# Adwait Jog

## Academic Positions

Jan'23 - **Associate Professor (with Tenure) and Anita Jones Faculty Fellow**, *University of Virginia*, Charlottesville, VA

Aug'21–Dec'22 **Adina Allen Associate Professor (with Tenure)**, *William & Mary*, Williamsburg, VA

Aug'15–July'21 **Assistant Professor**, *William & Mary*, Williamsburg, VA

## Education

2009–2015 **Ph.D.**, *Pennsylvania State University (Penn State)*, University Park, PA, USA
Computer Science and Engineering

2005–2009 **B.Tech**, *National Institute of Technology (NIT)*, Rourkela, Orissa, India
Electronics and Instrumentation Engineering

## Professional Experience

2009–2015 **Graduate Research/Teaching Assistant**, Penn State, PA

Summer 2013 **Graduate Research Intern**, *NVIDIA Research*, Santa Clara, CA

Summer 2012 **Graduate Research Intern**, *Intel Labs*, Hillsboro, OR

Summer 2011 **Graduate Technical Intern**, *Intel*, Hillsboro, OR

## Awards and Honors

2024 **Best Paper Nomination at the IEEE International Conference on Cluster Computing (CLUSTER)**
Among top 3 papers out of 39 accepted ones!

2023 **Elevated to IEEE and ACM Senior Member**
Honor bestowed only to those who have made significant contributions to the profession.

2023 **Recognized for Outstanding Research and Scholarship**
Invited to the 2023 Research Achievement Award Ceremony

2022 **Plumeri Award for Faculty Excellence**
William & Mary's highest honors

2021 **Google Research Scholar Award**
One out of six awardees worldwide in the systems area

2021 **Nominated for Graduate Student Mentoring Award**
Received commendation letter from the William & Mary Dean's office

2018 **NSF CAREER Award**
One of the most prestigious NSF awards in support of early-career faculty

2014 **Outstanding Graduate Research Assistant Award, CSE, Penn State**

2013 **NVIDIA Graduate Fellowship Finalist**

2013 **Best Paper Nomination at the International Conference on Parallel Architectures and Compilation Techniques (PACT)**

## Awarded Funding

NSF **CORE Medium**, *Collaborative Research: SHF: Medium: Enabling GPU Performance Simulation for Large-Scale Workloads with Lightweight Simulation Methods*, Award Amount: $379,800, PI at UVA, lead-PI at WM, Duration: 2024-2028

DOE **ASCR EXPRESS**, *Modeling the Memory-Compute Gap in Large-scale Superconductive Systems*, Award Amount: $86,500, PI at UVA (sub-award from Rochester), Duration: 2023-2025
Went through a rigorous peer-review process which included a pre-application review followed by a standard full-proposal review.

NSF **Faculty Early Career Development Program (CAREER)**, *CCF/SHF: Addressing Scalability Challenges in Designing Next-generation GPU-based Heterogeneous Architectures*, Award Amount: $466,000 (including REU supplement at UVA, sole-PI, Duration: 2018-2026 (Transferred to UVA from 2023 onwards)

Google **Research Scholar Award**, *Enabling Efficient Sharing of Emerging GPUs*, Award Amount: $60,000, PI, Duration: 2021-onwards (Transferred to UVA from 2023 onwards)
One out of six awardees worldwide in the systems area.

DOE/ANL **Subcontract**, *Efficient State Capture on Heterogeneous Architectures*, Award Amount: about $20,000, PI: Jog, Duration: Spring 2021

NSF **CORE Program**, *CCF/SHF: Small: Enabling and Analyzing Accuracy-aware Reliable GPU Computing*, Award Amount: $449,999, PI, Duration: Aug 2017 - July 2021

NSF **Computing Research Initiation Initiative (CRII)**, *CCF/SHF: Design and Analysis of Processing-Near-Memory Enabled GPU Architecture*, Award Amount: $175,000, sole-PI, Duration: 2017-2020

NVIDIA **Equipment Donation**, *Quadro P6000 (2018), Jetson TX1 (2017), Tesla K40 (2016), GeForce Titan (2013)*
Access End Date: Dec 2022

W&M **Summer Research Award**, *Award amount: $4,300 each year (2019,2018)*

W&M **Reves Faculty International Conference Travel Grant**, *Award amount each year: $400 each year (2018,2017,2015)*

## Publications

As per Google Scholar, total citations: 3200+, h-index: 25

Underlined students are/were my advisees at the time the work was performed.

### Peer-reviewed Publications

[1] Amel Fatima, Yang Yang, Yifan Sun, Rachata Ausavarungnirun, and Adwait Jog. NetCrafter: Tailoring Network Traffic for Non-Uniform Bandwidth Multi-GPU Systems. In *the Proceedings of 52nd International Symposium on Computer Architecture*

**(ISCA)**, *Tokyo, Japan, Accepted, Acceptance rate: 127/570 ~22%*, June 2025.

[2] Ying Li, Yuhui Bao, Gongyu Wang, Xinxin Mei, Pranav Vaid, Anandaroop Ghosh, Adwait Jog, Darius Bunandar, Ajay Joshi, and Yifan Sun. TrioSim: A Lightweight Simulator for Large-Scale DNN Workloads on Multi-GPU Systems. In *the Proceedings of 52nd International Symposium on Computer Architecture* **(ISCA)**, *Tokyo, Japan, Accepted, Acceptance rate: 127/570 ~22%*, June 2025.

[3] Yang Yang, Mohammad Sonji, and Adwait Jog. Dissecting performance overheads of confidential computing on gpu-based systems. In *IEEE International Symposium on Performance Analysis of Systems and Software* **(ISPASS)**, *Ghent, Belgium, Accepted, Acceptance rate: 28/99 ~28%*, May 2025.

