Energy Insights In a Box

Extensible Data Analytics System

Abdul Rehman

Advisor: Prateek Sharma

Design Goals

Train/Update on cloud across different data-sets

Pretrained + Fine-tuned Foundation Models

Continuous Analytics [Forecasting, anomaly detection, missing vals]

Various Models
- Easy to Develop, Deploy and Finetune

MQTT Broker

Time-series Database

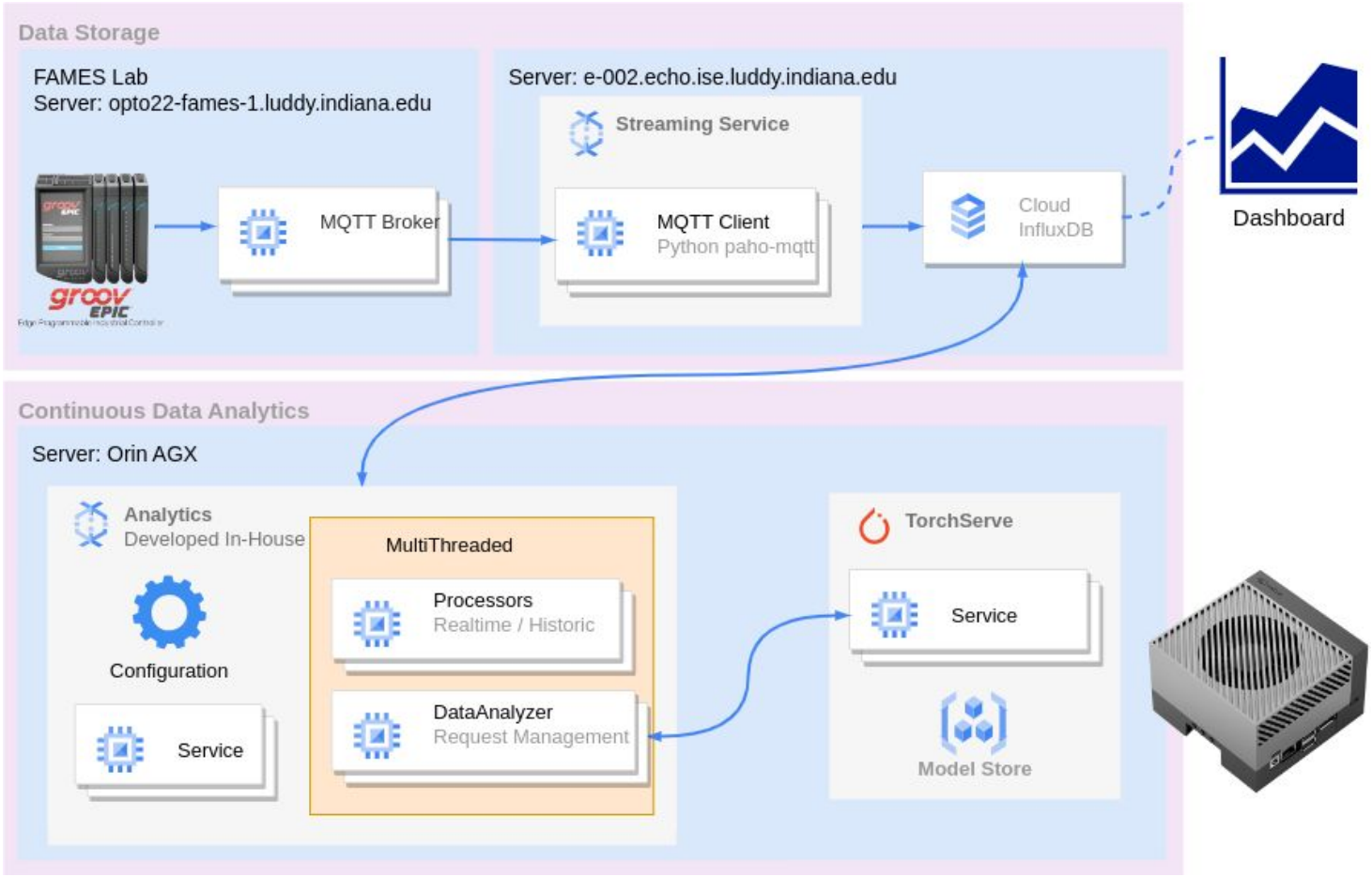
Dashboard

Easy to Visualize

Energy Insights In A Box

Why?
- Privacy. No data leaves premises
- Cheap: $500-1000 per month on AWS vs. total $2000 for Orin AGX
- Real time analytics using DNNs

Implementation



Demo

- Setup
  - **Data Storage Layer:** Real Time data streaming to InfluxDB from sensors.
  - **Continuous Data Analytics Layer:** Hosted on Jetson, predicting values for 54 sensors in real time using Recurrent Neural Network and ARIMA.

- InfluxDB Dashboard
  - Showcases current and predicted values for select sensors.

Analytics Pipeline: Extensible Json Configurations

- specify sensor to model correspondence
- Specify data analyzer - depends on endpoint
- InfluxDB bucket to fetch data from
- setup window size and time step
- processor historic or real time data


```
{
  "version": 10,
  "sensor": [
     "Power_Meters/CNC_Machines/ST10_old/PhB-TruePowerWatts/Value"
  ],
  "models": [
    "arima"
  ],

  "dataanalyzer": "tserve_arima",
  "influxdb_bucket": "einsights_bucket",

  "window_size": 121,
  "time_step": 1,
  "processor": "historic",
  "t_start": "2023-12-20 04:44:00",
  "t_end": "2023-12-20 15:54:24"
}
```

Implementation

- ~5K Lines of Code
  - MQTT - Service
  - Data Analytics
    - Analytics Services
    - Torch Serve

- Open Source Codebase

Performance and Scalability

For 54 sensors being predicted using arima model and RNN models

- Network bandwidth consumption on Jetson ~800 Kbytes per sec
- CPU utilization on Jetson ~25 %
- GPU utilization ~40 %

InfluxDB storage increases by ~7 Kbytes per second

- it would take 866 days to fill up 500 Gb of space which is ~2.3 years

Future Work

- Extend it to 1000 metrics
  - Distributing the analysis across edge devices.
  - Throttling of requests.

- Implement the Conversational AI interface for the dashboard queries.
  - Example
    - Is the power usage going to cross 5000 watts tomorrow?

- Price aware scheduling of Model Inference.

Thank You