# Haseeb Ashfaq

haseeb.luminite@gmail.com | New York City, NY, 10009
Anticipated Graduation Date: 11/2026

## EXPERIENCE

**Google** **May 2025 – August 2025**
*Software Engineering Intern* California, USA

- Part of AI and Infrastructure organization, worked on improving robustness of ML infrastructure
- Developed a tool, *hstprof*, for profiling GPU/TPU workloads using high frequency network telemetry
- *hstprof* enabled fine grained view of the network traffic of ML training jobs, reporting utilization at 100 microseconds level
- Used *hstprof* to analyze Gemini training cluster to investigate long tail latencies
- Found imbalance of packet queues across memory banks of switches where one bank had 10x higher load than the other

**Nokia** **June 2023 – August 2023**
*Networking Research Intern* New Jersey, USA

- Developed (in C++, Unix) a streaming service for AR/VR content for heterogeneous networks
- Implemented a resource-efficient transcoding mechanism for volumetric videos that achieved 75% CPU savings
- Developed an encoder/decoder for point cloud data that can tolerate packet losses in the network which enabled utilizing unreliable transport protocol (UDP) instead of TCP for point cloud streaming
- Implemented a mixed-reliability transmission protocol using QUIC streams and datagrams (with Cloudflare's Quiche)

**Systems Group NYU** **June 2022 – June 2023**
*Graduate Research Assistant* New York, USA

- Developed a trace-aware access control for microservices, implemented via Istio Envoy proxies.
- Designed and implemented a special priority queue, LOQ, for cloud hosted financial exchanges, that enhances a matching engine's throughput by up to 150% and lowers latency by 90%.
- Developed a cloud-native multicast service for market data that achieves 50% lower latency and better scalability than AWS TGW-based multicast. Prototyped in C++ and evaluated on AWS and GCP.

**PosterMyWall** **June 2020 – August 2021**
*Software Engineer (Full Time)* Lahore, PK

- Designed and implemented, in PHP and JS, an access control system for internal tools of the company
- Setup CI/CD pipeline along with testing infrastructure using TeamCity and AWS
- Automated AWS-hosted development infrastructure, shortening the testing cycle time by more than 50%
- Secured the product website by eliminating critical vulnerabilities (XSS, CSRF, IDOR) and did backend development
- A recommendation letter from my manager describing me as an exceptional engineer is available on LinkedIn

## EDUCATION

**PhD and MS, Computer Science** **Sept. 2021 – May 2026**
*New York University, New York, USA* GPA: 4.0/4.0
Research Interests: Distributed Systems, Networks, Cloud Computing, Financial Technologies, AI Infrastructure

**Bachelor of Science, Computer Science** **Sept. 2016 – May 2020**
*Lahore University of Management Sciences, Lahore, Pakistan* GPA: 3.7/4.0
Courses: Algorithms, Data Structures, Distributed Systems, Computer Networks, Machine Learning

## SELECT PROJECTS

**Network Support For Scalable Cloud Hosted Financial Exchanges**

- Implemented a low latency market data service that achieves less than 1-microsecond latency difference across receivers
- Utilized kernel bypass and zero-copy packet replication techniques to enable fast packet processing, implemented in C++
- Utilized eBPF/XDP and eBPF/TC for efficient packet processing when using Linux kernel

**Codesign Of Tensors Encoding And Transcoding For Decentralized ML**

- Designed and implemented a mechanism for packing tensors in network packets that enable a resource efficient dissemination mechanism, akin to Scalable Video Codec but for tensors

- Enabled utilizing overlay multicast for distributing training data across geo-distributed heterogeneous clients
- Reduced memory utilization by 30% and increased throughput of data dissemination by 25%

## RESEARCH PAPERS

**Design and Implementation of a Scalable Financial Exchange in the Public Cloud**
Accepted for publication by **ACM Sigcomm'25**, [Arxiv Link], Cited by **Jane Street** in their research paper

**A Scalable and Fair Multicast for Financial Exchanges in the Cloud**
ACM Sigcomm Demos & Posters (Presented a poster in Sydney, Australia) [Link]

**QuEST: Fast, Expressive, and Cheap Analytics for Distributed Traces Using Cloud Storage**
*CloudDB, a VLDB workshop* [Link]

**To Block or Not To Block: Accelerating Mobile Webpages On-The-Fly Through JavaScript Classification**
*ICTD 2022 (Presented the paper in Seattle, Washington)* [Link]

**Using Application Layer Banner Data To Automatically Identify IoT Devices**
*ACM Sigcomm CCR 2020* [Link]

## INVITED TALKS

I have been invited to give talks about my work on low latency and scalable systems in the cloud.

**Rutgers University:** Network support for cloud hosted financial exchanges. 30/10/2024
**Google:** How to build an ultra-fast and scalable financial exchange on the public cloud? 12/03/2024

## AWARDS, FELLOWSHIPS AND SERVICES

**National Science Foundation (NSF) Travel Grant**
Funds for traveling to ACM Sigcomm 2024 in Sydney, Australia

**Outstanding Student Research Award**
*Granted by Nokia Bell Labs during Global Student Program 2023*

**HotNets Travel Grant**
*Funds for traveling to ACM HotNets in Boston, United States*

**Patent: A Method To Enable Fast Transmission And Processing Of 3D Telepresence Data Encoded As Octrees**
*Approved by Nokia's internal board, In submission to USPTO, Received Monetary Award from Nokia, Link*

**Reviewer for ACM Journal on Computing and Sustainable Societies (JCSS)**
*Served as a reviewer for research articles submitted to ACM JCSS*

**MacCracken Fellowship**
*Granted by New York University for a Ph.D. in Computer Science*

## SKILLS

C/C++, Python, PHP, SQL, Go, Javascript, React/React Native, Rust, AWS, Debugging, Testing, DPDK, eBPF, Linux, Kubernetes, Docker, Istio, Microservices, Congestion Signalling (CSig), High Frequency Network Telemetry, System Design

## MISC.

**LinkedIn:** https://www.linkedin.com/in/haseeb-ashfaq-66248213b
**Personal Site**: https://haseebashfaq.com
**Legal Name**: Muhammad Haseeb

**GitHub:** https://github.com/HaseebLUMS
**Phone:** +1 (646) 240-6375