Applying dataflow analysis to dimension buffers for guaranteed performance in networks on chip
Hansson, M.A.; Wiggers, M.; Moonen, A.J.M.; Goossens, K.G.W.; Bekooij, M.J.G.

Published in:
2nd ACM/IEProceedings of the 2nd ACM/IEEE International Symposium on Networks-on-Chips (NOCS 2008) 7
- 11 April 2008, Newcastle upon Tyne, UK

DOI:
10.1109/NOCS.2008.4492742

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 author's 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.

Citation for published version (APA):

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

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Applying Dataflow Analysis to Dimension Buffers for Guaranteed Performance in Networks on Chip

Andreas Hansson\(^1\), Maarten Wiggers\(^2\), Arno Moonen\(^1\), Kees Goossens\(^3,4\) and Marco Bekooij\(^1\)

\(^1\)Eindhoven University of Technology, Eindhoven, The Netherlands
\(^2\)Twente University of Technology, Enschede, The Netherlands
\(^3\)Delft University of Technology, Delft, The Netherlands
\(^4\)Research, 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, 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 \( \beta_p \) and \( \beta_c \).

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...
An efficient on-chip network interface offering guaranteed communication in the Æthereal network on chip: Concepts, architectures, and implementations.

Applying dataflow analysis to dimension buffers for latency-rate servers

An OCP compliant network adapter for GALS-based Intellectual Property (IP) modules and the router network.

Efficient Computation of Buffer Capacities for Cyclo-Static Dataflow

Table 1. Buffer sizes for mobile phone system

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Run time(s)</th>
<th>Tot. buf. (words)</th>
<th>Impr. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical [5]</td>
<td>0.05</td>
<td>1025</td>
<td>ref</td>
</tr>
<tr>
<td>Dataflow approx. [10]</td>
<td>0.78</td>
<td>721</td>
<td>30</td>
</tr>
<tr>
<td>Dataflow exact [4]</td>
<td>547</td>
<td>680</td>
<td>34</td>
</tr>
</tbody>
</table>

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

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 presented 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