for new interconnect technologies such as Radio Frequency (RF), Optics and 3D. This paper presents a new routing mechanism based on the WiNoCoD architecture, a hierarchical RF interconnect with dynamic bandwidth allocation for Chip Multi Processors (CMP). WiNoCoD consists of three levels of hierarchy with an RF NoC at the highest level, a wired mesh NoC and a crossbar at the lowest level. In this paper, we propose a new organisation for the manycore supporting the RF-NoC, and introduce a new routing algorithm that can choose between wired and RF networks in order to optimize the use of RF. We show that the new architecture improves performance in terms of latency and saturation. Platform simulations show a latency reduction for lower transaction rates of up to 10% compared to the original WiNoCoD architecture and up to 30% compared to a wired mesh NoC. Furthermore, smart routing pushes back the saturation threshold up to 20%.

**Index Terms**—Integrated circuit interconnections, Network on Chip, NoC, Radio Frequency, RF, Dynamically Reconfigurable, Multi-Core, Many-Core

## I. INTRODUCTION

Conventional Network-on-Chips (NoCs) are interconnects using routers linked together to form a grid and used to connect processor nodes inside a chip. They have been developed to resolve communication issues that classical communication mediums, such as bus and crossbar, present in the context of Chip Multi Processors (CMP). Indeed, bus interconnects have latency issues when too many initiators want access for communications. Furthermore, physical limitations such as wire length and clock frequency limit the bandwidth of the bus. Crossbars improve latency and bandwidth at the cost of a lot more chip surface dedicated to the interconnect. Moreover, the bandwidth grows with the size of the network, even if the maximum length of the wires used to connect the routers can have an impact on the topology of the network.

However, in conventional NoCs, latency limitations arise when the number of cores increases: the further two cores are apart, the more intermediate routers will have to be crossed by transmitted data to reach the destination. Additionally, wired mesh networks are not particularly suited for broadcast type transactions and dynamic reconfiguration.

Technological solutions such as 3D, Optics and Radio Frequency (RF) are being developed to alleviate those issues.

In [1], we introduced the WiNoCoD architecture (Wired RF-based Network On Chip Reconfigurable On Demand) as an RF solution using a hierarchical RF NoC with dynamic allocation of communication resources.

In this paper, we present a new routing mechanism for the WiNoCoD architecture. It is based on a new organisation of the manycore supporting the RF-NoC and introduces a routing mechanism that chooses between the RF and the wired networks, according to the distances between sources and destinations. The rest of the paper is organised as follows. Section 2 presents an overview of related works. Section 3 presents the RF NoC architecture. In order to evaluate our new solution, section 4 details the experimental setup used to evaluate the RF NoC, and section 5 presents and analyses the obtained results. Finally, section 6 concludes the paper.

## II. RELATED WORKS

Following ITRS recommendations [2], new technologies are being explored to solve intra-chip communication issues such as latency and bandwidth. 3D solutions try to reduce distances while optic and RF solutions try to reduce transit time. Furthermore, architectural solutions are explored to optimize the use of communication resources thanks to dynamic allocation.

### A. Technological Solutions

1) *3D NoC*: With 3D chips, stacked layers are used to increase the connectivity of the NoC routers. Vertical connections are added between layers and can be used to transmit data [3]. By switching from a large 2D chip to several smaller stacked layers, the maximum distance between two nodes in a network of size  $N$  is reduced from  $\sqrt{N}$  to  $\sqrt[3]{N}$ . This type of connection can also be used to connect specialised layers in heterogeneous System-on-Chip (SoC). However, larger 3D NoCs encounter the same problems as 2D NoCs in terms of latency and saturation. Mixed systems using 3D and optics [4] or 3D and RF [5] aim to solve those problems. Nevertheless, heat dissipation remains a problem in 3D architectures [6].

2) *Optical NoC*: Optical NoCs allow for much greater transmission speed than regular wired NoCs [7]. They are composed of emitters that convert electrical signals into optical ones; a physical optical guide where optical waves travel; receivers that translate optical waves into electrical signals [8].

