# HARSHITH UMESH

Boston, MA | +1(857) 799-0566 | umesh.h@northeastern.edu | www.harshithumesh.com | linkedin.com/in/harshith-umesh | github.com/Harshith-umesh

# **EDUCATION**

### Northeastern University

- Master of Science in Computer Science
  - Graduate Teaching Assistant: Networks and Distributed Systems Jan 2024 - Present
  - Guided over 160 students in understanding concepts, projects and labs.

- Relevant Coursework: Distributed Systems, MLops, Algorithms, Program Design, Cloud Computing, DBMS

#### **PES University**

Bachelor of Technology in Computer Science and Engineering

- Research publication: Four Layer Authentication with Honeypot and Cloud Data Encryption, IEEE Xplore
- Research publication: Setting Up and Exploration of Security in a Hybrid Cloud, Springer

### **TECHNICAL SKILLS**

Languages: Python, Java, Golang, Bash, C, SQL, HTML, CSS, Javascript

Cloud Technologies: Docker, Kubernetes, Openshift, Ansible, Podman, AWS, GCP, Terraform, Packer, RabbitMO, ArgoCD, RAFT Machine Learning: Tensorflow, DVC, MLflow, Huggingface Transformers, Vertex AI, Pandas, Matplotlib, Snorkel, Jupyter, Scikit, ELK Tools: Linux, Git, Apache airflow, Jenkins, Grafana, Kibana, Flask, Django, Nginx, Prometheus, Junit, pytest, Postman Databases: MySQL, PostgreSQL, MongoDB, Elasticsearch, SQLite, Ceph

#### WORK EXPERIENCE

# **Red Hat**

# Software Engineer Intern

Performance and Scalability - Ansible Saas

- Developed an end-to-end Jenkins pipeline, automating Ansible SaaS performance and scalability tests, reducing manual effort by 10+ hours per deployment, and increasing testing consistency across various configurations.
- Automated provisioning and configuration of execution nodes for Ansible Automation Platform, reducing environment setup time by 40% and enabling QE and Dev teams to run tests autonomously.
- Migrated Perf scale tests to Ansible Testing Framework (ATF), improving test coverage by 80%, reducing test debugging time by 50% and standardizing the testing framework across teams.

# Software Engineer

#### Performance and Scalability - Openshift - R&D

- Researched and Developed tools to improve the performance of Red Hat OpenShift.
- Contributed significantly to the development of two CNCF sandbox projects "Kube-burner" and "Arcaflow".
- Revamped the OpenShift Performance CI by migrating from a legacy system to a GitOps model powered by Apache Airflow.
- Architected and implemented an ETL pipeline to move data and metrics stored on elasticsearch to an RDBMS, which is consumed for • data science, machine learning and data visualization applications.
- Authored and maintained scripts written in python and bash, which use boto3 to identify and delete idle or zombie OpenShift clusters and • all associated resources created on AWS, helping reduce operational costs by 45%.
- Lead maintainer of the tooling and infrastructure used by the team, deployed argoCD using helm to monitor the health of all services used by the team. Integrated it with slack to notify of failures, reducing downtimes by 75%.
- Conducted analysis and benchmarking of OCP and Kubernetes control and data plane components across 6 minor releases.
- Developed an end-to-end CI/CD pipeline using Jenkins to streamline the OpenShift installer development process; any major code change triggers an installation of an OpenShift cluster on AWS, Azure or GCP and returns a pass or fail report.
- Developed and deployed an internal object storage file server "Snappy" using podman containers to store internal data as an alternative to AWS S3 reducing cloud spending costs by 20%.
- Designed and deployed Analytics dashboards using Grafana and Kibana to help visualize the performance of benchmarking tools such as FIO, Uperf, HammerDB, Sysbench and Prometheus on OpenShift.

## PROJECTS

# Scalable User Management Application on GCP using Terraform - Github link

- Developed and deployed a Python Flask RESTful application for user management with CRUD operations and health checks.
- Utilized GCP infrastructure with Terraform and Packer for IAC and image creation. The architecture includes VPCs, CloudSQL, Cloud • DNS, SSL certificates, and customer-managed encryption keys. Integrated Google Pub/Sub and Cloud Functions for event handling ( Mailgun verification emails).
- Configured Systemd for API service initialization and set up firewalls for network security. Enabled scalability and high availability with • managed instance groups, load balancing, and auto-scaling.

Implemented CI/CD pipeline using GitHub Actions for automated testing, VM image building, and rolling updates. •

#### MLops Pipeline for Sentiment Analysis using LLMs - Github link

- Developed and deployed an end-to-end MLOps pipeline for sentiment analysis of Amazon customer reviews using fine-tuned LLMs (DistilBERT, RoBERTa, ALBERT) with TensorFlow, automated workflows with Apache Airflow, and CI/CD pipelines with Jenkins and Docker for scalable model training and deployment on Vertex AI.
- Implemented monitoring and drift detection with Grafana and Evidently AI, and built an interactive Streamlit Application hosted on • Kubernetes for real-time sentiment analysis and visualization of customer insights.

#### DBaaS for a Rideshare Application - Github link

- Built a highly available database as a service(DBaaS) for a carpooling 'RideShare' application using REST APIs and python-flask.
- SOLite is the database used. The application is run inside docker containers hosted on AWS EC2 instances. AWS load balancer and zookeeper are used to maintain high availability.
- Docker SDK and RabbitMQ are used for scaling up/down containers, and Postman for testing the API endpoints.

# Sept 2023 - May 2025

Boston, Massachusetts

Aug 2017 - May 2021 Bangalore, India

**Bangalore**, India

Jan 2021 - Aug 2023

May 2024 - Aug 2024

Boston, MA