Robotic Computing on FPGAs

Synthesis Lectures on Computer Architecture

Editor
Natalie Enright Jerger, ***University of Toronto***

Editor Emerita
Margaret Martonosi, ***Princeton University***

Founding Editor Emeritus
Mark D. Hill, ***University of Wisconsin, Madison***

Synthesis Lectures on Computer Architecture publishes 50- to 100-page books on topics pertaining to the science and art of designing, analyzing, selecting, and interconnecting hardware components to create computers that meet functional, performance, and cost goals. The scope will largely follow the purview of premier computer architecture conferences, such as ISCA, HPCA, MICRO, and ASPLOS.

Quantum Computer System: Research for Noisy Intermediate-Scale Quantum Computers
Yongshan Ding and Frederic T. Chong
2020

A Primer on Memory Consistency and Cache Coherence, Second Edition
Vijay Nagarajan, Daniel J. Sorin, Mark D. Hill, and David Wood
2020

Innovations in the Memory System
Rajeev Balasubramonian
2019

Cache Replacement Policies
Akanksha Jain and Calvin Lin
2019

The Datacenter as a Computer: Designing Warehouse-Scale Machines, Third Edition
Luiz André Barroso, Urs Hölzle, and Parthasarathy Ranganathan
2018

Principles of Secure Processor Architecture Design
Jakub Szefer
2018

General-Purpose Graphics Processor Architectures
Tor M. Aamodt, Wilson Wai Lun Fung, and Timothy G. Rogers
2018

Compiling Algorithms for Heterogenous Systems
Steven Bell, Jing Pu, James Hegarty, and Mark Horowitz
2018

Architectural and Operating System Support for Virtual Memory
Abhishek Bhattacharjee and Daniel Lustig
2017

Deep Learning for Computer Architects
Brandon Reagen, Robert Adolf, Paul Whatmough, Gu-Yeon Wei, and David Brooks
2017

On-Chip Networks, Second Edition
Natalie Enright Jerger, Tushar Krishna, and Li-Shiuan Peh
2017

On-Chip Photonic Interconnects: A Computer Architect's Perspective
Christopher J. Nitta, Matthew K. Farrens, and Venkatesh Akella
2013

Optimization and Mathematical Modeling in Computer Architecture
Tony Nowatzki, Michael Ferris, Karthikeyan Sankaralingam, Cristian Estan, Nilay Vaish, and David Wood
2013

Security Basics for Computer Architects
Ruby B. Lee
2013

The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, Second Edition
Luiz André Barroso, Jimmy Clidaras, and Urs Hölzle
2013

Shared-Memory Synchronization
Michael L. Scott
2013

Resilient Architecture Design for Voltage Variation
Vijay Janapa Reddi and Meeta Sharma Gupta
2013

Multithreading Architecture
Mario Nemirovsky and Dean M. Tullsen
2013

Performance Analysis and Tuning for General Purpose Graphics Processing Units (GPGPU)
Hyesoon Kim, Richard Vuduc, Sara Baghsorkhi, Jee Choi, and Wen-mei Hwu
2012

Automatic Parallelization: An Overview of Fundamental Compiler Techniques
Samuel P. Midkiff
2012

Phase Change Memory: From Devices to Systems
Moinuddin K. Qureshi, Sudhanva Gurumurthi, and Bipin Rajendran
2011

Multi-Core Cache Hierarchies
Rajeev Balasubramonian, Norman P. Jouppi, and Naveen Muralimanohar
2011

A Primer on Memory Consistency and Cache Coherence
Daniel J. Sorin, Mark D. Hill, and David A. Wood
2011

Dynamic Binary Modification: Tools, Techniques, and Applications
Kim Hazelwood
2011

Quantum Computing for Computer Architects, Second Edition
Tzvetan S. Metodi, Arvin I. Faruque, and Frederic T. Chong
2011

High Performance Datacenter Networks: Architectures, Algorithms, and Opportunities
Dennis Abts and John Kim
2011

Processor Microarchitecture: An Implementation Perspective
Antonio González, Fernando Latorre, and Grigorios Magklis
2010

Transactional Memory, Second Edition
Tim Harris, James Larus, and Ravi Rajwar
2010

Computer Architecture Performance Evaluation Methods
Lieven Eeckhout
2010

Introduction to Reconfigurable Supercomputing
Marco Lanzagorta, Stephen Bique, and Robert Rosenberg
2009

On-Chip Networks
Natalie Enright Jerger and Li-Shiuan Peh
2009

The Memory System: You Can't Avoid It, You Can't Ignore It, You Can't Fake It
Bruce Jacob
2009

Fault Tolerant Computer Architecture
Daniel J. Sorin
2009

The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines
Luiz André Barroso and Urs Hölzle
2009

Computer Architecture Techniques for Power-Efficiency
Stefanos Kaxiras and Margaret Martonosi
2008

Chip Multiprocessor Architecture: Techniques to Improve Throughput and Latency
Kunle Olukotun, Lance Hammond, and James Laudon
2007

Transactional Memory
James R. Larus and Ravi Rajwar
2006

Quantum Computing for Computer Architects
Tzvetan S. Metodi and Frederic T. Chong
2006

