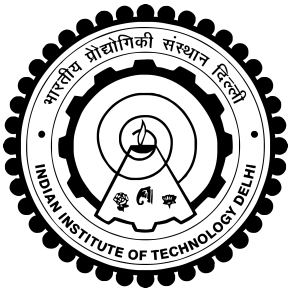
APPLICATION MAPPING ONTO RECONFIGURABLE COARSE-GRAINED ARRAYS

MANSUREH S. MOGHADDAM



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
AUGUST 2015

APPLICATION MAPPING ONTO RECONFIGURABLE COARSE-GRAINED ARRAYS

by

MANSUREH S. MOGHADDAM

Department of Computer Science and Engineering

Submitted
in fulfillment of the requirements of the degree of
Doctor of Philosophy

to the

Indian Institute of Technology Delhi
AUGUST 2015

To My Parents

Certificate

This is to certify that the thesis titled **"APPLICATION MAPPING ONTO RECONFIGURABLE COARSE-GRAINED ARRAYS"** being submitted by Ms. **Mansureh S. Moghaddam** to the **Indian Institute of Technology Delhi**, for the award of the degree of **Doctor of Philosophy**, is a record of bona-fide research work carried out by her under our supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.

M. Balakrishnan
Professor
Department of Computer Science
and Engineering
Indian Institute of Technology Delhi
New Delhi, India 110 016

Kolin Paul
Associate Professor
Department of Computer Science
and Engineering
Indian Institute of Technology Delhi
New Delhi, India 110 016

Date:

New Delhi

Acknowledgements

I acknowledge my research supervisors Prof. M. Balakrishnan and Dr. Kolin Paul for their support throughout the course of the thesis. They introduced me to the art of technical writing and helped me improve my writing abilities with utmost patience. I never forget them sitting next to me and correcting portions of the paper in the process of teaching me. I specially thank them for the freedom and encouragement they have given me from the very beginning of my work that shaped my research. I should specially thank prof. M. Balakrishnan, whose insightful comments and questions helped me to think critically and improve my work, without his help it would have been impossible to debug my ILP-based mapping solution.

My sincere thanks to Dr. Marcus M. Edvall with whom I had many fruitful discussions on using TOMLAB/CPLEX to optimize large scale (M)ILP problems. I also thank Dr Todor Stefanov and Mohamed A. Bamakhrama (www.liad.nl) for their help and suggestions on using DAEDALUS framework.

Many thanks to my colleagues, Surender Kumar Verma, Madhulika Madhur and Mehdi Niktabar who were always ready to help and more than eager for any discussion.

I am grateful to my family for their unconditional love and support. My parents always encouraged me to do whatever interests me and without their support, I would not have been the same.

AUGUST 2015 Mansureh S. Moghaddam

Abstract

Coarse-Grained Reconfigurable Architectures (CGRAs) have become popular in recent times as the increased transistor densities have enabled greater integration of increasingly complex "compute cores". These devices pack massive compute power and can be effectively used to build efficient solutions for applications which have a significant degree of parallelism. In many cases, these CGRAs are also partially reconfigurable. Clearly to make effective use of these highly "parallel compute platforms", a good mapping flow is required to map the parallelism that is present in a target application. We propose an **optimal** mapping flow which also exploits the partial reconfiguration property of modern CGRAs. Hence this algorithm is not only suitable for applications which can be accommodated in the available silicon but also for larger applications (or a set of applications) which need more area than that provided by the CGRA platform. In this work partitions of the application are mapped partially and dynamically (during the life time of the application) onto the platform. While in most mapping flows, "*scheduling*", "*binding*" and "*place & route*" steps are carried out in sequence, our proposed mapping flow integrates all three steps in one "*mapping*" step using an Integer Linear Problem formulation. The algorithm is implemented using TOMLAB/CPLEX toolbox and we assess its efficacy on a set of 40 synthetic task graphs and some real life multimedia applications.

