Special Issue on Top Picks From the 2022 Computer Architecture Conferences

Christopher Batten, *Cornell University, Ithaca, NY, 14853, USA*
Jae W. Lee, *Seoul National University, Seoul, 08826, South Korea*

It is our pleasure to introduce the *IEEE Micro* Special Issue on Top Picks from the 2022 Computer Architecture Conferences. This special issue represents some of the very best recent research in the field of computer architecture in terms of novelty and potential for long-term impact. The 12 articles in this special issue were selected by a committee of experts from both industry and academia using a rigorous selection process over a period of three months. The committee also selected an additional 12 articles to be recognized with an honorable mention (see "Honorable Mentions").

SELECTION PROCESS

This year's selection committee (SC) consisted of 46 experts in computer architecture. The committee included 20% women, 28% junior researchers (assistant professors or Ph.D. recipients within the past seven years), 22% from industry, 22% from outside the United States, and 43% having not served on the committee in the past seven years (see "Selection Committee").

Similar to last year, we expanded the scope of a "computer architecture conference" to include both traditional conferences in the field as well as other top conferences in related areas. This recent change reflects the interdisciplinary nature of modern computer architecture research. We received a record-number 130 submissions, including nine submissions for papers from conferences in related areas. Authors submitted a three-page extended abstract, which included a two-page summary of the article and a one-page argument for its potential long-term impact. One paper was desk rejected due to noncompliance with the formatting requirements.

As in previous years, we adopted a two-round review process. In the first round, four reviews were assigned to each submission. After the review deadline, an online discussion period was started to determine which papers had at least one reviewer who felt the paper was worthy of either a top pick or an honorable mention selection. Sixty-seven submissions advanced to the second round and received four additional reviews. At the end of the second review period, a discussion lead was assigned to each submission to manage online discussion and bring the reviewers to a consensus on which papers to discuss in the SC meeting, and how to rank these papers to form a discussion order. Ultimately, 43 papers were selected as candidates for discussion at the SC meeting.

The SC meeting was held online using Zoom in January 2023. We utilized Zoom's breakout room features to carefully manage conflicts of interest. As there is a hard limit of 12 papers that can be selected as top picks, the discussion order was critically important. The papers to be discussed were first sorted by the reviewers before the meeting into three categories: *top pick, top pick or honorable mention*, and *maybe*. Within each category, we utilized both a coarse-grained merit score and fine-grained rankings by asking each reviewer to rank each paper relative to the other papers in his or her pile. This ranking mechanism was reintroduced this year to produce a discussion order that more accurately reflected the preferences of the SC.

We took the following structured approach to discussion and voting during the SC meeting. 1) The discussion lead provided a brief summary of the paper. 2) A champion among the reviewers argued for selection as either a top pick or an honorable mention, followed by the other reviewers contributing their thoughts. 3) The paper was then opened for discussion by the rest of the SC. 4) The reviewers then voted via voice over Zoom to select either yes/top pick, yes/honorable mention, or no/not selected. 5) Finally, an SC-wide vote was used sparingly if neither the yes nor the no votes reached a supermajority (two thirds of the votes) or there were fewer than five reviewers present. We also had a process for an honorable mention paper to be upgraded to a top pick paper at specific points during the SC meeting if there were remaining top pick slots. However, there was no such upgrade request this year.

The SC meeting ended with the selection of 12 submissions as top picks and another 12 submissions as

0272-1732 © 2023 IEEE
Digital Object Identifier 10.1109/MM.2023.3278069
Date of current version 29 June 2023.

honorable mentions. Among the 12 top picks papers, four appeared at the International Symposium on Computer Architecture (ISCA), three at the International Symposium on Microarchitecture (MICRO), two at the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), and two at the IEEE International Symposium on High-Performance Computer Architecture (HPCA). Finally, one paper was selected from the USENIX Security Symposium.

