



Special Interest Group on Design Automation
ACM/SIGDA E-NEWSLETTER, Vol. 52, No. 3

SIGDA - The Resource for EDA Professionals

This newsletter is a free service for current SIGDA members and is added automatically with a new SIGDA membership.
Circulation: 2,700

Online archive: <http://www.sigda.org/newsletter>

SIGDA News

(1) [U.S. Chops Russia's Access to Integrated Circuits](#)

The U.S. government is tightening controls on the semiconductor supply chain to restrict Russian access to chip technology used in military equipment, in response to Russia's invasion of Ukraine.

(2) [Google AI Boosts Plasma Control for Fusion Energy](#)

Google has applied AI in an effort to manage plasma within a nuclear fusion reactor. DeepMind Technologies, Google's British AI subsidiary, used its machine learning expertise to manage a tokamak, a circular nuclear fusion reactor, in partnership with EPFL's Swiss Plasma Center (École Polytechnique Fédérale de Lausanne). The findings, published in the journal Nature, may provide new avenues for developing fusion as a sustainable energy source.

(3) [Intel Will Rely on TSMC for its Rebound](#)

Intel is increasing its reliance on erstwhile rival Taiwan Semiconductor Manufacturing Co. (TSMC) in its attempt to boost sales and eventually regain dominance as the world leader in manufacturing scale and chip process technology.

(4) [Bosch Ups Investments in Reutlingen, Dresden Fabs](#)

Addressing persistent chip shortages, Bosch announced plans to invest an additional \$296 million to expand its semiconductor manufacturing capacity. The investment follows the German company's previously pledged investments aimed at current

Messages from the EiCs

Dear ACM/SIGDA members,

We are excited to present to you March E-Newsletter. We encourage you to invite your students and colleagues to be a part of the SIGDA newsletter. The newsletter covers a wide range of information from the upcoming conferences to technical news and activities of our community. Get involved and contact us if you want to contribute articles or announcements.

The newsletter is evolving. Please let us know what you think.

Happy reading!

Debjit Sinha, Keni Qiu,
Editors-in-Chief,
SIGDA E-News

expansion efforts as demand increases for its MEMS and other devices.

(5) [AI Accelerators Enter IoT SoCs](#)

Silicon Labs' latest families of wireless-enabled SoCs for IoT applications for the first time include a hardware AI/ML accelerator. The upgrade is indicative of the growing popularity of AI/ML techniques for a variety of IoT markets, including smart home, medical and industrial. Dedicated AI/ML hardware on-chip improves power consumption, critical to many IoT applications, even bringing AI/ML within reach for more power-sensitive IoT applications.

(6) [Taking a RISC: Expanding Chip Options](#)

After almost two years, the pandemic is still affecting many industries. The initial "stay home" impact of Covid hit air travel, hotels, entertainment, retail and other labor-intensive sectors such as manufacturing. Semiconductor manufacturing has been especially hard hit, leading to a well-publicized shortage of semiconductor chips, particularly in the automotive industry. A study from 2021 pegged the automotive industry's lost revenue from the chip shortage at \$210 billion.

What Is

What is Hardware Accelerator and Neural Network Co-Design?

Contributing authors: Yi Sheng¹, Lei Yang², Weiwen Jiang¹, Yiyu Shi³, and **Jingtong Hu**⁴

¹George Mason University, ²University of New Mexico, ³University of Notre Dame, ⁴University of Pittsburgh

Editors: Muhammad Shafique and Xun Jiao

Machine learning has been achieving great success for various tasks including image classification, image segmentation, and language modeling. However, most models are designed with the aim of maximizing software performance only (e.g., network accuracy) and do not take into consideration the hardware devices that will be finally used to implement the neural network. This will potentially

SIGDA EC

Yiran Chen,
Chair

Sudeep Pasricha,
Vice Chair and Conference
Chair

X. Sharon Hu,
Past chair

Yu Wang,
Award Chair

Wanli Chang,
Finance Chair

Yuan-Hao Chang,
Technical Activity Chair

Jingtong Hu,
Education Chair

Preeti Ranjan Panda,
Communication Chair

SIGDA

E-News

Editorial

Board

Debjit Sinha, co-EiC

Keni Qiu, co-EiC

Xiang Chen, AE for News

Yanzhi Wang,
AE for Local chapter news

lead to excessive latencies beyond specifications, rendering the resulting architectures useless. To address these issues, co-design is presented as a method to solve this problem.

