

Computer Communications and Networks

Series editor

A.J. Sammes

Centre for Forensic Computing

Cranfield University, Shrivenham campus

Swindon, UK

The **Computer Communications and Networks** series is a range of textbooks, monographs and handbooks. It sets out to provide students, researchers, and non-specialists alike with a sure grounding in current knowledge, together with comprehensible access to the latest developments in computer communications and networking.

Emphasis is placed on clear and explanatory styles that support a tutorial approach, so that even the most complex of topics is presented in a lucid and intelligible manner.

More information about this series at <http://www.springer.com/series/4198>

Veljko Milutinović • Jakob Salom
Nemanja Trifunovic • Roberto Giorgi

Guide to DataFlow Supercomputing

Basic Concepts, Case Studies,
and a Detailed Example

 Springer

Veljko Milutinović
School of Electrical Engineering
University of Belgrade
Serbia

Nemanja Trifunovic
Maxeler Technologies Inc.
Palo Alto, CA, USA

Jakob Salom
MISANU
Belgrade, Serbia

Roberto Giorgi
University of Siena
Italy

ISSN 1617-7975 ISSN 2197-8433 (electronic)
Computer Communications and Networks
ISBN 978-3-319-16228-7 ISBN 978-3-319-16229-4 (eBook)
DOI 10.1007/978-3-319-16229-4

Library of Congress Control Number: 2015939052

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

This book is meant to support teaching of the DataFlow subject. University courses fully or partially dedicated to the DataFlow subject are important for the following reason: More and more Big Data are present in all kinds of research or commercial challenges. Consequently, the DataFlow paradigm is getting importance, since it has been proven that it is the most suitable computing paradigm for Big Data. It offers superior speedups (depending on the application, from about 20 to about 200, even 2000 in some isolated cases) as well as power savings (typically about 20 times); it brings size reduction, too. A recent study by researchers of the Tsinghua University in China reveals that for Shallow Water Weather Forecast (a Big Data problem), on the 1U level, compared to Tianhe-2 (at the time of writing of this book, rated #1 on the Top 500 SuperComputers list, which compares SuperComputers based on Linpack, a small data benchmark), Maxeler (a DataFlow machine) demonstrates the speedup of 14. With all the above in mind, the book is divided into four chapters: The first one is of an introductory nature. The second gives an overview of the related research. The third represents a case study. The fourth one is oriented to the ease of use and covers the issues of importance for WebIDE (a web-based integrated development environment). The work on this book was partially supported by the MISANU project #44006.

Belgrade, Serbia
Belgrade, Serbia
Palo Alto, CA, USA
Siena, Italy

Veljko Milutinović
Jakob Salom
Nemanja Trifunovic
Roberto Giorgi

Contents

1	The DataFlow Paradigm	1
1.1	Introduction	1
1.2	The Concept	4
1.3	The Programming Paradigm	6
1.4	The Compilation	8
1.5	Comparisons	9
1.6	DataFlow Hardware (A Maxeler Example)	12
1.7	Application Types	13
1.8	Application Examples	15
1.9	Acceleration Is Hard	17
1.10	Open Research Problems	20
1.11	Enabler Mechanisms for DataFlow Supercomputing	23
1.12	Getting Started	25
1.13	OPEX	26
1.14	Computer Size Matters (1U, 2U, 4U)	26
1.15	Bottom Line	27
1.16	OpenSPL	30
1.17	WebIDE	31
1.18	Instead of the Conclusion	32
	References	38
2	Selected Case Studies	41
2.1	Classification of Selected Examples	41
2.1.1	Examples for Each Leaf Within the Main Classification	42
2.2	Presentation of Examples	43
2.2.1	Examples from Exascale Fundamental Drivers Leaf #1: Finance	44
2.2.2	Examples from Exascale Fundamental Drivers Leaf #2: Mathematics	50
2.2.3	Examples from Exascale Science and Technology Drivers Leaf #1: Biology	55

