

Panel: Coarse-Grained Reconfigurable Arrays and their Opportunities as Application Accelerators

Coarse-grained reconfigurable arrays (CGRAs) are programmable logic devices having large configurable logic blocks (often ALU-like blocks) and datapath-style (word-wide) interconnect. In comparison with FPGAs, the coarse-grained nature of CGRAs implies significantly less silicon area overhead dedicated to programmability. The value proposition of CGRAs is that they are closer to custom ASICs from the perspectives of power, performance and area (PPA), yet they retain some of the programmability of fine-grained FPGAs. While CGRAs emerged as a research topic about 25 years ago in the 1990s, they are garnering much attention today for two reasons: 1) The slowing of Moore's law and Dennard scaling has meant that standard processors are no longer getting faster, leading to today's "accelerator era" where customized hardware is used to raise application performance and power efficiency. CGRAs are one available means of realizing such compute accelerators. 2) The rise of, and increasing ubiquity of ML, as enabled by specialized hardware. Deployment of ML in edge or cloud environments requires bespoke hardware, and recent years have seen a large number of startup companies (e.g. Cerebras, Tenstorrent, Preferred Networks), as well as large multinational companies (e.g. Google, Baidu, Microsoft) developing custom hardware for ML inference and training. Such custom hardware closely resembles CGRAs: arrays of coarse-grained blocks (typically MAC-based blocks interspersed with programmable interconnect.

With the rising visibility of CGRA research, and the lack of a single CGRA-related forum, we believe it is a timely moment to have a panel discussion on CGRAs at IEEE ASAP 2021, gathering together some of the world's leading CGRA researchers to share their latest results and discuss future directions.

Moderator: Prof. Jason Anderson, University of Toronto, Canada

Panelists: Dr. Kentaro Sano, RIKEN, Japan

Prof. Shouyi Yin, Tsinghua University, China

Dr. Cheng Tan, Pacific Northwest National Lab (PNNL), USA

Prof. Hideharu Amano, Keio University, Japan

Prof. Masato Motomura, Tokyo Institute of Technology, Japan

Prof. Yasuhiko Nakashima, Nara Institute of Science and Technology (NAIST), Japan

Bio for panelists:

Dr. Kentaro Sano, RIKEN, Japan

Dr. Sano leads a group at RIKEN exploring the potential for CGRAs in the datacenter. He had a paper at ASAP 2020 on a template-based CGRA architecture with an emphasis on supporting different memory styles (<https://ieeexplore.ieee.org/document/9153262>) and also recently co-authored a survey paper on CGRAs (<https://arxiv.org/abs/2004.04509>).

Prof. Shouyi Yin, Tsinghua University, China

Prof. Yin has been researching CGRAs for years, and recently co-authored a survey paper in ACM Computing Surveys (<https://dl.acm.org/doi/abs/10.1145/3357375>). Prof. Yin's most recent work includes optimizations to improve CGRA memory bandwidth (<https://ieeexplore.ieee.org/document/9371661>).

Dr. Cheng Tan, Pacific Northwest National Lab (PNNL), USA

Dr. Tan is a CGRA researcher at PNNL and recently proposed the OpenCGRA framework -- an open-source framework for CGRA architecture exploration (<https://ieeexplore.ieee.org/abstract/document/9283606>).

Prof. Hideharu Amano, Keio University, Japan

Prof. Amano is a long-time CGRA researcher and has published many papers on the theme. Most recently, he has been working on low-power CGRAs, as well as on CGRA mapping algorithms (<https://ieeexplore.ieee.org/document/9149647>).

Prof. Masato Motomura, Tokyo Institute of Technology, Japan and**Prof. Yasuhiko Nakashima, Nara Institute of Science and Technology (NAIST), Japan**

Profs. Motomura and Nakashima are currently exploring CGRAs for ML acceleration. Prof. Motomura led the development of the NEC Dynamically Reconfigurable Processor (DRP) -- a commercial CGRA.

Prof. Jason Anderson, University of Toronto, Canada

Prof. Anderson's group proposed the open-source CGRA-ME (modelling and exploration) framework to facilitate research on CGRA architectures and design methodologies, with the initial paper on the framework appearing in ASAP 2017 (<https://ieeexplore.ieee.org/document/7995277>). Recent work includes architecture-agnostic approaches for CGRA mapping (<https://ieeexplore.ieee.org/document/8735511>).