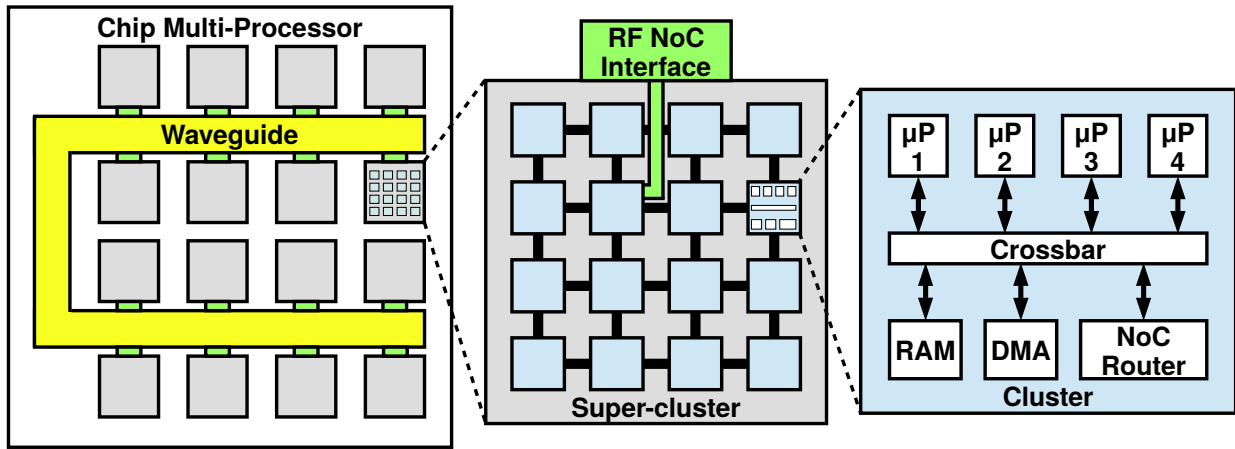


Fig. 1. Strict CMP hierarchical architecture

If these NoCs allow for greater bandwidth and lower latency, they can require different semi-conductor technologies than CMOS. Furthermore, they suffer from greater thermal sensitivity and typically require an external light source.

3) *RF NoC*: Similar to optical NoCs, RF NoCs enable faster communication to solve bandwidth and latency problems [9]. However, they have the benefit of being fully compatible with current CMOS technologies. Radio waves can be transmitted using a waveguide [10] or antennas [11]. Architectures using antennas are more flexible than those using waveguides but suffer from more sensitivity to interference and increased power consumption [12].

### B. Communication Resources Allocation

The above technological solutions aim to improve both latency and bandwidth. However, on-chip traffic is not necessarily constant and uniform in both space and time. Dealing with this spatio-temporal heterogeneity is one of the major limitations of classical wired NoCs. Adding priority to some messages does not remove the obligation to travel through intermediate routers for node-to-node communication. Moreover, reassigning unused bandwidth to parts of the chip that need it is a complex task on both 2D and 3D wired NoCs.

However, both optical and RF NoCs can establish direct communication links between any source and destination on the chip. Overlying such network on top of a classical NoC allows to create shortcuts thanks to the physical medium (waveguide, antennas, frequency bands, etc.) of this top-level network. Moreover, this extra bandwidth can be dynamically reallocated to handle heterogeneous traffic. Therefore, three types of allocation mechanisms have been proposed: Multiple Write Single Read (MWSR), Single Write Multiple Read (SWMR) and Multiple Write Multiple Read (MWMR).

Indeed, in optical and RF NoCs, the bandwidth is divided in channels that can be allocated according to the communication needs of the nodes. In [13], an MWSR communication scheme is presented where each receiver node has an associated

channel used to receive data from other nodes. However, MWSR does not support broadcast and requires additional logic to avoid data collision from multiple sources writing to the same destination.

In [8], an SWMR scheme is implemented. Each channel is used by a single emitter node to communicate. Each receiver node has to decode every channel to identify the data intended for it. SWMR supports broadcasting but requires more power as a node writes on all available channels.

The MWMR scheme allows for multiple channels to be assigned to a single node for transmitting while simultaneously allowing for each node to receive data from different channels and nodes [14]. This scheme requires arbitration on both emission and reception and consumes the most energy. However, it is the most flexible scheme, allowing for advanced reconfiguration mechanisms and supports broadcast.

### III. PROPOSED ARCHITECTURE

As said in the previous section, RF communication has a number of advantages, especially in terms of propagation time, which is about 0.5 ns for 100 mm in silicon. At the scale of a chip, it can be considered constant regardless of source and destination. Furthermore, RF waves transmission and reception times are also constant for every source and destination. This means an RF NoC is able to transmit data in constant time, unlike router-based NoCs where delay increases along with the distance between source and destination. Still, even if constant, this time is greater than the one required to cross a single router. It is then relevant to use the RF NoC only from a certain distance of transmission. Another benefit of using RF communication is the ability to dynamically allocate communication resources in order to maximise bandwidth occupation.

The WiNoCoD hierarchical architecture has therefore been set up, with the RF NoC as the highest communication level. In [15], we evaluated the physical characteristics of WiNoCoD for a CMP with 4 096 cores assuming a 22 nm target technology. The complete RF NoC contributes to 5.5% of the overall

power consumption which is 180 W, and 4.5% of the overall area which is 920 mm<sup>2</sup>. Those overall physical characteristics being comparable with those of the IBM POWER8 which are 190 W and 649 mm<sup>2</sup>.