[4] Rishabh Jain, Vivek M. Bhasi, Adwait Jog, Anand Sivasubramaniam, Mahmut T. Kandemir, and Chita R. Das. Pushing the Performance Envelope of DNN-based Recommendation Systems Inference on GPUs. In *the 57th IEEE/ACM International Symposium on Microarchitecture* **(MICRO)**, *Austin, TX*, pages 1217–1232, *Acceptance rate: 113/497 ~23%*, November 2024.

[5] Anna Schmedding, Lishan Yang, Adwait Jog, and Evgenia Smirni. Aspis: Lightweight neural network protection against soft errors. In *IEEE International Symposium on Software Reliability Engineering* **(ISSRE)**, *Tsukuba, Japan*, pages 248–259, *Acceptance rate: 53/205 ~26%*, October 2024.

[6] Lishan Yang, George Papadimitriou, Dimitrios Sartzetakis, Adwait Jog, Evgenia Smirni, and Dimitris Gizopoulos. GPU Reliability Assessment: Insights Across the Abstraction Layers. In *IEEE International Conference on Cluster Computing* **(CLUSTER)**, *Kobe, Japan*, pages 1–13, *Acceptance rate: 39/138 ~28%*, **Best Paper Nomination**, September 2024.

[7] Lishan Yang, George Papadimitriou, Dimitrios Sartzetakis, Adwait Jog, Evgenia Smirni, and Dimitris Gizopoulos. Probing weaknesses in GPU reliability assessment: A cross-layer approach. In *IEEE International Symposium on Performance Analysis of Systems and Software* **(ISPASS)**, *Indianapolis, IN, USA*, pages 331–333, May 2024.

[8] Ying Li, Yifan Sun, and Adwait Jog. Path Forward Beyond Simulators: Fast and Accurate GPU Execution Time Prediction for DNN Workloads. In *the Proceedings of 56th International Symposium on Microarchitecture* **(MICRO)**, *Toronto, Canada*, pages 380–394, *Acceptance rate: 101/424 ~24%*, October 2023.

[9] Ying Li, Yuhui Bao, Pranav Vaid, Gongyu Wang, Adwait Jog, Darius Bunandar, Ajay Joshi, and Yifan Sun. TraceSim: a Lightweight Simulator for Large-Scale DNN Workloads on Multi-GPU Systems. In *1st Workshop on Computer Architecture Modeling and Simulation (CAMS) held with 56th International Symposium on Microarchitecture (MICRO), Toronto, Canada*, October 2023.

[10] Rishabh Jain, Scott Cheng, Vishwas Kalagi, Vrushabh Sanghavi, Samvit Kaul, Meena Arunachalam, Kiwan Maeng, Adwait Jog, Anand Sivasubramaniam, Mahmut T. Kandemir, and Chita R. Das. Optimizing CPU Performance for Recommendation

Systems At-Scale. In *the Proceedings of 50th International Symposium on Computer Architecture* **(ISCA)**, *Orlando, Florida*, pages 77:1–77:15, *Acceptance rate: 79/372 ~21%*, June 2023.

[11] Hongyuan Liu, Sreepathi Pai, and Adwait Jog. Asynchronous Automata Processing on GPUs. In *the Proceedings of the ACM on Measurement and Analysis of Computing Systems* **(SIGMETRICS)**, *Orlando, Florida, USA*, pages 27:1–27:27, June 2023.

[12] Ying Li, Yifan Sun, and Adwait Jog. POSTER: A Regression Model for DNN Execution Time Prediction on GPUs. In *the Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software* **(ISPASS)**, *Raleigh, NC*, pages 343–345, April 2023.

[13] Khoa Ho, Hui Zhao, Adwait Jog, and Saraju Mohanty. Improving GPU Throughput through Parallel Execution Using Tensor Cores and CUDA Cores. In *the Proceedings of 2022 IEEE Computer Society Annual Symposium on VLSI* **(ISVLSI)**, *Pafos, Cyprus*, pages 223–228, July 2022.

[14] Hongyuan Liu Bogdan Nicolae, Sheng Di, Franck Cappello, and Adwait Jog. Accelerating DNN Architecture Search at Scale using Selective Weight Transfer. In *the Proceedings of 23rd IEEE Cluster 2021 International Conference* **(CLUSTER)**, *Virtual Event*, pages 82–93, *Acceptance rate: 48/163 ~29%*, Sept 2021.

[15] Gurunath Kadam, Evgenia Smirni, and Adwait Jog. Data-centric Reliability Management in GPUs. In *the Proceedings of 51st International Conference on Dependable Systems and Networks* **(DSN)**, *Virtual Event*, pages 271–283, *Acceptance rate: 48/295 ~16%*, June 2021.

[16] Lishan Yang, Bin Nie, Adwait Jog, and Evgenia Smirni. SUGAR: Speeding Up GPGPU Application Resilience Estimation with Input Sizing. In *the Proceedings of the ACM on Measurement and Analysis of Computing Systems (POMACS Journal) at* **SIGMETRICS** *conference, Virtual Event*, pages 1–29, *Acceptance rate: ~12%*, June 2021.

[17] Lishan Yang, Bin Nie, Adwait Jog, and Evgenia Smirni. Enabling Software Resilience in GPGPU Applications via Partial Thread Protection. In *the Proceedings of International Conference on Software Engineering* **(ICSE)**, *Virtual Event*, pages 1248–1259, *Acceptance rate: 138/615 ~22%*, May 2021.

