FROM THE EDITOR-IN-CHIEF

Environmentally Sustainable Computing

Lizy Kurian John, *The University of Texas at Austin, Austin, TX, 78712, USA*

With mobile and wearable computing, with electronic devices sensing and controlling many everyday activities, with connectivity becoming pervasive, with cloud and data centers, with pervasive smartphones, and with computing augmented with Internet of Things, a significant carbon footprint is generated not only by the usage of the devices but also by their manufacturing. With consumers upgrading devices such as cellphones every year, manufacturing carbon is becoming significant in comparison to the carbon footprint of daily usage. Carole-Jean Wu notes in her guest editorial in this issue: "A recent study found that, in 2019, manufacturing carbon contributed to roughly half of cloud computing's carbon footprint, whereas manufacturing carbon accounted for over 70% of consumer electronics." While progress in computing has enhanced everyday human lives, it is our responsibility to make sure that computing stays environmentally responsible.

Taking on the sustainability challenge, *IEEE Micro* presents in the first issue of 2023 several articles relevant to environmentally sustainable computing. Carole-Jean Wu of Meta AI guest edited the Special Issue on Environmentally Sustainable Computing that yielded three articles. Several interesting topics appear in these articles. Peruse through the pages and you will see papers on understanding the environmental impact of computer systems, on the sustainability design space of hardware, and on the rising environmental impact of carbon emissions from computing on autonomous vehicles. The article on design space of hardware posits that graphics processing units may be less sustainable than emerging in-memory processing technologies for artificial intelligence processing at the edge. Read the guest editorial written by Carole-Jean Wu to understand more about the community undertakings on environmentally sustainable computing and specifically about the three chosen articles.

In addition, we also selected a few articles from the 2022 Cool Chips conference, a conference focusing on low power designs. The Cool Chips conference has been taking place in Japan every year since 1998. Several industry and academic researchers present their work on power- and energy-efficient computing at the Cool Chips conference annually, and we have included a few selected articles from Cool Chips 2022 here in this issue. These articles very appropriately enrich the environmental sustainability theme. Ryusuke Egawa of Tokyo Denki University and Yasutaka Wada of Meisei University, Japan guest edited the Special Issue on Cool Chips. Out of the 12 presentations at the Cool Chips conference, they selected a few for submitting written articles to *IEEE Micro*. After a review process, two articles were selected for inclusion in this issue. In the first article, "TCN-CUTIE: A 1,036-TOp/s/W, 2.72-μJ/Inference, 12.2-mW All-Digital Ternary Accelerator in 22-nm FDX Technology," Scherer el al. describe a flexible, fully digital ternary neural network accelerator in a RISC-V-based system-on-chip. In the second article, "A Scalable Body Bias Optimization Method Toward Low-Power CGRAs," Kojima et al. explore fine body bias control considering the power and area overhead of an on-chip body bias generator and present design options that save 66% of energy consumption. Read the guest editorial written by Egawa and Wada, which provides a short description of the two articles selected for the Special Issue on Cool Chips.

These articles are followed by two papers on contemporary commercial products. The International Symposium on Computer Architecture (ISCA) has started the tradition of including a special industry session, and *IEEE Micro* selected two articles from the submissions for inclusion in this Issue. Lisa Hsu of Microsoft, who was the program chair for the ISCA 2022 industry session, coordinated the Special Issue on Commercial Products.

In the first commercial products article, "Enterprise-Class Multilevel Cache Design: Low Latency, Huge Capacity, and High Reliability," Berger et al. from IBM describe the innovative cache architecture of the IBM z series Telum Processor. The article focuses on the innovations to provide low latency, large capacity, and reliable L2 caching, in this 8 core processor with semiprivate L2 caches. The L2 caches use a novel

0272-1732 © 2023 IEEE
Digital Object Identifier 10.1109/MM.2022.3229504
Date of current version 13 January 2023.

horizontal cache persistence algorithm, and also serve as systemwide L3 and L4 caches to provide high performance for enterprise applications, yielding 11% per-core performance improvement over the prior z15 hardware, on real-world enterprise applications.

In the second commercial products article, "Data Movement Accelerator Engines on a Prototype Power10 Processor," Sugawara et al. of IBM present the active messaging engine (AME), for offloading operations related to data movement from the main central processing unit. By supporting offloading of general messaging libraries for processing, sending, and receiving on-node and off-node messages, AME enables truly asynchronous progress to overlap computation and communication of multicores. AME is a tiny but fully programmable 64-bit processing engine, optimized for communication. AME design goals include small size and low power, with a purpose-built instruction set, but relatively low instructions per cycle.