In this section, we first describe the WiNoCoD architecture and then detail changes proposed to allow optimization of the use of communication resources. We then focus on communication management itself within the architecture, first presenting the RF channels allocation algorithm, then focusing on the new smart routing mechanism that aims to judiciously dispatch data between the wired and the RF networks.

#### A. CMP Architecture and RF-NoC Interface

The hierarchical architecture of the CMP used in WiNoCoD is composed of super-clusters which contain clusters, which themselves contain multiple cores and various peripherals as shown in Fig. 1.

The first level in the hierarchy is the cluster. A cluster contains  $P$  processors, each with an instruction and a data cache, along with a local RAM and a DMA. These components are connected with a crossbar that can access a grid-router, allowing inter-cluster communications on the second hierarchy level. The second level is the super-cluster. A cluster is connected to four of its neighbours to form a  $M \times M$  grid. One router in each super-cluster (typically the one at the center) is connected to the RF waveguide using a specialised RF interface. The third and final level is the RF NoC. It is used to connect the different super-clusters together via the waveguide using the RF interface. The waveguide is the structure that supports radio-waves transmission from emitter to receiver.

Within this scheme, two types of organisations have been defined and tested, the original strict version [15] and a new flexible one which we introduce in this work:

1) *Strict CMP Hierarchy*: In the original strict organisation, super-clusters can only communicate using the RF NoC through the waveguide, as shown in Fig. 1. This has the advantage of simplifying routing since there is only one path between two clusters in different super-clusters: the RF waveguide. However, this topology increases the latency between physically neighbouring clusters when they are located in different super-clusters: indeed, they have to pass across the CMP using RF, which has a higher transmission time than crossing a single router. Furthermore, this organisation increases the number of communications through the RF NoC, which may artificially create bottlenecks in the network and therefore increase global latency.

2) *Supple CMP Hierarchy*: We introduce a new organisation, pictured in Fig. 2. The supple hierarchy adds wired connections between adjacent clusters in different super-clusters. A 2D-mesh that spans the entire CMP is thus created and data can now travel either as if it was on a classical wired NoC mesh or by using the RF waveguide. However, the routers of the clusters now have to decide between routing data using the wired mesh or using the RF NoC. This is achieved by modifying the router and adding time and distance

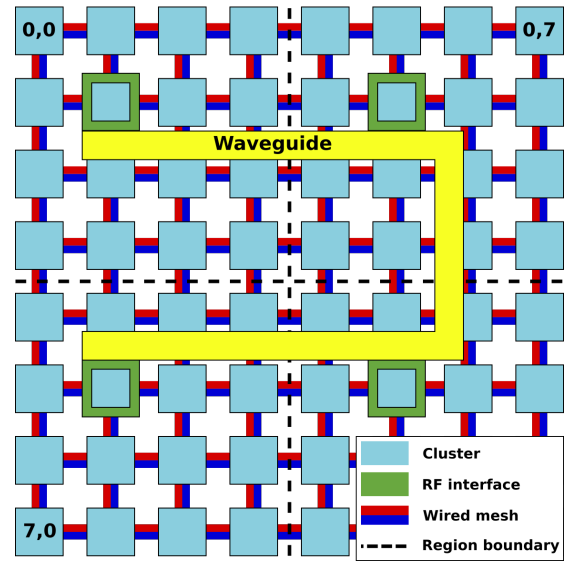


Fig. 2. Supple CMP hierarchical architecture

considerations between the emitter and the receiver on the mesh, at the cost of more hardware complexity.

3) *RF NoC Interface Architecture*: In both strict and supple organisations, each cluster features an RF interface that allows access to the waveguide (see Fig. 3). The interface translates data packets coming from the wired network into radio waves. The RF NoC uses Orthogonal Frequency Division Multiple Access (OFDMA) to encode data. This type of scheme splits the bandwidth into sub-channels that can be assigned and reassigned to the various super-clusters using a dynamic bandwidth allocation algorithm. It allows multiple super-clusters to simultaneously transmit on the different sub-channels of the bandwidth, depending on their needs. The RF interface converts the flow units (flits) into RF signals and its main components are:

- 1) An arbiter that queues data flits coming from the wired network and translates them into flits that will circulate on the RF network. It also send the filling state of its internal FIFO which is used in the dynamic allocation algorithm.
- 2) A coder handles data encoding into symbols using BPSK, QPSK or 16-QAM modulation schemes.
- 3) A demultiplexer performs the placement of data symbols corresponding to the allocated sub-channels of the RF bandwidth and the filling state symbols on the dedicated service sub-channel of the super-cluster.
- 4) An  $N$ -point IFFT produces the real and imaginary parts of the signal to be sent, with  $N$  the sub-channels total number.
- 5) An emitter that converts the IFFT outputs from digital domain to analog domain and then sends the signal on the waveguide.
- 6) The waveguide.

