

Mohammad Sonji

Computer Science PhD Candidate

Research Interests

My research is broadly in computer systems, focusing on GPU architecture/kernel optimization, high-performance computing, and interconnect/memory systems.

Education

Aug 2024 - **Doctor of Philosophy, Computer Science, University of Virginia (UVA)**

May 2029 : CGPA: 3.9

Advisor: **Dr. Adwait Jog**

Jan 2022 - **Master of Science, Computer Science, American University of Beirut (AUB)**

Dec 2023 : CGPA: 3.78, Thesis title: Predicting If Executing Applications Are Near Completion

Advisor: **Dr. Izzat El Hajj**

2018 - 2021 : **Bachelor of Science, Computer Science, Beirut Arab University (BAU)**

CGPA: 3.76, Class rank: 3rd, Graduated with honors.

Experience

Research

Aug 2024 - **Research Assistant, UVA**

present : I joined **Dr. Adwait Jog's** Insight Computer Architecture Lab as a research assistant.

Jun 2022 - **Research Assistant, AUB & Hewlett Packard Labs, HPE**

Jul 2024 : I joined **Dr. Izzat El Hajj's** team in collaboration with Hewlett Packard Labs (HPE) led by **Dr. Dejan Mijojic** and his team.

Teaching

Teaching Assistant, UVA

Spring 26 : CS6501: GPU Architectures

Fall 25 : CS4444: Introduction to Parallel Computing (CUDA)

Teaching Assistant, AUB

Fall 23-24 : CMPS221: Computer Organization & Design

Spring 22-23 : CMPS202: Intermediate Programming with Data Structures

Fall 22-23 : CMPS224/CMPS396AA: GPU Computing (CUDA)

Spring 21-22 : CMPS200: Introduction to Programming

CMPS212: Intermediate Programming with Data Structures

Selected Academic Projects

2022 : **GPU Computing**

I designed a CUDA application in C/C++ for computing the Jaccard similarity among vertices in a graph, implementing four distinct versions of the code, each integrating one or more novel optimizations.

I was able to achieve the fastest execution time among my classmates.

2022 : **Compiler Construction**

Using the LLVM compiler and its Clang frontend, I implemented the following tasks in C++:

- Source-to-source compiler as a recursive AST visitor in Clang
- Code generator as a non-recursive AST visitor in Clang
- Aggressive Dead Code Elimination optimization.

Publications

UVA

ISPASS'25 **Dissecting Performance Overheads of Confidential Computing on GPU-based Systems**, Yang Yang, Mohammad Sonji, Adwait Jog , In the Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Ghent, Belgium, 2025

AUB

ICPE'26 **Are We There Yet? Predicting if Executing Applications are Near Completion**, M. Sonji, M. Baydoun, S. Diab, A. Nassereldine, P. Bruel, A. Dhakal, R. Enriquez, G. Rattihalli, D. Tootaghaj, G. Renaud, B. Chapman, F. Salem, E. Frachtenberg, D. Milojevic, I. Hajj, In the Proceedings of ACM/SPEC International Conference on Performance Engineering (ICPE), Florence, Italy, 2026

iWAPT'25 **Predicting Performance Variability**, M. Baydoun, M. Sonji, P. Bruel, D. Milojevic, E. Frachtenberg, I. El Hajj, In the Proceedings of IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 2025

GrAPL'25 **Serverless Graph Analytics on Multi-Instance GPU**, M. Sonji, M. Baydoun, A. Dhakal, G. Rattihalli, D. Milojevic, I. El Hajj, In the Proceedings of IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 2025

Awards

2024 : **Fellowship**, UVA

2024 : **Henri Qais Naccache Best Thesis Award**, AUB

2018-2021 : **Scholarships**, BAU

Technical skills

Parallel, Distributed, & GPGPU Programming: CUDA, MPI, Pthreads

Programming Languages: C, C++, Python, Bash

Systems & Tools: Linux, Docker, Kubernetes, Knative, Git

NVIDIA GPU Partitioning: Multi-Instance GPU (MIG), Multi-Process Service (MPS)

LLM Inference, vLLM, HuggingFace

Certificates

2025 : **Basic Responsible Conduct of Research Course**, UVA, CITI PROGRAM

2022 : **Physical Science Responsible Conduct of Research**, AUB, CITI PROGRAM

2020,2019 : **Lebanese Collegiate Programming Contest**, ACM, LCPC

Extracurricular Activities

2018 - 2020 : **Volunteer**, Red Cross youth sector