ARSHIA AKHAVAN

Undergraduate Student of Computer Engineering

@ letmemakenewone@gmail.com +98 933 457 4957 Tehran, Iran
in arshia-akhavan ArshiAAkhavan

EDUCATION

B.Sc. in Computer Engineering

Sharif University of Technology

Sep 2018 – ongoing

RESEARCH INTERESTS

- Operating System Design and Implementations
- Distributed Systems and Parallel Computing
- Programming Languages and Formal Verification
- Cloud Computing
- High-Performance Computing
- Computer Networks

RESEARCH EXPERIENCE

Enhancing BeeGFS Metadata IO Latency

HPC Lab

Oct 2022 – currently Sharif University of Technology

My research project aimed to improve the performance of the parallel filesystems employed in our university's High-Performance Computing (HPC) center, specifically focusing on enhancing metadata management within the BeeGFS system. To achieve this, I initiated the project by categorizing the various workloads utilized on the HPC and conducting a comprehensive analysis of their metadata latency. Subsequently, I am now implementing a caching mechanism on the client nodes to ameliorate metadata latency issues. After successful implementation and deployment of the MVP in the HPC environment, our Implementation resulted in a notable reduction of overall I/O latency.

Support for Ordered List on Big buckets in Ceph

HPC Lab

May 2022 – currently Sharif University of Technology

My research project focuses on addressing a critical challenge in modern data pipelines, where Apache Spark serves as the processing engine. Historically, Spark applications have relied on Apache HDFS as their storage backend, but HDFS falls short in terms of cloud-native adaptability and scalability for large-scale operations. As a solution Spark adds support for the S3 protocol as an alternative.
Although Cloud-native, it presents a bottleneck due to the extensive list operations inherent in Spark's shuffling between stages. The core issue arises from S3's ordered listing, placing significant strain on S3 providers like Ceph, especially when dealing with buckets containing a multitude of files. This particular issue hinders the migration of Big Data applications from HDFS to S3 providers. My research goal is to mitigate this problem within the Ceph ecosystem by implementing caching mechanisms and preprocessing techniques, ultimately facilitating a smoother transition for Big Data applications to leverage the benefits of S3 while maintaining efficient and scalable data processing.

INTERNSHIPS

Semantic memory garbage collection for concurrent lock-free data structures

Kaist CP

May 2023 – Aug 203 Kaist University

My internship focuses on **memory model for atomic operations** and Design of **lock-free data structures** for high concurrency workloads. I start by learning and implementing different models of Locks and then went on implementing lock-free data structures suck as **Hash Map based on Split-Ordered List** and other data structures. As the next and necessary step, I studied different methods for **Semantic Memory Reclamation** for these data structures and learn many schemes such as **hazard pointers** and **EBR**. I also was given the chance to work with professor Kang's recent published scheme, HP++ which was a improvement over Hazard pointers. it was also a good opportunity for me to get better in rust programming language

PROJECTS

CS162 Berkeley's Pintos Project

Developed an operating system kernel, addressing challenges like system call implementation, filesystem management, and process scheduling, showcasing expertise in OS design and low-level programming. (git)

Matrix Vector multiplication using AVX

Develop a library for multiplication of sparse Matrices on top of Intel's AVX vector processors to further increase the performance (git)

TCP network stack

Implement a TCP FSM with Rust using the Linux kernel's TUN/TAP based on RFC 793

C minus Compiler

Implement a Compiler for C minus language (git)

Hash Map based on Split-ordered list

Implement a Concurrent HashMap based on Split-ordered list paper in Rust (git) (paper)

Hazard pointers

Hazard pointer implementation using Rust (git)

INDUSTRIAL EXPERIENCE (SUMMERIZED)

Software Engineer

Sotoon

Jan 2023 – Currently Tehran, Iran

Sotoon is a B2B cloud provider which focuses on different aspects of Cloud Computing, such as hardware abstraction, Providing storage servers, DNS, CDN, Bigdata solution and managed k8s cluster. I worked in the K8s as a Service team. Experiences:

- **Hiding Control-plane** from customers and fully manage it.
- Providing **Additional cluster plugins** such as DNS and CNI as an automated solution.
- Develop **our own CSI** to better interact with our storage team.
- Developing **AutoScaler** for k8s cluster to reduce costumers costs.
- Develop our own **CCM** module to better interact with the underlying compute infrastructure

Data Platform Engineer

Digikala

Aug 2021 – Jan 2023 Tehran, Iran

Digikala is the largest e-commerce company in Iran with 40+ million users. Experiences:

- Developing **Karavan**, a **Parallel Processing Engine** for data pipelines on top of **Apache Spark**.
- **Reducing Computation time** of our Pipelines by **revising our data retrieval models and Queries**.
- **Reducing latency** of our API by locating **critical paths** in the codebase of the framework and **optimizing** them.