© Springer Nature Switzerland AG 2022
Reprint of original edition © Morgan & Claypool 2021

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

Robotic Computing on FPGAs

Shaoshan Liu, Zishen Wan, Bo Yu, and Yu Wang

ISBN: 978-3-031-00643-2 paperback
ISBN: 978-3-031-01771-1 ebook
ISBN: 978-3-031-00068-3 hardcover

DOI 10.1007/978-3-031-01771-1

A Publication in the Springer series
SYNTHESIS LECTURES ON COMPUTER ARCHITECTURE

Lecture #56
Series Editor: Natalie Enright Jerger, *University of Toronto*
Editor Emerita: Margaret Martonosi, *Princeton University*
Founding Editor Emeritus: Mark D. Hill, *University of Wisconsin, Madison*
Series ISSN
Print 1935-3235 Electronic 1935-3243

Robotic Computing on FPGAs

Shaoshan Liu
PerceptIn

Zishen Wan
Georgia Institute of Technology

Bo Yu
PerceptIn

Yu Wang
Tsinghua University

SYNTHESIS LECTURES ON COMPUTER ARCHITECTURE #56

ABSTRACT

This book provides a thorough overview of the state-of-the-art field-programmable gate array (FPGA)-based robotic computing accelerator designs and summarizes their adopted optimized techniques. This book consists of ten chapters, delving into the details of how FPGAs have been utilized in robotic perception, localization, planning, and multi-robot collaboration tasks. In addition to individual robotic tasks, this book provides detailed descriptions of how FPGAs have been used in robotic products, including commercial autonomous vehicles and space exploration robots.

KEYWORDS

robotics, FPGAs, autonomous machines, perception, localization, planning, control, space exploration, deep learning

Contents

Preface

In this book, we provide a thorough overview of the state-of-the-art FPGA-based robotic computing accelerator designs and summarize their adopted optimized techniques. The authors combined have over 40 years of research experiences of utilizing FPGAs in robotic applications, both in academic research and commercial deployments. For instance, the authors have demonstrated that, by co-designing both the software and hardware, FPGAs can achieve more than 10× better performance and energy efficiency compared to the CPU and GPU implementations. The authors have also pioneered the utilization of the partial reconfiguration methodology in FPGA implementations to further improve the design flexibility and reduce the overhead. In addition, the authors have successfully developed and shipped commercial robotic products powered by FPGAs and the authors demonstrate that FPGAs have excellent potential and are promising candidates for robotic computing acceleration due to its high reliability, adaptability, and power efficiency.

The authors believe that FPGAs are the best compute substrate for robotic applications for several reasons. First, robotic algorithms are still evolving rapidly, and thus any ASIC-based accelerators will be months or even years behind the state-of-the-art algorithms. On the other hand, FPGAs can be dynamically updated as needed. Second, robotic workloads are highly diverse, thus it is difficult for any ASIC-based robotic computing accelerator to reach economies of scale in the near future. On the other hand, FPGAs are a cost effective and energy-effective alternative before one type of accelerator reaches economies of scale. Third, compared to systems on a chip (SoCs) that have reached economies of scale, e.g., mobile SoCs, FPGAs deliver a significant performance advantage. Fourth, partial reconfiguration allows multiple robotic workloads to time-share an FPGA, thus allowing one chip to serve multiple applications, leading to overall cost and energy reduction.

Specifically, FPGAs require little power and are often built into small systems with less memory. They have the ability of massively parallel computations and to make use of the properties of perception (e.g., stereo matching), localization (e.g., simultaneous localization and mapping (SLAM)), and planning (e.g., graph search) kernels to remove additional logic so as to simplify the end-to-end system implementation. Taking into account hardware characteristics, several algorithms are proposed which can be run in a hardware-friendly way and achieve similar software performance. Therefore, FPGAs are possible to meet real-time requirements while achieving high energy efficiency compared to central processing units (CPUs) and graphics processing units (GPUs). In addition, unlike the application-specific integrated circuit (ASIC) counterparts, FPGA technologies provide the flexibility of on-site programming and re-programming without going through re-fabrication with a modified design. Partial Recon-

figuration (PR) takes this flexibility one step further, allowing the modification of an operating FPGA design by loading a partial configuration file. Using PR, part of the FPGA can be reconfigured at runtime without compromising the integrity of the applications running on those parts of the device that are not being reconfigured. As a result, PR can allow different robotic applications to time-share part of an FPGA, leading to energy and performance efficiency, and making FPGA a suitable computing platform for dynamic and complex robotic workloads. Due to the advantages over other compute substrates, FPGAs have been successfully utilized in commercial autonomous vehicles as well as in space robotic applications, for FPGAs offer unprecedented flexibility and significantly reduced the design cycle and development cost.

This book consists of ten chapters, providing a thorough overview of how FPGAs have been utilized in robotic perception, localization, planning, and multi-robot collaboration tasks. In addition to individual robotic tasks, we provide detailed descriptions of how FPGAs have been used in robotic products, including commercial autonomous vehicles and space exploration robots.

Shaoshan Liu
June 2021