Shail Dave

sites.google.com/view/shail

RESEARCH INTERESTS

- **Design Automation:** Design space formulation and exploration; Hardware/software codesign; ML-based automation.
- **Computer Architecture:** Hardware accelerators; Execution modeling; Near-data processing; Microarchitecture.
- **Embedded Systems:** Resource-efficient execution; TinyML; Compiler, runtime for distributed, time-sensitive apps.

Envisioned Research & Development Objective: "Enable Agile, Sustainable, and Learning-Assisted Automated Exploration of Efficient Domain-Specific Accelerator Systems through Research, Development, and Demonstration at Scale" [Brief Position Paper] [Short Talk]
(Please reach out to me for any questions/details about the vision and prospectus, as some of my novel directions has led to some works-in-progress and new project proposals, some of which are not publicly available.)

EDUCATION

Ph.D. in Computer Engineering (Computer Systems) **2017–Present**
Anticipated graduation: Spring 2023 (Available for research internship in Spring 2023) CGPA: 4.00/4.00
Ira A. Fulton School of Engineering, Arizona State University, Tempe, Arizona
Dissertation: Agile and Sustainable Methodology for Designing Efficient Accelerators.
Advisor: Prof. Aviral Shrivastava, Graduate Program Chair, Computer Science and Engineering.
Committee Members: Prof. Aviral Shrivastava, Prof. Tony Nowatzki, Prof. Jae-sun Seo, Prof. Fengbo Ren, and Prof. Baoxin Li.

Master of Science, Computer Engineering (Electrical Engineering) **December 2016**
Ira A. Fulton School of Engineering, Arizona State University, Tempe, Arizona CGPA: 4.00/4.00
Thesis: Scalable Register File Architecture and Compiler for CGRA Accelerators
Committee: Prof. Aviral Shrivastava, Prof. Umit Ogras, and Prof. Fengbo Ren.

Bachelors of Engineering, Electronics and Communication Engineering **July 2014**
L. D. College of Engineering, Gujarat Technological University, Gujarat CGPA: 8.80/10.00
Final-year Project: End-to-end System Development for Precise Temperature Control of Cryogenic Cooler Systems of Next Generation Geostationary Satellites.

PROFESSIONAL EXPERIENCE

Research Associate, Compiler Microarchitecture Lab, *Tempe, AZ* **January 2016 – Present**

- Research techniques and develop tools for efficiently executing critical applications, such as machine learning or time-sensitive apps, on hardware accelerators in resource-efficient manner and compiler-aware hardware codesign.
- Developed state-of-the-art techniques and open-source tools for (1) modeling hardware accelerators and analyzing their execution costs (power, performance, area), (2) comprehensive design space formulation for finding efficient designs, and (3) quickly exploring pareto-optimal hardware/software designs from vast space (in seconds/minutes).
- Prospectus Overview Video

Research Intern, *Mathworks*, *Natick, MA* **Summer 2018**

- Project: Loop Optimizations for Target-Aware Code Generation (Group: Code Efficiency, Embedded Coder)
- Established target-aware code generation research infrastructure. Developed compiler transformations to realize performance improvements by embedding target platform specific features in the code generation environment. Evaluations were performed successfully at scale for avionics and automotive apps, demonstrating up to an order-of-magnitude higher performance.

ASIC Verification Engineer (Intern), *SanDisk*, *Milpitas, CA* **Summer 2015**

- Worked with architect and technical leads of ASIC team to define and develop an automated tool that helped to make the development of Enterprise Solid State Drive Controller ASIC more efficient.
- Successfully performed module level verification for key modules of novel large complex ARM based SoC controller architecture, including developing verification plan, improving functional coverage, developing constrained random test cases, and verifying interesting corner cases in UVM based verification environment.

Research Intern, *Space Application Centre, Indian Space Research Organization,* *India* **Spring 2014**

- Collaborated with a team of senior research scientists and group director and achieved precise temperature control of cryogenic coolers (±0.1^{0}C at -123^{0}C) for next generation geostationary imaging satellites.
- Developed end-to-end system for plant modeling and precise temperature control of cryogenic cooler systems.
- Implemented digital PID controller on FPGAs for precise temperature control.
- Demonstrated the end-to-end system through two real-world system prototypes in the real-time environment under electro-mechanical disturbances.

