HAMED SEYEDROUDBARI

US Citizen

hamed@gatech.edu | +1 (424) 777-5443 | hamedsey.github.io

OBJECTIVE

Seeking a full-time opportunity to research and develop forms of in-network computing to accelerate latency-sensitive datacenter workloads.

EDUCATION

2019 – pres. Ph.D. - Electrical and Computer Engineering – **GPA 4.0**

- Georgia Institute of Technology, Atlanta, GA
- Research Advisor: Dr. Alexandros Daglis
- Topic: Leveraging SmartNICs to Accelerate Latency-Critical Online Services.

2015 - 2019 B.S. - Computer Engineering – **GPA 3.93**

- California State University, Northridge (CSUN), Northridge, CA

EMPLOYMENT

2020 – pres. Graduate Research Assistant – Georgia Tech, Atlanta, GA

Exploring in-network compute offloads onto SmartNICs

1. Load Balancing µs-scale RPCs (published at HPCA'23)
   - Insight: µs-scale RPCs require immediate load imbalance detection and remediation mechanism that operates on a per-packet granularity.
   - Adaptive load balancing intelligently steers packets to user space queues at line rate, improving throughput under SLO and reducing tail latency.
   - Implemented on Mellanox Innova Flex-4 FPGA SmartNIC
   - Evaluated on RDMA (UD/UC/RC) microbenchmark and Masstree-RDMA key-value store with various service time distributions.

2. Data Movement (on ArXiv)
   - Insight: Moving excess packets across PCIe increases PCIe contention and amplifies PCIe interface latency.
   - Shallow network functions only require packet header for processing.
   - Implemented FPGA emulation of mechanism to slice packet into header and payload on ingress, sending only the header to the host, storing payload on the NIC. Payload is spliced back to header on egress.
   - Enabling payload slicing eliminates PCIe data movement bottlenecks between the NIC and server, reducing tail response latency.

3. Notification mechanism to assist CPU in handling RDMA connections at scale
   - Insight: RDMA's polling mechanism of handling many connections presents a triangle tradeoff between idle polling, inter-core synchronization, and load imbalance overheads, which are prohibitive in handling µs-scale services.
   - Designed, implemented SmartNIC mechanism to notify cores of active connections, and balance connections across CPU cores while maintaining connection affinity.

2024 Research Scientist – Microsoft (Remote)

- Optimized RDMA networks to extract peak network utilization for AI workloads by investigating loss recovery mechanisms when packets arrive mis-ordered.
- Developed simulator in SystemC to simulate and analyze NIC-based loss recovery mechanisms (i.e., go-back-n, selective retransmit) for a RDMA/RoCE network.
- Analyzed performance by tuning simulator knobs (e.g., receiver queue depth, memory latency, link latency, timeout) and observing metrics such as flow completion time and tail latency.
- Proposed NIC architectural optimizations to enable support for load balancing, congestion control, recovery from multi-path packet losses.

2022 Research Intern (Systems) – Arm Research, Austin, TX

- Developed high performance RDMA-based microbenchmark to identify system bottlenecks (outperforming state-of-the-art perftest from NVIDIA UCX library).
- Characterized performance of BlueField-2 DPU in terms of peak sustainable throughput and communication latency and gauged its potential as a candidate for accelerating latency critical workloads.
- Demonstrated the Bluefield-2 DPU can boost server's peak sustained throughput by 66% without incurring additional cost for cloud service provider.

TECHNICAL SKILLS

Programming Language	: C, C++, SystemC, Python
HDL	: Verilog, SystemVerilog
Toolchain	: Xilinx Vivado, HLS, Vitis
NIC Platforms	: Mellanox Innova Flex-4 SmartNIC, Bluefield-2 DPU
Computer Networks	: Ethernet, RDMA, RoCE, IBVerbs, BGP, TCP/IP, UDP

FELLOWSHIPS/AWARDS

2019	President's Fellowship, Georgia Tech
2019	Outstanding Graduating Senior Award, CSUN
2018	Robert Sprague Foundation Scholarship, CSUN
2017, 2018	University Scholarship, CSUN

PUBLICATIONS

Sassy: SmartNIC-Assisted Notification Delivery for µs-scale RDMA Workloads (In review)
H. Seyedroudbari, A. Daglis

Turbo: SmartNIC-enabled Dynamic Load Balancing of µs-scale RPCs
H. Seyedroudbari, S. Vanavasam, A. Daglis
IEEE International Symposium on High-Performance Computer Architecture (HPCA 2023)

NFSlicer: Data Movement Optimization for Shallow Network Functions
A. Sarma, **H. Seyedroudbari,** H. Gupta, U. Ramachandran, A. Daglis
ArXiv, March 2022