

# Contents

<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xix</b>
<b>Abbreviations and Acronyms</b>	<b>xxi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Thesis Contributions . . . . .	4
1.2 Thesis Outline . . . . .	6
<b>2 Background and Related Work</b>	<b>9</b>
2.1 The Cache Hierarchy . . . . .	10
2.1.1 Cache Coherence Protocols . . . . .	11
2.1.1.1 Invalidiation-based vs Update-based Protocols . . . . .	11
2.1.1.2 Steady States at L1 Cache Controllers . . . . .	12
2.1.1.3 Snoopy and Directory Protocols . . . . .	16
2.1.1.4 Directory Implementation . . . . .	18
2.1.2 Block Mapping Policies in Shared Banked LLCs . . . . .	23
2.1.3 Power Implications . . . . .	25
2.2 The Network-on-Chip . . . . .	26
2.2.1 NoCs Topology . . . . .	27
2.2.2 The Switch . . . . .	28
2.2.3 Data Units . . . . .	29
2.2.4 Switching . . . . .	30
2.2.5 Flow Control . . . . .	32
2.2.6 Arbitration . . . . .	33
2.2.7 Routing . . . . .	33
2.2.7.1 Implementation of a Routing Algorithm . . . . .	34
2.2.7.2 Unicast, Multicast and Broadcast Messages . . . . .	36
2.2.8 NoC and Cache Coherence . . . . .	38
2.3 Evaluation Platform . . . . .	39
2.3.1 gMemNoCsim . . . . .	39
2.3.2 Graphite . . . . .	41
2.3.3 Sniper . . . . .	41
2.3.4 CACTI . . . . .	42
2.3.5 Orion-2 . . . . .	42
2.3.6 Xilinx ISE . . . . .	42

<b>3 Network-Level Optimizations</b>	<b>43</b>
3.1 Introduction . . . . .	44
3.2 The Gather Network . . . . .	46
3.2.1 Description of a Logic Block . . . . .	47
3.2.2 GN Wiring Layout . . . . .	49
3.2.3 Implementation Analysis . . . . .	50
3.2.4 Sequential Implementation of the Gather Network . . . . .	52
3.3 GN Applied to Hammer Protocol . . . . .	54
3.3.1 Reseting The GN Wires . . . . .	55
3.4 GN Applied to Directory Protocol . . . . .	56
3.4.1 Reseting the GN Wires . . . . .	57
3.4.2 Protocol Modifications . . . . .	58
3.5 GN Performance Evaluation . . . . .	59
3.5.1 Directory Protocol with GN . . . . .	60
3.5.2 Hammer Protocol with GN . . . . .	62
3.5.3 Sequential Gather Network . . . . .	67
3.6 Conclusions . . . . .	70
<b>4 Runtime Home Mapping</b>	<b>73</b>
4.1 Introduction . . . . .	74
4.2 Runtime Home Mapping . . . . .	78
4.2.1 Avoiding Multiple LLC Misses . . . . .	83
4.2.2 Adapting the GN Module to Support RHM . . . . .	84
4.2.3 Mapping Algorithm . . . . .	88
4.2.4 Replacements in L1 Cache . . . . .	91
4.3 Optimizations to RHM . . . . .	92
4.3.1 Block Migration . . . . .	92
4.3.2 Block Replication . . . . .	94
4.3.3 RHM and Broadcast-based Coherence Protocols . . . . .	98
4.3.3.1 Broadcast Network . . . . .	99
4.3.4 Merging Hammer Protocol and RHM . . . . .	99
4.3.5 Parallel Tag Access . . . . .	102
4.4 Evaluation . . . . .	104
4.4.1 Performance . . . . .	105
4.4.2 Performance Conclusions . . . . .	107
4.4.3 Energy . . . . .	109
4.4.4 Parallel Tag Access . . . . .	110
4.5 Conclusions . . . . .	111
<b>5 pNC: Partitioned NoC and Cache Hierarchy</b>	<b>113</b>
5.1 Introduction . . . . .	114
5.2 NoC and Cache Hierarchy Substrate . . . . .	116
5.2.1 pNC: LBDR and RHM Support to Virtualization . . . . .	116
5.2.2 LBDR Regions . . . . .	120
5.2.3 Memory Controller Design . . . . .	121
5.2.4 Mapping Algorithm . . . . .	122
5.3 Evaluation . . . . .	123