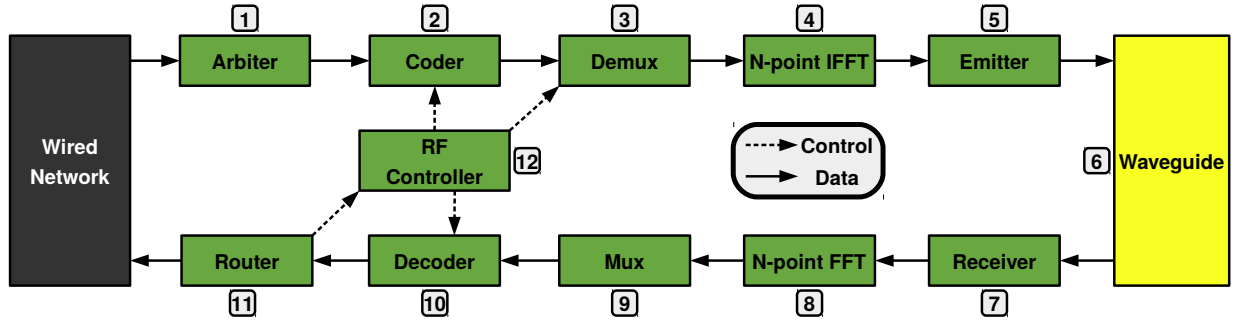


Fig. 3. Functional diagram of the RF interface

- 7) A receiver that collects the signal sent on the waveguide and then converts it back in the digital domain.
- 8) An  $N$ -point FFT that divides the signal into the  $N$  original sub-carriers.
- 9) A multiplexer that reconstructs the emitted data flits and the service flit by selecting symbols on sub-channels of the bandwidth.
- 10) A decoder that decodes data according to the chosen modulation schemes.
- 11) A router that transmits data flits to the wired network, only if the flit is intended for the given super-cluster, and the service flit to the RF controller.
- 12) An RF controller that implements the dynamic allocation algorithm. To do so, it receives the filling state information of all the clusters thanks to the dedicated service sub-channel of the RF bandwidth. It also indicates to the demultiplexer which sub-channels are available to send data and to the coder and the decoder which modulation scheme is currently used according to the allocation algorithm.

### B. Communications Resources Management

Routing in the wired network is performed using a *X-first* algorithm: routing is first done in the  $X$  dimension and once the correct column is reached, the packet is vertically routed in the  $Y$  dimension. This routing is deadlock-free since restrictions on data routing prevents circular dependencies in routers. In this subsection, we detail how the RF interface dynamically allocates RF channels to super-clusters. In order to limit the saturation of the RF network and optimize the use of RF, we then introduce the routing mechanism that chooses between the two networks.

1) *RF Dynamic Bandwidth Algorithm*: In order to optimize the bandwidth utilisation, a dynamic allocation algorithm was devised exploiting the OFDMA reconfiguration potential [15]. This algorithm is distributed in all the RF controllers to avoid latency and extra communications between RF interfaces. To do so, each super-cluster must periodically transmit its filling state using its dedicated part of the service sub-channel of the RF NoC. Thanks to the architecture's intrinsic broadcast support, sending this information to all the nodes takes the same effort as sending them to a single node.

The dynamic allocation algorithm is divided into four states (see Fig. 4):

- 1) At the start of the RF cycle  $t$ , each super-cluster  $i$  sends its filling state  $B_i^t$  on the RF-NoC. The RF controllers then collect all the emitted filling states.
- 2) The controllers compute the sum  $S^t$  of the filling states  $B_i^t$  of all the  $N$  super-clusters connected to the waveguide:  $S^t = \sum_{j=1}^N B_j^t$ .
- 3) The controllers then compute  $A_i^{t+1}$ , the number of sub-channels to be allocated to node  $i$  at cycle  $t+1$ , with  $M$  the total number of sub-channels:  $A_i^{t+1} = M \times \frac{B_i^t}{S^t}$ .
- 4) The allocated sub-channels are then registered in the demultiplexer. They can now be used by the super-cluster to transmit data through the RF NoC.

Data can be continuously transmitted over the RF network, even while the controller computes a new allocation configuration. When computation is finished, the new available sub-channels are simply provided to the demultiplexer.