In a multi-processing fabric like CGRA, tiles work in parallel and in a pipelined manner. Therefore processing the application task graph to redistribute the tasks to the nodes can be taken as a preprocessing step to the mapping problem. We deal with this problem in the thesis by proposing algorithms for rebalancing. We implement the proposed mapping flow on a Xilinx XUPV5 FPGA board. This platform is used to implement a 2D mesh of processing elements (tiles) alongwith a main controller. The main controller provides new content to the tiles. The context switch between snapshots is also controlled by the controller using two levels of handshaking. Further, to validate our approach, we integrate it to a mapping flow named Daedalus.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BN	**B**est **N**eighbour
BSB	**B**ase **S**ystem **B**uilder
BSP	**B**oard **S**upport **P**ackage
CA	**C**luster **A**gent
CCG	**C**ombined **C**ore **G**raph
CDCM	**C**ommunication **D**ependence and **C**omputation **M**odel
CDM	**C**ommunication **D**ependence **M**odel
CF	**C**ompact **F**lash
CGRA	**C**oarse **G**rained **R**econfigurable **A**rray
CGRM	**C**oarse **G**rained **R**econfigurable **M**odule
CSO	**C**ontext **S**witch **O**verhead
CTG	**C**ommunication **T**ask **G**raph
CTM	**C**ore **T**ile **M**apping
DAG	**D**irected **A**cyclic **G**raph
DCT	**D**iscrete **C**osine **T**ransform
DFG	**D**ata **F**low **G**raph
DMA	**D**irect **M**emory **A**ccess
DVS	**D**ynamic **V**oltage **S**caling
EDK	**E**mbedded **D**evelopment **K**it
ESL	**E**lectronic **S**ystem **L**evel
FF	**F**irst **F**ree
FFT	**F**ast **F**ourier **T**ransform
FIFO	**F**irst **I**n **F**irst **O**ut

FPGA	**F**ield **P**rogrammable **G**ate **A**rray
FSL	**F**ast **S**erial **L**ink
FU	**F**unctional **U**nit
GA	**G**enetic **A**lgorithm
GA	**G**lobal **A**gent
ICAP	**I**nternal **C**onfiguration **A**ccess **P**ort
IP	**I**ntellectual **P**roperty
ILP	**I**nteger **L**inear **P**roblem
KPN	**K**ahn **P**rocess **N**etwork
LACG	**L**ogical **A**pplication **C**ommunication **G**raph
LED-DN	**L**ow **E**nergy **C**onsumption **D**ependences **N**eighbourhood
LS	**L**ink **S**hutdown
MAC	**M**inimum **A**verage **C**hannel
mCSM	**m**any **C**ore **S**witch **M**apping
MILP	**M**ixed **I**nteger **L**inear **P**roblem
MIMD	**M**ultiple **I**nstruction **M**ultiple **D**ata
MIQP	**M**ixed **I**nteger **Q**uadratic **P**rogramming
MMC	**M**inimum **M**aximum **C**hannel
MPSoC	**M**ulti **P**rocessor **S**ystem **O**n **C**hip
MT-ADRES	**M**ulti **T**hreaded **ADRES**
NCR	**N**ear **C**onvex **R**egion
NN	**N**earest **N**eighbour
NoC	**N**etwork **O**n **C**hip
NRE	**N**on **R**ecurring **E**ngineering
PCM	**P**artial **C**onfiguration **M**anager
PDH	**P**ath-based **D**eterministic **H**euristic
PE	**P**rocessing **E**lement
PL	**P**ath **L**oad
PLB	**P**rocessor **L**ocal **B**us
PPN	**P**olyhedral **P**rocess **N**etwork
PRH	**P**ath-based **R**andomization **H**euristic

PSO	**P**article **S**warm **O**ptimization
QEA	**Q**uantum-inspired **E**valuation **A**lgorithm
QoS	**Q**uality **O**f **S**ervice
QP	**Q**uadratic **P**rogramming
RTL	**R**egister **T**ransfer **L**anguage
SANLP	**S**tatic **A**ffine **N**ested **L**oop **P**rogram
SDF	**S**ynchronous **D**ata **F**low
SDK	**S**oftware **D**evelopment **K**it
SNN	**S**mart **N**earest **N**eighbour
ST-TA	**S**imulated **A**nnealing with **T**iming **A**djustment
TGFF	**T**ask **G**raph **F**or **F**ree
TMTSC	**T**opology **M**apping and **T**raffic **S**urface **C**reating
TSM	**T**ile **S**witch **M**apping
UART	**U**niversal **A**synchronous **R**eceiver **T**ransmitter
WCTG	**W**eighted **C**ommunication **T**ask **G**raph
XMD	**X**ilinx **M**icroprocessor **D**ebug
XPS	**X**ilinx **P**latform **S**tudio