



Universiteit  
Leiden  
The Netherlands

## **On the power efficiency, low latency, and quality of service in network-on-chip**

Wang, P.

### **Citation**

Wang, P. (2020, February 12). *On the power efficiency, low latency, and quality of service in network-on-chip*. Retrieved from <https://hdl.handle.net/1887/85165>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/85165>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/85165> holds various files of this Leiden University dissertation.

**Author:** Wang, P.

**Title:** On the power efficiency, low latency, and quality of service in network-on-chip

**Issue Date:** 2020-02-12

# **On the Power Efficiency, Low Latency, and Quality of Service in Network-on-Chip**

Peng Wang



# **On the Power Efficiency, Low Latency, and Quality of Service in Network-on-Chip**

## **PROEFSCHRIFT**

ter verkrijging van  
de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus Prof.mr. C.J.J.M. Stolker,  
volgens besluit van het College voor Promoties  
te verdedigen op woensdag 12 februari 2020  
klokke 11:15 uur

door

Peng Wang  
geboren te Liaoning, China  
in 1990

<b>Promotor:</b>	Dr. Todor P. Stefanov	Universiteit Leiden
<b>Second-Promotor:</b>	Prof. Dr. Joost N.Kok	Universiteit Leiden
<b>Promotion Committee:</b>	Prof. Dr.-ing Diana Göhringer	Technische Universität Dresden
	Prof. Dr. Georgi Gaydadjiev	Rijksuniversiteit Groningen
	Prof. Dr. Kees G.W. Goossens	Technische Universiteit Eindhoven
	Prof. Dr. Aske Plaat	Universiteit Leiden
	Prof. Dr. Harry A.G. Wijshoff	Universiteit Leiden
	Prof. Dr. Fons J. Verbeek	Universiteit Leiden

On the Power Efficiency, Low Latency, and Quality of Service in Network-on-Chip  
Peng Wang. -  
Dissertation Universiteit Leiden. - With ref. - With summary in Dutch.

Copyright © 2019 by Peng Wang. All rights reserved.

This dissertation was typeset using L<sup>A</sup>T<sub>E</sub>X in Linux.  
Cover designed by Jian Zhang

ISBN: 978-90-9032677-1  
Printed in the Netherlands.

# Contents

<b>Contents</b>	<b>v</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Abbreviations</b>	<b>1</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Design Trends in Network-on-Chip . . . . .	3
1.1.1 Low Power Consumption . . . . .	3
1.1.2 Low Latency . . . . .	4
1.1.3 Advanced Quality of Service . . . . .	5
1.2 Contradictions between Design Trends . . . . .	6
1.3 Problem Statement . . . . .	7
1.3.1 Problem 1 . . . . .	7
1.3.2 Problem 2 . . . . .	8
1.4 Contributions of The Thesis . . . . .	9
1.5 Thesis Outline . . . . .	12
<b>2 Background</b>	<b>13</b>
2.1 Network-on-Chip . . . . .	13
2.1.1 Network Topologies . . . . .	14
2.1.2 Routing Approaches . . . . .	16
2.1.3 Flow Control Approaches . . . . .	18
2.1.4 Router Architecture in NoCs . . . . .	20
2.2 Power Consumption Analysis . . . . .	23
2.3 Conventional Power Gating in A NoC . . . . .	24

<b>3</b>	<b>Duty Buffer Based Power Gating Approach</b>	<b>27</b>
3.1	Problem Statement . . . . .	27
3.2	Contributions . . . . .	28
3.3	Related work . . . . .	29
3.4	DB-based Approach . . . . .	30
3.4.1	Input Port with Duty Buffer . . . . .	31
3.4.2	Power Gating on VCs . . . . .	32
3.4.3	Power Gating Scheme . . . . .	33
3.5	Experimental Results . . . . .	36
3.5.1	Evaluation on Synthetic Workloads . . . . .	37
3.5.2	Evaluation on Real Application Workloads . . . . .	39
3.6	Discussion . . . . .	41
<b>4</b>	<b>D-bypass Power Gating Approach</b>	<b>43</b>
4.1	Problem Statement . . . . .	43
4.2	Contributions . . . . .	45
4.3	Background . . . . .	46
4.4	Related Work . . . . .	47
4.5	D-bypass Approach . . . . .	48
4.5.1	Extended Router Structure . . . . .	48
4.5.2	An Example of the Reservation Process . . . . .	50
4.5.3	Power Gating Conditions . . . . .	52
4.6	Experimental Results . . . . .	53
4.6.1	Evaluation on Synthetic Workloads . . . . .	54
4.6.2	Evaluation on Real Application Workloads . . . . .	57
4.7	Discussion . . . . .	60
<b>5</b>	<b>EVC-based Power Gating Approach</b>	<b>61</b>
5.1	Problem Statement . . . . .	61
5.2	Contributions . . . . .	62
5.3	Background . . . . .	63
5.4	Related Work . . . . .	64
5.5	Our EVC-based Power Gating . . . . .	66
5.5.1	Distribution of Virtual Bypass Paths . . . . .	66
5.5.2	Extended Router Structure . . . . .	67
5.5.3	Power Gating Scheme . . . . .	68
5.5.4	Example of Our Power Gating Approach . . . . .	69
5.5.5	Resolving Starvation . . . . .	71
5.6	Experimental Results . . . . .	71
5.6.1	Evaluation on Synthetic Workloads . . . . .	73