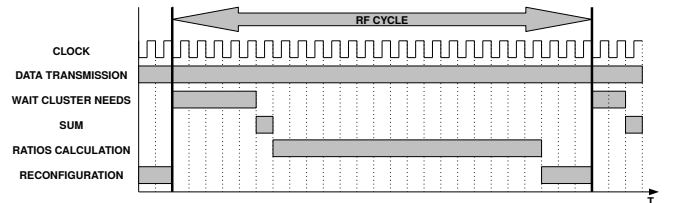


Fig. 4. RF controller allocation scheme timing diagram

2) *Smart Routing*: The new supply architecture requires a routing algorithm that chooses between the wired and the RF networks. In order to do so, we developed a routing algorithm which estimates the routing time between the current router and the destination. It takes  $A$  cycles to cross a single router and  $B$  cycles to cross the RF NoC. So, to go from a local node  $(x_l, y_l)$  to a destination node  $(x_d, y_d)$  using only the wired NoC, you need the following number of cycles  $t_{NoC}$ :

$$t_{NoC} = A \times (|x_d - x_l| + |y_d - y_l|) \quad (1)$$

To travel using the RF NoC, the algorithm first routes the flit from the local cluster  $(x_l, y_l)$  to the local RF interface cluster

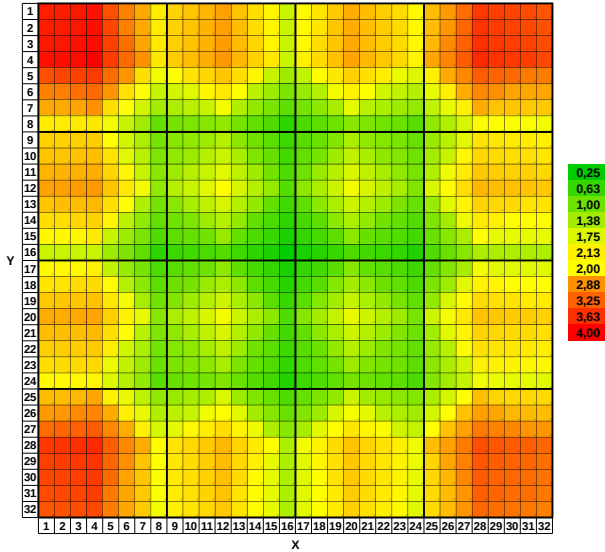


Fig. 5. Ratio of use between RF and wires for each cluster.

$(x_{l_{RF}}, y_{l_{RF}})$  using the wired network. The flit then travels through the waveguide to the destination super-cluster. It is received by the destination RF interface cluster  $(x_{d_{RF}}, y_{d_{RF}})$ . Finally, it is routed to the destination cluster  $(x_d, y_d)$  using the wired network. We therefore need to add these three transmission times to obtain the total number of cycles  $t_{RF}$  to pass through the RF:

$$\begin{aligned}
 t_{RF} = & A \times (|x_{l_{RF}} - x_l| + |y_{l_{RF}} - y_l|) \\
 & + B \\
 & + A \times (|x_d - x_{d_{RF}}| + |y_d - y_{d_{RF}}|)
 \end{aligned} \quad (2)$$

If the cycle number  $t_{NoC}$  is less than or equal to  $t_{RF}$ , the routing will be done over the wired network and otherwise over the RF network. Without smart routing and for homogeneous traffic, a packet is  $\times 15$  more likely to use RF over the wired network, thus probably creating saturation issues. Smart routing allows to level out this ratio, as can be seen on Fig. 5, which shows the ratio between packets using RF and packets using only the wired network. As found by simulation, the maximum ratio is now 4. Clusters on the edges of the chip, or at the center of super-clusters are more likely to require long distance transmission. Therefore, smart routing grants them a better access to the RF NoC.

#### IV. EXPERIMENTAL METHODOLOGY

In order to simulate the WiNoCoD architecture, test platforms were modelled in SystemC CABA (Cycle Accurate - Bit Accurate) using the SoCLib library of components [16]. We aimed to evaluate the new supply organisation and measure the impact of the communication resources optimization features: dynamic RF bandwidth allocation and smart routing between RF and wired networks. To do so, the models must provide the RF bandwidth distribution among the super-clusters, and also represent data travelling on these RF channels. In the RF-NoC interface architecture, on the emitter side, this information is available at the output of the demultiplexer stage (or at the input of the multiplexer stage on the receiver side). As a

result, for simulation speed purposes, we chose for our models to abstract the RF front-end and the waveguide, therefore, in our platforms, an "ideal digital waveguide" connects the RF interfaces demultiplexers with the multiplexers. More information about the RF front end and its performances can be found in [17].

Another simplification was the substitution of CPU cores inside the cluster with traffic generators. Four kinds of traffics were tested on the architectures:

- Homogeneous traffic (Fig. 6) in which a cluster sends a transaction to another randomly chosen cluster.
- Heterogeneous traffic (Fig. 7) in which all clusters send transactions to the same cluster thus creating a hot-spot.
- Broadcast traffic (Fig. 8) in which all the clusters can generate broadcast type transactions.
- Mixed traffic (Fig. 9) which aims to approximate the overall traffic of real applications. It is composed of 5% of broadcasts, 10% of heterogeneous transactions and 85% of homogeneous traffic. These average rates have been estimated by Jerger et al. [18] for broadcast and Chiu et al. [19] for hot-spots.