SELECTED PAPERS

Security

Hardware security continues to be a critical concern, with researchers exploring both new security vulnerabilities as well as new security defense mechanisms. "PACMAN: Attacking ARM Pointer Authentication With Speculative Execution"[A1] is an article from the Massachusetts Institute of Technology that was originally published in ISCA. This article proposes a novel attack to bypass pointer authentication by using speculative execution to avoid program crashes when probing signed pointers. This work illustrates the potential of closely studying the interaction between speculative execution attacks and memory corruption vulnerabilities in future hardware security research. "Hertzbleed: Turning Power Side-Channel Attacks Into Remote Timing Attacks on x86"[A2] is an article from The University of Texas at Austin, University of Illinois at Urbana-Champaign, and the University of Washington that was originally published at the USENIX Security Symposium. This article demonstrates how traditional power side-channel attacks, which require physical proximity to the target system or access to low-level software power management interfaces, can be transformed into timing side-channel attacks by exploiting the fact that dynamic voltage and frequency scaling adjustments can depend on secret data. This new approach enables completely remote power side-channel attacks, even against constant-time cryptography and will likely inspire a new line of attacks and defense mechanisms.

The article "There's Always a Bigger Fish: A Clarifying Analysis of a Machine-Learning-Assisted Side-Channel Attack,"[A3] is from the Massachusetts Institute of Technology and was originally published at ISCA. This article revisits recent work using machine learning (ML) to discover side-channel attacks in complex modern systems. The article presents a comprehensive analysis to demonstrate that a previously ML-derived attack is actually due to a very different side channel than previously thought. This work illustrates the need for thoroughly analyzing ML-derived security attacks to truly understand the underlying vulnerability. "Revizor: Testing Black-Box CPUs Against Speculation Contracts"[A4] is an article from Microsoft, Technische Universitat Dresden, and Technion that was originally published at ASPLOS. This article explores using speculation contracts to model expected information leaks, and then uses random testing to detect unexpected information leaks that could suggest security vulnerabilities. This is one of the first techniques capable of testing realistic microarchitectures for speculation contract compliance.

Memory Systems

Our community is making progress on improving memory systems both by revisiting classic ideas and exploring the unique challenges of modern systems. The "Effective Mimicry of Bélády's MIN Policy"[A5] article is from the University of Texas at Austin and was originally published at HPCA. This article advances the state of the art in the well-studied area of cache-replacement policies with a simple yet effective approach to emulating Bélády's MIN policy. The proposed technique exclusively uses a more accurate reuse time predictor for all eviction decisions and is thus able to approach the performance of Bélády's impractical MIN policy with modest hardware overheads. "Revisiting Residue Codes for Modern Memories"[A6] is an article from Columbia University that was originally published at MICRO. This article creatively applies residue codes, which are traditionally used for reliable computation, to the problem of error correction in modern memory systems. The work explores how to overcome the challenges associated with this novel approach as well as the potential benefits in terms of storage overhead. The "HeteroGen: Automatic Synthesis of Heterogeneous Cache Coherence Protocols"[A7] article is from the University of Edinburgh and Duke University and was originally published in HPCA. This article seeks to reduce the significant complexity associated with heterogeneous cache-coherent memory systems by developing a tool to automatically generate heterogeneous cache-coherence protocols that satisfy a precisely defined compound consistency model.

Data Center and Cloud Computing

Data center and cloud computing continue to offer new challenges and opportunities when deploying workloads at scale. The "Online Code Layout Optimizations via OCOLOS"[A8] article is from the University of Pennsylvania, University of Michigan, Intel, and the University of California, Santa Cruz and was originally published in MICRO. This article discusses a new framework that combines profile-guided optimization

