Digital Twin for Plant Health Monitoring
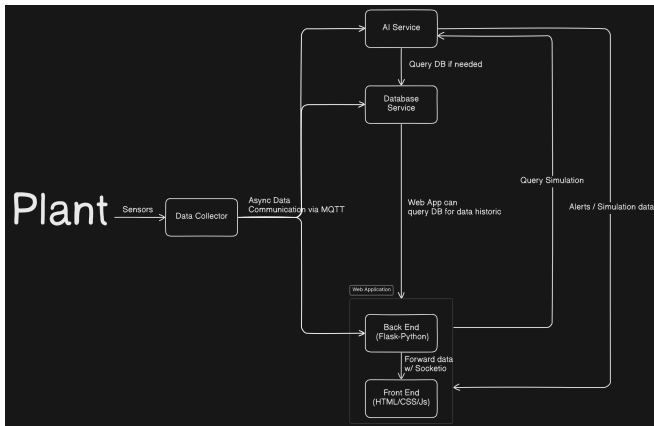
Rayan Contuliano Bravo

Université Libre de Bruxelles

03/04/2025

Current System

- **MQTT** protocol: For ease of use, lightweight and decoupling of systems modules.
- **Sqlite** Ease of use and serverless.
- **Flask** application for DB, AI service and web application because easy to query by REST API endpoints

Scalability of current system

If we had a system with $+1000$ sensors :

- ▶ **Clustering** of RPI, each RPI would be handling a subset of sensors with a Sensor Manager
- ▶ **Migrate** into a cloud-based Broker instead of a local one.
- ▶ Integration of a **Streaming Platform** like **Kafka** and a **Streaming Processing** framework like **Spark Streaming** for real-time transformation/analytics before the data hits the DB, WebApp, AI service
- ▶ Use **Time-series Databases** as they are designed to handle high cardinality and time-series data and have compression algorithms.

Scalability of current system (2)

Currently the systems scale well for a lower-to-moderate scale project. However, with +1000 sensors, it would become ineficient. SensorManager would really be struggling with constant polling + data handling would be difficult.

Maybe use event-based data retrieving instead of sequential constant polling of the sensors

Adaptability of current system

If you want to create a new sensor :

- ▶ Create the corresponding class that inherit from the Abstract Sensor class
- ▶ Create an MQTT topic to publish the data. Consumers like WebApplication, Db, Ai service would be able to access the data, just write the corresponding code to handle those specific data (analysis, vizualisation, etc.).

If you would like to create a new service:

- ▶ Just listen to the MQTT topic your service needs and write your code.

Data Management

We need to avoid storing all the data as the amount will become rapidly huge. Then it would become very difficult to store and analyze.

- ▶ Keep high granularity data for a short amount of time (few hours / 1-2 days)
- ▶ Downsample raw data as time passes by summarizing (e.g average) and keep thoses summaries for a longer period (weeks/months)
- ▶ Use the help of Time-series databases

Plant Watering

- **Water plant** 5v + pvc tube. amazon
- **Relay module** 5V: amazon
- **Battery case** to support the relay, because it can put a lot of stress of RPI.