Harshit Gupta

PhD Candidate

harshitg@gatech.edu | 4045129792 | https://sites.cc.gatech.edu/~hgupta40/ | hagupta | harshitgupta1337

EDUCATION

Ph.D. in Computer Science, Georgia Institute of Technology, Atlanta 2016 – 08/2022
- Ph.D. dissertation title: "Mechanisms for Compute and Data Orchestration for Geo-Distributed Situation-Awareness Applications on Edge Infrastructure"
- Awarded Outstanding Graduate Research Assistant award in 2019

B.Tech. + M.Tech. (Dual Degree) in Computer Science and Engineering, Indian Institute of Technology, Kharagpur 2011 – 2016
- M.Tech. thesis title: "FogSim: A Toolkit for Modelling and Simulation of Resource Management Techniques in Internet of Things, Edge and Fog Computing Environments"

SKILLS

Edge Computing | Cloud Computing | Publish-Subscribe Systems | Application Schedulers | Key-Value Stores
Functions-as-a-Service | Network Function Virtualization | Software Defined Networking | Distributed System Verification
Distributed Monitoring | DPDK | Kubernetes | Docker | Linux | C++ | Java | Python | Bash | SQL

WORK EXPERIENCE

AT&T Research Labs (Bedminster, NJ, USA) : Research Intern 2018
Mentors: Dr. Abhigyan Sharma, Alex Zelezniak
- Built an **application-agnostic utilization estimator for Virtual Network Functions (VNFs)** at AT&T's datacenters.
- Experimentally verified **high correlation of hardware (PMU) events with network load** served by diverse VNFs.
- Used **Linear Regression** over PMU event data **to predict VNF utilization** with less than 10% error.
- Implemented an end-to-end power manager using **DVFS frequency-scaling** that **cuts VNF power usage** by **30%**.
- Published findings in the 11th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 2019).

Tools used: Intel **Data Plane Development Kit (DPDK)**, **Support Vector Machines**, **Python, C++**

Amazon Development Center (Hyderabad, India) : Software Developer Intern 2014
Mentor: Vikas Vishwanatham
- **Migrated the backing data store** of **Amazon's product forecasts to AWS DynamoDB** from an indigenous store.
- Designed and implemented a **workflow using AWS Data Pipeline** to transfer data **from AWS S3 to DynamoDB**.
- **Updated the interface** of module consuming product forecasts to **extract data from DynamoDB**.

Tools used: **AWS DynamoDB, AWS Data Pipeline, Java**

RESEARCH EXPERIENCE

Control Plane for Publish-Subscribe Systems on Geo-Distributed Edge Infrastructure. 2020 – 2021
Advisor: Dr. Umakishore Ramachandran, Georgia Institute of Technology.
- Implemented an adaptive topic partitioning among brokers to guarantee end-to-end message delivery latency.
- Leveraged a peer-to-peer network coordinate system (Vivaldi) for scalable network proximity estimation.
- Implemented the proposed control-plane in a popular open-source publish-subscribe system - Apache Pulsar.
- Led collaboration with School of ECE at GA Tech to deploy the proposed system atop a homegrown 4G + Mobile Edge Computing testbed built using OpenAirInterface and Kubernetes.

Tools used: **Apache Pulsar**, **HashiCorp Serf**, **Containernet, Ryu SDN Controller, Open vSwitch, Docker, Java**

Control-Plane for Application Scheduling on Geo-Distributed Edge Infrastructure. 2019 – 2021
Advisor: Dr. Umakishore Ramachandra, Georgia Institute of Technology
- Designed a hybrid control-plane architecture that can schedule both multi-client and client-local applications.
- Central controller schedules multi-client applications using global allocation state of entire infrastructure.
- Multiple distributed schedulers manage client-local applications on each edge-location using local state.
- Optimistic concurrency control for synchronizing allocation state among centralized and distributed schedulers.

Tools used: **Kubernetes, MongoDB, Microsoft Azure Compute, C++, Python**

Geo-Distributed Data-Store for Situation-Awareness Applications. 2018 – 2019
Advisor: Dr. Umakishore Ramachandran, Georgia Institute of Technology.
- Proposed a key-value store serving situation-awareness applications operating in a geo-distributed setting.
- Developers can tune spatio-temporal context of data-items that guides data placement over the infrastructure.
- Tunable consistency guarantees to ensure low-latency data access with high consistency for contextually relevant clients, while ensuring tolerance from geographically correlated failures.

Tools used: **Apache Cassandra, GeoHash, Yahoo Cloud Serving Benchmark, Java, Python**

PUBLICATIONS

ePulsar: Control Plane for Publish-Subscribe Systems on Geo-Distributed Edge Infrastructure. 2021
6th ACM/IEEE Symposium on Edge Computing (SEC), 2021.
Authors: Harshit Gupta, Tyler C. Landle, Umakishore Ramachandran.

OneEdge: An Efficient Control Plane for Geo-Distributed Infrastructures. 2021
12th ACM Symposium on Cloud Computing (SoCC), 2021.
Authors: **Harshit Gupta**, Enrique Saurez, Alexandros Daglis and Umakishore Ramachandran.

Power Management of Polled Network IO using Hardware Performance Counters. 2019
USENIX Workshop on Hot Topics in Cloud Computing (HotCloud).
Authors: **Harshit Gupta**, Abhigyan Sharma and Umakishore Ramachandran.

FogStore: A Geo-Distributed Key-Value Store Guaranteeing Low Latency for Strongly Consistent Access. 2018
12th ACM International Conference on Distributed and Event-based Systems (DEBS).
Authors: **Harshit Gupta** and Umakishore Ramachandran.

STTR: A System for Tracking All Vehicles All the Time At the Edge of the Network. 2018
12th ACM International Conference on Distributed and Event-based Systems (DEBS).
Authors: Zhuangdi Xu, **Harshit Gupta** and Umakishore Ramachandran.

TEACHING EXPERIENCE

Teaching Assistant, Georgia Institute of Technology 2021
Systems Issues in Cloud Computing (Instructor: Prof. Umakishore Ramachandran)
- Designed course content for module on Network Functions Virtualization (NFV)
(content for lecture slides, weekly system implementation workshops and module project)
- NFV module content is now live on Coursera
- Weekly discussions with student teams to assess progress, provide feedback on system design and debugging tips

Teaching Assistant, Georgia Institute of Technology 2020
Distributed Systems (Instructor: Prof. Ada Gavrilovska)
- Introduced term project on distributed system verification through the open-source tool DSLabs

TECHNICAL TALKS

- ACM DEBS 2018
- Fog World Congress 2017
- ACM SoCC 2021
- IETF Computing in the Network Research Group summit 2021
- ACM/IEEE SEC 2021
- USENIX HotCloud 2019