HONORABLE MENTIONS

"SupermarQ: A Scalable Quantum Benchmark Suite," by Teague Tomesh, Pranav Gokhale, Victory Omole, Gokul Subramanian Ravi, Kaitlin N. Smith, Joshua Viszlai, Xin-Chuan Wu, Nikos Hardavellas, Margaret R. Martonosi, and Frederic T. Chong (HPCA 2022).
"FlexDriver: A Network Driver for Your Accelerator," by Haggai Eran, Maxim Fudim, Gabi Malka, Gal Shalom, Noam Cohen, Amit Hermony, Dotan Levi, Liran Liss, and Mark Silberstein (ASPLOS 2022).
"Taurus: A Data Plane Architecture for Per-Packet ML," by Tushar Swamy, Alexander Rucker, Muhammad Shahbaz, Ishan Gaur, and Kunle Olukotun (ASPLOS 2022).
"A Software-Defined Tensor Streaming Multiprocessor for Large-Scale Machine Learning," by Dennis Abts, Garrin Kimmell, Andrew Ling, John Kim, Matthew Boyd, Andrew Bitar, Sahil Parmar, Ibrahim Ahmed, Roberto DiCecco, David Han, John Thompson, Michael Bye, Jennifer Hwang, Jeremy Fowers, Peter Lillian, Ashwin Murthy, Elyas Mehtabuddin, Chetan Tekur, Thomas Sohmers, Kris Kang, Stephen Maresh, and Jonathan Ross (ISCA 2022).
"EDAM: Edit Distance Tolerant Approximate Matching Content Addressable Memory," by Robert Hanhan, Esteban Garzón, Zuher Jahshan, Adam Teman, Marco Lanuzza, and Leonid Yavits (ISCA 2022).
"Mixed-Proxy Extensions for the NVIDIA PTX Memory Consistency Model," by Daniel Lustig, Simon Cooksey, and Olivier Giroux (ISCA 2022).
"Tiny but Mighty: Designing and Realizing Scalable Latency Tolerance for Manycore SoCs," by Marcelo Orenes-Vera, Aninda Manocha, Jonathan Balkind, Fei Gao, Juan L. Aragón, David Wentzlaff, Margaret Martonosi (ISCA 2022).
"Lukewarm Serverless Functions: Characterization and Optimization," by David Schall, Artemiy Margaritov, Dmitrii Ustiugov, Andreas Sandberg, and Boris Grot (ISCA 2022).
"SNS's Not a Synthesizer: A Deep-Learning-Based Synthesis Predictor," by Ceyu Xu, Chris Kjellqvist, and Lisa Wu Wills (ISCA 2022).
"ANT: Exploiting Adaptive Numerical Data Type for Low-Bit Deep Neural Network Quantization," by Cong Guo, Chen Zhang, Jingwen Leng, Zihan Liu, Fan Yang, Yunxin Liu, Minyi Guo, and Yuhao Zhu (MICRO 2022).
"Automatic Domain-Specific SoC Design for Autonomous Unmanned Aerial Vehicles," by Srivatsan Krishnan, Zishen Wan, Kshitij Bhardwaj, Paul Whatmough, Aleksandra Faust, Sabrina Neuman, Gu-Yeon Wei, David Brooks, and Vijay Janapa Reddi (MICRO 2022).
"Datamime: Generating Representative Benchmarks by Automatically Synthesizing Datasets," by Hyun Ryong Lee and Daniel Sánchez (MICRO 2022).

with online layout transformations in the context of data center workloads using unmanaged languages. This work can potentially open the door for a variety of dynamic/continuous optimizations on unmanaged languages that were previously only possible for managed languages. "IOCost: Block Input–Output Control for Containers in Datacenters"[A9] is an article from Meta and Carnegie Mellon University that was originally published in ASPLOS. This article introduces a new, general mechanism to quantify, control, and divide input–output (IO) resources for containers in the data center. This work is already being adopted by multiple cloud providers and will likely inspire a new line of research in IO resource isolation.

Building Real Systems

Building real computer systems enables exploration of the software and hardware implications of a complete end-to-end application and also enables discovery of new opportunities to improve the system design process. "EyeCoD: Eye Tracking System Acceleration via FlatCam-Based Algorithm and Hardware Co-Design"[A10] is an article from the Georgia Institute of Technology, Rice University, and Meta that was originally published in ISCA. This article describes a prototype system for front-end eye tracking in augmented-/virtual-reality applications and includes creative co-design of a lensless camera, region-of-interest prediction algorithm, and dynamically scheduled hardware accelerator. A chip

