Data Processing on FPGAs

Synthesis Lectures on Data Management

Editor
M. Tamer Özsu, ***University of Waterloo***

Synthesis Lectures on Data Management is edited by Tamer Özsu of the University of Waterloo. The series will publish 50- to 125 page publications on topics pertaining to data management. The scope will largely follow the purview of premier information and computer science conferences, such as ACM SIGMOD, VLDB, ICDE, PODS, ICDT, and ACM KDD. Potential topics include, but not are limited to: query languages, database system architectures, transaction management, data warehousing, XML and databases, data stream systems, wide scale data distribution, multimedia data management, data mining, and related subjects.

Data Processing on FPGAs
Jens Teubner and Louis Woods
2013

Perspectives on Business Intelligence
Raymond T. Ng, Patricia C. Arocena, Denilson Barbosa, Giuseppe Carenini, Luiz Gomes, Jr. Stephan Jou, Rock Anthony Leung, Evangelos Milios, Renée J. Miller, John Mylopoulos, Rachel A. Pottinger, Frank Tompa, and Eric Yu
2013

Semantics Empowered Web 3.0: Managing Enterprise, Social, Sensor, and Cloud-based Data and Services for Advanced Applications
Amit Sheth and Krishnaprasad Thirunarayan
2012

Data Management in the Cloud: Challenges and Opportunities
Divyakant Agrawal, Sudipto Das, and Amr El Abbadi
2012

Query Processing over Uncertain Databases
Lei Chen and Xiang Lian
2012

Foundations of Data Quality Management
Wenfei Fan and Floris Geerts
2012

Incomplete Data and Data Dependencies in Relational Databases
Sergio Greco, Cristian Molinaro, and Francesca Spezzano
2012

Business Processes: A Database Perspective
Daniel Deutch and Tova Milo
2012

Data Protection from Insider Threats
Elisa Bertino
2012

Deep Web Query Interface Understanding and Integration
Eduard C. Dragut, Weiyi Meng, and Clement T. Yu
2012

P2P Techniques for Decentralized Applications
Esther Pacitti, Reza Akbarinia, and Manal El-Dick
2012

Query Answer Authentication
HweeHwa Pang and Kian-Lee Tan
2012

Declarative Networking
Boon Thau Loo and Wenchao Zhou
2012

Full-Text (Substring) Indexes in External Memory
Marina Barsky, Ulrike Stege, and Alex Thomo
2011

Spatial Data Management
Nikos Mamoulis
2011

Database Repairing and Consistent Query Answering
Leopoldo Bertossi
2011

Managing Event Information: Modeling, Retrieval, and Applications
Amarnath Gupta and Ramesh Jain
2011

Fundamentals of Physical Design and Query Compilation
David Toman and Grant Weddell
2011

Methods for Mining and Summarizing Text Conversations
Giuseppe Carenini, Gabriel Murray, and Raymond Ng
2011

Probabilistic Databases
Dan Suciu, Dan Olteanu, Christopher Ré, and Christoph Koch
2011

Peer-to-Peer Data Management
Karl Aberer
2011

Probabilistic Ranking Techniques in Relational Databases
Ihab F. Ilyas and Mohamed A. Soliman
2011

Uncertain Schema Matching
Avigdor Gal
2011

Fundamentals of Object Databases: Object-Oriented and Object-Relational Design
Suzanne W. Dietrich and Susan D. Urban
2010

Advanced Metasearch Engine Technology
Weiyi Meng and Clement T. Yu
2010

Web Page Recommendation Models: Theory and Algorithms
Sule Gündüz-Ögüdücü
2010

Multidimensional Databases and Data Warehousing
Christian S. Jensen, Torben Bach Pedersen, and Christian Thomsen
2010

Database Replication
Bettina Kemme, Ricardo Jimenez-Peris, and Marta Patino-Martinez
2010

Relational and XML Data Exchange
Marcelo Arenas, Pablo Barcelo, Leonid Libkin, and Filip Murlak
2010

User-Centered Data Management
Tiziana Catarci, Alan Dix, Stephen Kimani, and Giuseppe Santucci
2010

Data Stream Management
Lukasz Golab and M. Tamer Özsu
2010

Access Control in Data Management Systems
Elena Ferrari
2010

An Introduction to Duplicate Detection
Felix Naumann and Melanie Herschel
2010

Privacy-Preserving Data Publishing: An Overview
Raymond Chi-Wing Wong and Ada Wai-Chee Fu
2010

Keyword Search in Databases
Jeffrey Xu Yu, Lu Qin, and Lijun Chang
2009

© Springer Nature Switzerland AG 2022
Reprint of original edition © Morgan & Claypool 2013

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.

