

Applying dataflow analysis to dimension buffers for guaranteed performance in networks on chip

Citation for published version (APA):

Hansson, M. A., Wiggers, M., Moonen, A. J. M., Goossens, K. G. W., & Bekooij, M. J. G. (2008). Applying dataflow analysis to dimension buffers for guaranteed performance in networks on chip. In *2nd ACM/IEE Proceedings of the 2nd ACM/IEEE International Symposium on Networks-on-Chips (NOCS 2008) 7 - 11 April 2008, Newcastle upon Tyne, UK* (pp. 211-212). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/NOCS.2008.4492742>

DOI:

[10.1109/NOCS.2008.4492742](https://doi.org/10.1109/NOCS.2008.4492742)

Document status and date:

Published: 01/01/2008

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Applying Dataflow Analysis to Dimension Buffers for Guaranteed Performance in Networks on Chip

Andreas Hansson¹, Maarten Wiggers², Arno Moonen¹, Kees Goossens^{3,4} and Marco Bekooij⁴

¹Eindhoven University of Technology, Eindhoven, The Netherlands

²Twente University of Technology, Enschede, The Netherlands

³Delft University of Technology, Delft, The Netherlands

⁴Research, NXP Semiconductors, Eindhoven, The Netherlands

m.a.hansson@tue.nl

Abstract

A Network on Chip (NoC) with end-to-end flow control is modelled by a cyclo-static dataflow graph. Using the proposed model together with state-of-the-art dataflow analysis algorithms, we size the buffers in the network interfaces. We show, for a range of NoC designs, that buffer sizes are determined with a run time comparable to existing analytical methods, and results comparable to exhaustive simulation.

1 Introduction

A growing number of applications, often with *real-time requirements*, are integrated on the same System on Chip (SoC), in the form of hardware and software Intellectual Property (IP). Applications are split into *tasks*, and it is the onus of the interconnect to facilitate the real-time requirements of the inter- and intra-task communication.

Networks on Chip (NoC) offer *latency and throughput guarantees* [6, 8]. The guarantees depend on the arbitration in the routers and Network Interfaces (NI), but also on the NI buffers. These *decoupling buffers* absorb differences in speed and burstiness between the IP and the NoC [5], and thereby hide network internals, such as packetisation, arbitration, and *end-to-end flow control* [2, 9]. If the NI buffers are not sufficiently large, the guarantees are violated. The size must, however, be minimised, as the buffers are a major contributor to NoC power and silicon area [3].

Existing approaches to dimension NI buffers [3, 5] are based on linear bounds [5], resulting in a low run time but large buffers, or exhaustive simulation [3], with smaller buffers but a run time of several days for larger SoC designs. In this work, we model the NoC and the IP using a *dataflow graph*. In contrast to [3, 5], that are based on *network calculus*, dataflow analysis cannot only dimension the buffers given the temporal requirements, but also *determine the temporal behaviour of the SoC for given buffer sizes*, e.g. to analyse if new applications fit on an existing NoC.

As the main contributions of this paper, we: 1) show how to construct a dataflow graph for a NoC communication channel, 2) use this model with state-of-the-art dataflow analysis techniques [10] to dimension the NI buffers. The run time is comparable to existing analytical methods, and the results are comparable to exhaustive simulation.

Section 2 describes the proposed channel model. In Section 3, we apply dataflow analysis [4, 10] to determine conservative bounds on the NI buffer sizes. Finally, conclusions are drawn in Section 4. More details are found in [7].

2 Channel model

We use Cyclo-Static Dataflow (CSDF) [1] models to compute buffer sizes. A CSDF graph is a directed graph, consisting of *actors* connected by edges. An actor has distinct *phases* of execution, and synchronises by communicating *tokens* over edges. An actor is enabled to fire when tokens are available on all its input edges and transitions from phase to phase in a cyclic fashion.

The proposed channel model is shown in Figure 1. In the figure, $n \times 1$ denotes a vector of ones of length n , and the italic symbol 1 denotes a vector of ones of appropriate length. The Response Times (RT) [1] of the individual actors appear above and below the actors. Similar to [3, 5], the model is based on the notion of a producer and consumer, connected by a *forward* channel that carries data and a *reverse* channel that carries *end-to-end flow-control* credits. The buffers of the channel are represented by β_p and β_c .

Our method *allows any CSDF model of the IP*, but to enable a comparison with existing models, the IP behaviour is described by a period of p_p and p_c cycles, and a burst size of b_p and b_c words, for producer and consumer, respectively. The model reflects that only one word can be produced per cycle, thereby reducing the resulting buffer sizes.

In this work, we model the *Æthereal* NoC [6] that uses time-division multiplexing (TDM) to provide latency and throughput guarantees. The model has five parameters, the period of the TDM wheel p_n , and four parameters related