WHILE PROGRESS IN COMPUTING HAS ENHANCED EVERYDAY HUMAN LIVES, IT IS OUR RESPONSIBILITY TO MAKE SURE THAT COMPUTING STAYS ENVIRONMENTALLY RESPONSIBLE.

This issue also contains a general interest feature article, "RDMA Congestion Control: It Is Only for the Compliant," by Snyder et al. of Duke University. Remote direct memory access (RDMA) enables offlocating of networking tasks to hardware and allows applications to communicate without invoking system software on the data path. The authors describe how RDMA generally enables low latency and low CPU utilization, but points to performance isolation concerns in a shared cloud environment. For example, congestion control algorithms can render the network vulnerable to performance hacking attacks, which give the attacker extra bandwidth in the network. The authors illustrate the need for incorporating performance isolation into the design and evaluation of congestion control algorithms.

In addition to the aforementioned eight papers, this Issue features an article from our Micro Law columnist Joshua Yi, titled "Does Academic Research Drive Industrial Innovation in Computer Architecture?—Analyzing Citations to Academic Papers in Patents." Yi examines the number of citations to articles published in top computer architecture conferences, such as Architectural Support for Programming Languages and Operating System (ASPLOS), High Performance Computer Architecture, ISCA, and International Symposium on Microarchitecture (MICRO). Read the entire article to find interesting tidbits, for example, the fact that even though ISCA started after MICRO, ISCA papers had 130 patent citations to conferences held between 1973 and 1990 while MICRO only had four citations to proceedings during that timeframe. Other observations include that one paper in the first ISCA (ISCA 1973), "DAP—A Distributed Array Processor" authored by Stewart F. Reddaway, was cited 51 times. Also noted is that an ASPLOS 2000 paper "OceanStore: An Architecture for Global-Scale Persistent Storage" by John Kubiatowicz et al. was cited 1,070 times.

In addition, this issue also features a Micro Economics column by Shane Greenstein of Harvard Business School titled, "The Modern Digital Operating Model." Electronic commerce relies on understanding customer arrival patterns, their browsing behavior, and the path to a purchase. Measurement of such data helps decision making from the business side. Greenstein describes the types of experimentation and learning that businesses conduct in order to understand and design the path to commercial success. Digital record keeping, electronic communications, analytical methods, experimentation, measurement, and iterative improvement are all pieces of the modern digital enterprise, and Greenstein discusses the crucial elements of the modern digital operating model, such as what should go into a checklist.

I take this opportunity to express my gratitude to our guest editors Drs. Wu, Egawa, Wada, and Hsu for proposing and editing their special issues. I am sure that readers will enjoy these articles and will benefit from them.

More than 50 cities in China, the United States, Philippines, South Korea, Canada, Spain, Italy, Germany, and the United Kingdom boast of the availability of 5G communications in 2022. Major U.S. carriers, such as AT&T, T-Mobile, and Verizon, have all launched versions of 5G broadly enough to tout the claim of nationwide service in the United States. Data and videos are exchanged and communicated at high speeds; more and more computing can be done at fingertips wherever you are. This certainly is a time to wonder—Are our computing trends environmentally sustainable? The awareness created by this special issue hopefully leads to action on solving

environmental and sustainability problems while at the same time keeping economic growth. I hope this special issue is thought provoking for our readers and leads to more work on sustainable computing.

THE AWARENESS CREATED BY THIS SPECIAL ISSUE HOPEFULLY LEADS TO ACTION ON SOLVING ENVIRONMENTAL AND SUSTAINABILITY PROBLEMS WHILE AT THE SAME TIME KEEPING ECONOMIC GROWTH.

IEEE Micro is interested in submissions on any aspect of chip/system design or architecture. Submit articles to regular issues and special issues for 2023. Stay tuned for the Call for Papers of an upcoming special issue on TinyML.

This is my last year as Editor-in-Chief of *IEEE Micro*. IEEE is seeking applications and nominations for the new Editor-in-Chief. Apply or nominate colleagues for these important positions: https://www.computer.org/press-room/2022-news/ieee-computer-society-publications-seek-applications-for-2024-editors-in-chief

Happy New Year! May the New Year 2023 bring you glad tidings and interesting readings from *IEEE Micro*! I wish all readers a very happy 2023 both professionally and personally.

LIZY KURIAN JOHN is a Truchard Foundation chair and professor in the Electrical and Computer Engineering Department, University of Texas at Austin, Austin, TX, 78712, USA. She is a Fellow of IEEE, ACM, and the National Academy of Inventors. Contact her at ljohn@ece.utexas.edu.