SELECTION COMMITTEE

- Jung Ho Ahn, Seoul National University
- Murali Annavaram, University of Southern California
- Rajeev Balasubramonian, University of Utah
- Nathan Beckmann, Carnegie Mellon University
- Abhishek Bhattacharjee, Yale University
- David Brooks, Harvard University
- Reetuparna Das, University of Michigan
- Yongshan Ding, Yale University
- Mattan Erez, University of Texas at Austin
- Babak Falsafi, École polytechnique fédérale de Lausanne
- Mingyu Gao, Tsinghua University
- Saugata Ghose, University of Illinois at Urbana-Champaign
- Antonio Gonzalez, Universitat Politécnica de Catalunya
- Tae Jun Ham, Google
- Yipeng Huang, Rutgers University
- Koji Inoue, Kyushu University
- Aamer Jaleel, Nvidia
- Nuwan Jayasena, AMD
- Timothy Jones, University of Cambridge
- John Kim, Korea Advanced Institute of Science and Technology
- Martha Kim, Columbia University
- Nam Sung Kim, University of Illinois at Urbana-Champaign
- Tushar Krishna, Georgia Institute of Technology
- Jane Li, University of Pennsylvania
- Celine Lin, Georgia Institute of Technology
- Heiner Litz, University of California, Santa Cruz
- Brandon Lucia, Carnegie Mellon University
- Martin Maas, Google
- Onur Mutlu, ETH Zürich
- Tony Nowatzki, University of California, Los Angeles
- Gennady Pekhimenko, University of Toronto
- Michael Pellauer, Nvidia
- Dmitry Ponomarev, Binghamton University
- Moinuddin Qureshi, Georgia Institute of Technology
- Parthasarathy Ranganathan, Google
- Minsoo Rhu, Korea Advanced Institute of Science and Technology/Meta
- Andre Seznec, Intel
- Viji Srinivasan, IBM
- Steven Swanson, University of California, San Diego
- Caroline Trippel, Stanford University
- Paul Whatmough, ARM
- Lisa Wu Wills, Duke University
- Yuan Xie, University of California, Santa Barbara.
- Mengjia Yan, Massachusetts Institute of Technology
- Jun Yang, University of Pittsburgh
- Cliff Young, Google

tape-out of a small-scale accelerator is used to validate cycle-level models of a complete eye-tracking system. "Toward Developing High-Performance RISC-V Processors Using Agile Methodology"[A11] is an article from the Institute of Computing Technology of the Chinese Academy of Sciences that was originally published in MICRO. This article describes a suite of new testing, verification, and debugging tools that address the weaknesses of open source agile hardware development tool flows. The potential for these tools is demonstrated through multiple tape-outs of a high-performance, out-of-order superscalar RISC-V processor.

Sustainability

There is an emerging subfield within computer architecture that explores sustainability as a key metric for computer system design. "Architectural CO_2 Footprint Tool: Designing Sustainable Computer Systems With an Architectural Carbon Modeling Tool"[A12] is an article from Harvard and Meta that was originally published in ISCA. This article describes a new tool that helps computer architecture researchers quantitatively evaluate the carbon footprint of various design choices. This tool is then used in several case studies to illustrate how optimizing for carbon can potentially result in different solutions compared to optimizing for performance and efficiency.

ACKNOWLEDGMENTS

We thank Lizy Kurian John, *IEEE Micro* editor-in-chief, for inviting us to serve as SC cochairs and guest editors for the Special Issue on Top Picks from the 2022

Computer Architecture Conferences. We are grateful for her invitation, guidance, and support throughout the process. We also thank the SC members for their hard work in selecting the best research in computer architecture in 2022, especially for working through the winter holidays. Sudhanva Gurumurthi and Radu Teodorescu shared their valuable experience as the SC cochairs and guest editors for the Special Issue on Top Picks last year. We greatly appreciate their insights and guidance as well as the materials and Zoom scripts they provided for running the online SC meeting. We would also like to thank Daniel Jiménez and Hyesoon Kim for their advice based on their experiences chairing Top Picks for 2020 and 2019 Conferences, respectively. Ulya Karpuzcu kindly offered to provide a few last-minute reviews for a committee member who had a personal emergency; we are grateful for her assistance. We would like to give special thanks to Peitian Pan, a doctoral student at Cornell, and Yeonhong Park and Seong Hoon Seo, doctoral students at Seoul National University, who were tremendously helpful in running HotCRP and Zoom during the SC meeting. Finally, we would like to thank all the authors who submitted papers to Top Picks this year. We greatly appreciate their excellent contributions to the field of computer architecture.