A version of WiNoCoD was written using the Distributed Scalable Programmable Interconnect Network (DSPIN) [20]. This router network is already described using SoCLib and thus provides the 2D mesh framework needed by WiNoCoD. The DSPIN routers were modified and a new communication port was added in order to route data toward the RF network. Note that it is also possible to interface WiNoCoD with different protocols in order to be linked with different sets of IP cores.

For the experimental parameters, we chose to simulate three different architectures, each featuring 1024 clusters:

- A 2D  $32 \times 32$  mesh NoC based on DSPIN routers,
- A strict WiNoCoD architecture of 16 super-clusters, each including  $8 \times 8$  clusters, with and without dynamic reconfiguration,
- A supple WiNoCoD architecture of  $32 \times 32$  clusters, with and without dynamic reconfiguration.

The number of cycles to cross a single DSPIN router was set to  $A = 3$  since their finite-state machines include 3 states. The number of cycles to cross the waveguide depends on the bandwidth characteristics and the CPU core frequency. We assume a bandwidth of 20 GHz divided into 512 sub-carriers and cores running at 1 GHz which induces  $B = 25$  cycles to cross the RF NoC [17].

#### V. RESULTS ANALYSIS

At this point, we need to carefully distinguish "RF-cycles", corresponding to the OFDMA period of 25 ns and used to represent the injection rate, from "cycles" corresponding to the 1 GHz clock frequency of the cores and used to evaluate the latency of communications. Fig. 6 shows that the wired mesh network has a constant latency of 139 cycles for injection rate up to 1 transaction per cluster per RF-cycle. The mesh saturation is not shown on the graph, the saturation occurring for injection rates around 4 transaction/cluster/RF-cycle. The



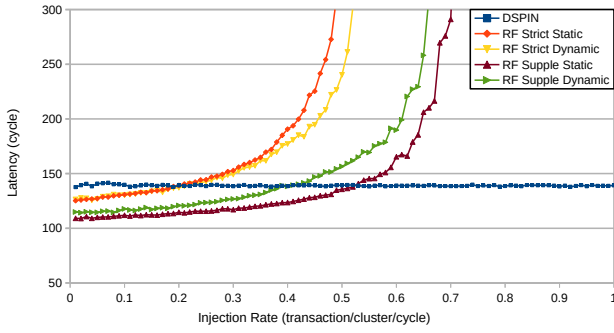


Fig. 6. Communications latency for homogeneous traffic

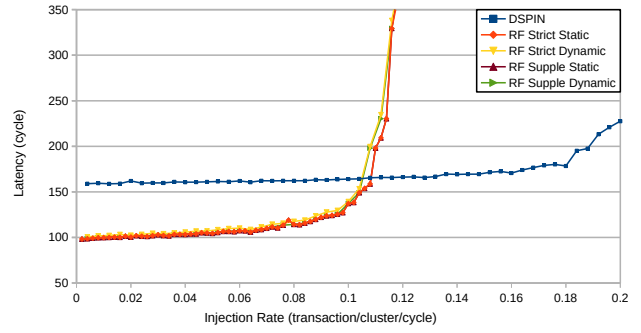


Fig. 8. Communications latency for broadcast traffic

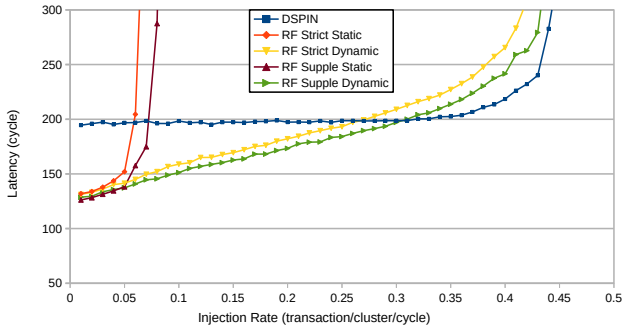


Fig. 7. Communications latency for heterogeneous traffic

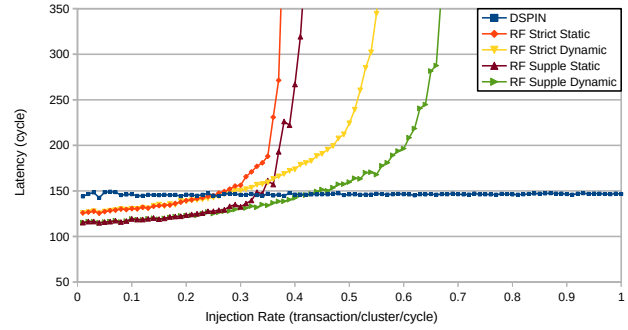


Fig. 9. Communications latency for mixed traffic

strict WiNoCoD architecture starts out with a latency 10% lower at 127 cycles for lower injection rates. It saturates for injection rates of about 0.5 transaction/cluster/RF-cycle. The supple WiNoCoD architecture has a latency 18% lower than that of the wired mesh at 114 cycles also for lower injection rates.

