

Project Title: Network Telemetry Pipeline

Supervisor Team:

Nasim Ferdosian
Himanshu Agrawal
James Lim

Project Description

These days every business and enterprise is seeking ways to collect and integrate data from sources distributed through their communication network and analyze them to gain business or network insights for efficient management and competitive advantage. A data pipeline is a set of actions that ingest raw data from distributed sources and move the data to a destination for storage and then analysis.

The focus of this project is to build a network data pipeline to transport network data from network communication platforms and database sources to data warehouses for use by analytics tools. A network data pipeline contains the following components such as: Data Collector, Data Processing, Data Storage, Data Query, and Data Visualization.

A student can build this network data pipeline by using and interfacing the available technologies and solutions for different components of the pipeline for example: CloudFlare and Cisco network telemetry as the Data Collector, Apache Kafka as the Streaming Storage Platform, and the open-source Grafana platform as the Data Visualization.

Specific Investigation Goals

This project is considered for students interested in **Big data** and **data analytics** solutions in **Internet of Things**. Within this context, the students will have the opportunity to:

- Acquire practical knowledge and hands-on experience on *data analysis, processing, storing, and monitoring tools and techniques*.
- Get hands-on experience with network testbed environments containing Cisco network equipment and perform experimental evaluation of the proposed model;
- Equip with the ability to envision, develop, launch and grow innovations, and provide the prospective experiences to be an attractive recruit for innovation-focused companies and organizations.