Adwait Jog

Contact email: ajog@virginia.edu

APPOINTMENTS

University of Virginia, Associate Professor and Anita Jones Faculty Fellow, Computer Science, Jan 2023-Present
William & Mary, Adina Allen Associate Professor, Department of Computer Science, Aug 2021-Dec 2022
William & Mary, Assistant Professor, Department of Computer Science, Aug 2015-July 2021

EDUCATION

Ph.D. in Computer Science and Engineering, Pennsylvania State University, PA, USA, 2015
B.Tech. in Electronics and Instrumentation Engineering, NIT Rourkela, India, 2009

SELECTED PUBLICATIONS

1. Yang Yang and Jog, A. (2026). LÆGIS: Pinpointing and Addressing Performance Overheads of GPU-based Confidential Computing. In *the Proceedings of 53rd International Symposium on Computer Architecture **(ISCA)**, Raleigh, NC, Acceptance rate: 161/845 ∼19%*
2. Amel Fatima, Yang Yang, Sun, Y., Ausavarungnirun, R., and Jog, A. (2025). NetCrafter: Tailoring Network Traffic for Non-Uniform Bandwidth Multi-GPU Systems. In *the Proceedings of 52nd International Symposium on Computer Architecture **(ISCA)**, Tokyo, Japan, Acceptance rate: 127/570 ∼22%*
3. Li, Y., Bao, Y., Wang, G., Mei, X., Vaid, P., Ghosh, A., Jog, A., Bunandar, D., Joshi, A., and Sun, Y. (2025). 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, Acceptance rate: 127/570 ∼22%*
4. Yang Yang, Mohammad Sonji, and Jog, A. (2025). Dissecting performance overheads of confidential computing on gpu-based systems. In *IEEE International Symposium on Performance Analysis of Systems and Software **(ISPASS)**, Ghent, Belgium, Acceptance rate: 28/99 ∼28%*
5. Jain, R., Bhasi, V. M., Jog, A., Sivasubramaniam, A., Kandemir, M. T., and Das, C. R. (2024). 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%*
6. Ying Li, Sun, Y., and Jog, A. (2023). 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%*
7. Hongyuan Liu, Pai, S., and Jog, A. (2023). 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
8. Mohamed Assem Ibrahim, Kayiran, O., Eckert, Y., Loh, G. H., and Jog, A. (2021). 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%*
9. Mohamed Assem Ibrahim, Kayiran, O., Eckert, Y., Loh, G. H., and Jog, A. (2020). 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%*
10. Hongyuan Liu, Pai, S., and Jog, A. (2020). 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%*
11. Mohamed Assem Ibrahim, Hongyuan Liu, Kayiran, O., and Jog, A. (2019). 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%*
12. Hongyuan Liu, Mohamed Assem Ibrahim, Kayiran, O., Pai, S., and Jog, A. (2018). 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%*
13. Haonan Wang, Fan Luo, Mohamed Assem Ibrahim, Kayiran, O., and Jog, A. (2018). 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%*

14. Kayıran, O., Jog, A., Kandemir, M. T., and Das, C. R. (2013). 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%*
15. Jog, A., Kayiran, O., Mishra, A. K., Kandemir, M. T., Mutlu, O., Iyer, R., and Das, C. R. (2013b). 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%*
16. Jog, A., Kayiran, O., Chidambaram Nachiappan, N., Mishra, A. K., Kandemir, M. T., Mutlu, O., Iyer, R., and Das, C. R. (2013a). 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%*

AWARDS

1. UVA Outstanding Researcher Award, 2024
2. Senior Member of IEEE and ACM, 2023
3. Plumeri Award for Faculty Excellence from William & Mary, 2022
4. Google Research Scholar Award, 2021
5. NSF CAREER Award, 2018

STUDENT ADVISING

More details are available on our lab website: https://insight-cal.github.io/

1. **Current graduate students:** Y. Yang (Ph.D.), M. Sonji (Ph.D.)
2. **Graduated Ph.D. students:** 1) **Haonan Wang**, graduated in Aug 2020; now tenure-track Assistant Professor at SJSU, 2) **Mohamed Assem Ibrahim** (graduated in May 2021; now Senior Member of Technical Staff at AMD Research), 3) **Gurunath Kadam** (graduated in May 2021; now Senior Deep Learning Performance Architect at NVIDIA), 4) **Hongyuan Liu** (graduated in Jan 2022; now Assistant Professor at Stevens Institute of Technology), 5) **Amel Fatima** (graduated in May 2026; Member of Technical Staff at AMD).

MAJOR OPEN SOURCE SOFTWARE

1. Automata Acceleration on GPUs: https://github.com/insight-cal-uva/gpunfa-artifact
2. Multiple Application Framework for GPUs: https://github.com/insight-cal-uva/mafia
3. Superfast Confidential Computing with GPUs: https://github.com/insight-cal-uva/laegis-isca26-artifact

SERVICE TO PROFESSION

1. **Editor/Leadership:** IISWC 2026 PC co-chair; Associate Editor, IEEE Micro; Associate Editor, IEEE Transactions on Cloud Computing (TCC), IEEE Executive Committee Member of Technical Community on Microprogramming & Microarchitecture (TCuARCH)
2. **Major Conference Program Committee Member:** ISCA (2026, 2024, 2023, 2021); HPCA (2026, 2025, 2023, 2021, 2020, 2018), MICRO (2023, 2021, 2019), ASPLOS (2025, 2023, 2022)
3. **Organization committee member of conferences:** ASPLOS 2026 (Workshops/Tutorial co-chair), HPCA 2025 (Workshops/Tutorial Chair), SIGMETRICS 2023 (SRC Judge), PACT 2020 (ACM SRC chair), GPGPU 2019, 2020 (Workshop co-chair), MICRO 2019 (Publicity co-chair), CGO 2019 (Publicity chair), ASPLOS 2018 (Local arrangements co-chair), Min-Move 2017, 2018 (Founder and Workshop Co-Chair)
4. **Journal Reviewer for** IEEE Computer Architecture Letters (CAL), IEEE Transactions on Computers (TC), IEEE Transactions on Parallel and Distributed Systems (TPDS), ACM Architecture and Code Optimization (TACO), ACM Transactions on Embedded Computing (TECS), ACM Transactions on Design Automation of Electronic Systems (TODAES)
5. **Other:** Served on multiple NSF panels, departmental/university committees, and external tenure and promotion evaluator. Invited talks in academia (NCSU, GWU, Pittsburgh, VCU, Utah, Riverside, Merced, IITs, IISc) and industry (AMD, CC Summit, Cisco, IBM, Intel, Microsoft, Nvidia)