

Vichara Nimnaka Wijetunga

✉ vnwijetu@uwaterloo.ca 🏠 nimwijetunga.github.io 🐙 github.com/nimwijetunga in linkedin.com/in/vichara-w

Programming Skills

Languages: Python, Java, Scala, C++, JavaScript, PHP, SQL

Data & Infrastructure: Spark, Airflow, Kafka, Hive, Hadoop, Parquet, Kubernetes, Docker, Nginx, Terraform, PostgreSQL

AWS: EMR, DynamoDB, DataPipeline, NAT, Route53, SQS, SNS, S3, VPC, ElastiCache, RDS

Experience

University of Waterloo

Waterloo, ON

Undergraduate Research Assistant (Prof. Zahedi), Large-Scale Distributed Mobile Computing January 2020 - Present

- Leveraging a network of mobile devices (as worker nodes) to perform **distributed computations** using Apache **Spark**
- Refactoring Spark Worker Core, including the **RPC** (written in Scala), so it can be compiled using the **Android DVM**

SnapTravel

Toronto, ON

Data & Infrastructure Engineering Intern (medium.com/@nim.wijetunga) September - December 2019

- Implemented an **EMR** (AWS **Elastic MapReduce**) Deployment Workflow, utilizing custom **Airflow Operators** to submit **Spark** Jobs, and **CI/CD** to deploy Spark resources to **S3**, gaining full adoption by the Data Engineering Team
- Migrated a CPU intensive Python ETL to **PySpark SQL**, using **Hive** Tables to store data, improving runtimes by **80%**
- Improved Hotel Catalogue Data Ingestion Speeds by **85%+**, utilizing **AWS SQS (Queue)** to stream data across **distributed** Airflow workers and Python **Concurrency** (Asyncio/Asyncpg) to improve I/O bottlenecks
- Built **Datadog** Dashboards using **Terraform** to monitor **RDS & ElastiCache** services, helping identify **15+** incidents
- Migrated Airflow Deployment to Astronomer (Airflow IaaS), using **Terraform** for provisioning **NAT** Instances (with **iptables** - Port Forwarding, **Nginx** - Reverse Proxy) to access ElastiCache clusters inside the AWS **Private Subnet**

SnapTravel

Toronto, ON

Backend Engineering Intern (medium.com/@nim.wijetunga) January - April 2019

- Transformed **100M+** rows of logs, pushing our Data Warehouse (**Snowflake**) size from **500GB to over 9TB**
- **Visualized** ETL data, using **PeriscopeData**, into **SOV** (Share of Voice) and **Spread-to-Bar** reports
- Experimented with Python-Flask **rate-limiting** to reduce the chance of a **TOTP brute force** attack by **92%**
- Implemented a **Redis Caching Layer**, providing **2x** the speed compared to direct **Postgres** access

The Weather Network

Toronto, ON

Full-Stack Developer Intern May - August 2018

- Introduced **data compression** for **PHP** endpoints, reducing API request time by **0.5s** (client side)

Education, Courses & Awards

University of Waterloo

Waterloo, ON

Bachelor of Software Engineering (3rd Year); CGPA: 3.8, Dean's Honours List 2018 September 2017 - April 2022

Kafka (Udemy) 🐘: Brokers, Partitions, Producers/Consumers, **Zookeeper**, Kafka **Connect** and Kafka **Streams**

Awards: PennApps 2018 🏆 Best Campus Life Hack Runner-Up, DeltaHacks 2018 🏆 Best Healthcare Hack Runner-Up

Extracurricular Activities

UWBlueprint, Infrastructure Developer 🐙: Containerized a Flask and PostgreSQL Application using **docker-compose**, and experimented with **Kompose** to deploy the app to a **Kubernetes** Cluster