SARA MAHDIZADEH SHAHRI

Carnegie Mellon University
Department of Electrical and Computer Engineering
4720 Forbes Avenue, Pittsburgh, PA 15213

Google Scholar
Email: smahdiz@cmu.edu
Web: https://smahdizadeh.github.io

BRIEF BIOGRAPHY

My work bridges computer architecture and software systems, demonstrating the importance of that bridge in enabling efficient hyperscale web services with improved user experience via solutions that span the compute stack.

Today, it is critical to improve the experience of users who use modern hyperscale web services like web search and social media, to reduce user abandonment. Traditionally, the web systems that support these services typically adopt a "performance-first" approach. Indeed, my earlier work presents hardware solutions to address key hardware performance bottlenecks imposed by hyperscale services.

My research shows that building web systems with a "performance-first" approach often inadvertently affects certain user demographic groups disproportionately, hurting user experience. As examples, my work shows that modern web schedulers and databases often make performance optimizations (e.g., prioritizing requests, approximating responses) to improve the average latency, precipitating disproportionately poor responses for user groups in the tail end of the spectrum. To improve user experience, my work systematically identifies disproportionate responses across user groups, introduces metrics to measure them, and develops cross-stack systems to mitigate such disproportionate behaviors.

Looking ahead, I plan to extend this approach to other key components of large-scale web systems, including machine learning inference, to identify and mitigate instances where performance optimizations inadvertently degrade outcomes for certain user groups.

My research has been recognized with the 2025 Benjamin Garver Lamme/Westinghouse Graduate Fellowship, 2024 CyLab Presidential Fellowship, 2023 K&L Gates Presidential Fellowship, 2023 CMU College of Engineering Presidential Fellowship, 2023 Boeing Scholarship, 2022 Carnegie Institute of Technology Dean's Fellowship, and 2021 Rackham Merit Ph.D. Fellowship.

EDUCATION

Ph.D., Electrical and Computer Engineering — *Carnegie Mellon University*
Advisor: Prof. Akshitha Sriraman — *Aug 2022 - Present*
CyLab Presidential Fellowship; K&L Gates Presidential Fellowship; CMU CoE Presidential Fellowship
Dissertation title: Introducing Equitable Web Systems

Ph.D., Computer Science and Engineering — *University of Michigan*
Advisor: Prof. Baris Kasikci — *Aug 2021 - Aug 2022*
GPA: 4 out of 4
Rackham Merit Fellowship

M.Sc., Computer Science and Engineering — *Pennsylvania State University*
Advisor: Dr. Aasheesh Kolli — *Aug 2018 - Dec 2020*
GPA: 4 out of 4

B.Sc., Computer Engineering — *Sharif University of Technology*
Advisor: Prof. Hamid Sarbazi-Azad — *Sep 2013 - Feb 2018*
GPA: 18.51 out of 20
Ranked second in the Computer Hardware Engineering discipline

AWARDS AND HONORS

Benjamin Garver Lamme/Westinghouse Graduate Fellowship — *2025*
Awarded $100,000 towards tuition and stipend

CyLab Presidential Fellowship — *2024*
Awarded $50,000 towards tuition and stipend

K&L Gates Presidential Fellowship — *2023*
Awarded $108,000 towards tuition and stipend

CMU College of Engineering Presidential Fellowship *2023*
Awarded $50,000 towards tuition and stipend

Boeing Scholarship *2023*
Awarded partial tuition and stipend support

Carnegie Institute of Technology Dean's Fellowship *2022*
Awarded $83,000 towards tuition, stipend, and travel

University of Michigan Rackham Merit Fellowship *2021*
Awarded $92,000 towards tuition, stipend, and travel

Ranked 2^{nd} in Computer Hardware Engineering *2018*

Selected among the top 7 replacement policies in the 2^{nd} CRC *2017*
Secured a top place in the Cache Replacement Championship (CRC) co-located with ISCA 2017

