Wibheda: Framework for Data Dependency-aware Multi-constrained Hardware-Software Partitioning in FPGA-based SoCs for IoT Devices

Deshya Wijesundera*, Alok Prakash†, Thilina Perera*, Kalindu Herath* and Thambipillai Srikanthan†
Nanyang Technological University, Singapore 639798
Email: *{deshyase001,pere0004,kalindub001}@e.ntu.edu.sg, †{alok,astsrikan}@ntu.edu.sg

***Abstract*—The increasing popularity of FPGA-based system-on-chip (SoC) devices for Internet of Things (IoT) applications calls for hardware-software partitioning solutions optimized for performance under stringent area and power constraints. In this work, we propose Wibheda, a heuristic based framework for data dependency-aware multi-constrained hardware-software partitioning at fine-granularity that can be employed to partition designs for FPGA-based SoCs used in IoT. Wibheda, evaluated on 6 applications from the popular CHStone benchmark suite has been shown to find solutions with 98.7% accuracy within several milliseconds compared to several minutes or hours in an existing state-of-the-art work and an exhaustive approach respectively.**

I. Introduction

Modern FPGAs are not only suited for accelerating critical parts of an application, but also for realizing an entire System-on-Chip (SoC), constituting processors, programmable logic, memory subsystems, etc. However, efficient partitioning of the application between the processor and programmable logic is crucial to exploit the benefits offered by both worlds. Partitioning decisions must typically be made early in the design of a product. However, finding the optimally partitioned solution is an NP-complete problem [1]. This has encouraged many researchers to explore heuristic based approaches.

The selection of granularity at which to partition an application poses another challenge in the partitioning process. A coarse-grained approach implements large sections of code where only a small fraction of the code may actually execute frequently enough to provide meaningful acceleration in execution time, resulting in reduced return on the investment of hardware area. A fine-grained approach provides much greater control on the accelerated code segments resulting in higher acceleration in execution time for the same hardware space, but leads to a much more complex design space exploration. This also, incurs higher data communication cost. Thus, accurate modelling of data communication costs is important in a fine-grained approach.

At the same time, the increasing popularity of Internet of Things (IoT) devices necessitates designs with extremely tight constraints in terms of size (area), power consumption, costs, etc. Thus, reconfigurable solutions provide a favorable design platform for such devices. This has resulted in FPGA vendors offering FPGAs specifically targeted for IoT [2] [3]. Such systems can benefit immensely from intelligent fine-grained acceleration to improve performance in highly resource constrained environments. This is also evident from Xilinx's initiative for industrial IoT solutions focusing on software programmability and hardware acceleration [4].

II. Proposed Framework

We propose Wibheda, a framework for rapid data dependency-, area- and power constraint- aware HW-SW partitioning at a fine-grained(basic block) level that can be applied to applications of varying size and complexity. The main contributions of this work are, a methodology for analysis of data communication cost between basic blocks and memory components and a scalable heuristic formulation to select the most profitable HW-SW partitioning considering (i) data communication cost of basic blocks and memory components, (ii) area constraints in terms of look-up-table (LUT), digital signal processing (DSP) block and flip-flop (FF) and (iii) power constraints.

III. Results and Discussion

The runtime of Wibheda is in the order of milliseconds while that of the state-of-the-art (SoA) [5], is in the order of minutes. Averaging across 6 applications from the CHStone benchmark suite Wibheda shows an average estimation error of only 1.3% in comparison to the SoA work which has an error of 16.5%. We also used 3 different (LUT, FF, DSP and power) constraints, each representing a latest FPGA device targetted for IoT applications [2] [3] to validate Wibheda in a system level design. The average difference in performance for the 6 applications across the 3 experiments is only 0.27%.

Acknowledgment

This research project is partially funded by the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme with the Technical University of Munich at TUM-CREATE.

References

[1] P. Arato *et al.*, "Hardware-software partitioning in embedded system design," in *ISISP '03*, 2003.
[2] Lattice Semiconductor, "iCE40 Ultra/UltraLite/UltraPlus-Lattice Semiconductor," http://bit.ly/2n9SEhR, 2017.
[3] A. Shilov, "Intel Announces Cyclone 10 FPGAs for IoT Devices," http://bit.ly/2m30SUX, 2017.
[4] Xilinx, "Industrial IoT Solutions Powered by Xilinx," http://bit.ly/2Dj04Dl, 2018.
[5] A. Prakash *et al.*, "Rapid Memory-Aware Selection of Hardware Accelerators in Programmable SoC Design," *TVLSI*, 2017.