5.6.2	Evaluation on Real Application Workloads . . . . .	75
5.7	Discussion . . . . .	77
<b>6</b>	<b>Energy-Efficient Confined-interference Communication</b>	<b>79</b>
6.1	Problem Statement . . . . .	79
6.2	Contributions . . . . .	81
6.3	Background . . . . .	82
6.3.1	Surf-routing . . . . .	82
6.3.2	BLESS-routing . . . . .	83
6.4	Related Work . . . . .	83
6.5	Surf-Bless Routing Approach . . . . .	85
6.5.1	Wave Pattern in Surf-Bless . . . . .	85
6.5.2	Router Architecture . . . . .	87
6.5.3	Surf-Bless routing algorithm . . . . .	88
6.6	Experimental Results . . . . .	89
6.6.1	Evaluation on Synthetic Workloads . . . . .	90
6.6.2	Transfer of Multiple Class Packets . . . . .	94
6.7	Discussion . . . . .	97
<b>7</b>	<b>Summary and Conclusion</b>	<b>99</b>
	<b>Bibliography</b>	<b>105</b>
	<b>List of Publications</b>	<b>113</b>
	<b>Samenvatting</b>	<b>117</b>
	<b>Acknowledgements</b>	<b>119</b>
	<b>Curriculum Vitae</b>	<b>121</b>



# List of Figures

1.1	Contributions outline. . . . .	9
2.1	An example of a many-core system. . . . .	14
2.2	Classical network topologies. . . . .	15
2.3	Deadlock caused by cyclic dependency of packets. . . . .	17
2.4	A timeline example of a credit-based flow control. . . . .	19
2.5	Router architecture. . . . .	20
2.6	Router pipeline. . . . .	22
2.7	Power consumption in a $8 \times 8$ 2D mesh NoC. . . . .	23
2.8	Conventional NoC power gating. . . . .	24
2.9	Wakeup process. . . . .	25
3.1	The NoC structure and extended input port and output port. . . . .	31
3.2	Controllers for input and output ports. . . . .	34
3.3	The interactive process between an output port in the upstream router and the input port in the corresponding downstream router. . . . .	35
3.4	Average packet latency across full range of workloads. . . . .	37
3.5	The average packet latency. . . . .	39
3.6	The breakdown of the NoC power consumption. . . . .	40
4.1	Node-Router Decoupling. . . . .	46
4.2	Extended router structure in D-bypass. . . . .	49
4.3	Example of the reservation process. . . . .	50
4.4	Packet latency across different injection rates. . . . .	55
4.5	Power consumption across different injection rates. . . . .	56
4.6	Execution time. . . . .	57
4.7	Average packet latency. . . . .	58
4.8	Breakdown of the NoC power consumption. . . . .	59
5.1	Express virtual channel. . . . .	64

5.2	Extended EVC-based power gating approach. . . . .	67
5.3	An example of our power gating approach. . . . .	70
5.4	Latency across different injection rates. . . . .	73
5.5	Power consumption across different injection rates. . . . .	74
5.6	Execution time. . . . .	75
5.7	Average network latency. . . . .	76
5.8	Power consumption. . . . .	77
6.1	Confined-interference communication on a NoC. . . . .	80
6.2	The wave pattern in Surf-routing. . . . .	83
6.3	Wave pattern in Surf-Bless routing. . . . .	86
6.4	Router architecture in Surf-Bless. . . . .	87
6.5	Non-interference between domains. . . . .	91
6.6	Energy consumption across different number of domains. . . . .	92
6.7	Latency across different number of domains. . . . .	93
6.8	Application execution time. . . . .	94
6.9	NoC packet latency. . . . .	95
6.10	NoC energy consumption. . . . .	96
7.1	Packet latency (PL) and power consumption (PC) at low traffic workloads (l), medium traffic workloads (m), and high traffic workloads (h). . . . .	100

# List of Tables

1.1	NoCs on real chips. . . . .	3
3.1	Parameters. . . . .	36
3.2	Area of router components. . . . .	41
4.1	Parameters. . . . .	54
5.1	Parameters used in experiments. . . . .	72
6.1	Parameters. . . . .	90

