2008 14th IEEE International Conference on Parallel and Distributed Systems

A Performance Model of Communication in the Quarc NoC

M. Moadeli[1], W. Vanderbauwhede[1], A. Shahrabi[2]
1: Department of Computing Science
University of Glasgow
Glasgow, UK
Email: {mahmoudm, wim}@dcs.gla.ac.uk
2 : School of Engineering and Computing
Glasgow Caledonian University
Glasgow, UK
Email: a.shahrabi@gcal.ac.uk

Abstract

Networks On-Chip (NoC) emerged as a promising communication medium for future MPSoC development. To serve this purpose, the NoCs have to be able to efficiently exchange all types of traffic including the collective communications at a reasonable cost. The Quarc NoC is introduced as a NOC which is highly efficient in performing collective communication operations such as broadcast and multicast. This paper presents an introduction to the Quarc scheme and an analytical model to compute the average message latency in the architecture. To validate the model we compare the model latency prediction against the results obtained from discrete-event simulations.

1 Introduction

The Network-on-Chip (NoC) concept is an emerging communication-centric architecture for future complex System-on-chip (SoC) design providing scalable, energy efficient and reliable communication. In a NoC-based system, different components such as computation elements, memories and specialized IP blocks exchange data using a network as a communication infrastructure.

Designing a flexible on-chip communication network for a NoC platform, which can provide the desired bandwidth and at the same time be reused across many applications, is a challenging task which requires trading-off between a number of cross-cutting concerns such as performance, cost and size. In addition to the technology in which the hardware is implemented, the topology, switching method, routing algorithm and the traffic pattern are some other key factors which have direct impact on the performance of a NoC platform.

To meet these challenges, research carried out in the field has proposed the idea of using a packet switched communication network for on-chip communication. A packet switched NoC consists of an interconnection of many routers that connect IPs together to form a given topology in order to enable a large number of units (cores) to communicate with each other. The underlying topology of this architecture is the key element of on-chip network, since it provides a low latency communication mechanism and, when compared to traditional bus-based approaches, resolves physical limitations due to wire latency providing higher bandwidth and parallelism.

Deterministic routing and wormhole switching are regarded as the dominant routing and switching mechanism in the NoC domain [18]. Those options mainly originate from the resource constraints at intermediate routers [5, 18].

Most recent proposed NoC architectures have been founded on top of ring, fat-tree or 2D mesh topologies as they have an area efficient layout on a two dimensional surface which is most suitable for NoC design. Nostrum [14], Æthereal [6], and Xpipes [13] are some examples of architectures used for on-chip networks. The Spidergon NoC [12] is also one of the ring-based architectures proposed recently.

By adopting wormhole switching, deterministic routing and homogeneous, low-degree routers; the Spidergon scheme aimed to address the demand for a fixed and optimized network on-chip architecture to realize cost effective MPSoC development. However, the edge-asymmetric property of the Spidergon causes the number of messages that cross each physical link varies severely, resulting in an unbalanced traffic on network channels and, thus, leading to poor performance of the whole network. This situation

1521-9097/08 $25.00 © 2008 IEEE
DOI 10.1109/ICPADS.2008.54

is even exacerbated when the network is under bursty traffic as a result of some operations such as broadcast.

While preserving all features of the Spidergon, the Quarc scheme introduces an extra physical link to the cross link of the Spidergon to separate right-cross-quarter from left-cross-quarter to balance the traffic. It also employs an all-port router architecture to reduce the message blocking latency during collective communication operations. The Quarc NoC's features result in a NoC that is highly efficient in exchanging all types of traffic. In particular as paper shows the Quarc NoC is highly efficient for performing collective communication operations.

The rest of the paper is organized as follows. Section 2 introduces the Quarc NoC. It then investigates the architecture of the switches. Routing discipline, including unicast and broadcast, is also presented in this section. Section 3 presents the traffic analysis and the analytical model to compute average message latency in the Quarc NoC. The validation and analysis are presented in Section 4. Finally, we make concluding remarks in Section 5.

2 Quarc: A NoC Architecture

The topology of an on-chip network specifies the structure in which routers connect the IPs together. Fat tree, mesh, torus and variations of rings are among the topologies introduced or adopted for the NoC domain.

Typically, a particular topology is selected in order to trade-off between a number of cross-cutting measures such as performance and cost. A number of important characteristics that affect the decision on adopting a particular topology are network diameter, the highest degree of nodes in the network, regularity, scalability and synthesis cost for an architecture.