2.2.4	Examples from Exascale Science and Technology Drivers Leaf #2: Geophysics	59
2.2.5	Examples from Exascale Science and Technology Drivers Leaf #3: Meteorology	62
2.2.6	Examples from Exascale Engineering and Innovation Drivers Leaf #5: Oil Industry	65
2.2.7	Additional Research Papers	70
	References	70
3	An Example Application: Fourier Transform	73
3.1	Introduction	73
3.1.1	About the Cooley-Tukey Algorithm	73
3.1.2	About the Fast Fourier Transform Algorithm	73
3.1.3	Overview of Different Fast Fourier Transform Algorithms	74
3.2	The Radix 2 Cooley-Tukey Algorithm	75
3.3	Mathematical Background	76
3.4	Pseudo Code of the Radix 2 Cooley-Tukey Algorithm	77
3.5	Original Implementation	78
3.5.1	In-Place Implementation of the Cooley-Tukey Algorithm	78
3.6	Analysis of the Accelerated Version of the Cooley-Tukey Algorithm	79
3.7	Implementational Details	79
3.7.1	Using DFE for Calculating the Fast Fourier Transform from C\C++ Code	81
3.8	Benchmarking Program	82
3.9	Performance Measurements of the Accelerated Algorithm and Analysis of the Results	82
3.9.1	Assumptions and Limitations	82
3.9.2	Types of Experiments	82
3.9.3	Comparison of the Performance of the Accelerated Algorithm with Other Publicly Available Implementations	83
3.9.4	Comparison of the Performance of Accelerated Algorithm with Original Implementation	87
3.9.5	Graph Drawing Script	93
3.9.6	Results Overview and Performance Predictions	94
3.9.7	Analysis of the Bottlenecks	94
3.10	Possible Modifications and Improvements	94
3.11	Conclusion	95
3.12	Original Implementation Code	96
3.12.1	MaxJ Code of Kernel that Calculates Fast Fourier Transform	96

- 3.12.2 MaxJ Manager Code of Cooley-Tukey
Radix 2 Algorithm 97
- 3.12.3 MaxJ Manager Code of Bit Reverse Kernel 99
- 3.12.4 C++ Code of Radix 2 In-Place Cooley-Tukey
Algorithm 100
- 3.12.5 C++ Benchmark Code..... 102
- 3.12.6 Script for Filtering Raw Data About Average
Execution Speed of the Experiments..... 104
- 3.12.7 Python Script for Automatic Graph Generation 104
- References 106
- 4 Using the WebIDE 107**
 - 4.1 The WebIDE System Overview 107
 - 4.2 Installation Instructions..... 107
 - 4.2.1 WebIDE Dependencies 107
 - 4.2.2 Installation Steps 108
 - 4.3 Running WebIDE 108
 - 4.4 Security Notes 109
 - 4.4.1 Note1 109
 - 4.4.2 Note2 109
 - 4.5 Deployment Notes 109
 - 4.5.1 NGINX 109
 - 4.5.2 Gunicorn..... 109
 - 4.6 User Administration 110
 - 4.7 Login 110
 - 4.7.1 Login Credentials 112
 - 4.8 Project Explorer..... 112
 - 4.9 Creating a New Project 112
 - 4.10 Deleting an Existing Project 113
 - 4.11 WebIDE with Welcome Windows 115
 - 4.12 WebIDE with Side-by-Side Code Editors 115
 - 4.13 Code Editor 116
 - 4.14 Build Output 117
 - 4.15 Folder Explorer 118
 - 4.16 Image Viewer 119
 - 4.17 PDF Viewer 120
 - 4.18 CSV Viewer 121
 - 4.19 Terminal 121
 - 4.20 WebIDE Change Log 122
- Epilogue 123**
- Postscript 125**
- Index 127**

Computer and Communications networks are key infrastructures of the information society with high socio-economic value as they contribute to the correct operations of many critical services (from healthcare to finance and transportation). Internet is the core of today's computer-communication infrastructures. This has transformed the Internet, from a robust network for data transfer between computers, to a global, content-rich, communication and information system where contents are increasingly generated by the users, and distributed according to human social relations. Next-generation n... A computer network links computers together via a combination of infrared light signals, radio wave transmissions, telephone lines, television cables, and satellite links. The challenge for computer scientists has been to develop protocols (standardized rules for the format and exchange of messages) that allow processes running on host computers to interpret the signals they receive and to engage. The field of networking and communication includes the analysis, design, implementation, and use of local, wide-area, and mobile networks that link computers together. The Internet itself is a network that makes it feasible for nearly all computers in the world to communicate.

of Computer and Communication Networks; and (iii) Mobile and Wireless Communications. However Help Your Kids with Computer Science: A Unique Visual Step-by-Step Guide to Computers, Coding. 258 Pages 2018 41.17 MB 71,978 Downloads New! Science: A Unique Visual Step-by-Step Guide to Computers, Coding, and Communication Dorling Kindersley ... in Communication Networks: From Wired Technologies to Cloud Computing and the Internet of Things Load more similar PDF files. PDF Drive investigated dozens of problems and listed the biggest global issues facing the world today. Let's Change The World Together. Pdfdrive:hope Give books away. A computer network is a group of computers that use a set of common communication protocols over digital interconnections for the purpose of sharing resources located on or provided by the network nodes. The interconnections between nodes are formed from a broad spectrum of telecommunication network technologies, based on physically wired, optical, and wireless radio-frequency methods that may be arranged in a variety of network topologies.