

Amir Noohi

PhD Candidate
University of Edinburgh

amirmnoohi@gmail.com | +44 7774 222 216 | noohi.org | linkedin.com/in/amirmnoohi | github.com/amirmnoohi

EDUCATION

PhD, Computer Science

Jan 2023 - Dec 2026

The University of Edinburgh, Edinburgh, UK
Advisor: Dr. Antonio Barbalace & Dr. Boris Grot

Master of Science, Computer Science

Sep 2019 - Aug 2021

Isfahan University of Technology, Isfahan, Iran
Advisor: Dr. Massoud Reza Hashemi
GPA: 4/4 - 19.01/20 · Ranked 1st

Bachelor of Science, Computer Engineering

Sep 2015 - Jun 2019

Isfahan University of Technology, Isfahan, Iran
Advisor: Dr. Elham Mahmoud Zadeh
GPA: 3.8/4 - 18.51/20 · Ranked 1st

PUBLICATIONS

- Amir Noohi, Antonio Barbalace, Artemiy Margaritov. "Warden: Sandboxing Every Agent Tool at Near-Fork Cost." ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2027. (Under Review)
- Amir Noohi, Artemiy Margaritov, Antonio Barbalace. "Aquifer: Following Live KV Production for Distributed Prefix Reuse." ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2027. (Under Review)
- Amir Noohi, Antonio Barbalace. "TENURE: Residency Leases for Tiered LLM KV Caches." ACM Symposium on Cloud Computing (SoCC), 2026. (Under Revision)
- Amir Noohi, Mostafa Derispour, Rodrigo Caetano de Oliveira Rocha, Antonio Barbalace, Artemiy Margaritov. "SagePTE: Co-Locating Guest and Host PTEs to Accelerate Virtual Memory Translation." IEEE/ACM International Symposium on Microarchitecture (MICRO), 2026.
- Amir Noohi, Bitia Asoodeh, Antonio Barbalace. "Towards a solution to the management scaling paradox in distributed LLM inference." Proceedings of the 6th Workshop on Machine Learning and Systems (EuroMLSys), 2026.
- Amir Noohi, Saba Ebrahimi, Mostafa Derispour, Antonio Barbalace. "NTSM: OS-Level Memory Middleware for Scaling AI Without Boundaries." ACM/IFIP International Middleware Conference (Middleware), 2026.
- Amir Noohi, Mostafa Derispour, Antonio Barbalace. "RMAI: Rethinking Memory for AI (Inference) In-Kernel Remote Shared Memory as a Software Alternative to CXL." Proceedings of the 5th Workshop on Machine Learning and Systems (EuroMLSys), 2025.
- Abbas Khater, Amir Noohi, Massoud Reza Hashemi, Zeinab Zali. "An NFV-Based Framework for Autonomous Deployment of New Protocols in SDN Networks." IEEE Access, 2024.

RESEARCH INTERESTS

- Computer Architecture
- CPU Architecture
- Systems for ML
- Operating Systems
- Distributed Systems
- HW/SW Co-design
- Virtualization

EXPERIENCE

CPU Researcher, Huawei Technologies Research & Development (UK) Ltd

May 2026 - Present

Researching CPU architecture optimizations for AI agentic workloads.

CPU Architect, Huawei Technologies Research & Development (UK) Ltd

Oct 2024 - Apr 2026

Leveraged the 128-bit extended PTEs developed during my internship to accelerate virtualized address translation.

EXPERIENCE (CONTINUED)

CPU Architect, Huawei Technologies Research & Development (UK) Ltd	Jul 2024 - Oct 2024
Extended Linux kernel page table entries (Level 3 entries) to 128-bit and modified QEMU to support this feature.	
CEO & Founder, HIGHHOST	Aug 2011 - Jan 2023
Founded and operated a comprehensive web hosting company offering dedicated servers, virtual private servers (VPS), shared hosting, cPanel hosting, reseller hosting, and domain registration services.	
Research Assistant, Isfahan University of Technology	Sep 2019 - Sep 2022
Designed and implemented a novel NFV-based framework for autonomous protocol deployment in SDN networks.	
Technical Product Manager, IAI Core Center	Feb 2020 - Jul 2020
Led the development of a patient-nurse communication platform for Iran Zamin Clinic, consisting of Android and web applications.	
Cloud Engineer, IUT ECE Department	Dec 2019 - Feb 2020
Infrastructure: Architected and deployed a production-grade private cloud serving 100+ physical servers across 5 data center racks.	
Software Developer, IUT ECE Department	Apr 2018 - Jul 2019
Developed the backend for Jinaro project using Laravel (PHP), implementing RESTful APIs and database architecture.	

