A Cloud-based ML-facilitating Data Pipeline for Wireless Spectrum Radio Data

Zhen Jiang, Victor Li, Shruti Satrawada, Sivani Voruganti, Gen Yang
Advisors: Prof. Anant Sahai, Josh Sanz

SpectrumX is an NSF Spectrum Innovation Center, funded via Award 2132700

What is Spectrum?

Spectrum is a resource consisting of the range of electromagnetic radiation frequencies used to transmit information wirelessly.

The **radio frequency spectrum** powers all the communication around us—from cell phones to WiFi and more.

Project Objective

The Problem:

Wireless Spectrum data's complex and dynamic nature requires a **custom pipeline** to handle the massive volume and diversity of the data while ensuring quality, reliability, and privacy.

Our Solution:

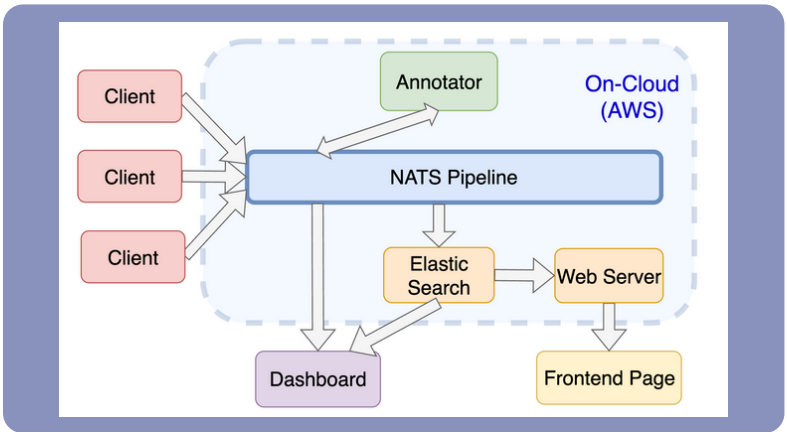
- An **end-to-end data pipeline** that collects, visualizes and analyzes spectrum radio data.
- Detailed **tutorials** for diverse audiences to build their own custom spectrum pipelines.
- An **example airplane tracker application** to showcase how our pipeline can be adapted to process diverse types of spectrum data.

Unique Values: Adaptable & Affordable

Our end-to-end data pipeline is:

- **replicable & affordable**: individuals can build their own pipeline following our start-from-scratch tutorials without the need for company-level staff resources
- **adaptable & open-source:** we aim to make traditionally under-utilized spectrum data more accessible to researchers, hobbyists, and the general public at large to apply towards various applications

Data Pipeline Architecure

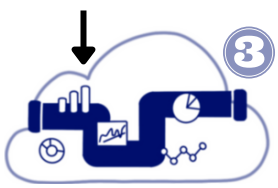


Our pipeline utilizes the **NATS.IO** system for data transmission, an **Elasticsearch** database for AI/ML application integration, & **Prometheus** for status monitoring, all within a distributed **AWS cloud-based environment**.

Airplane Tracker Application Example

1. All airplanes emit Spectrum Radio signals with geographical information

2. Sense & collect signals using radio devices in different geographical regions (SF, Denver)



3. **Data Pipeline**: Process raw signals and store in database

AI/ML Integrations

Visualize live airplane locations on **Interactive Map**

Dashboard to visualize metrics of our system

Using our pipeline, users can build various spectrum applications such as our airplane tracker, radio content analyzers, and signal monitors.