[18] Mohamed Assem Ibrahim, Onur Kayiran, Yasuko Eckert, Gabriel H. Loh, and Adwait Jog. Analyzing and Leveraging Decoupled L1 Caches in GPUs. In *the Proceedings of 27th International Symposium on High Performance Computer Architecture* **(HPCA)**, *Virtual Event*, pages 467–478, *Acceptance rate: 63/258 ~24%*, Feb 2021.

[19] Lishan Yang, Bin Nie, Adwait Jog, and Evgenia Smirni. Practical Resilience Analysis of GPGPU Applications in the Presence of Single- and Multi-Bit Faults. *IEEE Transactions on Computers,* **(TC)**, 70(1):30–44, Jan 2021.

[20] Mohamed Assem Ibrahim, Onur Kayiran, Yasuko Eckert, Gabriel H. Loh, and Adwait Jog. Analyzing and Leveraging Shared L1 Caches in GPUs. In *the Proceedings of*

*29th International Conference on Parallel Architectures and Compilation Techniques* **(PACT)**, *Virtual Event*, pages 161–173, *Acceptance rate: 35/135 ~25%*, Oct 2020.

[21] Bin Nie, Adwait Jog, and Evgenia Smirni. Characterizing Accuracy-Aware Resilience of GPGPU Applications. In *the Proceedings of 20th International Symposium on Cluster, Cloud and Internet Computing* **(CCGgrid)**, *Virtual Event*, pages 111–120, *Acceptance rate: 66/234 ~28%*, May 2020.

[22] Hongyuan Liu, Sreepathi Pai, and Adwait Jog. Why GPUs are Slow at Executing NFAs and How to Make them Faster. In *the Proceedings of 23rd International Conference on Architectural Support for Programming Languages and Operating Systems* **(ASPLOS)**, *Virtual Event*, pages 251–265, *Acceptance rate: 86/479 ~18%*, March 2020.

[23] Gurunath Kadam, Danfeng Zhang, and Adwait Jog. BCoal: Bucketing-based Memory Coalescing in GPUs. In *the Proceedings of 26th International Symposium on High Performance Computer Architecture* **(HPCA)**, *San Diego, CA*, pages 570–581, *Acceptance rate: 48/248 ~19%*, Feb 2020.

[24] Mohamed Assem Ibrahim, Hongyuan Liu, Onur Kayiran, and Adwait Jog. Enhancing Bandwidth Utilization via Efficient Inter-core Communication in GPUs. In *the Proceedings of 28th International Conference on Parallel Architectures and Compilation Techniques* **(PACT)**, *Seattle, WA*, pages 258–271, *Acceptance rate: 26/126 ~21%*, Sept 2019.

[25] Haonan Wang and Adwait Jog. Exploiting Latency and Error Tolerance of GPGPU Applications for an Energy-efficient DRAM. In *the Proceedings of 49th International Conference on Dependable Systems and Networks* **(DSN)**, *Portland, OR*, pages 362–374, *Acceptance rate: 54/252 ~21%*, June 2019.

[26] Haonan Wang, Mohamed Assem Ibrahim, Sparsh Mittal, and Adwait Jog. Address-Stride Assisted Approximate Value Prediction in GPUs. In *the Proceedings of 33rd International Conference on Super Computing* **(ICS)**, *Phoenix, Arizona*, pages 184–194, *Acceptance rate: 45/193 ~23%*, June 2019.

[27] Ashutosh Pattnaik, Xulong Tang, Onur Kayiran, Adwait Jog, Asit Mishra, Mahmut T. Kandemir, Anand Sivasubramaniam, and Chita R. Das. Opportunistic Computing in GPU Architectures. In *the Proceedings of 46th International Symposium on Computer Architecture* **(ISCA)**, *Phoenix, Arizona*, pages 210–223, *Acceptance rate: 62/365 ~17%*, June 2019.

[28] Xulong Tang, Ashutosh Pattnaik, Onur Kayiran, Adwait Jog, Mahmut Taylan Kandemir, and Chita R. Das. Quantifying Data Locality in Dynamic Parallelism in GPUs. In *the Proceedings of the ACM on Measurement and Analysis of Computing Systems (POMACS) at* **SIGMETRICS** *conference, Phoenix, Arizona*, pages 39:1–39:24, *Acceptance rate: 50/317 ~15%*, June 2019.

[29] Hongyuan Liu, Mohamed Assem Ibrahim, Onur Kayiran, Sreepathi Pai, and Adwait Jog. Architectural Support for Efficient Large-Scale Automata Processing. In *the*

*Proceedings of 51st International Conference on Micro-Architecture* **(MICRO)**, *Fukuoka, Japan*, pages 908–920, *Acceptance rate: 74/351 ~21%*, Oct 2018.

[30] Lishan Yang Bin Nie, Evgenia Smirni, and Adwait Jog. Fault Site Pruning for Practical Reliability Analysis of GPGPU Applications. In *the Proceedings of 51st International Conference on Micro-Architecture* **(MICRO)**, *Fukuoka, Japan*, pages 749–761, *Acceptance rate: 74/351 ~21%*, Oct 2018.

[31] Rachata Ausavarungnirun, Vance Miller, Joshua Landgraf, Saugata Ghose, Jayneel Gandhi, Adwait Jog, Christopher J. Rossbach, and Onur Mutlu. MASK: Redesigning the GPU Memory Hierarchy to Support Multi-application Concurrency. In *the Proceedings of 23rd International Conference on Architectural Support for Programming Languages and Operating Systems* **(ASPLOS)**, *Williamsburg, VA*, pages 503–518, *Acceptance rate: 56/319 ~17%*, March 2018.