HONORS AND AWARDS

- Completion Fellowship, Graduate College, Arizona State University – 2022
  (merit-based university-wide selection, 1-year funding awarded to top outgoing PhD students. ~1-2 nominations out of 100 computer engineering PhD students, subjected to further ASU-wide selection process.)
- ACM SIGBED Student Research Competition, First Runner-Up – 2022
- Outstanding Research Award, Graduate and Professional Students Association, Arizona State University – 2022
  (About 10-15 students awarded from 12,000+ graduate students and 4,000+ doctoral students.)
- Doctoral Fellowship, School of Computing and Augmented Intelligence (SCAI), ASU – 2019
- Research Spotlight, IEEE Eta Kappa Nu (HKN) – 2022 (Also featured in IEEE Bridge Volume 118 and social media).
- Richard Newton Young Student Fellowship, 53rd Annual Design Automation Conference (DAC) – 2016
- Outstanding Computer Engineering Teaching Assistant Award, School of Computing and AI (SCAI), ASU – 2018
- Student lead, ASU team, NSF/Intel joint research center for Computer Assisted Programming for Heterogeneous Architectures (CAPA) – 2018–2021 *(Our group was among total three hubs supported nation-wide. Hubs: MIT/Stanford/University of Washington; UCLA/Cornell; ASU/Penn State/Lehigh.)*
- Invited for Inaugural Google Systems Innovation Summit – 2021
  (*Invitations limited to top global PhD students in computing systems.*)
- PhD Forum Presentations at DAC 2022 and IPDPS 2022
- Certificate of Appreciation (for Outstanding Service as Social Media Chair),
  ACM/IEEE Embedded Systems Week – 2021, 2022
- Engineering Grad Fellowship, Ira A. Fulton Schools of Engineering, ASU – 2018, 2019, 2020

Competitive Travel Grant Awards:

- Graduate College, ASU – 2019
- School of Computing and Augmented Intelligence, ASU – 2018, 2019, 2022
- Computer Engineering Program, ASU – Spring 2018
- Graduate and Professional Student Association (GPSA), ASU – Summer 2016, Spring 2018, Fall 2019
- ACM Special Interest Group on Design Automation (SIGDA) – 2018, 2022
- NSF travel award, ACM Special Interest Group on Embedded Systems (SIGBED) – 2019
- IEEE Computer Society Technical Committee on Parallel Processing (TCPP) – 2022

RESEARCH IMPACT

Citations: They are more than 190[1] and have been increasing exponentially (nearly doubled every year), and almost all are for my first-authored works. My h-index and i-10 index are 7. My research is regularly refereed or evaluated by some of the top global experts (including from MIT, Stanford University, UC Berkeley, UIUC, NUS, ETH Zurich, Tsinghua, SJTU, Georgia Tech, UCLA, Purdue, CMU, Harvard, USC, NTU, Peking, University of Tokyo) in their papers published at topmost venues in machine learning, computer architecture, design automation, embedded systems, parallel computing, and code optimization, such as conferences like *ICLR, ICML, NeurIPS, IJCNN, ISCA, MICRO, ASPLOS, HPCA, SIGMETRICS, DAC, ICCAD, ESWEEK, DATE, CGO, PACT*, journals like *IEEE TPDS, JMLR, PMLR, ACM CSUR, IEEE TCAS-I, IEEE TCAD, IEEE Sensors*, *ACM TECS, IEEE TVLSI, ACM TACO*, books like *Morgan Claypool synthesis lectures,* US patents, and master's and PhD dissertations at top universities, including *NUS, ETH Zurich, Georgia Tech, Harvard, UCLA, Purdue, ASU, and UCSB*.

[1] Based on my profiles on Google Scholar, Semantics Scholar, ResearchGate, etc.

Research Usability: Citations of my papers by new works describe my techniques as state-of-the-art prior work and many build upon it or compare against it even quantitatively. Some citing works are industrial, including from NVIDIA Research, Google Brain, Samsung Research, Facebook AI Research, Intel AI, Xilinx, EdgeCortix, GraphCore, PNNL, Sandia National Labs, LBNL, RIKEN, SimpleMachines, Neural Magic, SambaNova, Robert Bosch, etc. Besides, my research is studied in the classes at top universities (e.g., Utah, UT Austin). My first-authored article at Proceedings of the IEEE (flagship IEEE publication and leading venue since past 110 years) describes most recent advances in efficient processing of recent machine/deep learning models on hardware accelerators, which is well received and widely studied by computer hardware, computing systems, and machine learning communities, including senior researchers. Further, my research frameworks are downloaded and used by tens of groups world-wide, including in their research development and dissertations of the students. My frameworks on GitHub receive more than 30–60 stars in a few years and several downloads every month.

