Aditya K Kamath

Website: akkamath.github.io

Email: akkamath@cs.washington.edu

Research direction: My projects revolve around reducing the impact of data movement. I'm currently working on (1) improving LLM inference latencies by designing better attention kernels and (2) reducing CPU-GPU data transfer in GNNs and DLRMs.

ACADEMIC QUALIFICATIONS

Year	Degree	Institute, City
2021 - Present	Ph.D. in Computer Science	University of Washington, Seattle
2015 – 2019	B.Tech. in Computer Science	National Institute of Technology Karnataka, Surathkal

PROFESSIONAL EXPERIENCE

University of Washington | Graduate Research Assistant

(Sep 2021 – Ongoing)



- Working on reducing data movement in contemporary applications [ISCA '24].
- Wrote CUDA kernels to reduce data movement, and worked with the gem5 simulator.
- Part of the Computer Systems Lab advised by <u>Prof. Simon Peter</u>.

Microsoft Research | Research Intern

(Jun 2024 – Ongoing)



- Working on improving LLM inference by designing better attention kernels for hybrid batches containing chunked prefills and decodes.
- Mentored by <u>Dr. Ashish Panwar</u>.

AMD Research | GPU-Centric Collectives Distributed Systems Research Intern

(Jun 2022 – Sep 2022)



- Worked on improving GPU-initiated collective communication.
- Improved ROC SHMEM's All-to-All communication collective using CUDA/HIP.
- Worked with the parallel and distributed programming team.

Indian Institute of Science | Research Assistant

(Jun 2019 - Aug 2021)



- Worked on enhancing race detection in GPUs. [ISCA '20, SOSP '21]
- Applied NVM to parallel architectures, i.e., GPU-enhanced persistent KVS and DB. [ASPLOS '22, '23]
- Worked under the guidance of Prof. Arkaprava Basu.

NOTABLE PUBLICATIONS

1. [ISCA '24] (MC)^2 : Lazy MemCopy at the Memory Controller [Paper] [Video]

Aditya K Kamath, Simon Peter

51st IEEE/ACM International Symposium on Computer Architecture

2. [ISCA '24] Scalable, Programmable and Dense: The HammerBlade Open-Source RISC-V Manycore
Dai Cheol Jung, Max Ruttenberg, Paul Gao, Scott Davidson, Daniel Petrisko, Kangli Li, Aditya K Kamath, et. al.

51st IEEE/ACM International Symposium on Computer Architecture

3. [ASPLOS '23] Scoped Buffered Persistency Model for GPUs [Paper] [Video]

Shweta Pandey*, Aditya K Kamath*, Arkaprava Basu

28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems

4. [ASPLOS '22] GPM: Leveraging Persistent Memory from a GPU [Paper] [Video]

Shweta Pandey*, Aditya K Kamath*, Arkaprava Basu

27th ACM International Conference on Architectural Support for Programming Languages and Operating Systems

5. [SOSP '21] iGUARD: In-GPU Advanced Race Detection [Paper] [Video]

Aditya K Kamath and Arkaprava Basu

ACM SIGOPS 28th Symposium on Operating Systems Principles

6. [ISCA '20] ScoRD: A Scoped Race Detector for GPUs [Paper] [Video]

Aditya K. Kamath*, Alvin A George*, Arkaprava Basu

47th IEEE/ACM International Symposium on Computer Architecture

*Authors contributed equally

TEACHING EXPERIENCE

Undergraduate Teaching Assistant at NITK Surathkal

(2018 - 2019)

- Taught a lesson on the functioning of a cache and modern cache replacement policies.
- Taught a lesson on Persistent Memory and possible future uses.
- Taught a lesson on importance of simulation in systems research, and how to use Intel PIN tool for tracing.
- Designed a project for students to create a working cache simulator.

VOLUNTEER SERVICE

- Grad Admission Reader at University of Washington: Reviewed applications of graduate school applicants.
- Pre-Application Mentorship Program (2022, 2023) at University of Washington: Guided students from historically
 marginalized groups through the graduate application process, revising their SOP and resume.
- **Head Placement Coordinator** at **NITK**: Responsible for directing the entire NITK campus hiring process for 2019. Managed dozens of Placement Coordinators and coordinated with HRs of hundreds of companies.
- **Co-Head** of **Algorithms Group** of Web Enthusiasts' Club at **NITK**: Organised competitive coding events in college. Gave talks on the basics of algorithms and optimisations.

TECHNICAL SKILLS

Programming Languages: C, C++, CUDA, Python

• Simulator Experience: gem5, GPGPU-Sim, SST, ns-3, ChampSim

Relevant Courses: Computer Organization and Architecture, High Performance Computing, Heterogeneous

Parallel Computing, Data Structures and Algorithms, Operating Systems