[32] Haonan Wang, Fan Luo, Mohamed Assem Ibrahim, Onur Kayiran, and Adwait Jog. Efficient and Fair Multi-programming in GPUs via Effective Bandwidth Management. In *the Proceedings of 24th International Symposium on High Performance Computer Architecture* **(HPCA)**, *Vienna, Austria*, pages 247–258, *Acceptance rate: 54/260 ~20%*, Feb 2018.

[33] Gurunath Kadam, Danfeng Zhang, and Adwait Jog. RCoal: Mitigating GPU Timing Attack via Subwarp-based Randomized Coalescing Techniques. In *the Proceedings of 24th International Symposium on High Performance Computer Architecture* **(HPCA)**, *Vienna, Austria*, pages 156–167, *Acceptance rate: 54/260 ~20%*, Feb 2018.

[34] Hengyu Zhao, Colin Weinshenker, Mohamed Assem Ibrahim, Adwait Jog, and Jishen Zhao. Layer-wise performance bottleneck analysis of deep neural networks. *Architectures for Intelligent Machines* **(AIM)** *Workshop in conjunction with International Conference on Parallel Architectures and Compilation Techniques (PACT), Hillsboro, OR*, Sep 2017.

[35] Nandita Vijaykumar, Gennady Pekhimenko, Adwait Jog, Saugata Ghose, Abhishek Bhowmick, Rachata Ausavarangnirun, Chita Das, Mahmut T Kandemir, Todd C Mowry, and Onur Mutlu. A Framework for Accelerating Bottlenecks in GPU Execution with Assist Warps. *Advances in GPU Research and Practice*, pages 373–415, Sept 2017.

[36] Sparsh Mittal, Rajendra Bishnoi, Fabian Oboril, Haonan Wang, Mehdi Tahoori, Adwait Jog, and Jeffrey Vetter. Architecting SOT-RAM Based GPU Register File. In *the Proceedings of IEEE Computer Society Annual Symposium on VLSI* **(ISVLSI)**, *Bochum, Germany*, pages 38–44, *Acceptance rate: 67/212 ~32%*, July 2017.

[37] Xulong Tang, Ashutosh Pattnaik, Huaipan Jiang, Onur Kayiran, Adwait Jog, Sreepathi Pai, Mohamed Assem Ibrahim, Mahmut Kandemir, and Chita Das. Controlled Kernel Launch for dynamic parallelism in GPUs. In *the Proceedings of 23rd International Symposium on High Performance Computer Architecture* **(HPCA)**, *Austin, TX*, pages 649–660, *Acceptance rate: 50/224 ~22%*, Feb 2017.

[38] Sparsh Mittal, Haonan Wang, Adwait Jog, and Jeffrey Vetter. Design and Analysis of Soft-Error Resilience Mechanisms for GPU Register File. In *the Proceedings of 30th International Conference on VLSI design and 16th International Conference on Embedded Systems* **(VLSID)**, *Hyderabad, India*, pages 409–414, *Acceptance rate: 71/292 ~24%*, Jan 2017.

[39] Nandita Vijaykumar, Kevin Hsieh, Gennady Pekhimenko, Samira Khan, Ashish Shrestha, Saugata Ghose, Adwait Jog, Phillip B Gibbons, and Onur Mutlu. Zorua: A Holistic Approach to Resource Virtualization in GPUs. In *the Proceedings of 49th International Symposium on Micro Architecture* **(MICRO)**, *Taipei, Taiwan*, pages 1–14, *Acceptance rate: 61/288 ~21%*, Oct 2016.

[40] Robert Risque and Adwait Jog. Characterization of Quantum Workloads on SIMD Architectures. In *the Proceedings of International Symposium on Workload Characterization* **(IISWC)**, *Providence, RI*, pages 1–9, *Acceptance rate: 21/71 ~29%*, Oct 2016.

[41] Ashutosh Pattnaik, Xulong Tang, Adwait Jog, Onur Kayiran, Asit K Mishra, Mahmut T Kandemir, Onur Mutlu, and Chita R Das. Scheduling Techniques for GPU Architectures with Processing-in-Memory Capabilities. In *the Proceedings of 25th International Conference on Parallel Architectures and Compilation Techniques* **(PACT)**, *Haifa, Israel*, pages 31–44, *Acceptance rate: 31/139 ~22%*, Sept 2016.

[42] Onur Kayiran, Adwait Jog, Ashutosh Pattnaik, Rachata Ausavarungnirun, Xulong Tang, Mahmut T Kandemir, Gabriel H Loh, Onur Mutlu, and Chita R Das. $\mu$C-States: Fine-grained GPU Datapath Power Management. In *the Proceedings of 25th International Conference on Parallel Architectures and Compilation Techniques* **(PACT)**, *Haifa, Israel*, pages 17–30, *Acceptance rate: 31/139 ~22%*, Sept 2016.

[43] Adwait Jog, Onur Kayiran, Ashutosh Pattnaik, Mahmut T Kandemir, Onur Mutlu, Ravishankar Iyer, and Chita R Das. Exploiting Core Criticality for Enhanced GPU Performance. In *the Proceedings of 42nd International International Conference on Measurement and Modeling of Computer Systems* **(SIGMETRICS)**, *Antibes Juan-Les-Pins, France*, pages 351–363, *Acceptance rate: 28/208 ~13%*, June 2016.

[44] Adwait Jog, Onur Kayiran, Tuba Kesten, Ashutosh Pattnaik, Evgeny Bolotin, Niladrish Chatterjee, Stephen W Keckler, Mahmut T Kandemir, and Chita R Das. Anatomy of GPU Memory System for Multi-application Execution. In *the Proceedings of 1st International Symposium on Memory Systems* **(MEMSYS)**, *Washington, DC*, pages 223–234, Oct 2015.

