# ARCHITECTURE EXPLORATION OF FPGA BASED ACCELERATORS FOR BIOINFORMATICS APPLICATIONS

**B. SHARAT CHANDRA VARMA**


**AMAR NATH SHASHI KHOSLA SCHOOL OF INFORMATION TECHNOLOGY**
**INDIAN INSTITUTE OF TECHNOLOGY DELHI**
**JANUARY 2015**

# ARCHITECTURE EXPLORATION OF FPGA BASED ACCELERATORS FOR BIOINFORMATICS APPLICATIONS

by

**B. SHARAT CHANDRA VARMA**

Amar Nath and Shashi Khosla School of Information Technology

Submitted

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

to the


**Indian Institute of Technology Delhi**
**January 2015**

# Certificate

This is to certify that the thesis titled **Architecture Exploration of FPGA based Accelerators for Bioinformatics Applications** being submitted by **B. Sharat Chandra Varma** for the award of **Doctor of Philosophy** in **Amar Nath and Shashi Khosla School of Information Technology** is a record of bona-fide work carried out by him under our guidance and supervision at the **Amar Nath and Shashi Khosla School of Information Technology, Indian Institute of Technology Delhi**. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma.

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

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

# Acknowledgments

I would like to thank my supervisors Dr. Kolin Paul and Prof. M. Balakrishnan for their continuous support and expert guidance during the entire duration of my PhD. Their doors were always kept open anytime day or night for me. The interactions with them helped me grow both on the professional and personal front. I would like to thank them for being patient and sharing their immense knowledge with me. I thank them for supporting me in all aspects during my stay at IIT Delhi. I would like to thank Prof.G. S. Visweswaran, Prof. P. R. Panda and Prof. Anshul Kumar for their valuable feedback.

I would also like to thank the staff of Philips and FPGA lab, IIT Delhi, Mrs. Vandana Ahluwalia and Mr. S. D. Sharma for their help with the lab resources.

I would like to thank Prof. Dominique Lavenier for letting me do internship in the SYMBIOSE lab, IRISA, Rennes, France. I am grateful to him for providing me guidance on the project. The work done at Rennes has became an important part of this thesis. I would like to thank Pierre Peterlongo for his valuable inputs on Mapsembler. I would like to thank Prof. Olivier Temam, INRIA, France and Prof. Thambipillai Srikantan, NTU, Singapore for allowing me to spend time in their labs before starting my PhD. The research environment helped me to develop more interest towards research.

I wish to thank all those who have helped me directly and indirectly to reach this special goal of mine. There have been a lot of people who have contributed their bit for me to achieve this goal.

I would like to thank my parents and my grandparents for being patient with me and unconditionally supporting me to do whatever I wished to do. This would not have been possible without their support and belief in me. I would like to thank my in-laws for praying for me through ups and downs during my PhD. I would also like to thank them for trusting me and allowing me to marry their beloved daughter Tenssy, during my PhD study. I thank them for taking care of my daughter Meghana and allowing me and my wife to progress in life. I am grateful to my wife and daughter for having patiently waited

and prayed for me. My daughter has been very kind and adapted to all the circumstances during my PhD. I thank all my family members who helped me by taking care of my daughter and letting me concentrate on my research.

I would like to thank Neeraj Goel for helping me both personally and professionally during my initial days. I also thank my seniors Aryabattha Sahu, Anant Vishnoi, Krishnaih and Silpa who helped me with their suggestions. I wish to thank my colleagues Arun Parakh, Sohan Sharma, Vaibhav Jain for their helpful inputs. I would like to thank Manish Jaiswal for helping me in hardware designing. I thank Gayathri, Rajshekar, Prathmesh, Sandeep Chandran, Eldose Peter who brought back joy in my life when things were not going so well. I would like to thank Dip Narayan Guha, Magesh and Mohan for helping me in my PhD admission process. If it were not for you guys, I would not have been admitted here. I also wish to thank my close friends Wayeez and Akarsh for being with me throughout my journey of life. I would like to thank other members of Symbiose team and my Indian friends in Rennes; Ajay Kattepur, Yogesh Karpate, Srinivas Pinnisetty, Raghu Ram, and Arulnambi for making my stay at Rennes fun and memorable. I thank my friends from Manipal; Aravind M. V., Ashwath Roa, Dharmendra and Gopal Karamore for encouraging me to join PhD.

Above all, I would like to thank God for giving me all these wonderful people and holding my hand in each and every step of my life.

**B. Sharat Chandra Varma**

# Abstract

*Bioinformatics applications take very long time to execute on general purpose computers and hence accelerators are used to reduce the overall execution time. FPGA based accelerators have proven to be a good candidate for accelerating these applications. Hardware implementation of the compute intensive kernels is done on FPGAs and multiple copies of such kernels are run in parallel to get speedups over software implementations. In the past, FPGAs had only programmable LUTs and programmable interconnects. In modern FPGAs, the digital circuitry common to many applications are being embedded as hard embedded blocks (HEBs) to use the silicon area effciently.*

*We developed a methodology for carrying out design space exploration of FPGA based accelerators for bioinformatics applications. We illustrate the use of our methodology with two important bioinformatics applications that take large amount of time when implemented in software: Protein docking and Genome assembly. We identified the kernels in the protein docking application and mapped it onto efficient hardware. We show that speedups of upto 12x can be achieved over software implementation. We studied genome assembly application and found that direct hardware implementation of the application is not possible due to FPGA resource constraints. We developed a novel method to pre-process the input and reduce the overall execution time. We show speedups of upto 8x can be achieved using accelerators implemented on FPGAs.*

*Further, our methodology was used to identify HEBs and evaluate them for accelerating both protein docking and* de novo *genome assembly. 3D-FFT is at the core of "FTDock", a popular implementation of protein docking application. "Butterfly" HEBs could accelerate 3D-FFT computations by over 1953x. Overall, specifically identified HEBs show an acceleration of upto 17x for the two bioinformatics applications. For speeding up the design space exploration, models at higher level of abstraction were used. Using these high level models, performance estimates for platforms with multi-FPGAs with different interconnect topologies was made possible.*

# Contents