The following results highlight the two main characteristics of WiNoCoD: the RF NoC and the dynamic reconfiguration. With the RF NoC, latency improvements are achieved for all traffic types at low injection rates. In particular, Fig. 8 shows that thanks to OFDMA, the latency of broadcast transactions is decreased by 36%, from an average of 159 cycles on the DSPIN architecture to 101 cycles for the WiNoCoD architectures. For heterogeneous traffic and low injection rates, Fig. 7 shows a 30% improvement in latency between the DSPIN architecture and both WiNoCoD architectures. For homogeneous and mixed traffic, latency for low injection rates is reduced by 9% and 13% respectively as shown in Fig. 6 and Fig. 9 respectively.

Additionally, the new smart routing scheme is able to further reduce latency by dispatching data in a more optimized way between the wired and the RF networks. For homogeneous traffic and low injection rates, Fig. 6 shows that latency decreases by 10% between the original strict architecture and the new supple one, from 127 cycles to 114 cycles, for homogeneous traffic. For mixed traffic and low injection rates,

Fig. 9 shows a 20% reduction in latency between the strict architecture and the supple one.

Furthermore, the new routing scheme pushes back the saturation threshold. Fig. 6 shows that the saturation point is pushed back by 30%, from an injection rate of 0.5 transaction/cluster/cycle to 0.65. The same effect can be seen in Fig. 9, the saturation point goes from 0.5 for the strict architecture to 0.6 for the supple. Overall, comparing the strict and supple architectures, both using dynamic bandwidth allocation, we see that the saturation threshold is pushed back by 20%.

As for smart routing, the RF dynamic allocation algorithm benefits vary with the type of generated traffic. For example, in Fig. 8, comparisons between static and dynamic versions of the architecture show that RF bandwidth reconfiguration does not improve performances for broadcast. However, for others traffics such as heterogeneous traffic, Fig. 7 shows that dynamic reconfiguration allows to greatly improve the saturation issue, as it goes from an injection rate of 0.05 for static configurations to 0.4 for the dynamic ones.

In the end, for mixed traffic, Fig. 9 shows that dynamic reconfiguration allows to push back the saturation threshold. With the strict architecture, it is improved by up to 43% compared to a fixed allocation of the RF channels. The saturation point goes from 0.35 to 0.5. With the supple architecture, the saturation point is improved by up to 50%, from 0.4 with a static configuration to 0.6 with a dynamic allocation.

Finally, comparing WiNoCoD with the DSPIN architecture,

Fig. 9 shows that on one hand, the supple WiNoCoD architecture outperforms DSPIN for injection rates up to 0.45. On the other hand, the saturation threshold for DSPIN occurs at a higher rate than WiNoCoD. However, the conducted experiments only take into account a limited spatial and temporal communication heterogeneity, since all clusters follow the same behavior throughout simulation duration. This would not necessarily be the case if we were to run real software applications. So, it would be interesting to run such experiments to evaluate up to which point WiNoCoD's lower saturation threshold is penalising.

## VI. CONCLUSION

This paper proposed a new routing mechanism for the WiNoCoD architecture and studied its impact in terms of latency and saturation. A new, more flexible architecture with smarter transaction routing has been defined. In order to compare our new architecture with the original WiNoCoD architecture and with a more conventional wired NoC, SystemC models of platforms have been developed. We studied a CMP with 1024 clusters in which the RF NoC represents 5.5% of the overall power consumption and 4.5% of the overall area of the chip. Simulations were performed using synthetic traffic generators with several profiles: homogeneous, heterogeneous, broadcast and mixed. For homogeneous traffic, the new, more flexible architecture allows to improve latency by up to 10% compared to the original strict architecture and up to 30% compared to a wired 2D NoC architecture based on DSPIN routers. For mixed traffic, latency is decreased by 20% between the strict architecture and the new supple one. The supple WiNoCoD architecture also allows to push back the saturation threshold by 20% compared to the strict architecture. Future work aims at refining the routing algorithm by including other parameters such as RF and wired networks saturation consideration, in order to further optimize the use of the RF NoC. Work is also being carried out to test real software on the architecture to evaluate and compare more precisely our improvements.