[45] Nandita Vijaykumar, Gennady Pekhimenko, Adwait Jog, Abhishek Bhowmick, Rachata Ausavarungnirun, Chita Das, Mahmut T Kandemir, Todd C Mowry, and Onur Mutlu. A Case for Core-assisted Bottleneck Acceleration in GPUs: Enabling Flexible Data Compression with Assist Warps. In *the Proceedings of 42nd International Symposium on Computer Architecture* **(ISCA)**, *Portland, OR*, pages 41–53, *Acceptance rate: 58/305 ~19%*, June 2015.

[46] Evgeny Bolotin, Zvi Guz, Adwait Jog, Steve Keckler, and Michael Parker. Approach to Adaptive Allocation of Shared Resources in Computer Systems, June 2015. US Patent 2015/0163324 A1, Assigned to NVIDIA.

[47] Onur Kayiran, Nachiappan Chidambaram Nachiappan, Adwait Jog, Rachata Ausavarungnirun, Mahmut T Kandemir, Gabriel H Loh, Onur Mutlu, and Chita R Das. Managing GPU Concurrency in Heterogeneous Architectures. In *the Proceedings of 47th International Symposium on Micro Architecture* **(MICRO)**, *Cambridge, UK*, pages 114–126, *Acceptance rate: 53/273 ~19%*, Dec 2014.

[48] Wei Ding, Mahmut Kandemir, Diana Guttman, Adwait Jog, Chita R Das, and Praveen Yedlapalli. Trading Cache Hit Rate for Memory Performance. In *the Proceedings of 23rd International Conference on Parallel Architectures and Compilation Techniques* **(PACT)**, *Edmonton, Alberta, Canada*, pages 357–368, *Acceptance rate: 37/144 ~25%*, August 2014.

[49] Adwait Jog, Evgeny Bolotin, Zvika Guz, Mike Parker, Stephen W Keckler, Mahmut T Kandemir, and Chita R Das. Application-aware Memory System for Fair and Efficient Execution of Concurrent GPGPU Applications. In *the Proceedings of 7th Workshop on General Purpose Computing using GPUs* **(GPGPU)**, *co-located with 19th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), Salt Lake City, UT*, pages 1–10, *Acceptance rate: 12/27 ~44%*, March 2014.

[50] Onur Kayıran, Adwait Jog, Mahmut T Kandemir, and Chita Ranjan Das. Neither More Nor Less: Optimizing Thread-level Parallelism for GPGPUs. In *the Proceedings of 22nd International Conference on Parallel Architectures and Compilation* **(PACT)**, *Edinburgh, Scotland*, pages 157–166, *Acceptance rate: 36/208 ~19%*, Sept 2013.

[51] Adwait Jog, Onur Kayiran, Asit K Mishra, Mahmut T Kandemir, Onur Mutlu, Ravishankar Iyer, and Chita R Das. Orchestrated Scheduling and Prefetching for GPGPUs. In *the Proceedings of 40th International Symposium on Computer Architecture* **(ISCA)**, *Tel Aviv, Israel*, pages 332–343, *Acceptance rate: 56/288 ~19%*, June 2013.

[52] Adwait Jog, Onur Kayiran, Nachiappan Chidambaram Nachiappan, Asit K Mishra, Mahmut T Kandemir, Onur Mutlu, Ravishankar Iyer, and Chita R Das. OWL: Cooperative Thread Array Aware Scheduling Techniques for Improving GPGPU Performance. In *the Proceedings of 18th International Conference on Architectural Support for Programming Languages and Operating Systems* **(ASPLOS)**, *Houston, TX*, pages 395–406, *Acceptance rate: 44/191 ~23%*, March 2013.

[53] Adwait Jog, Asit K Mishra, Cong Xu, Yuan Xie, Vijaykrishnan Narayanan, Ravishankar Iyer, and Chita R Das. Cache Revive: Architecting Volatile STT-RAM Caches for Enhanced Performance in CMPs. In *the Proceedings of 49th Design Automation Conference* **(DAC)**, *San Francisco, CA*, pages 243–252, *Acceptance rate: 168/741 ~22%*, June 2012.

## Invited Scholarly Talks

**CPU+GPU Security Workshop, Sponsored by IIT Bombay**
- IIT Bombay, India, Dec 2024 (in-person, duration: 1-day of external activity)

**Towards Efficient, Reliable, and Secure GPU Computing**
- University of Louisville, KY, USA, November 2023 (remote)

**GPU Computing and Beyond**
- IIT Bombay, India, March 2023 (in-person)
- IIT Indore, India, March 2023 (in-person)
- Systems Seminar, University of Virginia, USA, Feb 2023 (in-person)

**Low-overhead Reliable and Secure GPU Computing: Are we there yet?**
- AMD Research, Dec 2021 (remote)
- UC Merced, CA, Aug 2021 (remote)

**Breaking the Memory Wall in Current and Emerging Accelerators**
- IBM T.J. Watson Research Center, NY, USA, March 2019
- Microsoft Research, India, Dec 2018
- Intel Labs, India, Dec 2018
- Indian Institute of Science, Bangalore, India Dec 2018
- University of Pittsburgh, PA, Oct 2018

**Understanding and Addressing Throughput and Security Trade-offs in GPU-based Systems**
- HPCA SechArch workshop 2019, Washington DC, Feb 2019
- Intel, Hillsboro, May 2018
- Cisco, April 2018 (remote)

**Memory Bandwidth Management for Enhanced Throughput and Security in GPUs**
- North Carolina State University, April 2018
- George Washington University, April 2018

**Breaking the Memory Bandwidth Wall in GPUs**
- Virginia Commonwealth University (VCU), Feb 2016
- Indian Institute of Science, Bangalore, India, Dec 2015

**The Future of Parallel Computing with GPUs**
- William & Mary, Feb 2015
- University of Utah, Mar 2015
- Temple University, Mar 2015
- AMD Research, Mar 2015
- UC Riverside, Apr 2015
- Intel Labs, Apr 2015

## Open-Source Artifacts

– Artifact for ISPASS 2025 paper: Dissecting Performance Overheads of Confidential Computing on GPU-based Systems. This artifact can be found here: `https://github.com/insight-cal-uva/hcc-ispass25-artifact`. This artifact was evaluated by the ISPASS 2025 committee and received a badge.