The topology of the Quarc NoC is quite similar to that of the Spidergon NoC. Therefore, the next section presents a brief description of the Spidergon NoC, followed by introduction of the Quarc NoC.

2.1 The Spidergon NoC

The Spidergon NoC [12] is a network architecture which has been recently proposed by STMicroelectronics [17]. The objective of the Spidergon topology has been to address the demand for a fixed and optimized topology to realize low cost multi-processor SoC implementation. In the Spidergon topology an even number of nodes are connected by unidirectional links to the neighboring nodes in clockwise and counter-clockwise directions plus a cross connection for each pair of nodes. Each physical link is shared by two virtual channels in order to avoid deadlock. Fig. 1 depicts a Spidergon topology of size 16 and its layout on a chip.

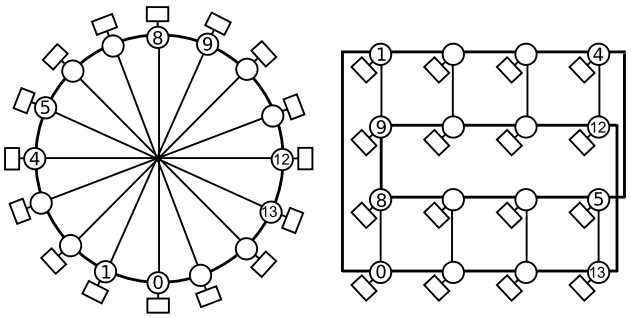

Figure 1. The Spidergon topology and the on chip layout.

The key characteristics of the this topology include good network diameter, low node degree, homogeneous building blocks (the same router to compose the entire network), vertex symmetry and simple routing scheme. Moreover, the Spidergon scheme employs packet-based wormhole routing which can provide low message latency at a low cost. Furthermore, the actual layout on-chip requires only a single crossing of metal layers.

In the Spidergon NoC, two links connecting a node to surrounding neighboring nodes carry messages destined for half of nodes in the network, while the node is connected to the rest of the network via the cross link. Therefore, the cross link can become a bottleneck. Also, since the router at each node of the Spidergon NoC is a typical one-port router, the messages may block on occupied injection channel, even when their required network channels are free. Moreover, performing broadcast communication in a Spidergon NoC of size N using the most efficient routing algorithm requires traversing $N - 1$ hops.

2.2 The Quarc

The Quarc NoC improves on the Spidergon scheme by making following changes: *(i)* adding an extra physical link to the cross link to separate right-cross-quarter from left-cross-quarter, *(ii)* enhancing the one-port router architecture to an all-port router architecture and *(iii)* enabling the routers to absorb-and-forward flits simultaneously. The Quarc preserves all features of the Spidergon including the wormhole switching and deterministic shortest path routing algorithm, as well as the efficient on-chip layout.

The resulting topology for an 8-node NoC is represented in Fig. 2.

Unlike the Spidergon NoC, in the Quarc architecture a messages will be blocked only when its requested network resources are occupied. This feature significantly enhances the performance of the network by reducing the waiting time at source node. Moreover, adding another physical

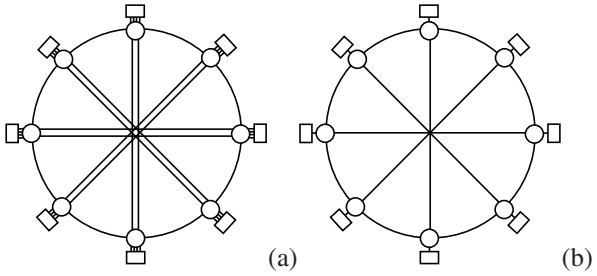

Figure 2. Quarc topology (a) vs Spidergon (b)

link to the cross network links improves access to the cross-network nodes. And last but not the least, the effect of the modification manifests itself most clearly when performing broadcast or multicast communication operations. In the Spidergon NoC, deadlock-free broadcast can only be achieved by consecutive unicast transmissions. The NoC switches must contain the logic to create the required packets on receipt of a broadcast-by-unicast packet. In contrast, the broadcast operation in the Quarc architecture is a true broadcast, leading to much simpler logic in the switch fabric; furthermore, the latency for broadcast traffic is dramatically reduced.