Secured a place in the top 25% of entries in the FPGA National Contest *2016*

Ranked 201^{st} in the National University Entrance Exam *2013*

PEER-REVIEWED CONFERENCE/JOURNAL PUBLICATIONS

- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Title omitted to maintain anonymity.* ACM International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS '26**). 2026. (Under submission).
- Shixin Song, Tanvir Ahmed Khan, **Sara Mahdizadeh Shahri**, Akshitha Sriraman, Niranjan K Soundararajan, Sreenivas Subramoney, Daniel A Jimnez, Heiner Litz, Baris Kasikci. *Thermometer: Profile-guided BTB Replacement for Data Center Applications.* In proceedings of the 49^{th} International Symposium on Computer Architecture (**ISCA 2022**). Jun 2022. [link]
  Acceptance rate: 67/400 = 16.8%.

  Introduces the first Branch Target Buffer replacement technique to achieve near-ideal BTB performance by using program context information to inform BTB replacement decisions
- Akshay Krishna Ramanathan, **Sara Mahdizadeh Shahri**, Yi Xiao, Vijaykrishnan Narayanan. *Achieving Crash Consistency by Employing Persistent L1 Cache.* In proceedings of Design, Automation & Test in Europe Conference & Exhibition (**DATE 2022**). Mar 2022. [link]

  Achieving crash consistency through a 3D ferroelectric L1 cache design with near-zero performance overhead
- **Sara Mahdizadeh Shahri**, Armin Vakil Ghahani, Aasheesh Kolli. *(Almost) Fence-less Persist Ordering.* In proceedings of International Symposium on Microarchitecture (**MICRO 2020**). Oct 2020. [link]
  Acceptance rate: 82/424 = 19.3%.

  Lightweight x86 persistency extensions for efficient, fence-free ordering
- Seyed Armin Vakil Ghahani, **Sara Mahdizadeh Shahri**, Mohammad-Reza Lotfi-Namin, Mohammad Bakhshalipour, Pejman Lotfi-Kamran, and Hamid Sarbazi-Azad. *Cache Replacement Policy Based on Expected Hit Count.* In IEEE Computer Architecture Letters (**CAL 2017**). Oct 2017. [link]

  Reimagines cache replacement with a predictive policy driven by expected hit count

PEER-REVIEWED WORKSHOP PUBLICATIONS & POSTERS

- **Sara Mahdizadeh Shahri**, Martin Prammer, Jignesh Patel, Akshitha Sriraman. *Studying Differences Across User Query Groups In Vector Similarity Search.* USENIX Symposium on Operating Systems Design and Implementation (**Poster at OSDI 2025**). Jul 2025.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Studying Demographic Bias in Web Scheduling Systems.* The ACM Symposium on Operating Systems Principles (**Poster at SOSP 2024**). Nov 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Studying Demographic Bias in Web Scheduling Systems.* The 2nd Workshop on Hot Topics in System Infrastructure (**HotInfra 2024**) held in conjunction with the ACM Symposium on Operating Systems Principles (**SOSP**). Nov 2024.

- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Studying Demographic Bias in Web Scheduling Systems.* The SOSP Doctoral Workshop 2024 (**SySDW 2024**) held in conjunction with the ACM Symposium on Operating Systems Principles (**SOSP**). Nov 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Demographic Bias in Web Scheduling Systems.* The ACM Student Research Competition (**SRC 2024**) held in conjunction with the ACM Symposium on Operating Systems Principles (**SOSP**). Nov 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Studying Demographic Bias in Web Scheduling Systems.* Career Workshop for Inclusion and Diversity in Computer Architecture (**CWIDCA 2024**) held in conjunction with the IEEE/ACM International Symposium on Microarchitecture (**MICRO**). Nov 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Demographic Bias in Web Scheduling Systems.* **CyLab Partners Conference** Poster Session. Sep 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Studying Demographic Bias in Web Scheduling Systems.* USENIX Symposium on Operating Systems Design and Implementation (**Poster at OSDI 2024**). 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Studying Demographic Bias in Data Center Systems.* Workshop on Ethical System and Architecture Design (**HotEthics 2024**) held in conjunction with the International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS**). Apr 2024.
- **Sara Mahdizadeh Shahri**, Akshitha Sriraman. *Demographic Bias in Data Center Systems.* Young Architect Workshop (**YArch 2024**) held in conjunction with the International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS**). Apr 2024.
- **Sara Mahdizadeh Shahri**, Sahana Rangarajan, and Akshitha Sriraman. *Lifting the Systems Ostrich's Head from the Sand: Studying Demographic Bias in Data Center Systems.* Workshop on Hot Topics in System Infrastructure (**HotInfra 2023**) held in conjunction with the International Symposium on Computer Architecture (**ISCA**). Jun 2023.
- Sahana Rangarajan, **Sara Mahdizadeh Shahri**, Jaylen Wang, Pratyush Patel, and Akshitha Sriraman. *Designing Equitable Data Center Scheduling Systems.* Career Workshop for Inclusion and Diversity in Computer Architecture (**CWIDCA 2022**) held in conjunction with with the IEEE/ACM International Symposium on Microarchitecture (**MICRO**). Oct 2022.
- Zefeng Wang, **Sara Mahdizadeh Shahri**, Vyas Sekar, Assane Guane, and Akshitha Sriraman. *Designing Web Applications for Rural Communities.* Career Workshop for Inclusion and Diversity in Computer Architecture (**CWIDCA 2022**) held in conjunction with with the IEEE/ACM International Symposium on Microarchitecture (**MICRO**). Oct 2022.
- **Sara Mahdizadeh Shahri**, Shixin Song, Tanvir Ahmed Khan, Akshitha Sriraman, and Baris Kasikci. *Web Applications: Past, Present, Future.* Career Workshop for Inclusion and Diversity in Computer Architecture (**CWIDCA 2021**) held in conjunction with with the IEEE/ACM International Symposium on Microarchitecture (**MICRO**). Jun 2021.
- **Sara Mahdizadeh Shahri**, Aasheesh Kolli. *Delivering Correct and Fast Persistency Guarantees.* The First Young Architect Workshop (**YArch 2019**) held in conjunction with the IEEE International Symposium on High-Performance Computer Architecture (**HPCA**). Feb 2019. [link]

PROFESSIONAL EXPERIENCE

Graduate Research Assistant, Carnegie Mellon University *Aug 2022 - Present*
Advisor: Prof. Akshitha Sriraman

Introducing equality as a first-order hardware/software system design concern and designing the data center computing stack to enable efficient and impartial hyperscale web systems

Graduate Research Assistant, University of Michigan *Aug 2021 - Aug 2022*
Advisor: Prof. Baris Kasikci

Characterizing performance bottlenecks of emerging data center applications on modern processors to address these bottlenecks and make open-source web services more representative of real-world web services

Software Engineer Intern, Google *May 2022 - Aug 2022*
Team: Cloud Technical Infrastructure
Supervisors: Shay Gal-on, Tao Chen

Enabling early insights in the process of architecting future hardware by providing a framework for projecting the performance bottlenecks of applications on a new platform

Software Engineer Intern, Google *May 2021 - Aug 2021*
Team: Cloud Dataflow
Supervisors: Aaron Li, Yuta Labur

Improving the container startup latency for workers in Google Dataflow Service by initiating startup of containers right away, alleviating the need for entire container images to be pulled locally

Graduate Research Assistant, Pennsylvania State University *Aug 2018 - Aug 2021*
Advisor: Dr. Aasheesh Kolli

Architecting new hardware to enable fast recoverable data structures by leveraging emerging non-volatile memory technologies

Undergraduate Research Assistant, Sharif University of Technology *Sep 2016 - Feb 2018*
Advisor: Prof. Pejman Lotfi-Kamran, Prof. Hamid Sarbazi-Azad

Designing new cache replacement policies to improve the performance of modern processors