– Artifact for ISSRE 2024 paper: Aspis: Lightweight Neural Network Protection Against Soft Errors. This artifact can be found here: `https://figshare.com/articles/code/Aspis/24757071`. This artifact was evaluated by the ISSRE 2024 evaluation committee and received a badge.

– Artifact for MICRO 2024 paper: Pushing the Performance Envelope of DNN-based Recommendation Systems Inference on GPUs. This artifact can be found here: `https://zenodo.org/records/13325108`. This artifact was evaluated by the MICRO 2024 evaluation committee and received a badge.

– Artifact for MICRO 2023 paper: Path Forward Beyond Simulators: Fast and Accurate GPU Execution Time Prediction for DNN Workloads. This artifact can be found here: `https://zenodo.org/records/8365078`. This artifact was evaluated by the MICRO 2023 evaluation committee and received all possible validation badges.

– Artifact for ISCA 2023 paper: Optimizing CPU Performance for Recommendation Systems At-Scale. This artifact can be found here: `https://zenodo.org/records/7957909`. This artifact was evaluated by the ISCA 2023 evaluation committee and received all possible validation badges.

– Artifact for ASPLOS 2020 paper: Why GPUs are Slow at Executing NFAs and How to Make them Faster. It is a fast NFA Engine for GPUs. This artifact can be found here: `https://github.com/insight-cal/gpunfa-artifact`. This artifact was evaluated by the ASPLOS 2020 evaluation committee and received all possible validation badges.

– Multiple Application Framework for GPU Architectures (MAFIA) is developed for supporting multiple application execution on GPUs. Currently, it supports 25 benchmarks from various benchmark suites (e.g., CUDA, Parboil, SHOC and Rodinia). From these benchmarks, one can construct 300 2-application workloads and 2300 3-application workloads. This artifact can be found here: `https://github.com/adwaitjog/mafia`

## Teaching at University of Virginia

Course syllabi and class-by-class schedule information is available here: `https://adwaitjog.github.io/teaching.html`

| | |
|---|---|
| Fall 2025 | CS 4444 – Intro to Parallel Computing (undergraduate course) |
| Spring 2025 | CS 6501 – GPU Architectures (special topics course) |
| Fall 2024 | CS 6354 – Computer Architecture (a required graduate course for CpE students) |
| Spring 2024 | CS 6501 – Modern Computer Architecture (special topics course) |
| Fall 2023 | CS 6354 – Computer Architecture (a required graduate course for CpE students) |
| Fall 2023 | CS 6190 – CS Perspectives Guest Lecturer (once) |
| Spring 2023 | Teaching Release. Guest lecturer (twice) in Prof. Skadron's Accelerators Course |

## Teaching at William & Mary

Course syllabi and class-by-class schedule information is available here: `https://adwaitjog.github.io/teaching.html`

| | |
|---|---|
| Fall 2022 | CSCI 424/524 – Computer Architecture (upper-level undergraduate course) |
| Spring 2022 | CSCI 424/524 – Computer Architecture (upper-level undergraduate course) |
| Fall 2021 | CSCI 780 – Non-conventional Computer Architecture (seminar graduate course) |
| Spring 2021 | CSCI 424/524 – Computer Architecture (upper-level undergraduate course) |

Fall 2020 CSCI 780 – Topics in Computer Architecture (seminar graduate course)
Spring 2020 CSCI 674 – GPU Architectures (upper-level graduate course)
Fall 2019 CSCI 424/524 – Computer Architecture (upper-level undergraduate course)
Spring 2019 CSCI 674 – GPU Architectures (upper-level graduate course)
Fall 2018 CSCI 424/524 – Computer Architecture (upper-level undergraduate course)
Spring 2018 CSCI 680 – GPU Architectures (upper-level graduate course)
Fall 2017 CSCI 424/524 – Computer Architecture (upper-level undergraduate course)
Spring 2017 CSCI 680 – Parallel Computer Architecture (upper-level graduate course)
Fall 2016 CSCI 424/524 – Computer Architecture (upper-level undergraduate course)
Spring 2016 CSCI 780 – GPU Architectures (seminar graduate course)
Fall 2015 CSCI 780 – Topics in Computer Architecture (seminar graduate course)

## Student Advising

### Ph.D. Students

2024– **Mohammad Sonji**, *Topic: Fine-grained GPU Partitioning*
University of Virginia

2023– **Yang Yang**, *Topic: Confidential Computing with GPUs*
University of Virginia

2023– **Amel Fatima**, *Topic: Multi-GPU/Chiplet Virtual Memory Management*
University of Virginia, Passed Dissertation Proposal in Spring 2025

2023–2025 **Divya Bagla**, *Topic: CXL Memory Systems*
University of Virginia, co-advised with Prof. Dwarkadas

2020-2022 **Ying Li**, *Topic: GPU Performance Simulation and Modeling*
moved to Prof. Yifan Sun's group at William & Mary in 2023