2.3 Routing algorithm

2.3.1 Unicast routing

For the Quarc, the surprising observation is that there is no routing required by the switch: flits are either destined for the local port or forwarded to a single possible destination. Consequently, the proposed NoC switch requires no routing logic. The route is completely determined by the port in which the packet is injected by the source. Of course, the NoC interface (transceiver) of the source processing element (PE) must make this decision and therefore calculate the quadrant as outlined above. However, in general the PE transceiver must already be NoC-aware as it needs to create the header flit and therefore look up the address of the destination PE. Calculating the quadrant is a very small additional action.

2.3.2 Broadcast operation

Collective communications operations have been traditionally adopted to simplify the programming of applications for parallel computers, facilitate the implementation of efficient communication schemes on various machines, and promote the potability of applications across different architectures [8]. These communication operations are particularly useful in applications which often require global data movement and global control in order to exchange data and synchronize the execution among nodes. The most widely used collective communication operations are *broadcast*, *multicast*, *scatter*, *gather* and *barrier synchronization*.

The support for collective communication may be implemented in *software* or/and *hardware*. The software-based approaches [7] rely on unicast-based message passing mechanisms to provide collective communication. They mostly aim to reduce the height of multicast tree and minimize the contention among multiple unicast messages.

In the tree-based scheme, the multicast problem is finding a Steiner tree with a minimal total length to cover all network nodes [3]. The tree operation introduces additional network resource dependencies which could lead to deadlock which is difficult to avoid if global information is not available. Hence, in wormhole-routed direct networks, the tree based multicast is usually undesirable, unless the messages are very short.

Broadcast and multicast traffic in Networks on Chip is an important research field that has not received much attention. A multicasting scheme for a circuit-switched network on chip proposed in [9]. Since the scheme relies on the global network state using global traffic information it is not easily scalable. Multicast operation is provided by Æthereal NoC [10]. However, Æthereal relies on a logical notion of global synchronicity which is not trivial to implement as the system scales. In [4] a multicast scheme in wormhole-switched NoCs is proposed. By this scheme, a multicast procedure consists of establishment, communication and release phase. A multicast group can request to reserve virtual channels during establishment and has priority on arbitration of link bandwidth.

Quarc Broadcast in the Quarc is elegant and efficient: The Quarc NoC adopts a BRCP (Base Routing Conformed Path) [1] approach to perform multicast/broadcast communications. BRCP is a type of path-based routing in which the collective communication operations follow the same route as unicasts do. Since the base routing algorithm in the Quarc NoC is deadlock-free, adopting BRCP technique ensures that the broadcast operation, regardless of the number of concurrent broadcast operations, is also deadlock-free.

To perform a broadcast communication the transceiver of the initiating node has to broadcast packet on each port of the all-port router. The transceiver tags the header flit of each of four packets destined to serve each branch as broadcast to distinguish it from other types of traffic. The transceiver also sets the destination address of each packet as the address of the last node that the flits stream may traverse according to the base routing. The receiving nodes simply check if the destination address at the header flit matches its local address. If so, the packet is received by the local node. Otherwise, if the header flit of the packet is tagged as broadcast, the flits of the packet at the same

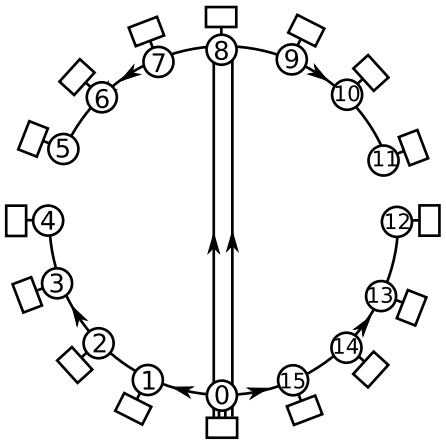

Figure 3. Broadcast in the Quarc NoC.

time are received by the local node and forwarded along the rim. This is simply achieved by setting a flag on the ingress multiplexer which causes it to clone the flits.

The broadcast in a Quarc NoC of size 16 is depicted in Fig. 3. Assuming that Node 0 initiates a broadcast, it tags the header flits of each stream as broadcast and sets the destination address of packets as 4, 5, 11 and 12 which are the address of the last node visited on left, cross-left, cross-right and right rims respectively. The intermediate nodes receive and forward the broadcast flit streams, while the destination node absorbs the stream.