Data Processing on FPGAs

Jens Teubner and Louis Woods

ISBN: 978-3-031-00721-7 paperback
ISBN: 978-3-031-01849-7 ebook

DOI 10.1007/978-3-031-01849-7

A Publication in the Springer series
SYNTHESIS LECTURES ON DATA MANAGEMENT

Lecture #35
Series Editor: M. Tamer Özsu, *University of Waterloo*
Series ISSN
Synthesis Lectures on Data Management
Print 2153-5418 Electronic 2153-5426

Data Processing on FPGAs

Jens Teubner
Databases and Information Systems Group, Dept. of Computer Science, TU Dortmund

Louis Woods
Systems Group, Dept. of Computer Science, ETH Zürich

SYNTHESIS LECTURES ON DATA MANAGEMENT #35

ABSTRACT

Roughly a decade ago, power consumption and heat dissipation concerns forced the semiconductor industry to radically change its course, shifting from sequential to parallel computing. Unfortunately, improving performance of applications has now become much more difficult than in the good old days of frequency scaling. This is also affecting databases and data processing applications in general, and has led to the popularity of so-called *data appliances*—specialized data processing engines, where software and hardware are sold together in a closed box. Field-programmable gate arrays (FPGAs) increasingly play an important role in such systems. FPGAs are attractive because the performance gains of specialized hardware can be significant, while power consumption is much less than that of commodity processors. On the other hand, FPGAs are way more flexible than hard-wired circuits (ASICs) and can be integrated into complex systems in many different ways, e.g., directly in the network for a high-frequency trading application. This book gives an introduction to FPGA technology targeted at a database audience. In the first few chapters, we explain in detail the inner workings of FPGAs. Then we discuss techniques and design patterns that help mapping algorithms to FPGA hardware so that the inherent parallelism of these devices can be leveraged in an optimal way. Finally, the book will illustrate a number of concrete examples that exploit different advantages of FPGAs for data processing.

KEYWORDS

FPGA, modern hardware, database, data processing, stream processing, parallel algorithms, pipeline parallelism, programming models

Contents

Preface . . . . . **xiii**

1 Introduction **1**

1.1 Moore's Law and Transistor-Speed Scaling 1
1.2 Memory Wall and Von Neumann Bottleneck 1
1.3 Power Wall 2
1.4 Multicore CPUs and GPUs 3
1.5 Specialized Hardware 4
1.6 Field-Programmable Gate Arrays (FPGAs) 5
1.7 FPGAs for Data Processing 5
1.7.1 Stream Processing 5
1.7.2 Big Data 6
1.7.3 Cloud Computing 6
1.7.4 Security 6

2 A Primer in Hardware Design **9**

2.1 Basic Hardware Components 9
2.1.1 Combinational Logic 9
2.1.2 Sequential Logic 10
2.1.3 Asynchronous sequential logic 11
2.1.4 Synchronous sequential logic 11
2.2 Hardware Programming 12
2.2.1 Hardware Description Languages (HDLs) 12
2.3 Circuit Generation 14
2.3.1 Logical Design Flow (Synthesis) 14
2.3.2 Physical Design Flow 15

3 FPGAs **17**

3.1 A Brief History of FPGAs 17
3.2 Look-up Tables—The Key to Re-Programmability 18
3.2.1 LUT Representation of a Boolean Function 18

3.2.2 Internal Architecture of an LUT . . . 18
3.2.3 LUT (Re)programming . . . 18
3.2.4 Alternative Usage of LUTs . . . 19
3.3 FPGA Architecture . . . 20
3.3.1 Elementary Logic Units (Slices/ALMs) . . . 20
3.4 Routing Architecture . . . 21
3.4.1 Logic Islands . . . 21
3.4.2 Interconnect . . . 22
3.5 High-Speed I/O . . . 22
3.6 Auxiliary On-Chip Components . . . 24
3.6.1 Block RAM (BRAM) . . . 24
3.6.2 Digital Signal Processing (DSP) Units . . . 25
3.6.3 Soft and Hard IP-Cores . . . 26
3.7 FPGA Programming . . . 26
3.7.1 FPGA Design Flow . . . 27
3.7.2 Dynamic Partial Reconfiguration . . . 28
3.8 Advanced Technology and Future Trends . . . 30
3.8.1 Die Stacking . . . 30
3.8.2 Heterogeneous Die-Stacked FPGAs . . . 31
3.8.3 Time-Multiplexed FPGAs . . . 31
3.8.4 High-level Synthesis . . . 32

4 FPGA Programming Models . . . 33
4.1 Re-Build, Parameterize, or Program the Hardware Accelerator? . . . 33
4.1.1 Re-Building Circuits at Runtime . . . 34
4.1.2 Parameterized Circuits . . . 35
4.1.3 Instruction Set Processors on top of FPGAs . . . 37
4.2 From Algorithm to Circuit . . . 38
4.2.1 Expression → Circuit . . . 38
4.2.2 Circuit Generation . . . 39
4.2.3 High-Level Synthesis . . . 40
4.3 Data-Parallel Approaches . . . 41
4.3.1 Data Parallelism . . . 42
4.4 Pipeline-Parallel Approaches . . . 43
4.4.1 Pipeline Parallelism in Hardware . . . 43
4.4.2 Pipelining in FPGAs . . . 45
4.4.3 Designing for Pipeline Parallelism . . . 46

