

Vichara Nimnaka Wijetunga

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

Programming Skills

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

Data & Infrastructure: Spark, Airflow, Kafka, Hive, Zookeeper, Docker, Nginx, Terraform, PostgreSQL, Datadog

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

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

Scala & Spark (Udemy): Scala Basics, Spark **RDDs/Dataframes**, Spark **SQL**, **MLlib**, Spark **Streams** and Spark on **EMR**

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

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

Experience

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 **3** Data Engineers
- Integrated **EMR** with **Zeppelin Notebooks** and SnapTravel Internal Tools, improving Analyst workflow speeds by **50%**
- Migrated a CPU intensive Python/Airflow **ETL** to **PySpark SQL**, using **key salting** to reduce runtimes by **90%+**
- Improved Hotel Catalogue Data Ingestion Speeds by **85%+**, utilizing **AWS SQS (Queue)** to stream data between **distributed** Airflow Tasks, and **Apache Hive** temporary tables to copy data to **DynamoDB**
- 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

- Designed and Implemented three key **ETLs: Supplier Availability, Supply/Demand** and **Expired Rates**
 - * 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
- Implemented a **2 Factor Authentication (2FA)** Microservice
 - * Experimented with Python-Flask **rate-limiting** to reduce the chance of a **TOTP brute force** attack by **92%**
- Increased **conversion** for the **Business Traveller** Program by **10%** after implementing its first perk: **Uber Vouchers**
 - * Implemented a **Redis Caching System**, providing **double** the speed compared to **Postgres**

The Weather Network

Toronto, ON

Full-Stack Developer Intern

May - August 2018

- Implemented **AMP** (Google's Accelerated Mobile Pages), reducing forecast page load times by **1.5s** (mobile)
- Introduced data aliasing for **PHP** endpoints, reducing API request time by **0.5s** (client side)
- Implemented the **MVC** design pattern for backend API's, using the **CodeIgniter** (PHP) framework
- Developed test suites (employed **TDD**) using **Jasmine**, **Mocha** and **PHPUnit**, maintaining **80%** test coverage
- Built **RESTful API** endpoints for displaying news articles provided by the **CMS**

Extracurricular Activities

UWBlueprint, Infrastructure Developer 🏆: Containerized a Flask and PostgreSQL Application using **docker-compose**, and built a **CI/CD** integration with **Github Workflows** to deploy the app to a Staging **Heroku** Environment