3 The Analysis Method

The objective of this section is to introduce the model developed to evaluate the average message latency in the interconnection networks employing wormhole switching. We define the latency as the time from the generation of the message at source node until the last flit of the message is absorbed by the destination node. The model uses some assumptions that are widely used in the literature [16, 2].

- Nodes are generating the messages independently and according to a Poisson process.
- Destination addresses are selected randomly.
- The arrival at each channel is approximated to be a Poisson process.
- Messages are all the same size and larger than the network diameter.
- The adopted routing is a shortest path deterministic routing algorithm.
- The network employs wormhole switching.

We view our network as a network of queues, where each channel is modeled as an $M/G/1$ queue. For an $M/G/1$ queue the average waiting time is [11]

$$W_{M/G/1} = \frac{\lambda\rho}{2(1-\lambda x)}(1+\frac{\sigma^2}{x^2}) \quad (1)$$

$$\rho = \lambda x \quad (2)$$

where λ is the mean arrival rate, x is the mean service time and σ^2 is the variance of the service time distribution.

$$\sigma = (x - msg) \quad (3)$$

where msg is the message length.

The service time at ejection channel equals to message length, msg, and service time at each intermediate channel from source to destination may be calculated as

$$\overline{x}_i^{in} = \sum_j ((1-\frac{\lambda_i^{in}}{\lambda_j}P_{i\to j})W_j + \overline{x}_j + 1)P_{i\to j} \quad (4)$$

where W_j is approximated using Eq. 1, x_j is the mean service time at channel j, λ_i^{in} is traffic rate from channel i to channel j, λ_j is traffic rate at channel j and finally $P_{i\to j}$ is the probability of taking channel j after channel i. The detailed explanation of the analytical model is presented in Moadeli et al. [15].

The Eq. 4 can be adopted to compute the service time at all channel from the ejection channel at destination back to the injection channel at source node. Averaging message latency over all nodes in the network yield the average message latency in the network.

3.1 Traffic Analysis

To evaluate the service and waiting time at each channel the detailed traffic information is required. Traffic on each channel is comprised of several incoming streams and will be transmitted to a number of successive channels. In this section the traffic on each equivalence class has been provided. As the traffic distribution is slightly different depending on whether the number of nodes is a factor of four ($N = 4x$) or only a factor of two ($N = 4x + 2$) we have presented them separately when required.

We denote by λ the traffic rate at which a node sends to each individual destination. Depending the destination address a packet may take any of the four injection channels. The rate on injection links for traffic heading to right surrounding link, left surrounding link, cross-right and cross-left are denoted by $\lambda_{inj-right}$, $\lambda_{inj-left}$, $\lambda_{inj-crs-right}$ and $\lambda_{inj-crs-left}$ respectively and equal to

$$\lambda_{inj-left} = \lambda_{inj-right} = \left\lceil \frac{N}{4} \right\rceil \lambda \quad (5)$$

$$\lambda_{inj-crs-right} = \left\lfloor \frac{N}{4} \right\rfloor \lambda \quad (6)$$

$$\lambda_{inj-crs-left} = (\left\lfloor \frac{N}{4} \right\rfloor - 1)\lambda \quad (7)$$

The traffic on each surrounding link is comprised of the traffic from three sources, i) cross network link, ii) previous link and iii) injection link and equals

$$\lambda_{surr} = \begin{cases} \left\lceil \frac{N}{4} \right\rceil^2 \lambda & N = 4x \\ (\left\lfloor \frac{N}{4} \right\rfloor^2 + \left\lfloor \frac{N}{4} \right\rfloor + 1)\lambda & N = 4x + 2 \end{cases} \quad (8)$$

The traffic on right, left and cross-right ejection links are denoted by $\lambda_{ej-right}$, $\lambda_{ej-left}$ and $\lambda_{ej-crs-right}$ respectively and equal to

$$\lambda_{ej-right} = \lambda_{ej-left} = (\frac{N}{2} - 1)\lambda \quad (9)$$

$$\lambda_{ej-crs-right} = \lambda \quad (10)$$

And finally, the traffic on right cross-network link, $\lambda_{cross-right}$ and left cross-network link, $\lambda_{cross-left}$, are equal to

$$\lambda_{cross-right} = \lambda_{inj-crs-right} \quad (11)$$

$$\lambda_{cross-left} = \lambda_{inj-crs-left} \quad (12)$$

4 Validation and Analysis

To validate the analytical model we have developed a discrete event simulator operating at flit level. Each simulation experiment is run until the network reaches its steady state, i.e. until a further increase in simulated network cycles does not change the collected statistics appreciably. The simulator operates on the same assumption as the analysis. Some of the assumptions are mentioned here. A network cycle is defined as the time required that a flit traverse between two adjacent router or between a router and an IP. The time consumed in the routers is also ignored in simulation. Messages are generated at each node according to a Poisson process. Also all messages are assumed to be of equal size.