SELECT INVITED TALKS

For every conference paper accepted (regular research papers, special sessions, workshops), we are required or invited to present a 15–30-minute talk. Likewise, for some journal articles, we are invited to present a talk. My research has also been presented into various invited talks from my collaborators at top industry research groups and different universities. Most importantly, my research on the tools and methodology for "accelerating computing" has been presented at the following premier events:

- NSF/Intel Annual CAPA day – 2018, 2019, 2020 (closed event)
  (Audience: NSF program managers, Intel researchers and fellows, researchers from top universities)
- Renegotiating the Levels of Abstraction for the Post Moore's Law Era, A Workshop at the Arm Research Summit, Austin, US – 2019
- Third Annual Future Chips Forum, "Reconfigurable Computing in Golden Age", Beijing Innovation Center for Future Chips and the Institute of Microelectronics, Tsinghua University – 2018
- ARM Research Summit, Cambridge, UK – 2018
- DAC-ROAD4NN: Annual International Workshop on Research Open Automatic Design for Neural Networks, co-located with the 59th Design Automation Conference (DAC 2022) – 2022 (usually 500 attendees)
- Annual IBM and IEEE CAS/EDS AI Compute Symposium – 2022
- 3rd Secure RISC-V Workshop (SECRISCV), co-located with 18th ACM/IEEE Embedded Systems Week – 2022

SELECT WORK FEATURED

My work is regularly posted, highlighted, or discussed by various experts on different forums and social media channels, including Twitter, Reddit, Hacker News, DeepAI, and arXiv-based platforms. My work has been showcased on following public pages:

- Intel Labs Select Publications
- insideHPC article on Energy-efficient Acceleration of Residual Neural Nets using CGRAs
- IEEE HKN Graduate Research Spotlight in IEEE Bridge, volume 119 and on various social media
- ARM Research Summit Talk on ARM Research YouTube Channel
- DAC recommended sessions from industry experts (session "The Art of Mapping for Accelerators")

PUBLICATIONS

My research areas include design automation, computer architecture, and embedded systems. My research is regularly published in (and referred by publications in) the top venues in these areas[2]. Information about impact assessment of the publication venues is available in Appendix A.

My most of the publications are first authored, involving my extreme involvement from conceptualization to artifact release for the publication. More information on my publications (author's version of the PDFs, presentation slides, posters, recorded YouTube talks, summary of the work, and open-source tools and artifacts) is available at: https://sites.google.com/view/shail/publications.

[2] Google Scholar publication metrics for computer hardware design

Peer-Reviewed, Referred Journal Papers

[J3] [Proceedings of the IEEE] Hardware Acceleration of Sparse and Irregular Tensor Computations of ML Models: A Survey and Insights. **Shail Dave**, Riyadh Baghdadi, Tony Nowatzki, Sasikanth Avancha, Aviral Shrivastava, Baoxin Li, in Proceedings of the IEEE, volume 109, issue 10, 2021, pages 1706-1752. (**20+ citations in a year) (impact factor: 14.91)**

- Collaboration with authors from NYU/MIT, UCLA, Intel Parallel Computing Lab, and ASU.
- *Key topics: Compact deep/machine learning models, their need, and sources of tensor compression; Efficient processing of compressed models with hardware and software; Quantitative and qualitative analysis of various techniques on performance, energy efficiency, storage, and chip area.*

[J2] [TACO] SPX64: A Scratchpad Memory for General-Purpose Microprocessors. Abhishek Singh, **Shail Dave**, PanteA Zardoshti, Robert Brotzman, Chao Zhang, Xiaochen Guo, Aviral Shrivastava, Gang Tan, Michael Spear, in ACM Transactions on Architecture and Code Optimization (TACO), Vol. 18, No. 1, 2021 [Presented by Invitation at HiPEAC 2021 - 16th International Conference on High-Performance Embedded Architectures and Compilers].