DevOps Engineer

Tapsell

Nov 2019 – Aug 2021 Tehran, Iran

Providing the infrastructure needed by other technical teams in the corporation.
Experiences:

- Optimizing Cassandra's **Disk usage** for different workloads
- Developing a fully automated backup, validity check, and restore tool for managing our backups in a cloud native infrastructure.

VOLUNTEER EXPERIENCE

Technical Chapter

Student's Scientific Chapter (SSC)

Aug 2020 – Aug 2021 Tehran, Iran

Maintain Technical resources owned by SSC and **supervise technical requirements** of events held by SSC

Chief of Infrastructure and System Team

Sharif AI Challenge 2021

Dec 2020 – May 2021 Tehran, Iran

Develop and deploy a **scalable infrastructure/software** for AI challenge's **high demanding game engine**. GitHub

Chief Technical Staff

ICPC

May 2023 Tehran, Iran

Build contest-customized OS images and distribute them on contestant hosts using **Network boot**.
Manage and configure hosts used by contestants and **monitor their connectivity**.
Restrict external access and host to host access of each contestant's host device.

Chief Technical Staff

ICPC west championship

Dec 2019 Tehran, Iran

Manage and configure hosts used by contestants and **monitor their connectivity**

TEACHING ASSISTANT EXPERIENCES

- Course Desgin - **Zero to Hero Linux** - Quera - link
- Head TA - **Operating System** - Dr. Mehdi Kharrazi - Spring 2023
- TA - **Operating System** - Dr. Mehdi Kharrazi - Spring 2022
- TA - **Computer Network** - Dr. Mehdi Jafari - Spring 2023
- TA - **Computer Network** - Dr. Mehdi Jafari - Fall 2022
- TA - **Automata Theory** - Dr. Ali Movaghar - Spring 2023
- Head TA - **Multi-Core Computing** - Dr. Hajar Falahati - Fall 2023
- Head TA - **Multi-Core Computing** - Dr. Hajar Falahati - Spring 2023
- TA - **Multi-Core Computing** - Dr. Hajar Falahati/Dr. Hamid Sarbazi - Spring 2022
- TA - **Computer Structure and Language** - Dr. Laleh Arshadi - Spring 2021
- TA - **Numerical Computation** - Dr. Samira Hossein ghorban/ Dr. Hamid Sarbazi - Spring 2023
- TA - **Design Of Database** - Dr. Mojtaba Varmazyar - Fall 2022
- TA - **Big Data** - Dr. Sharare Ali pour - Spring 2023
- Head TA - **Advanced Programming** - CE department - Spring 2021
- TA - **Advanced Programming** - CE department - Spring 2020
- Head TA - **Fundamentals of Programming** - Dr. Mohammad Amin Fazli - Fall 2022
- Head TA - **Fundamentals of Programming** - Dr. Mohammad Amin Fazli - Fall 2021
- TA - **Fundamentals of Programming** - Dr. Fakoori - Fall 2020
- TA - **Fundamentals of Programming** - CE department - Fall 2019

PROGRAMMING LANGUAGES

Rust	●●●●○
Go	●●●●○
C	●●●○○
Bash	●●●●●
Python	●●●●●
Java	●●●○○
SQL	●●●●●

TECHNOLOGIES

CNI · CSI · AVX · Cuda · OpenMP · ryon-rs · crossbeam-rs · k8s · Docker · Containerd · Docker-swarm · Nginx · HAproxy · GitlabCI · ArgoCD · Spark · Airflow · Ceph · minIO · Kafka · Redis · Cassandra · MongoDB · Ansible · terraform

LANGUAGES

English	●●●●○
French	●●○○○
Persian	●●●●●

SELECTED COURSES

- Operating Systems 20/20
- Computer Network 20/20
- Database Design 20/20
- Compiler Design 20/20
- Automata theory 20/20
- Computer Architecture 20/20
- Multi-Core Computing 20/20
- Real-time Systems 20/20