2017–2022 **Hongyuan Liu (graduated in Jan 2022; now Assistant Professor at Stevens Institute of Technology**, *Topic: Techniques for Accelerating Large-scale Automata Processing*, Internship at Intel Labs
Dissertation Link: `https://scholarworks.wm.edu/etd/1673275513/`

2016–2021 **Mohamed Assem Ibrahim (graduated in May 2021; now Member of Technical Staff at AMD Research)**, *Topic: Rethinking Cache Hierarchy and Interconnect Design for Next-generation GPUs*, Internship at AMD Research (twice)
Winner of the Distinguished Dissertation Award from William & Mary!, Dissertation Link: `https://scholarworks.wm.edu/etd/1627047836/`

2016–2021 **Gurunath Kadam (graduated in May 2021; now Member of Technical Staff at AMD Inc.)**, *Topic: Low-overhead Techniques for Improving GPU Security and Reliability*, Internship at Intel Labs
Dissertation Link: `https://scholarworks.wm.edu/etd/1627047873/`

2015–2020 **Haonan Wang (graduated in Aug 2020; now tenure-track Assistant Professor at San Jose State University)**, *Topic: Design and Analysis of Memory Management Techniques for Next-generation GPUs*, Internship at PNNL
Dissertation Link: `https://scholarworks.wm.edu/etd/1616444486/`

### M.S. Students

Fall 2019 **R. Xing**, *Topic: GPU Cache Management with Machine Learning*, MS Project, graduated

Spring 2017 **H. Wang**, *Topic: Value Approximation*, MS Project, graduated

Spring 2017 **R. Chai**, *Topic: DRAM Locality Analysis*, MS Project, graduated

Spring 2017 **C. Ames**, *Topic: DRAM Row-hammer Analysis*, MS Project, graduated

Summer 2016 **F. Luo**, *Topic: Parallelism Management in GPUs*, MS Project, graduated

### Undergraduate Students

2023– **Advising 50+ undergraduate students**

2023– **S. Kapa**, *Topic: Superconducting Architectures (involved in DOE project)*

2023– **P. Tangella**, *Topic: Superconducting Architectures (involved in DOE project)*

2023– **W. Kaiser**, *Topic: GPU Profiling and Benchmarking (NSF REU Student)*

2025– **B. Ruth**, *Topic: GPU Profiling and Benchmarking (NSF REU Student)*

2023–24 **H. Nguyen**, *Topic: GPU Programming (Dean's Fellowship)*

2020 **A. Chung**, *Topic: RISC-V Simulation with Python*, Monroe Scholar

2018 **S. Hong**, *Topic: GPU Security*, Charles Center Summer Research Scholar, graduated

2016–17 **C. Weinshenker**, *Topic: Chiplet-based Systems*, Honors thesis, graduated

2015–16 **R. Risque**, *Topic: Quantum Architectures*, Senior's thesis, graduated

## Professional Service

### Editorial Board / Leadership

May 2024– Associate Editor, IEEE Micro

Dec 2023– Associate Editor, IEEE Transactions on Cloud Computing (TCC)

2022– Executive Committee Member (`https://tc.computer.org/tcuarch/about/`), IEEE Computer Society Technical Committee on Microprogramming & Microarchitecture (TCuARCH)

### Tenure and Promotion Evaluator

2024, 2023 For several major universities in US

### Organizing Committees

2026 Workshop/Tutorial co-chair, ASPLOS 2026

2025 Workshop/Tutorial Chair, HPCA 2025

2023 ACM Student Research Competition (SRC) Judge, SIGMETRICS 2023

2023 Session Chair, SIGMETRICS 2023

2020 ACM Student Research Competition (SRC) Chair, PACT 2020

2020 Workshop co-chair, 13th Workshop on General Purpose Processing Using GPU (GPGPU 2020) @ PPoPP 2020

2020 Publicity co-chair, International Conference on Networking, Architecture and Storage (NAS)

2020 Session Chair, HPCA 2020

2019 Workshop co-chair, 12th Workshop on General Purpose Processing Using GPU (GPGPU 2019) @ ASPLOS 2019

2019 Publicity co-chair, International Symposium on Micro-architecture (MICRO)

2019 Publicity chair, International Symposium on Code Generation and Optimization (CGO)

2018 Local arrangements co-chair, International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)

2018 Workshop co-chair, 2nd Workshop on Minimizing Data Movement (Min-Move) @ASPLOS 2018

2017 Workshop co-chair, 1st Workshop on Minimizing Data Movement (Min-Move) @PACT 2017

2015 Proceedings Chair, International Symposium on Architectures for Networking and Communications Systems (ANCS)

### Funding Panel Reviewing

NSF Panelist: 2023 (once), 2022 (once), 2020 (once), 2019 (once), 2018 (two times), 2017 (four times), 2016 (once)

CCI/CVN: Workforce Development Reviewer: 2024

### Program Committees

2025– Program Committee Member, IEEE Micro Best Paper Award

2024– Program Committee Member, SIGMICRO nominations to Communications of the ACM's Research Highlights.