5.3.1	pNC Overhead . . . . .	124
5.3.2	Performance in Fault-Free Systems . . . . .	124
5.3.3	Performance in Faulty Systems . . . . .	128
5.4	Conclusions . . . . .	128
<b>6</b>	<b>Heterogeneous LLC Design</b>	<b>131</b>
6.1	Motivation . . . . .	132
6.2	Dynamic L2 Cache Line Allocation . . . . .	134
6.2.1	Replacement Policy . . . . .	137
6.2.2	Dynamic Power Techniques . . . . .	138
6.3	Performance Evaluation . . . . .	140
6.3.1	Benefits when Using MOESI Protocol . . . . .	143
6.4	Conclusions . . . . .	145
<b>7</b>	<b>Conclusions</b>	<b>146</b>
<b>A</b>	<b>Coherence Protocols</b>	<b>152</b>
A.1	Directory (MESI) . . . . .	152
A.2	Hammer . . . . .	155
A.3	Directory + RHM with Block Migration and Replication . . . . .	157
A.4	Hammer + RHM . . . . .	163
<b>B</b>	<b>Implementation of the Target CMP in an FPGA Board</b>	<b>167</b>
<b>References</b>		<b>175</b>



# List of Figures

1.1	Baseline CMP system.	5
1.2	Contributions of this thesis.	5
1.3	Final CMP system.	5
2.1	Baseline tile-based CMP system.	11
2.2	Simplified FSM for the MSI protocol.	13
2.3	Simplified FSM for the MOESI protocol.	14
2.4	Simplified FSM for the MESI protocol.	15
2.5	Snoopy protocol example.	17
2.6	Directory protocol example.	17
2.7	Simplified FSM for the LLC.	19
2.8	Requests for a private block in full-map directory protocols.	21
2.9	Requests for a shared block in full-map directory protocols.	22
2.10	Write request management in broadcast-based protocols.	22
2.11	A general overview of a network architecture.	26
2.12	A $4 \times 4$ 2-dimensional mesh.	28
2.13	General structure of a VC-less switch.	29
2.14	Data units.	30
2.15	LBDR logic and configuration bits for east output port.	35
2.16	Structure of gMemNoCsim.	40
3.1	Write request for a shared block in a Directory protocol.	44
3.2	Format of a short coherence message.	45
3.3	Gather Network (subnetwork for Tile 0).	46
3.4	Gathering ACKs in a broadcast-based protocol.	47
3.5	Logic block at Tile 5.	48
3.6	Control signals distribution (XY layout).	49
3.7	Control signals distribution (mixed layout).	50
3.8	Structure of a sequential GN module.	54
3.9	Logic at the inputs of each AND gate when the system implements Hammer coherence protocol.	55
3.10	Configuration of the Gather Network.	57
3.11	Logic at the inputs of each AND gate when the system implements Directory coherence protocol.	57
3.12	First alternative to let the GN work with directory-based protocols: acknowledgements are sent to the home L2.	58
3.13	Second alternative to let the GN work with directory-based protocols: the L1 invalidates the sharers.	59

3.14 Normalized execution time (SPLASH-2 applications). . . . .	61
3.15 Evaluation results with synthetic access traces. . . . .	62
3.16 Breakdown of network messages in Hammer protocol. . . . .	63
3.17 Normalized execution time (Hammer). . . . .	64
3.18 Normalized number of injected messages (GN signals are not included). .	64
3.19 Normalized store miss latency (Hammer). . . . .	65
3.20 Normalized load miss latency (Hammer). . . . .	65
3.21 Normalized NoC dynamic energy. . . . .	66
3.22 Normalized execution time with different GN delays. . . . .	66
3.23 Normalized execution time compared to a NoC with an high priority VC for the ACKs. . . . .	67
3.24 Normalized execution time with the two implementations of the Gather Network. . . . .	68
3.25 Number of conflicts per gather message received at destination node. .	68
3.26 Average GN latency (sequential implementation). . . . .	69
4.1 Average distance of L2 banks to their L1 requestors for different mapping policies. . . . .	75
4.2 RHM example (block is mapped on requestor's tile). . . . .	75
4.3 Runtime Home Mapping. Different scenarios. . . . .	76
4.4 Home search phase. . . . .	77
4.5 RHM global overview. From processor access to MC access. . . . .	79
4.6 RHM coherence actions (read request). . . . .	81
4.7 RHM coherence actions (write request). . . . .	82
4.8 Control info required for each version of RHM. . . . .	83
4.9 Switch with a GN module adapted to RHM. . . . .	85
4.10 GN input logic adapted to RHM. . . . .	86
4.11 GN central logic adapted to RHM. . . . .	87
4.12 GN output logic adapted to RHM. . . . .	88
4.13 GCN Mapping of IDs. . . . .	89
4.14 GN message format. . . . .	89
4.15 Mapping algorithm performed by the MC. . . . .	90
4.16 Block migration process. . . . .	93
4.17 Block replication. . . . .	95
4.18 Block replication process. . . . .	96
4.19 BN implementation. . . . .	99
4.20 Read request for a shared block in case of hit (left) or miss (right) in the local L2 bank. . . . .	100
4.21 Request hit in the local L2 bank. . . . .	101
4.22 Request miss in the local L2 bank. . . . .	102
4.23 Parallel Tag Access: motivation and implementation. . . . .	103
4.24 4-stages (left) and 3-stages (right) switches modified to allow parallel tag access. . . . .	104
4.25 Avg hop distance between L1 requestors and the tile where the data is found. . . . .	106
4.26 Percentage of hits in the L2 bank located in the tile's requestor. . . . .	106
4.27 Execution time normalized to the S-NUCA case. . . . .	107