- Collaboration with authors from Lehigh, Penn State, and ASU.
- *Key topics: Software-managed cache; Securing execution against side-channel attacks; Accelerating persistent transactions for non-volatile memory.*

[J1] [CODES+ISSS @ ESWEEK, TECS] dMazeRunner: Executing Perfectly Nested Loops on Dataflow Accelerators. **Shail Dave**, Youngbin Kim, Sasikanth Avancha, Kyoungwoo Lee, Aviral Shrivastava, in ACM Transactions on Embedded Computing Systems (TECS), Vol. 18, No. 5s, 2019 [Special Issue on ESWEEK 2019 - ACM/IEEE International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)] **(2nd most downloaded ESWEEK 2019 paper on ACM digital library and 2nd most cited CODES+ISSS 2019 paper. 30+ citations in 2 years)**.

- Collaboration with authors from Yonsei, Intel Parallel Computing Lab, and ASU.
- *Key topics: Defining comprehensive hardware/software design space for loop nests; Accelerator cost model for evaluating execution metrics for variations in hardware architecture, mappings, deep learning model layers; Search-space reduction techniques for getting efficient mappings in a few seconds; Generic algorithms for obtaining all unique data reuse scenarios for loop-orderings.*

Peer-Reviewed, Referred Conference Papers

[C5] [VTS] Towards an Agile Design Methodology for Efficient, Reliable, and Secure ML Systems. **Shail Dave**, Alberto Marchisio, Muhammad Abdullah Hanif, Amira Guesmi, Aviral Shrivastava, Ihsen Alouani, Muhammad Shafique, in Proceedings of the 40th IEEE VLSI Test Symposium (VTS), 2022 (Invited special session).

- Collaboration with authors from NYU, TU Wien, UPHF, and ASU.
- *Key topics: Agile design methodology for developing efficient neural processing unit architectures (NPUs); Mitigating reliability challenges such as soft errors, process variations, aging, and manufacturing defects; Securing AI systems with robustness to faults, adversarial attacks, and privacy challenges.*

[C4] [ICASSP] dMazeRunner: Optimizing Convolutions and GEMMs on Dataflow Accelerators. **Shail Dave**, Aviral Shrivastava, Youngbin Kim, Sasikanth Avancha, Kyoungwoo Lee, in Proceedings of the 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2020.

- Collaboration with authors from Yonsei, Intel Parallel Computing Lab, and ASU.
- *Key topics: Open-source framework for optimizing execution of deep learning models on hardware accelerators in a few seconds.*

[C3] [DAC] RAMP: Resource-Aware Mapping for CGRAs. **Shail Dave**, Mahesh Balasubramanian, Aviral Shrivastava, in Proceedings of the 55th Annual Design Automation Conference (DAC), 2018 **(Total citations: 50+ in 3 years; state-of-the-art mapping technique for executing general purpose tasks on spatial architectures)**.

- Presented in one of the recommended sessions of DAC 2018 by industry experts.
- *Key topics: Mapping optimizations for accelerating loops of general-purpose computing through software pipelining; Failure analysis for routing the data dependencies during mapping.*

[C2] [DATE] URECA: A Compiler Solution to Manage Unified Register File for CGRAs. **Shail Dave**, Mahesh Balasubramanian, Aviral Shrivastava, in Proceedings of the 21st International Conference on Design Automation and Test in Europe (DATE), 2018.

[C1] [DATE] LASER: A Hardware/Software Approach to Accelerate Complicated Loops on CGRAs. Mahesh Balasubramanian, **Shail Dave**, Aviral Shrivastava, Reiley Jeyapaul, in Proceedings of the 21st International Conference on Design Automation and Test in Europe (DATE), 2018.

- *Key topics: Mapping optimizations for accelerating imperfectly nested loops and loops with conditionals from general-purpose computing applications.*

Peer-Reviewed, Referred Workshops; Competitions; and Conference Demonstrations

[W4] [SNN - Sparse NN, collocated with ICML] Efficient Processing of Sparse and Compact DNN Models on Hardware Accelerators. Survey and Insights. Shail Dave and Aviral Shrivastava (Joint work with Riyadh Baghdadi @ MIT CSAIL and NYU, Tony Nowatzki @ UCLA, Sasikanth Avancha @ Intel Labs. and Baoxin Li @ ASU) [Summary Poster]

[W3] [DAC ROAD4NN 2022, University Demonstration, ACM Student Research Competition] Agile and Explainable Exploration of Hardware/Software Codesigns of Deep Learning Accelerators. (Joint work with Aviral Shrivastava @ASU and Tony Nowatzki @UCLA) [Teaser]