Destinations at each node are selected randomly and the traffic is uniform. The latency for a message is considered as the time a message is created in the source to the time when the last flit of the message is absorbed by the destination IP.

The model is compared against the simulation results for numerous configurations by changing the message length and also the network size. Fig. 4 compares the simulation results against the analysis for the networks of size 16, 32, 64 and 128 when the length of the message is set to 16, 32, 48 and 64.

The horizontal axis in the figures shows the message rate while the vertical axis describes the latency. As can be seen from the the figures the analytical model presents a good approximation of the network latency in the presence of the light and heavy traffic. In particular the figures reveal that the model predicts the network saturation points very accurately.

5 Conclusion

The Quarc introduced as a NoC scheme that is highly efficient in performing collective communication operations including broadcast and multicast. In this paper we have presented an introduction to the Quarc NoC and proposed an analytical performance model of the communication latency in the Quarc NoC. The simulation results have revealed that the model exhibit a very good degree of accuracy in predicting the average message latency in the Quarc scheme.

Our next objective is to present analytical model for broadcast and multicast in the Quarc NoC.

References

[1] D.K. Panda et al. Multidestination Message Passing Mechanism Conforming to Base Wormhole Routing Scheme. *Int'l Workshop on Parallel Computer Routing and Communication*, 1994.

[2] H. Hashemi-Najafabad , H. Sarbazi-Azad , and P. Rajabzadeh . An accurate performance model of fully adaptive routing in wormhole-switched two-dimensional mesh multicomputers. *Microprocessors and Microsystems*, 2007.

[3] Ju-Young Park. Construction of Optimal Multicast Trees Based on the Parameterized Communication Model. *Int'l Conf. on Parallel Processing*, 1996.

[4] Lu Zhonghai , Yin Bei , and A. Jantsch. Connection-oriented multicasting in wormhole-switched networks on chip. *IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures*, 2006.

[5] E. Bolotin, et. al. QoS architecture and design process for Networks-on- Chip. *Journal of Systems Arch*, 2004.

[6] E. Rijpkema, K. Goossens, and P. Wielage. Router Architecture for Networks on Silicon. *Progress , 2nd Workshop On Embedded Systems*, 2001.