On the software side, Neural Architecture Search (NAS) has become an important tool to design network architectures. It formulates the search space with the neural architecture hyperparameters, such as the number of channels, the kernel size, and what kind of neural operators to be used in each layer. On the hardware side, there are multiple computing platforms that can accelerate the neural networks, such as GPU, FPGA, Computing-in-Memory, and even Quantum computing. Target a computing platform, previous research efforts design dedicated accelerators for neural networks.

To introduce hardware into the AI system optimization loop, it seems straightforward to simply include an additional metric for the hardware in existing frameworks to describe the latency of a neural architecture on the target accelerator, such as FPGA. However, how to obtain such an evaluation metric is non-trivial. For one thing, the synthesis and implantation have a high overhead in time. What's worse, like the design flexibility of neural networks, the hardware accelerator also has design flexibilities, such as loop optimization settings.

The neural network and hardware accelerator co-design emerge to solve such a problem. It integrates the search space of hardware design, along with the search space of neural architecture, and an optimizer (e.g., meta-heuristic, reinforcement learning) can be applied to simultaneously explore these two coupled design spaces. FNAS, the first framework to support Hardware/Software Co-Exploration of Neural Architectures was proposed [1,2]. Works [3,4] propose tools for IoT and Autonomous systems. All these works target FPGA platforms. Followed by these works, a series of works targeting different computing platforms were proposed, including mobile platforms [5], Network-on-Chip [6], ASICs [7], and Computing-in-Memory [8]. It has also been extended to support different applications, such as medical images [9,10].

In summary, hardware and machine learning co-design are to find the optimal architecture having the maximized accuracy and satisfying the hardware requirements on target hardware devices.

References

[1] Jiang, Weiwen, Xinyi Zhang, Edwin H-M. Sha, Lei Yang, Qingfeng Zhuge, Yiyu Shi, and Jingtong Hu. "Accuracy vs. efficiency: Achieving both through fpga-implementation aware neural architecture search." In Proceedings of the 56th Annual Design Automation Conference 2019, pp. 1-6. 2019.

Pingqiang Zhou,
AE for Awards

Xun Jiao,
AE for What is

Muhammad Shafique,
AE for What is

**Rajsaktish
Sankaranarayanan,**
AE for Researcher spotlight

Xin Zhao,
AE for Paper submission

Ying Wang,
AE for Technical activities

Paper Deadlines

ISLPED'21 - ACM/IEEE Int'l Symposium on Low Power Electronics and Design

Boston University, Boston, MA
(Hybrid)

Deadline: March 18, 2022
(Abstract due: March 11, 2022)