[W2] [LATTE @ ASPLOS] Design Space Description Language for Comprehensive Exploration of Next-Gen Hardware Accelerators. **Shail Dave**, Aviral Shrivastava, in Second Workshop on Languages, Tools, and Techniques for Accelerator Design (LATTE), co-located with ACM 27th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS) 2022. **(Vision/Position paper)**

- *Key topics: Overcoming limitations of existing architecture design frameworks (that use template specified in architecture description language) with design space description language; Improving: (1) Comprehensiveness of design space (2) Reusability of design tools (3) Explainability of designs and their exploration, and (4) Exploration speed for making dynamic deployment feasible.*

[W1] [Demonstrations @ DATE, DAC] CCF: CGRA Compilation and Simulation Framework. Shail Dave, Aviral Shrivastava, in University Booth Demonstration at the 21st International Conference on Design Automation and Test in Europe (DATE), 2018 (also selected for university demonstration at 55th annual Design Automation Conference).

Other Peer-Reviewed Papers

[O1] [INROADS] Derivation of transfer function model based on Miniaturized cryocooler behavior. Jiten Bhatt, **Shail Dave**, Manish M. Mehta, and Nitin Upadhyay. INROADS, volume 5, no. 1s (2016): 336-340.

- Collaboration with senior researchers and group directors from Space Application Center, Indian Space Research Organization (undergraduate final project).
- *Key topics: Achieving precise temperature control of cryogenic cooler systems for next-gen geostationary satellites; End-to-end system that automatically fine-tunes the controller in presence of variations from mechanical and electrical components of the system.*

ArXiv and Under Progress/Submission

[S1] EXPLAINABLE-DSE: An Agile and Explainable Exploration of Hardware/Software Codesigns of Deep Learning Accelerators, Shail Dave, Tony Nowatzki, and Aviral Shrivastava.

[S2] Automating Cost Modeling and Characterization for Domain-Specific Accelerators.

Patents[3]

1. Hybrid and efficient approach to accelerate complicated loops on coarse-grained reconfigurable arrays (CGRA) accelerators, Mahesh Balasubramanian, Shail Dave, Aviral Shrivastava, and Reiley Jeyapaul. (Application number: US20200133672A1; Provisional patent filed in 2018).

TECHNICAL SKILLS

- **Programming Languages & HDL:** C, C++, Python, MIPS Assembly, Shell Scripting, Verilog, System Verilog.
- **Parallel Programming Libraries:** OpenMP, POSIX, MPI, OpenCL, CUDA.
- **Tools and Frameworks:** Gem5 Processor Simulator, LLVM Compiler (Front-End/Middle End), PyTorch, Apache TVM, Xilinx Vivado and Vivado HLS, Deep Learning Model Compression Tools, Altera ModelSim, Cadence and Synopsys ASIC design tools, MATLAB, GDB, Version management tools.
- **Relevant Courses:** Deep Learning, Hardware Acceleration, High Performance Computing, Computer Architecture, Distributed and Multi-Processing Operating Systems, VLSI Design, Hardware Design and Verification Languages.

RESEARCH PROJECTS

- Design Space Language and Framework for Exploring Next-Gen Hardware Accelerators — Spring 2022–Present
- Automated Cost Modeling for Arbitrary Dataflow Accelerators — Spring 2021 and Fall 2022
- Explainable Design Space Exploration of Hardware/Software Codesigns — Summer–Fall 2021
- Analyzing Sparse Tensor Computations of Machine Learning Models on Accelerators — Spring 2021–Fall 2021
- General-Purpose Scratchpad Memory for Secure, Persistent Execution — Fall 2019–Fall 2020
- Design modeling and exploration of dataflow accelerators for deep learning — Fall 2018–Fall 2019
- Compiler and simulator infrastructure for coarse-grain reconfigurable architectures (CGRAs) — Spring 2016–Fall 2018
- Mapping General-purpose loops on CGRA Accelerators — Spring 2016–Spring 2018

PROJECT GRANT PROPOSALS WRITING

Over last five years, I have taken a lead in writing quality, novel proposals for project grants, including for competitive special programs. Each NSF proposal is of usually 15 pages and other supplementary material, and others (from Semiconductor Research Corporation aka SRC, special calls from industry) are about 3-5 pages. Proposed NSF projects are of 250K-500 (small), 500-1200 (medium), or 1200-3000 (large), with projects lasting for 3-4 years. Industry proposals are roughly for 50k-90$ per year. I have assisted senior researchers (PIs) for conceptualization of novel projects, analysis of challenges and proposed tasks, and writing of the following proposals.