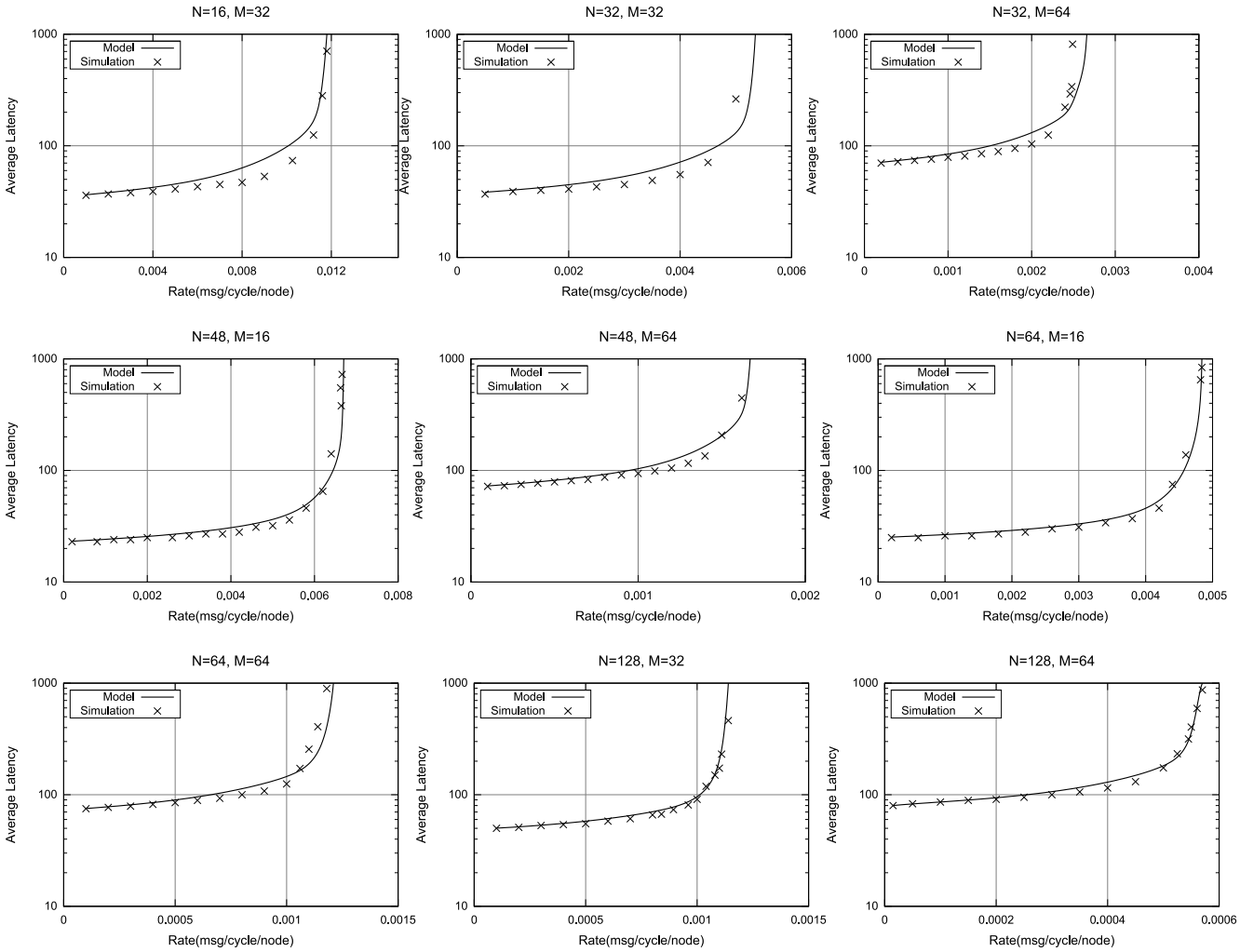

Figure 4. A comparison of the simulation results against the analytical model.

[7] Hong Xu et al. Optimal software multicast in wormhole-routed multistage networks. *IEEE Transactions on Parallel and Distributed Systems*, 1997.

[8] J. Duato et al. *Interconnection networks: An Engineering Approach*. Morgan Kaufmann, 2003.

[9] J. Liu, L.-R. Zheng, and H. Tenhunen . Interconnect intellectual property for network-on-chip. *Journal of System Architectures*, 2003.

[10] K. Goossens, J. Dielissen, and A. Radulescu. Aethereal network on chip: concepts, architectures, and implementations. *IEEE, Design and Test of Computers*, pages 414–421, 2005.

[11] L. Kleinrock. *Queueing Systems Volume I: Theory*. John Wiley and Sons, 1975.

[12] M. Coppola, R. Locatelli, G. Maruccia, L. Pieralisi, and A. Scandurra. Spidergon: a novel on-chip communication network. *Int'l Symposium on System-on-Chip*, 2004.

[13] M. Dall'Osso et al. xpipes: a Latency Insensitive Parameterized Network on-Chip Architecture for Multi-Processor SoCs. *Int'l Conf. on Computer Design table of contents*, 2003.

[14] M. Millberg, E. Nilsson, R. Thid, S. Kumar, and A. Jantsch. The nostrum backbone-a communication protocol stack for Networks on Chip. *17th Int'l Conf. on VLSI Design*, pages 693– 696, 2004.

[15] M. Moadeli, A. Shahrabi, W. Vanderbauwhede, and M. Ould-Khaoua. An analytical performance model for the Spidergon NoC. *21st IEEE Int'l Conf. on Advanced Information Networking and Applications*, pages 1014–1021, 2007.

[16] M. Moadeli et al. Communication Modeling of the Spidergon NoC with Virtual Channels. In *Int Conf on Parallel Processing*, 2007.

[17] STMicroelectronics. www.st.com.

[18] W. J. Dally and B. Towles. Route packets, not wires: On-chip interconnection networks. *Design Automation Conf. (DAC)*, pages 683–689, 2001.