Aug 1-3, 2022

<http://www.islped.org>

MDTS'22 - IEEE Microelectronics Design & Test Symposium

Virtual

Deadline: March 21, 2022

May 23-26, 2022

<http://natw.ieee.org>

ISVLSI'22 - IEEE Computer Society Annual Symposium on VLSI

Cyprus

Deadline: Mar 25, 2022

July 4-6, 2022

<http://www.ieee-isvlsi.org>

- [2] Jiang, Weiwen, Lei Yang, Edwin Hsing-Mean Sha, Qingfeng Zhuge, Shouzhen Gu, Sakyasingha Dasgupta, Yiyu Shi, and Jingtong Hu. "Hardware/software co-exploration of neural architectures." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 39, no. 12 (2020): 4805-4815.
- [3] Hao, Cong, Xiaofan Zhang, Yuhong Li, Sitao Huang, Jinjun Xiong, Kyle Rupnow, Wen-mei Hwu, and Deming Chen. "Fpga/dnn co-design: An efficient design methodology for 1ot intelligence on the edge." In *2019 56th ACM/IEEE Design Automation Conference (DAC)*, pp. 1-6. IEEE, 2019.
- [4] Hao, Cong, and Deming Chen. "Software/Hardware Co-design for Multi-modal Multi-task Learning in Autonomous Systems." In *2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, pp. 1-5. IEEE, 2021.
- [5] Niu, Wei, Zhenglun Kong, Geng Yuan, Weiwen Jiang, Jiexiong Guan, Caiwen Ding, Pu Zhao, Sijia Liu, Bin Ren, and Yanzhi Wang. "A Compression-Compilation Framework for On-mobile Real-time BERT Applications." *arXiv preprint arXiv:2106.00526* (2021).
- [6] Yang, Lei, Weiwen Jiang, Weichen Liu, H. M. Edwin, Yiyu Shi, and Jingtong Hu. "Co-exploring neural architecture and network-on-chip design for real-time artificial intelligence." In *2020 25th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 85-90. IEEE, 2020.
- [7] Yang, Lei, Zheyu Yan, Meng Li, Hyoukjun Kwon, Liangzhen Lai, Tushar Krishna, Vikas Chandra, Weiwen Jiang, and Yiyu Shi. "Co-exploration of neural architectures and heterogeneous asic accelerator designs targeting multiple tasks." In *2020 57th ACM/IEEE Design Automation Conference (DAC)*, pp. 1-6. IEEE, 2020.
- [8] Jiang, Weiwen, Qiuwen Lou, Zheyu Yan, Lei Yang, Jingtong Hu, Xiaobo Sharon Hu, and Yiyu Shi. "Device-circuit-architecture co-exploration for computing-in-memory neural accelerators." *IEEE Transactions on Computers* 70, no. 4 (2020): 595-605.
- [9] Yan, Xingang, Weiwen Jiang, Yiyu Shi, and Cheng Zhuo. "Ms-nas: Multi-scale neural architecture search for medical image segmentation." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 388-397. Springer, Cham, 2020.
- [10] Zeng, Dewen, Weiwen Jiang, Tianchen Wang, Xiaowei Xu, Haiyun Yuan, Meiping Huang, Jian Zhuang, Jingtong Hu, and Yiyu Shi. "Towards cardiac intervention assistance: hardware-aware neural architecture exploration for real-time 3D cardiac cine MRI segmentation." In *Proceedings of the 39th International Conference on Computer-Aided Design*, pp. 1-8. 2020.

IWLS'22 - International Workshop on Logic & Synthesis
Virtual conference
Deadline: April 18, 2022 (Abstracts due: April 11, 2022)
Jul 18-21, 2022
<https://www.iwls.org>

OSCAR'22 - First Workshop on Open-Source Computer Architecture Research
New York (co-located with ISCA 2022)
Abstract deadline: April 20, 2022
June 11, 2022
<https://oscar-workshop.github.io>

VLSI-SoC'22 - IFIP/IEEE Int'l Conference on Very Large Scale Integration
Patras, Greece
Deadline: April 25, 2022 (Abstracts due: April 18, 2022)
Oct 3-5, 2022
<http://www.vlsi-soc.com>

ICCAD'22 - IEEE/ACM Int'l Conference on Computer-Aided Design
Hybrid in-person and virtual conference
Deadline: May 23, 2022 (Abstracts due: May 16, 2022)
Oct 30 - Nov 3, 2022
<http://www.iccad.com>

BioCAS'22 - Biomedical Circuits and Systems Conference
Taipei, Taiwan
Deadline: June 10, 2022
Oct 13-15, 2022
<https://2022.ieee-biocas.org>