- *Explainable and Agile AI Hardware/Software Optimization With AI* (ASU) (SRC AIHW proposal)
- *Hardware-software Co-design for Delegated Homomorphic Computation* (ASU, Multiple PIs) (SRC proposal)
- *Sustainable Full System Stack Development for Next-Generation AI Hardware*
- *Concertina-A Language & Framework for Time-Sensitive Machine Learning Applications*. (CMU, ASU) (NSF proposal)
- *Automated Accident Detection Using Machine Learning* (ASU) (Arizona MCDOT proposal)
- PPoSS: *Towards Wide-scale, Time-sensitive, Federated Learning Applications*. (ASU multiple PIs) (NSF proposal)
- FoMR: *System-wide Reconfigurable Acceleration for Near-Data Computing*. (ASU and Lehigh) (NSF proposal)
- *Execute DNNs on Programmable Dataflow Accelerators: Where's your System Stack?* (ASU) (Facebook AI research)
- *Acceleration Beyond GPUs for High-Performance Computing*. (ASU) (SRC, NSF proposal)

[3] My recent research was supported through NSF/Intel program, which required the work to be public, so no patents could be filed from my research group during the 3–4-year timespan. For my most recent works on cutting-edge design methodologies for novel architectures, some are under filing process, and I plan to file a few more patents.

PROFESSIONAL SERVICE

Based on my research expertise and accomplishments, I am regularly invited to serve as an external/expert reviewer for the top/flagship conferences and journals in my research areas, including design automation, computer architecture, and embedded systems. Note that serving on the program committee (PC) of top conferences typically requires PhD, and therefore external experts including PhD students like myself are typically invited as an expert reviewer for a few papers (review load of 2-4 papers vs 10-15 for a PC member). My reviews are usually regarded by area editors or PC members as detailed, thorough, and carefully drafted. So, I am usually invited to continue as reviewer for these or other top venues, some of which I need to decline respectfully due to other commitments.

External Reviewer

- ACM Design Automation Conference (DAC) – 2018, 2019, 2020, 2022
- IEEE Real-Time Systems Symposium (RTSS) – 2019
- ACM/IEEE International Conference on Compilers, Architecture and Synthesis for Embedded Systems (CASES) – 2019
- IEEE/ACM International Conference on Hardware-Software Codesign & System Synthesis (CODES+ISSS) – 2016, 2017
- IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM) – 2019
- IEEE International Conference on Design Automation and Test in Europe (DATE) – 2016, 2017
- IEEE International Conference on Computer Design (ICCD) – 2019
- IEEE Asia South Pacific Design Automation Conference (ASP-DAC) – 2022
- IEEE International Conference on VLSI Design and Embedded Systems (VLSID) – 2017, 2018

Program Committee Member

- Second Workshop on Languages, Tools, and Techniques for Accelerator Design (LATTE), co-located with ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS) – 2022

Journal Referee

- IEEE Sensors Journal – 2022
- ACM Transactions on Embedded Computing Systems (TECS) – 2017, 2019, 2020, 2022
- ACM Transactions on Design Automation of Electronic Systems (TODAES) – 2019, 2020
- IEEE Transactions on Multi-Scale Computing Systems (TMSCS) – 2016, 2017
- Springer International Journal on Design Automation for Embedded Systems (DAES) – 2017

LEADERSHIP AND ORGANIZATIONAL ACTIVITIES

- Social Media Chair, ACM/IEEE Embedded Systems Week (ESWEEK) – 2021, 2022
- Graduate Vice Chair, IEEE Eta Kappa Nu Society, ASU Chapter – 2016-2017
- Invited Speaker, ASU International Student Orientation – 2016, 2017
  (Joint planning with ASU GPSA – Graduate and Professional Student Association and ISSC – International Students and Scholars Center)
- Logistics management, Distinguished Speaker Talks, School of Computing and Augmented Intelligence, ASU – 2017, 2018
- Early Student Career Panel, ASU International Student Orientation – 2016, 2017
- Special Session Organization, IEEE 40th VLSI Test Symposium – 2022
- Mentor, Computer Architecture Long-Term Mentoring (CALM) – 2022-Present
  Year-long mentoring program by ACM SIGARCH Computer Architecture Student Association (CASA)
- Mentor, Meet a Student Architect (MaSA) – 2022