TEACHING

University of Edinburgh	2023 - 2026
Lead Teaching Assistant — Operating Systems	
Lead Teaching Assistant — Programming for Data Science at Scale	
Teaching Assistant — Introduction to Computer Systems	
Lead Teaching Assistant — Operating Systems	
Lead Teaching Assistant — Programming for Data Science at Scale	
Lead Teaching Assistant — Operating Systems	
Teaching Assistant — Extreme Computing	
Teaching Assistant — Operating Systems	
Sharif University of Technology	2022 - 2023
Lecturer — Database	
Teaching Assistant — Data Mining	
Teaching Assistant — Advanced Software Engineering	
Isfahan University of Technology	2017 - 2020
Lead Teaching Assistant — Computer Networks	
Lead Teaching Assistant — Computer Networks	
Lead Teaching Assistant — Advanced Programming	
Teaching Assistant — Fundamental Programming	
Teaching Assistant — Computer Security	
Teaching Assistant — Computer Networks	
Lead Teaching Assistant — Advanced Programming	
Teaching Assistant — Basic Programming	
Lead Teaching Assistant — Advanced Programming	
Teaching Assistant — Basic Programming	
Teaching Assistant — Fundamental Programming	
Teaching Assistant — Basic Programming	
Teaching Assistant — Basic Programming	
Teaching Assistant — Basic Programming	
Teaching Assistant — Advanced Programming	

STUDENT MENTORSHIP

- **Hasan Alhussain** (MSc, University of Edinburgh) — Optimizing OS-level memory management for AI inference under memory pressure with CXL-based shared memory · Dec 2025 - Jun 2026
- **Arsh Kumar** (BSc, University of Edinburgh) — Optimizing OS-level memory management for AI inference under memory pressure with CXL-based shared memory · Nov 2025 - May 2026
- **Saba Ebrahimi** (MSc, Shahid Beheshti University) — Visiting intern at UoE working on thread migration project with Huawei · Aug 2024 - Dec 2025
- **Mostafa Derispor** (BSc, Isfahan University of Technology) — Distributed shared memory for AI workloads and SagePTE extended page table architecture · Jul 2023 - Nov 2025
- **Ali Farahani** (BSc, Sharif University of Technology) — Optimizing shuffle operations in distributed AI training through OS-level memory management, implementing zero-copy data movement and NUMA-aware buffer allocation to reduce inter-node communication overhead · Jan 2025 - May 2025
- **Zhengyi Chen** (UG Project, University of Edinburgh) — Coherency and consistency of modern DSM in Linux kernel · Sep 2023 - Mar 2024
- **Tom Blackburn** (UG Project, University of Edinburgh) — Comparison of user-space and kernel-space DSM · Sep 2023 - Jan 2024
- **Mohsen Ghasemi** (BSc, Sharif University of Technology) — Workload characterization and performance analysis for Distributed Shared Memory (DSM) systems, profiling memory access patterns, page fault rates, and coherence protocol overhead across AI and HPC workloads · Sep 2023 - Dec 2023
- **AmirHossein Barati** (BSc, Sharif University of Technology) — Design and implementation of kernel-level Distributed Shared Memory (DSM) subsystem, including page fault handlers, remote memory access protocols, and coherence mechanisms using RDMA for low-latency data transfer · Jun 2023 - Aug 2023

HONORS & AWARDS

- **Ph.D. Scholarship Award** — School of Informatics, University of Edinburgh, 2023
- **Ranked 1st in M.Sc.** — Isfahan University of Technology, 2021
- **Ranked 1st in B.Sc.** — Isfahan University of Technology, 2019
- **Direct Admission to M.Sc.** — Isfahan University of Technology, 2019
- **National Exceptional Talents Scholarship** — Iran's National Foundation for Exceptional Talents, 2019
- **Top 1% in University Entrance Exam** — Iranian National Exam, 2015
- **Physics Olympiad - Ranked 7th** — National Physics Olympiad, 2014
- **Math Olympiad - Ranked 5th** — National Math Olympiad, 2013

TALKS & SERVICE

- *"Q3 2026 Huawei Workshop"* — Huawei Technologies Research & Development, Q3 2026
- *"Q2 2026 Huawei Workshop"* — Huawei Technologies Research & Development, June 2026
- *"Towards a solution to the management scaling paradox in distributed LLM inference"* — EuroMLSys 2026 Workshop, May 2026
- *"Q4 2025 Huawei Workshop"* — Huawei Technologies Research & Development, December 2025
- *"RMAI: Rethinking Memory for AI Inference"* — EuroMLSys 2025 Workshop, March 2025
- *"Q4 2024 Huawei Workshop"* — Huawei Technologies Research & Development, December 2024
- Student Volunteer, EuroSys 2026 (May 2026)
- Artifact Evaluation Committee, EuroSys 2025 (Fall 2025)
- Artifact Evaluation Committee, EuroSys 2025 (Spring 2025)
- Shadow Program Committee, EuroSys 2025 (March 2025)

SKILLS

- **Languages:** C/C++, Python, Rust, PHP, Java, JavaScript/TypeScript, Bash/Shell, SQL, Assembly (x86/ARM64)
- **Systems Programming:** Linux Kernel Development, QEMU/KVM, Memory Management, Device Drivers
- **Virtualization & Cloud:** KVM/QEMU, VMware vSphere/ESXi, Docker/Containers, OpenStack, Xen, Hyper-V
- **Databases:** MySQL/MariaDB, PostgreSQL, MongoDB, Redis, SQLite
- **Networking:** Wireshark, tcpdump, iptables/nftables, BGP/OSPF, SDN/OpenFlow, Mininet, DPDK, NS2/NS3
- **ML & Data Science:** PyTorch, vLLM, Ray / Ray Serve, Hugging Face Transformers, DeepSpeed, NCCL, CUDA/cuDNN, LMCache, TensorFlow, Pandas/NumPy, Hadoop/Spark