APPENDIX: RELATED ARTICLES

A1. J. Ravichandran, W. T. Na, J. Lang, and M. Yan, "PACMAN: Attacking ARM pointer authentication with speculative execution," *IEEE Micro*, vol. 43, no. 4, pp. 11–18, Jul./Aug. 2023, doi: 10.1109/MM.2023.3273189.

A2. Y. Wang, R. Paccagnella, E. T. He, H. Shacham, C. W. Fletcher, and D. Kohlbrenner, "Hertzbleed: Turning power side-channel attacks into remote timing attacks on x86," *IEEE Micro*, vol. 43, no. 4, pp. 19–27, Jul./Aug. 2023, doi: 10.1109/MM.2023.3274619.

A3. J. Cook, J. Drean, J. Behrens, and M. Yan, "There's always a bigger fish: A clarifying analysis of a machine-learning-assisted side-channel attack," *IEEE Micro*, vol. 43, no. 4, pp. 28–36, Jul./Aug. 2023, doi: 10.1109/MM.2023.3273457.

A4. O. Oleksenko, C. Fetzer, B. Köpf, and M. Silberstein, "Revizor: Testing black-box CPUs against speculation contracts," *IEEE Micro*, vol. 43, no. 4, pp. 37–44, Jul./Aug. 2023, doi: 10.1109/MM.2023.3273009.

A5. I. Shah, A. Jain, and C. Lin, "Effective mimicry of Bélády's MIN policy," *IEEE Micro*, vol. 43, no. 4, pp. 45–52, Jul./Aug. 2023, doi: 10.1109/MM.2023.3275079.

A6. E. Manzhosov, A. Hastings, M. Pancholi, R. Piersma, M. T. I. Ziad, and S. Sethumadhavan, "Revisiting residue codes for modern memories," *IEEE Micro*, vol. 43, no. 4, pp. 53–61, Jul./Aug. 2023, doi: 10.1109/MM.2023.3273489.

A7. N. Oswald, V. Nagarajan, D. J. Sorin, V. Gavrielatos, T. X. Olausson, and R. Carr, "HeteroGen: Automatic synthesis of heterogeneous cache coherence protocols," *IEEE Micro*, vol. 43, no. 4, pp. 62–70, Jul./Aug. 2023, doi: 10.1109/MM.2023.3274993.

A8. Y. Zhang, T. A. Khan, G. Pokam, B. Kasikci, H. Litz, and J. Devietti, "Online code layout optimizations via OCOLOS," *IEEE Micro*, vol. 43, no. 4, pp. 71–79, Jul./Aug. 2023, doi: 10.1109/MM.2023.3274758.

A9. T. Heo et al., "IOCost: Block input–output control for containers in datacenters," *IEEE Micro*, vol. 43, no. 4, pp. 80–87, Jul./Aug. 2023, doi: 10.1109/MM.2023.3277783.

A10. H. You et al., "EyeCoD: Eye tracking system acceleration via FlatCam-based algorithm and hardware co-design," *IEEE Micro*, vol. 43, no. 4, pp. 88–97, Jul./Aug. 2023, doi: 10.1109/MM.2023.3274736.

A11. Y. Xu et al., "Toward developing high-performance RISC-V processors using agile methodology," *IEEE Micro*, vol. 43, no. 4, pp. 98–106, Jul./Aug. 2023, doi: 10.1109/MM.2023.3273562.

A12. U. Gupta et al., "Architectural CO_2 footprint tool: Designing sustainable computer systems with an architectural carbon modeling tool," *IEEE Micro*, vol. 43, no. 4, pp. 107–117, Jul./Aug. 2023, doi: 10.1109/MM.2023.3275139.

CHRISTOPHER BATTEN is a professor of electrical and computer engineering at Cornell University, Ithaca, NY, 14853, USA. His research focuses on the intersection of computer architecture, electronic design automation, and digital VLSI. Batten received his Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology. He is a Member of IEEE. Contact him at cbatten@cornell.edu.

JAE W. LEE is a professor of computer science and engineering at Seoul National University, Seoul, 08826, South Korea. His research interests include computer architecture and systems, with a recent emphasis on algorithm-hardware co-design for machine learning and data analytics. Lee received his Ph.D. degree in computer science from the Massachusetts Institute of Technology. He is a Senior Member of IEEE. Contact him at jaewlee@snu.ac.kr.