PROFESSIONAL ASSOCIATIONS

I hold student membership of and regularly participate into events (in various capacity) of the following organizations:

- Association for Computing Machinery (ACM)
- Institute of Electrical and Electronics Engineers (IEEE)
- New York Academy of Sciences (NYAS)

- IEEE Computer Society (IEEE CS)
- IEEE Eta Kappa Nu Honors Society (IEEE HKN)
- ACM Special Interest Group on Computer Architecture (SIGARCH)
- ACM Special Interest Group on Embedded Systems (SIGBED)
- IEEE Council on Design Automation (CEDA)
- IEEE Technical Committee on Computer Architecture (TCCA)
- Computer Architecture Student Association (CASA) (supported by: ACM SIGARCH / IEEE TCCA)

TEACHING

Course Developed, School of Computing and Augmented Intelligence, ASU

- Topics in Machine Learning Accelerator Design – Spring 2023

Guest Lecturer, School of Computing, Informatics and Decision Systems Engineering, ASU

- ASU 101 (The ASU Experience) – Fall 2018
- CSE 420/520 (Computer Architecture) – Fall 2018 (on ML accelerators), Fall 2021 (several tutorials on Gem5)

Teaching Assistant, School of Computing, Informatics and Decision Systems Engineering, ASU – 2016–2018, 2021

- CSE 100 (Principles of Programming with C++)
- CSE 330 (Operating Systems)
- CSE 420/520 (Computer Architecture I and II) **(CIDSE outstanding TA award in 2018)**

Instructional Aide, School of Mathematics and Statistical Sciences, Arizona State University – 2014–2016

- MAT 210 (Brief Calculus)
- MAT 142 (College Mathematics)
- MAT 117 (College Algebra)

OUTREACH AND MENTORING

Advising and Mentoring

During my research at ASU, I have regularly collaborated with or advised master's and PhD peer students, including under-represented students (about 40% of my advisees), in various capacities for their learning of cutting-edge technology topics and tools and professional growth. They are now senior research engineers or architecture engineers at international companies like Qualcomm, Intel, Apple, Cadence, ON semiconductor, MediaTech, Cirrus Logic, and one is now a PhD student at North Carolina State University.

I have now also signed up as a mentor for CALM: Computer Architecture Long-term Mentoring program and MaSA: Meet a Senior Architect, run by Computer Architecture Student Association (CASA) (supported by: ACM SIGARCH / IEEE TCCA). Through this global program, I plan to regularly mentor non-ASU students every year, especially juniors who do not have accesses to sufficient technical/professional resources, are less privileged, or under-represented.

Selected Outreach for Professional Growth

I regularly serve on student seminar series or panels, discussing about the research and professional growth opportunities. I have served as Vice chair of IEEE HKN ASU chapter, where I established "luncheon with faculty" program. It allowed students to meet faculty of their research interest from SECEE or SCAI and have discussions on a broad range of topics, beyond their courses. I also served for a few years as an invited speaker for ASU international student orientation, discussing various opportunities at ASU and how to build a strong professional profile. I have also served on various student panels, including GradLab podcast – IEEE HKN where I discussed navigating graduate school challenges and opportunities and at ASU 101, where I discussed with freshman students about computer science and engineering and their questions on various opportunities and career paths after their graduation. I have also assisted in hosting distinguished speakers for SCAI seminar series at ASU, including renowned professors from Stanford, University of Minnesota, KAIST, UT Austin, etc.

COLLABORATORS

I routinely collaborate with ASU and other experts from industry and academia on skills/knowledge that is complementary in developing novel projects and tools for cutting-edge research in my domain and even developing proposals for conceptualization/planning of the projects and in organizing special/topical sessions. Some of my recent collaborators are listed below.

