Guest Editor's Introduction

Chip Design 2020

Jaydeep P. Kulkarni
University of Texas at Austin

With slowing down of the dimensional scaling in advanced CMOS technologies coupled with the rise of custom hardware accelerators for energy-efficient computing in a variety of AI/ML, AR/VR applications, integrated circuits (ICs) are transforming rapidly to meet the power, performance, security, and connectivity demands of such systems and architectures. This special issue of *IEEE Micro* aimed at publishing some of the most significant research that can highlight the trends in IC design in 2020 and provide directions for the future IC design era. The scope of the paper submissions was very broad to cover many aspects of modern IC designs, including but not limited to state-of-the-art chip design in 2020 and challenges in novel logic and memory devices, EDA methodologies, and neural network accelerators.

This special issue of *IEEE Micro* aimed at publishing some of the most significant research that can highlight the trends in IC design in 2020 and provide directions for the future IC design era.

This special issue contains three articles. In the first article, Ankit *et al.* from Purdue University present circuits and architectures for in-memory computing (IMC)-based machine learning accelerators. Machine learning applications, especially deep neural networks (DNNs) have seen ubiquitous use in computer vision, speech recognition, and robotics. However, the growing complexity of DNN models has necessitated efficient hardware implementations. The key compute primitives of DNNs are matrix–vector multiplications, which leads to significant data movement between memory and processing units in today's von Neumann systems. A promising alternative would be colocating memory and processing elements, which can be further extended to performing computations inside the memory itself. The authors believe IMC is a propitious candidate for future DNN accelerators since it mitigates the memory wall bottleneck. They discuss various IMC primitives in both CMOS and emerging nonvolatile memory technologies. Subsequently, they describe how such primitives can be incorporated in standalone machine learning accelerator architectures. Finally, the authors analyze the challenges associated with designing such IMC accelerators, and explore future opportunities.

In the second article, Khailany *et al.* from NVIDIA Corporation present emerging EDA methodologies for accelerating chip design with machine learning. Recent advancements in machine learning provide an opportunity to transform chip design workflows. The authors review recent research applying techniques such as deep convolutional neural networks and graph-based neural networks in the areas of automatic design space exploration, power analysis, VLSI physical design, and analog design. They also present a

Digital Object Identifier 10.1109/MM.2020.3028179
Date of current version 21 October 2020.

Published by the IEEE Computer Society

future vision of an AI-assisted automated chip design workflow to aid designer productivity and automate optimization tasks.

In the third article, Keshavarzi *et al.* present a perspective on ferroelectronics for edge intelligence (EI). The future data-centric world with its many new applications demands EI. Based on the requirements of the smart systems that are widely deployed at the edge and distributed in the field for EI, research vectors on multiple fronts of materials, devices, circuits, and architecture are discussed. Both top-down and bottom-up approaches are needed. Challenges with Moore's Law scaling and limitations of the current von Neumann computing architectures are limiting the performance and energy-efficiency of conventional electronics. On the other hand, new promising discoveries of advanced CMOS-compatible HfO_2-based ferroelectric devices open the door for ferroelectronics. Ferroelectric building blocks integrated on advanced CMOS technology nodes will enable much needed computing capabilities to make EI a reality. Pursuing processing capability by IMC in data-flow architectures is at the core of ferroelectronics. This will enable building 1000× more compute-energy efficient small-system AI engines that offer the performance needed for EI. The data-centric world will rely on "smart" IoT devices at the edge, "smart" being defined as the ability to analyze data locally and to autonomously decide on a course of action. These smart IoT devices need to be able to address a wide range of new applications that demand higher performance and more capabilities supporting local embedded intelligence, real-time learning, and autonomy. They will drive the next phase of growth in the semiconductor industry. A micro Drone (uDrone) is an example of a smart IoT device that requires diverse computing capabilities for perception, path planning, localization, mapping, and optimization along with cooperative intelligence in swarms, sensor fusion. Applications of augmented reality in a small form factor ("smart glass") that have similar computationally demanding tasks as uDrones are discussed in this article as well.

We hope that these articles provide the reader insights into promising research directions for future chip designs ranging from ferroelectric devices, circuits, and architectures for IMC as well as AI-assisted automated chip design methodology.

In conclusion, this *IEEE Micro* Special Issue on Chip Design 2020 highlights the growing importance of AI/ML accelerator designs in coming years. We hope that these articles provide the reader insights into promising research directions for future chip designs ranging from ferroelectric devices, circuits, and architectures for IMC as well as AI-assisted automated chip design methodology.

Jaydeep P. Kulkarni is an Assistant Professor with the University of Texas at Austin. Contact him at jaydeep@austin.utexas.edu.