4.4.4 Turning a Circuit into a Pipeline-Parallel Circuit 47
4.5 Related Concepts 48

5 Data Stream Processing 51

5.1 Regular Expression Matching 51
5.1.1 Finite-State Automata for Pattern Matching 51
5.1.2 Implementing Finite-State Automata in Hardware 53
5.1.3 Optimized Circuit Construction 54
5.1.4 Network Intrusion Detection 56
5.2 Complex Event Processing 58
5.2.1 Stream Partitioning 58
5.2.2 Hardware Partitioner 59
5.2.3 Best-Effort Allocation 60
5.2.4 Line-Rate Performance 60
5.3 Filtering in the Data Path 61
5.3.1 Data Path Architecture in the Real World 61
5.4 Data Stream Processing 62
5.4.1 Compositional Query Compilation 64
5.4.2 Getting Data In and Out 67
5.5 Dynamic Query Workloads 67
5.5.1 Fast Workload Changes Through Partial Modules 68
5.6 Bibliographic Notes 69

6 Accelerated DB Operators 71

6.1 Sort Operator 71
6.1.1 Sorting Networks 71
6.1.2 BRAM-based FIFO Merge Sorter 73
6.1.3 External Sorting with a Tree Merge Sorter 74
6.1.4 Sorting with Partial Reconfiguration 75
6.2 Skyline Operator 76
6.2.1 Standard Block Nested Loops (BNL) Algorithm 77
6.2.2 Parallel BNL with FPGAs 78
6.2.3 Performance Characteristics 79

7 Secure Data Processing 83

7.1 FPGAs versus CPUs 83
7.1.1 Von Neumann Architecture 83

7.1.2 Trusted Platform Module (TPM) 83
7.2 FPGAs versus ASICs 84
7.3 Security Properties of FPGAs 84
7.3.1 Bitstream Encryption 85
7.3.2 Bitstream Authentication 85
7.3.3 Further Security Mechanisms 85
7.4 FPGA as Trusted Hardware 86
7.4.1 Fully Homomorphic Encryption with FPGAs 86
7.4.2 Hybrid Data Processing 86
7.4.3 Trusted Hardware Implementation 88

8 Conclusions 89

A Commercial FPGA Cards 93
A.1 NetFPGA 93
A.2 Solarflare's ApplicationOnload™ Engine 93
A.3 Fusion I/O's ioDrive 94

Bibliography 95

Authors' Biographies 101

Index 103

Preface

System architectures, hardware design, and programmable logic (specifically, *field-programmable gate arrays* or *FPGAs*) are topics generally governed by electrical engineers. "Hardware people" are in charge of embracing technological advantages (and turning them into improved performance), preferably without breaking any of the established hardware/software interfaces, such as instruction sets or execution models.

Conversely, computer scientists and software engineers are responsible for understanding users' problems and satisfying their application and functionality demands. While doing so, they hardly care how hardware functions underneath—much as their hardware counterparts are largely unaware of how their systems are being used for concrete problems.

As time progresses, this traditional separation between hard- and software leaves more and more potential of modern technology unused. But giving up the separation and building *hardware/software co-designed systems* requires that both parties involved understand each other's terminology, problems/limitations, requirements, and expectations.

With this book we want to help work toward this idea of co-designed architectures. Most importantly, we want to give the software side of the story—the database community in particular—a basic understanding of the involved hardware technology. We want to explain what FPGAs are, how they can be programmed and used, and which role they could play in a database context.

This book is intended for students and researchers in the database field, including those that have not had much contact with hardware technology in the past, but would love to get introduced to the field. At ETH Zürich/TU Dortmund, we have been teaching for several years a course titled "Data Processing on Modern Hardware." The material in this book is one part of that Master-level course (which further discusses also "modern hardware" other than FPGAs).

We start the book by highlighting the urgent need from the database perspective to invest more effort into hardware/software co-design issues (Chapter 1). Chapters 2 and 3 then introduce the world of electronic circuit design, starting with a high-level view, then looking at FPGAs specifically. Chapter 3 also explains how FPGAs work internally and why they are particularly attractive at the present time.

In the remaining chapters, we then show how the potential of FPGAs can be turned into actual systems. First, we give general guidelines how algorithms and systems can be designed to leverage the potential of FPGAs (Chapter 4). Chapter 5 illustrates a number of examples that successfully used FPGAs to improve database performance. But FPGAs may also be used to enable new database functionality, which we discuss in Chapter 7 by example of a database crypto

co-processor. We conclude in Chapter 8 with a wary look into the future of FPGAs in a database context.

A short appendix points to different flavors of FPGA system integration, realized through different plug-ins for commodity systems.

Jens Teubner and Louis Woods
June 2013