## REFERENCES

- [1] F. Drillet, M. Hamieh, L. Zerioul, A. Brière, E. Unlu *et al.*, "Flexible Radio Interface for NoC RF-Interconnect," in *17th Euromicro Conference on Digital System Design*. IEEE, 2014, pp. 36–41.
- [2] ITRS, "International Technology Roadmap for Semiconductors, 2013 version," <http://www.itrs.net/Links/2013ITRS/Home2013.htm>, 2013.
- [3] K. Banerjee, S. J. Souri, P. Kapur, and K. C. Saraswat, "3-D ICs: A novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration," *Proceedings of the IEEE*, vol. 89, no. 5, pp. 602–633, 2001.
- [4] Y. Ye, L. Duan, J. Xu, J. Ouyang, M. K. Hung *et al.*, "3D optical Networks-on-Chip (NoC) for Multi-Processor Systems-on-Chip (MP-SoC)," in *2009 IEEE International Conference on 3D System Integration*. IEEE, 2009, pp. 1–6.
- [5] J. Lee, M. Zhu, K. Choi, J. H. Ahn, and R. Sharma, "3D Network-on-Chip with Wireless Links through Inductive Coupling," in *2011 International SoC Design Conference*. IEEE, 2011, pp. 353–356.
- [6] H. Esmailzadeh, E. Blem, R. S. Amant, K. Sankaralingam, and D. Burger, "Power Challenges May End the Multicore Era," *Communications of the ACM*, vol. 56, no. 2, pp. 93–102, 2013.
- [7] I. O'Connor and F. Gaffiot, "On-Chip Optical Interconnect for Low-Power," in *Ultra Low-Power Electronics and Design*. Springer, 2004, pp. 21–39.
- [8] G. Kurian, J. E. Miller, J. Psota, J. Eastep, J. Liu *et al.*, "ATAC: A 1000-Core Cache-Coherent Processor with On-Chip Optical Network," in *Proceedings of the 19th international conference on Parallel architectures and compilation techniques*, ser. PACT '10. ACM, 2010, pp. 477–488.
- [9] M. F. Chang, V. P. Roychowdhury, L. Zhang, H. Shin, and Y. Qian, "RF/Wireless Interconnect for Inter- and Intra-Chip Communications," *Proceedings of the IEEE*, vol. 89, no. 4, pp. 456–466, 2001.
- [10] M. F. Chang, J. Cong, A. Kaplan, M. Naik, G. Reinman *et al.*, "CMP Network-on-Chip Overlaid With Multi-Band RF-Interconnect," in *2008 IEEE 14th International Symposium on High Performance Computer Architecture*. IEEE, 2008, pp. 191–202.
- [11] S. Deb, K. Chang, M. Cosic, A. Ganguly, P. P. Pande *et al.*, "CMOS Compatible Many-Core NoC Architectures with Multi-Channel Millimeter-Wave Wireless Links," in *Proceedings of the great lakes symposium on VLSI*, ser. GLSVLSI '12. ACM, 2012, pp. 165–170.
- [12] A. Karkar, T. Mak, K.-F. 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.
- [13] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi *et al.*, "Corona: System Implications of Emerging Nanophotonic Technology," in *Proceedings of the 35th Annual International Symposium on Computer Architecture*, ser. ISCA '08. IEEE Computer Society, 2008, pp. 153–164.
- [14] C. Xiao, F. Chang, J. Cong, M. Gill, Z. Huang *et al.*, "Stream Arbitration: Towards Efficient Bandwidth Utilization for Emerging On-Chip Interconnects," *ACM Transactions on Architecture and Code Optimization (TACO)*, vol. 9, no. 4, p. 60, 2013.
- [15] A. Brière, J. Denoulet, A. Pinna, B. Granado, F. Pêcheux *et al.*, "A Dynamically Reconfigurable RF NoC for Many-Core," in *Proceedings of the 25th edition on Great Lakes Symposium on VLSI*, ser. GLSVLSI '15. ACM, 2015, pp. 139–144.
- [16] LIP6, "Soclib Documentation," <http://www.soclib.fr/trac/dev>, 2006.
- [17] M. Hamieh, M. Ariaudo, S. Quintanel, and Y. Louët, "Sizing of the Physical Layer of a RF Intra-Chip Communications," in *21st IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. IEEE, 2014, pp. 163–166.
- [18] N. E. Jerger, L.-S. Peh, and M. Lipasti, "Virtual Circuit Tree Multicasting: A Case for On-Chip Hardware Multicast Support," in *2008 International Symposium on Computer Architecture*. IEEE, 2008, pp. 229–240.
- [19] G.-M. Chiu, "The odd-even turn model for adaptive routing," *IEEE Transactions on parallel and distributed systems*, vol. 11, no. 7, pp. 729–738, 2000.
- [20] I. M. Panades, A. Greiner, and A. Shebanyrad, "A Low Cost Network-on-Chip with Guaranteed Service Well Suited to the GALs Approach," in *2006 1st International Conference on Nano-Networks and Workshops*. IEEE, 2006, pp. 1–5.