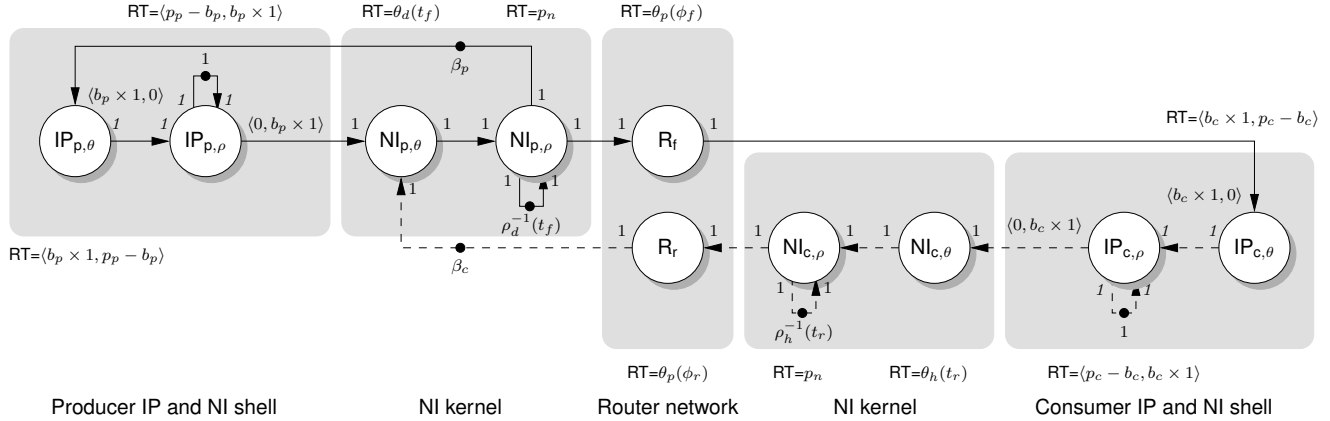


Figure 1. Data travelling in the forward direction (solid) and credits in the reverse direction (dashed).

Table 1. Buffer sizes for mobile phone system

Algorithm	Run time (s)	Tot. buf. (words)	Impr. (%)
Analytical [5]	0.05	1025	ref
Simulated [3]	6845	799	12
Dataflow approx. [10]	0.78	721	30
Dataflow exact [4]	547	680	34

to the allocated resources: the forward and reverse path, ϕ_f and ϕ_r , plus the time-slot allocation in the two directions, denoted t_f and t_r . The throughput and latency of the NI is determined by t_f and t_r . The functions θ_d , ρ_d^{-1} , θ_h and ρ_h^{-1} conservatively bound the *latency* and *rate* for data and credits respectively. The router network is modelled as a latency only, given by the path and the function θ_p .

The aforementioned bounding functions are determined by the NoC architecture, and include e.g. packetisation overhead, pipelining delay and arbitration. While the parameters and functions used in Figure 1 are specific for the *Æthereal* NoC, the model is applicable as long as the arbiters that are applied in the NIs and routers can be characterised as *latency-rate servers* [11], e.g. [2] and [8].

3 Experimental results

We compare the run time and buffer sizes derived using our approach with those of [3, 5]. Averaging over a set of 1000 randomly generated use-cases, each with 100 connections, we see a buffer-size reduction of 36% [10], 41% [3] and 44% [4] compared to [5]. The run time, using [10], is consistently below a second for all the different use-cases.

A phone SoC with telecom, multi-media and gaming constitutes our design example. The results are shown in Table 1. The dataflow-based methods result in improvements of more than 30% while the simulation only reduces the buffers by 12%. In addition, the run time of the dataflow approximation algorithm is four orders of magnitude lower.

4 Conclusions

The latency and throughput guarantees of Networks on Chip (NoC) depends on appropriately sized decoupling buffers in the network interfaces, situated between the Intellectual Property (IP) modules and the router network. Existing buffer-sizing methods are based on network calculus and rely on coarse linear bounds or exhaustive simulation, resulting in either large buffers or impractically long run times. In this work, we propose to capture the behaviour of the NoC and the IPs using a dataflow model. The presented model is an important step in enabling the use of dataflow analysis for NoC resource allocation. The proposed method is evaluated by comparing with existing buffer-sizing approaches on a range of SoC designs. Buffer sizes are determined with a run time comparable to existing analytical methods, and results comparable to exhaustive simulation. For larger SoC designs, where the simulation-based approach is not practical, our approach finishes in seconds.

References

- [1] G. Bilsen *et al.* Cyclo-Static Dataflow. *IEEE Tr. on Sig. Proc.*, 44(2), 1996.
- [2] T. Bjerregaard *et al.* An OCP compliant network adapter for GALS-based SoC design using the MANGO network-on-chip. In *Pr. SOC*, 2005.
- [3] M. Coenen *et al.* A buffer-sizing algorithm for networks on chip using TDMA and credit-based end-to-end flow control. In *Pr. CODES+ISSS*, 2006.
- [4] A. Dasdan. Experimental analysis of the fastest optimum cycle ratio and mean algorithms. *ACM TODAES*, 9(4), 2004.
- [5] O. P. Gangwal *et al.* Building predictable systems on chip: An analysis of guaranteed communication in the *Æthereal* network on chip. In *Dynamic and Robust Streaming In And Between Connected Consumer-Electronics Devices*. Kluwer, 2005.
- [6] K. Goossens *et al.* The *Æthereal* network on chip: Concepts, architectures, and implementations. *IEEE Des. and Test of Comp.*, 2005.
- [7] A. Hansson *et al.* Applying dataflow analysis to dimension buffers for guaranteed performance in networks on chip. Technical Report NXP-R-TN 2008/00013, NXP Semiconductors, 2008.
- [8] A. Jantsch. Models of computation for networks on chip. In *Pr. ACSD*, 2006.
- [9] A. Rădulescu *et al.* An efficient on-chip network interface offering guaranteed services, shared-memory abstraction, and flexible network programming. *IEEE Tr. on CAD of Int. Circ. and Syst.*, 2005.
- [10] M. Wiggers *et al.* Efficient Computation of Buffer Capacities for Cyclo-Static Dataflow Graphs. In *Pr. DAC*, 2007.
- [11] M. Wiggers *et al.* Modelling run-time arbitration by latency-rate servers in dataflow graphs. In *Pr. SCOPES*, 2007.