2024, '23, '21 Program Committee Member, International Symposium on Computer Architecture (ISCA)

2025, '23, '21, '20, '18 Program Committee Member, International Symposium on High-Performance Computer Architecture (HPCA)

2023, '21, '19 Program Committee Member, International Symposium on Micro-Architecture (MICRO)

2022 Program Committee Member, International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)

2022, '18 Program Committee Member, International Symposium on Performance Analysis of Systems and Software (ISPASS)

2021, '17 Program Committee Member, International Conference on Dependable Systems and Networks (DSN)

2025, '24, '19 Program Committee Member, International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)

2021 External Program Committee Member, International Symposium on Parallel Architectures and Compilation Techniques (PACT)

2020, '17 External Program Committee Member, International Symposium on Micro-Architecture (MICRO)

2020 External Program Committee Member, International Symposium on Computer Architecture (ISCA)

2025, '24, '23, '20 External Program Committee Member, International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)

2019, '17 Program Committee Member, International Symposium on High-Performance Parallel and Distributed Computing (HPDC)

2019, '18 Program Committee Member, International Conference on Computer Design (ICCD)

2023, '22, '18, '17, '16 Program Committee Member, Workshop on General Purpose Processing Using GPU (GPGPU)

2024,'23 Program Committee Member, International Workshop on Accelerating Analytics and Data Management Systems Using Modern Processor and Storage Architectures (ADMS), in conjunction with VLDB

2023 Program Committee Member, Workshop on Computer Architecture Modeling and Simulation (CAMS), in conjunction with MICRO

2016 Program Committee Member, International Conference on Parallel and Distributed Systems (ICPADS)

2016 Program Committee Member, International Conference on Supercomputing (ICS)

2016 Program Committee Member, International Conference on Parallel Processing (ICPP)

2016–2022 Program Committee Member, International Conference on Networking, Architecture and Storage (NAS)

### Journal Reviewing

IEEE Transactions on Computers (TC)

IEEE Computer Architecture Letters (CAL)

ACM Architecture and Code Optimization (TACO)

IEEE Transactions on Parallel and Distributed Systems (TPDS)

IEEE Transactions on Dependable and Secure Computing (TDSC)

ACM Transactions on Embedded Computing (TECS)

ACM Transactions on Design Automation of Electronic Systems (TODAES)

## Departmental and University-level Service

### Service to UVA (2023-onwards)

Distinguished Speaker Colloquium Committee [Chair] (2024–)

SEAS Reappointment Committee (2024–)

Computer Engineering Qualifying Exam (2024–)

Graduate Program Committee (2023–)

Distinguished Speaker Colloquium Committee Member (2023-2024)

Peer Review Committee (2023–)

Ph.D. Student Assessment Committee (2024–)

Graduate Admissions Committee (2023 calendar year)

Ph.D. Dissertation Committee (2023–) Member: A. Ahmed, L. Wu, O. Jaiyeoba, M. Baradaran, B. Gul, W. Choe, A. Kalita, S. Tajdari

### Service to W&M (2015-2022)

Web Presence Committee (chair: 2021–22)

Graduate Admissions Committee (2021–22, Spring 20, 2015–16)

Personnel Committee (2021–22)

Parking Appeals Committee Member (2021–22)

Undergraduate Research Committee Member (2021–22)

T. Jefferson Prize in Natural Philosophy, Committee Member, 2019

Faculty Hiring Committee (2019–20)

Graduate Research Symposium, Judge (2018–19)

Graduate Curriculum Committee (2019–21, 2017–18)

Colloquium Committee (2016–21, chair: 2018–19)

Awards and Prizes Committee (2017–19, chair: 2017–18)

Sanderson Award Committee Member for Undergraduate Mentoring (2016–17)

Pre-major Undergraduate Advising (2016–2021)

Ph.D. Dissertation Committee Member (in addition to my Ph.D. advisees): N. Carter, L. Ma, B. Nie, Z. Peng, R. Roy, L. Yang, T. Zhang

## Other Activities

| | |
|---|---|
| Fall 2024 | Mentor as a part of the Center for Diversity in Engineering (CDE) Mentorship Program, SEAS UVA |
| Fall 2024 | Served as a host for a week-long visit of Prof. Biswabandan Panda (and his student) from IIT Bombay, India at UVA. Trip Sponsored by IIT Bombay. |
| Fall 2023 | Served on a Panel on Future Faculty series for SEAS Graduate students. The panel was hosted by SEAS and moderated by Asst. Director of Graduate Success. |
| Fall 2024, 23 | Actively engaged with distinguished speakers: hosted one speaker, and met several distinguished speakers in person during their UVA visit. |
| Spring 2024, 23 | Actively participated in the Ph.D. student Recruiting Effort: took Zoom interviews of international applicants, met several prospective candidates in person during their UVA visit, nominated for prestigious UVA fellowships. |
| Spring 2024, 2023 | Actively participated in CS Faculty Recruiting Effort: hosted 3 candidates each year, interviewed many candidates in person, completed written reviews. |
| Spring 2022 | External Member on a Ph.D. Thesis Committee, IIT Guwahati, India |
| Spring 2021 | Judge, Cypher VI, W&M's 6th Annual Hackathon |
| Spring 2020 | Outreach Talk at University of Mary Washington |
| Spring 2019 | Hybrid Instructional Training (HIT), W&M, Spring 2019 |
| Summer 2018 | Summer School for Computer Architects (ACACES 2018) held at Fiuggi, Italy |
| 2018,2017 | Keynote Speaker, Focus on Future (FOF) Conference, School of Education, W&M |
| 2017 | Co-Editor/Committee Member, ETRI Journal, South Korea |

2017–18 Represented W&M at Center for Automata Processing, University of Virginia

2017–18 University (W&M) Teaching Project, Participant

## Professional Memberships

Senior Member of ACM and IEEE

Member of ACM Special Interest Group on Computer Architecture (ACM SIGARCH)

Member of ACM Special Interest Group on Programming Languages (ACM SIGPLAN)

Member of ACM Special Interest Group on Microarchitecture (ACM SIGMICRO)

Member of IEEE Computer Society

Member of IEEE Technical Community on Computer Architecture (IEEE TCCA)

Member of IEEE Technical Community on Microprogramming and Microarchitecture (IEEE TCuARCH)

Member of IEEE Technical Community on Dependable Computing and Fault Tolerance (TCFT)