# Haseeb Ashfaq

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

## EXPERIENCE

**Google**                                                          May 2025 – August 2025
*Software Engineering Intern, Ph.D.*                                        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 speed network telemetry
- *hstprof* enabled a fine grained view of the network traffic of ML jobs, reporting utilization at 100 microseconds timescale
- Used *hstprof* to analyze Gemini training clusters to investigate several regressions including packet drops and high latencies
- Found imbalance of packet queues across memory banks of switches which caused performance degradation

**Nokia Bell Labs**                                                 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

- 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% while retaining fairness among market participants
- 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

## EDUCATION

**PhD and MS, Computer Science**                                    Sept. 2021 – Aug. 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

## PUBLICATIONS

**Beyond Lamport, Towards Probabilistic Fair Ordering**
ACM HotNets'25
**Haseeb** Ashfaq, Jinkun Geng, Radhika Mittal, Aurojit Panda, Srinivas Narayana, Anirudh Sivaraman

**Network Support For Scalable And High Performance Cloud Exchanges**
ACM SIGCOMM '25, Cited by **Jane Street**, **Google** and **Microsoft Research**
**Haseeb** Ashfaq, Jinkun Geng, Daniel D-Cavalcanti, Xiyu, Hao, Ulysses Butler, Radhika Mittal, Srinivas Narayana, Anirudh Sivaraman

**ParserHawk: Hardware-aware parser generator using program synthesis**
ACM SIGCOMM '25
Xiangyu Gao, Jiaqi Gao, Karan Kumar, **Haseeb** Ashfaq, Enan Zhai, Bili Dong, Joseph Tassarotti, Srinivas Narayana, Anirudh Sivaraman

**Towards Efficient Transmission of 3D Point Clouds Via Adaptive Encoding and QUIC**
Emerging Multimedia Systems (EMS)'25, a SIGCOMM workshop
**Haseeb** Ashfaq, Eugene Chai, Matteo Varvello

**Codesign of Tensors Encoding And Transcoding**
Networks for AI Computing (NAIC)'25, a SIGCOMM workshop
Revant Teotia, **Haseeb** Ashfaq

**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 Bell Labs*
**Haseeb** Ashfaq, Eugene Chai, Matteo Varvello

**QuEST: Fast, Expressive, and Cheap Analytics for Distributed Traces Using Cloud Storage**
*CloudDB, a VLDB workshop*
Jessica Berg, **Haseeb** Ashfaq, Haiming Chen, Yaojia Ju, Anirudh Sivaraman, Ravi Netravali and Srinivas Narayana

---

## PROJECTS

**Design and Implementation of a Scalable Cloud Exchange**

- 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 transcoding (creating various bitrate versions) 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%

---

## INVITED TALKS

**HotNets'25:** Probabilistic Fair Ordering of Events. 11/18/2025
**Open Compute Project, OCPTAP:** Network Support For Scalable And High Performance Cloud Exchanges. 10/22/2025
**Google:** How to build an ultra-fast and scalable financial exchange on the public cloud? 12/03/2024
**Rutgers University:** Network support for cloud hosted financial exchanges. 10/30/2024

---

## AWARDS, FELLOWSHIPS AND SERVICES

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

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

**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 Speed Network Telemetry, System Design

---

## MISC.

**LinkedIn:** https://www.linkedin.com/in/haseeb-ashfaq-66248213b          **GitHub:** https://github.com/HaseebLUMS
**Personal Site**: https://haseebashfaq.com                                **Phone:** +1 (646) 240-6375
**Legal Name**: Muhammad Haseeb                                            **Available for full-time**: September 2026