HiPC'22 - IEEE Int'l Conference on High Performance Computing, Data, And Analytics
Deadline: June 24, 2022 (Abstracts due: June 10, 2022)
Dec 18-21, 2022
<http://www.hipc.org>

Who's Who

Xiaoming Chen

Associate Professor
Institute of Computing Technology, Chinese Academy of Sciences

Email: chenxiaoming@ict.ac.cn
Personal webpage: <http://people.ucas.edu.cn/~chenxm>
Research interests: EDA and computer architecture

Who's Who AEs:

Americas:

Yi-Chung Chen,
Tennessee State University
Email: ychen@tnstate.edu
Webpage <https://yichungchen84.github.io/>

Asia and Oceania:

Bei-Yu,
The Chinese University of Hong Kong
Email: byu@cse.cuhk.edu.hk
Webpage <http://www.cse.cuhk.edu.hk/~byu/>

Europe and Africa:

Hussam Amrouch,
University of Stuttgart
Email: amrouch@iti.uni-stuttgart.de
Webpage:
<https://www.iti.uni-stuttgart.de/en/institute/team/Amrouch/>

Upcoming Conferences

DATE'22 - Design Automation and Test in Europe
Antwerp, Belgium, and online
Mar 14-23, 2022
<http://www.date-conference.com>

ISPD'22 – ACM Int'l Symposium on Physical Design
Banff, Alberta, Canada
Mar 27 - 30, 2022
<http://www.ispd.cc>

ISQED'22 - Int'l Symposium on Quality Electronic Design
California
April 6-8, 2022
<http://www.isqed.org>

RTAS'22 - IEEE Real-Time and Embedded Technology and Applications Symposium
Milano, Italy
May 4-6, 2022
<http://2022.rtas.org>

FCCM' 22 - IEEE International Symposium On Field-Programmable Custom Computing Machines
New York
May 15–18, 2022
<https://www.fccm.org/>

ISCAS'22 – IEEE Int'l Symposium on Circuits and Systems
Austin, TX
May 28 - June 1, 2022
<http://iscas2022.org>

**GLSVLSI'22 – ACM Great Lakes
Symposium on VLSI**

Orange County, CA

June 6-8, 2022

<http://www.glsvlsi.org>

**ISCA'22 – Int'l Symposium on
Computer Architecture**

New York City, USA

June 11-15, 2022

<https://iscaconf.org/isca2022>

**HiPEAC'22: Int'l Conference on
High Performance Embedded
Architectures & Compilers**

Budapest, Hungary

June 20-22, 2022

<https://www.hipeac.net/2022/budapest>

**HOST'22 – IEEE Int'l Symposium
on Hardware-Oriented Security
and Trust**

Washington DC

Deadline: May 11, 2021 (Abstracts
due: Apr 27, 2021)

June 27-30, 2022

<http://www.hostsymposium.org>

**DAC'22 – Design Automation
Conference**

San Francisco, CA

July 10-14, 2022

<http://www.dac.com/>

Notice to authors

By submitting your article for distribution in this Special Interest Group publication, you hereby grant to ACM the following non-exclusive, perpetual, worldwide rights: to publish in print on condition of acceptance by the editor; to digitize and post your article in the electronic version of this publication; to include the article in the ACM Digital Library and in any Digital Library related services; and to allow users to make a personal copy of the article for noncommercial, educational or research purposes. However, as a contributing author, you retain copyright to your article and ACM will refer requests for republication directly to you.

This ACM/SIGDA E-NEWSLETTER is being sent to all persons on the ACM/SIGDA mailing list. To unsubscribe, send an email to listserv@listserv.acm.org with "signoff sigda-announce" (no quotes) in the body of the message. Please make sure to send your request from the same email as the one by which you are subscribed to the list.

To renew your ACM SIGDA membership, please visit <http://www.acm.org/renew> or call between the hours of 8:30am to 4:30pm EST at +1-212-626-0500 (Global), or 1-800-342-6626 (US and Canada). For any questions, contact acmhelp@acm.org.