

# Preface

Modern computing devices, ranging from smartphones and tablets up to powerful servers, rely on complex silicon chips that integrate inside them hundreds or thousands of processing elements. The design of such systems is not an easy task. Efficient design methodologies are needed that would organize the designer's work and reduce the risk for a low-efficiency system. One of the main challenges that the designer faces is how to connect the components inside the silicon chip, both physically and logically, without compromising performance. The network-on-chip (NoC) paradigm tries to answer this question by applying at the silicon chip level well established networking principles, after suitably adapting them to the silicon chip characteristics and to application demands. The routers are the heart and the backbone of the NoC. Their main function is to route data from source to destination, while they provide arbitrary connectivity between several inputs and outputs that allows the implementation of arbitrary network topologies.

This book focuses on the microarchitecture of NoC routers that together, with the network interfaces, execute all network functionalities. The routers implement the transport and physical layers of the NoC, and their internal organization critically affects the speed of the network in terms of clock frequency, the throughput of the network in terms of how many packets can the network service per clock cycle and, the network's area and energy footprint on the silicon die.

The goal of this book is to describe the complex behavior of network routers in a compositional approach following simple construction steps that can be repeated by any designer in a straightforward manner. The micro-architectural features presented in this book are built on top of detailed examples and abstracted models, when necessary, that do not leave any dark spots on the operation of the presented blocks and reveal the dependencies between the different parts of the router, thus enabling any possible future optimization. The material of each chapter evolves linearly, covering simpler cases before moving to more complex architectures.

Chapter 1 gives an overview of network-on-chip design at the system level and discusses the layered approach followed for transforming the abstract read and

write transactions between the modules of the system to actual bits that travel in parallel on the links of the network finding their path towards their final destination, using the routers of the network.

Chapter 2 deals with link-level flow control policies and associated buffering requirements for guaranteeing lossless and full throughput operation for the communication of a single sender and receiver pair connected with a simple point-to-point link. The discussion includes both simple ready/valid flow control as well as credit-based policies under a unified abstract flow control model. The behavior of both flow control policies when used in pipelined links is analyzed and analytical bounds are derived for each case. The chapter ends with the packetization process and the enhancement needed to link-level flow control policies for supporting multiword packets.

Chapter 3 departs from point-to-point links and discusses in a step-by-step manner the organization of many-to-one and many-to-many switched connections supporting either simple or fully unrolled datapaths. The interplay between arbitration, multiplexing and flow control is analyzed in detail using both credits and ready/valid protocols. The chapter ends with the design of a full wormhole (or virtual-cut through) router that includes also a routing computation module that allows routers to be embedded in arbitrary network topologies.

Chapter 4 departs from router microarchitecture and describes in detail the circuit-level organization of the arbiters and multiplexers used in the control and the datapath of the routers. A unified approach is presented that merges algorithmically the design of arbiters that employ various arbitration policies with that of multiplexing and allows the design of efficient arbiter and multiplexing circuits. Additionally, arbiters built on top of 2D relative priority state are also discussed in detail.

Chapter 5 dives deeper in the microarchitecture of a wormhole router and discusses in a compositional manner the pipeline alternatives of wormhole routers and their implementation/performance characteristics. Multiple pipelined organizations are derived based on two pipeline primitive modules. For each case, complete running examples are given that highlight the pipeline idle cycles imposed by the router's structural dependencies, either across packets or inside packets of the same input, and the way such dependencies are removed after appropriate pipeline modifications.

Chapter 6 introduces virtual channels together with the flow control mechanism and the buffering architectures needed to support their operation. Virtual channels correspond to adding lanes to a street network that allow cars (packets) to utilize in a more efficient manner the available physical resources. Lanes are added virtually and the packets that move in different lanes use the physical channels of the network in a time-multiplexed manner. The interplay of buffering, flow-control latencies and the chosen flow control mechanism (ready/valid or credits) are analyzed in detail in this chapter and the requirements of each configuration are identified.

Chapter 7 introduces the microarchitecture of routers that connect links that support multiple virtual channels. The design of virtual-channel-based switching

connections begins from a simple many-to-one switching module and evolves to a complete virtual-channel-based router. The operation of a virtual-channel-based router involves several tasks that are analyzed in detail together with their dependencies and their interaction with the flow-control mechanism.

Chapter 8 builds on top of Chap. 7 and presents the organization of high-speed allocators that speedup significantly the operation of a baseline single-cycle virtual-channel-based router. Multiple alternatives are presented that allow either the reduction of the needed allocation steps or their parallel execution that effectively reduces the hardware delay of the router.

Chapter 9 deals with the pipelined organization and microarchitecture of virtual-channel-based routers. The pipelined configurations of the virtual-channel-based routers are described in a modular manner, beginning from the description of the structure and operation of three primitive pipeline stages. Then, following a compositional approach, several multi-stage pipelined configurations are derived by connecting the presented primitive stages in a plug-and-play manner, which helps in understanding better the operation of complex organizations and their associated timing-throughput tradeoffs.

Overall, we expect system, architecture, circuit, and EDA researchers and developers, who are interested in understanding the microarchitecture of network-on-chip routers, the associated design challenges, and the available solutions, to benefit from the material of this book and appreciate the order of presentation that evolves in a step-by-step manner, from the basic design principles to sophisticated design techniques.

Xanthi, Greece  
June 2014

Giorgos Dimitrakopoulos  
Anastasios Psarras  
Ioannis Seitanidis

<http://www.springer.com/978-1-4614-4300-1>

Microarchitecture of Network-on-Chip Routers

A Designer's Perspective

Dimitrakopoulos, G.; Psarras, A.; Seitanidis, I.

2015, XIV, 175 p. 134 illus., 77 illus. in color., Hardcover

ISBN: 978-1-4614-4300-1