- Dr. Sasikanth Avancha, Intel Parallel Computing Lab
- Prof. Tony Nowatzki, UCLA
- Prof. Michael Spear, Lehigh University
- Prof. Xiaochen Guo, Lehigh University
- Prof. Riyadh Baghdadi, MIT CSAIL and New York University
- Prof. Gang Tan, Penn State University
- Prof. Kyongwoo Lee, Yonsei University
- Dr. Reiley Jeyapaul, ARM Research, Austin
- Dr. Partha Biswas, MathWorks Research
- Prof. Muhammad Shafique, New York University
- Prof. Jonathan Aldrich, Carnegie Mellon University
- Prof. Carlee Joe-Wong, Carnegie Mellon University
- Prof. Baoxin Li, Arizona State University
- Prof. Fengbo Ren, Arizona State University

APPENDIX A

PUBLICATION INFORMATION

My research areas include design automation, computer architecture, and embedded systems. My research is regularly published in (and referred by publications in) the top venues in these areas. Impact assessment of these venues is summarized in the table below, which shows h5-index citations collected with Google Scholar. In fact, these design automation venues are listed as top publication venues in "Computer Hardware Design" under Engineering and Computer Science category from google scholar publication metrics[4], while computer architecture venues are listed in "Computing Systems" categories.

Note that the papers published in the top-tier embedded systems and computer architecture conferences are typically double-blinded submissions and double-column, single-spaced papers of 10-12 pages. Each paper usually receives at least 4-5 detailed technical reviews (all or majority of them are from the Program Committee and the rest from invited experts). For design automation conferences, the same holds except that the papers are of 6-8 pages. For top journals in these areas, papers are of 12-15 pages and single-blinded submissions (double-column, single-spaced, published in IEEE) and 25 pages (single column, double-spaced, published in ACM). Typical acceptance rates of these conferences are 20%–25% and about 35% for journals[5].

Research Area	Publication		Acronym	Year[6]	h5-index[7]	Impact Factor
Design Automation	Conference	IEEE/ ACM Design Automation Conference	DAC	59	51	N/A
		IEEE Design, Automation, and Test in Europe	DATE	25	47	
		IEEE/ACM International Conference on Computer-Aided Design	ICCAD	41	38	

[4] Google Scholar publication metrics for computer hardware design: https://scholar.google.com/citations?view_op=top_venues&hl=en&vq=eng_computerhardwaredesign

[5] Journal acceptance ratio obtained from the CEDA DAWN webinar: https://ieee-ceda.org/presentation/webinar/dawn-publishing-eda-transactions-journals-and-magazines. Note that journals typically allow a major and a minor revision for quality papers, which makes their overall acceptance rate higher than the conferences.

[6] "Year" indicates that a conference or a journal is running since how many years.

[7] h5-index is the h-index for articles published in the last 5 complete years. It is the largest number h such that h articles published in 2016-2020 have at least h citations each. Taken from 2021 Google Scholar publication metrics.

	Journal	IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems	TCAD	41	51	2.87
Embedded Systems	Conference	ACM/IEEE Embedded Systems Week	ESWEEK	35	N/A[8]	
	Journal	ACM Transactions on Embedded Computing Systems	TECS	21	30	2.58
	Conference	IEEE International Conference on Acoustics, Speech and Signal Processing	ICASSP	47	96	N/A
Computer Architecture	Conference	ACM International Conference on Architectural Support for Programming Languages and Operating Systems	ASPLOS	27	56	N/A
		IEEE/ACM International Symposium on Computer Architecture	ISCA	49	55	
	Journal	ACM Transactions on Architecture and Code Optimization	TACO	19	26	1.89
Engineering	Journal	Proceedings of the IEEE	PIEEE	110	93	10.96

Note that most research in computing systems or computer hardware design requires developing one or more complicated artifacts, such as architectural simulators, compilers, or optimization frameworks, from scratch, which usually consumes several months. Therefore, it usually leads about 1-2 papers per every year or two, especially for first-authored works and when working in a small team or alone. For my research, which mostly focuses on developing specialized processors ("accelerators"), it exacerbates because developing tools from scratch for novel and cutting-edge techniques and designs is time-consuming, in contrast to leveraging off-the-shelf tools for general-purpose hardware like CPUs/GPUs (e.g., LLVM compiler, gem5 processor simulator) or black-box optimization frameworks like simulated annealing and genetic algorithm (e.g., from scikit python library or MATLAB).

[8] ESWEEK includes three major conferences, CASES, CODES+ISSS, and EMSOFT; These conferences were running separately before 2007, and some conferences like ISSS were established even as earlier in 1988. With journal-integrated publication model (for conference paper revisions) since 2017, the papers of ESWEEK conferences are now published in the IEEE TCAD and the ACM TECS at every alternate year.