INVITED TALKS

Studying Demographic Bias in Data Center Systems

- Workshop on Hot Topics in System Infrastructure (**HotInfra 2024**) *Nov 2024*
- Parallel Data Lab Annual Workshop & Retreat (**PDL 2024**) *Oct 2024*
- Young Architect Workshop (**YArch 2024**) *Mar 2024*
- Workshop on Ethical System and Architecture Design (**HotEthics 2024**) *Mar 2024*
- CMU - K&L Gates Conference on Ethics and AI *Jun 2023*

Thermometer: Profile-Guided BTB Replacement for Data Center Applications

- Google *Aug 2022*

Data Center Applications: Past, Present, Future

- ADA Annual Symposium 2022 *May 2022*
- Career Workshop for Inclusion and Diversity in Computer Architecture (CWIDCA) *Oct 2021*

(Almost) Fence-less Persist Ordering

- International Symposium on Microarchitecture (MICRO) *Oct 2020*

Delivering Correct and Fast Persistency Guarantees

- The First Young Architect Workshop (YArch) *Feb 2019*

TEACHING EXPERIENCE

Graduate Teaching Assistant, Pennsylvania State University

- Graduate Computer Architecture, Dr. Aasheesh Kolli *Fall 2019*

Undergraduate Teaching Assistant, Sharif University of Technology

- Digital Systems Design, Prof. Alireza Ejlali *Fall 2017*
- Computer Structure and Language, Dr. Hossein Asadi *Fall 2017*
- Computer Architecture, Prof. Hossein Asadi *Spring 2016*
- Logic Design, Prof. Alireza Ejlali *Spring 2016*
- Advanced Logic Design, Prof. Alireza Ejlali *Fall 2016*

PROFESSIONAL SERVICE (INVITED)

Program Committee Co-Chair

- ASPLOS Wild and Crazy Ideas (**WACI**), Apr 2026.

Workshop Co-founder

- Hot Topics in Ethical Computer Systems (**HotEthics**) at ASPLOS, Apr 2024.

External Review Committee Member

- ACM Transactions on Architecture and Code Optimization(**TACO**), 2025.
- Architectural Support for Programming Languages and Operating Systems (**ASPLOS**), 2022.

Artifact Evaluation Committee Member

- Architectural Support for Programming Languages and Operating Systems (**ASPLOS**), 2024.
- USENIX Symposium on Operating Systems Design and Implementation (**OSDI**), 2022.
- USENIX Annual Technical Conference (**ATC**), 2022.
- IEEE/ACM International Symposium on Microarchitecture (**MICRO**), 2022.

TECHNICAL SKILLS

Programming Languages	C/ C++, Python, Shell, Go, Verilog, Assembly
System and Infrastructure	Low-level Systems Programming, Kubernetes, Docker Linux Systems, Performance Characterization, Scripting
Databases and Monitoring	PostgreSQL, Prometheus, Grafana
Machine Learning and Statistical Methods	Kernel Regression, Kernel Ridge Regression, PCA, k-Means Clustering, Logistic Regression, Support Vector Machines, Gaussian Mixture Models, Statistical Modeling
Hardware Simulators	Gem5, DRAMsim2, CACTI, ChampSim
Tools and Frameworks	LLVM, DynamoRIO, Pin, Linux perf, Intel PMU tools, Intel PT, gRPC, Google Protobuf

REFERENCES

Academia:

Prof. Akshitha Sriraman (akshitha@cmu.edu)
Assistant Professor, Carnegie Mellon University

Prof. Jignesh Patel (jigneshp@andrew.cmu.edu)
Professor, Carnegie Mellon University

Prof. Carlee Joe-Wong (cjoewong@andrew.cmu.edu)
Professor, Carnegie Mellon University

Prof. Baris Kasikci (baris@cs.washington.edu)
Associate Professor, University of Washington

Industry:

Dr. Aasheesh Kolli (aasheesh@google.com)
Research Scientist, Google

Shay Gal-on (shayg@google.com)
Research Scientist, Google

Tao Chen (taoc@google.com)
Research Scientist, Google

Muhammad Talha Imran (timran@google.com)
Software Engineer, Google

Aaron Li (aaronleeiv@google.com)
Research Scientist, Google

Yuta Labur (ylabur@google.com)
Research Scientist, Google