4.28	Average load and store latency, normalized to the S-NUCA case. . . . .	108
4.29	NoC's energy consumption . . . . .	109
4.30	LLC's energy consumption . . . . .	110
4.31	Normalized reduction in broadcasts and execution time when using PTA. . . . .	111
5.1	Partitioned CMP system. . . . .	114
5.2	Virtualized CMP system to three applications. Resources are assigned to different applications. . . . .	115
5.3	Baseline system for the pNC approach. . . . .	117
5.4	GN signals in a virtualized environments. . . . .	117
5.5	pNC switch design. . . . .	118
5.6	Example of GCN connected with LBDR bits. . . . .	119
5.7	Processor Partitions and Home Partitions example. . . . .	121
5.8	PP, HP and faulty tables at the MC. . . . .	121
5.9	Mapping algorithm performed by the MC in pNC. . . . .	123
5.10	Normalized execution time. . . . .	125
5.11	Normalized L2 misses. . . . .	125
5.12	Home stealing configuration. . . . .	126
5.13	Normalized execution time and L2 misses (mixed applications). . . . .	127
5.14	Normalized execution time for each application of the three sets. . . . .	127
5.15	Normalized execution time and L2 misses (mixed applications) with faulty LLC banks. . . . .	128
6.1	Breakdown of actions performed by the LLC when an L1 request is received. . . . .	132
6.2	Percentage of stale and valid blocks replaced at the LLC. . . . .	133
6.3	LLC finite state machine (MESI protocol). . . . .	135
6.4	Different LLC configurations by changing the number of sets and the associativity of the L2 and directory structures. . . . .	136
6.5	Example of LLC reorganization. . . . .	137
6.6	Replacement policy. . . . .	138
6.7	Evolution of a block when dynamic power-off techniques are used. . . . .	139
6.8	Normalized execution time. MESI protocol with L2 banks with 512 sets. . . . .	141
6.9	Normalized execution time. MESI protocol with L2 banks with 256 sets. . . . .	141
6.10	Normalized LLC area occupancy. . . . .	142
6.11	Normalized L2 leakage. MESI protocol. . . . .	142
6.12	Normalized L2 leakage. MESI with sleep transistors. . . . .	143
6.13	Simplified FSM for the L2 cache (MOESI protocol). . . . .	144
6.14	Normalized execution time. MOESI (for 1:x ratio proposals) and MESI (for baseline). . . . .	144
6.15	Normalized L2 leakage. MOESI with sleep transistors (for 1:x ratio proposals) and MESI (for baseline). . . . .	145
B.1	Target system. . . . .	167
B.2	Tiled CMP overview. . . . .	168
B.3	Structure of a tile. . . . .	169
B.4	Structure of a cache module. . . . .	169
B.5	Caches/NI interface. . . . .	170
B.6	Breakdown of FPGA resources required by a tile. . . . .	171

B.7 Breakdown of FPGA resources required by L1 data cache. . . . .	172
B.8 Breakdown of FPGA resources required by an L2 bank. . . . .	172

# List of Tables

3.1	Area and delay for the switch modules . . . . .	51
3.2	Conventional 2D mesh critical path. . . . .	51
3.3	GN critical path. . . . .	52
3.4	Area and latency results of the sequential GN module. . . . .	54
3.5	Network and cache parameters (GN with Directory protocol). . . . .	61
3.6	Network and cache parameters (GN with Hammer protocol). . . . .	64
4.1	Network and cache parameters (RHM evaluation). . . . .	105
5.1	Network parameters (pNC evaluation). . . . .	124
5.2	Sets of applications executed in the CMP. . . . .	127
B.1	FPGA resource occupancy of a single tile. . . . .	172