

REAL TIME DATA AGGREGATION AND VISUALIZATION

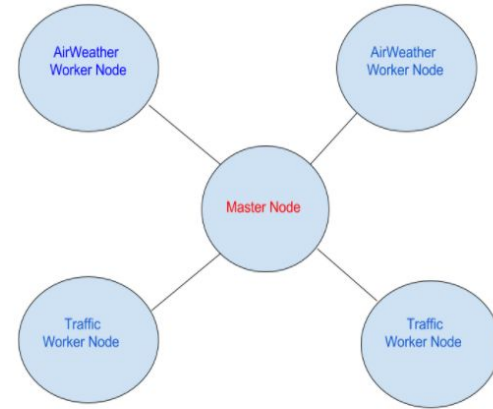
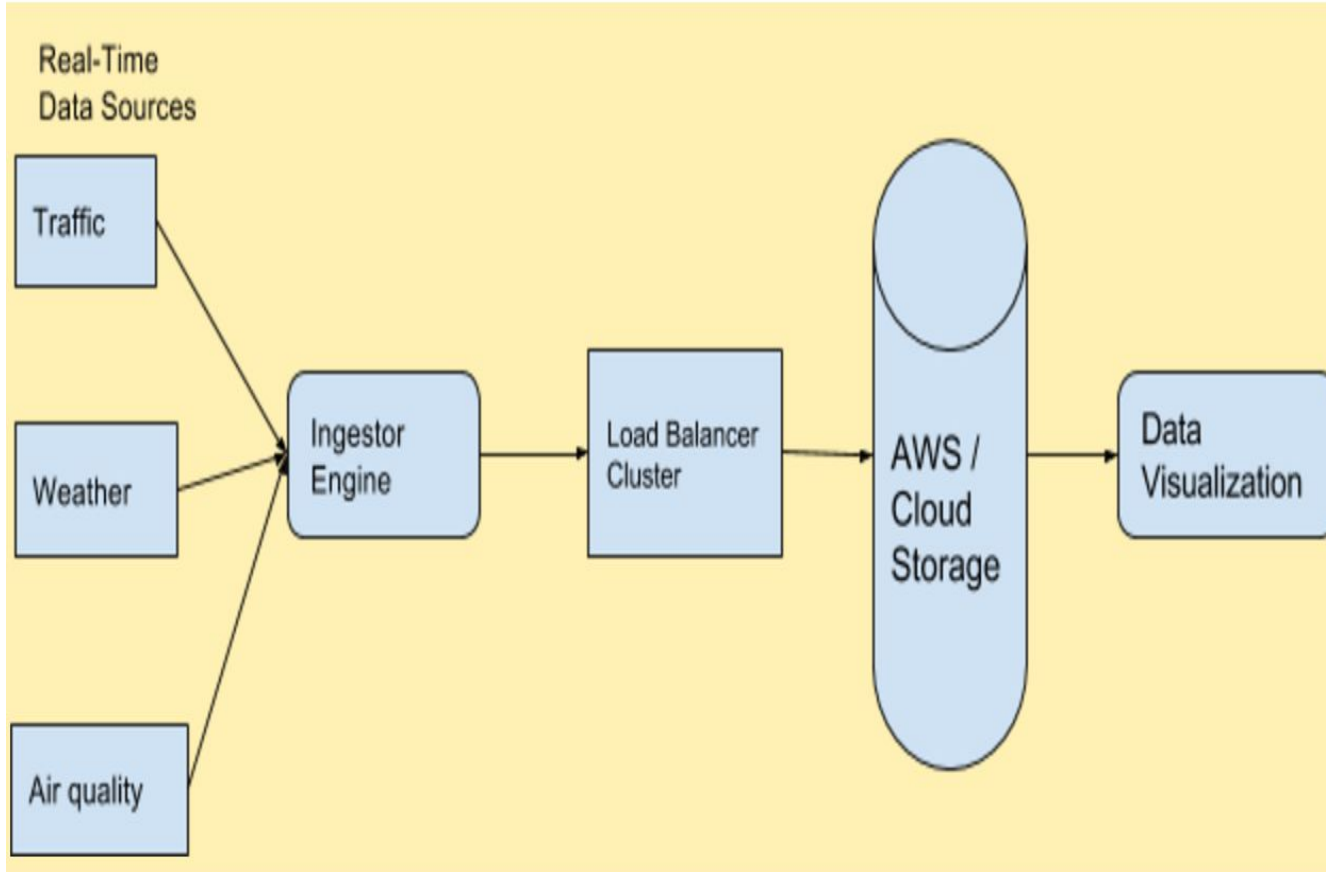
Group 12

Deepthi Devaiah D(65848248)

Kevin Sebastian(44322980)

Sachin B Raj(28568830)

Design

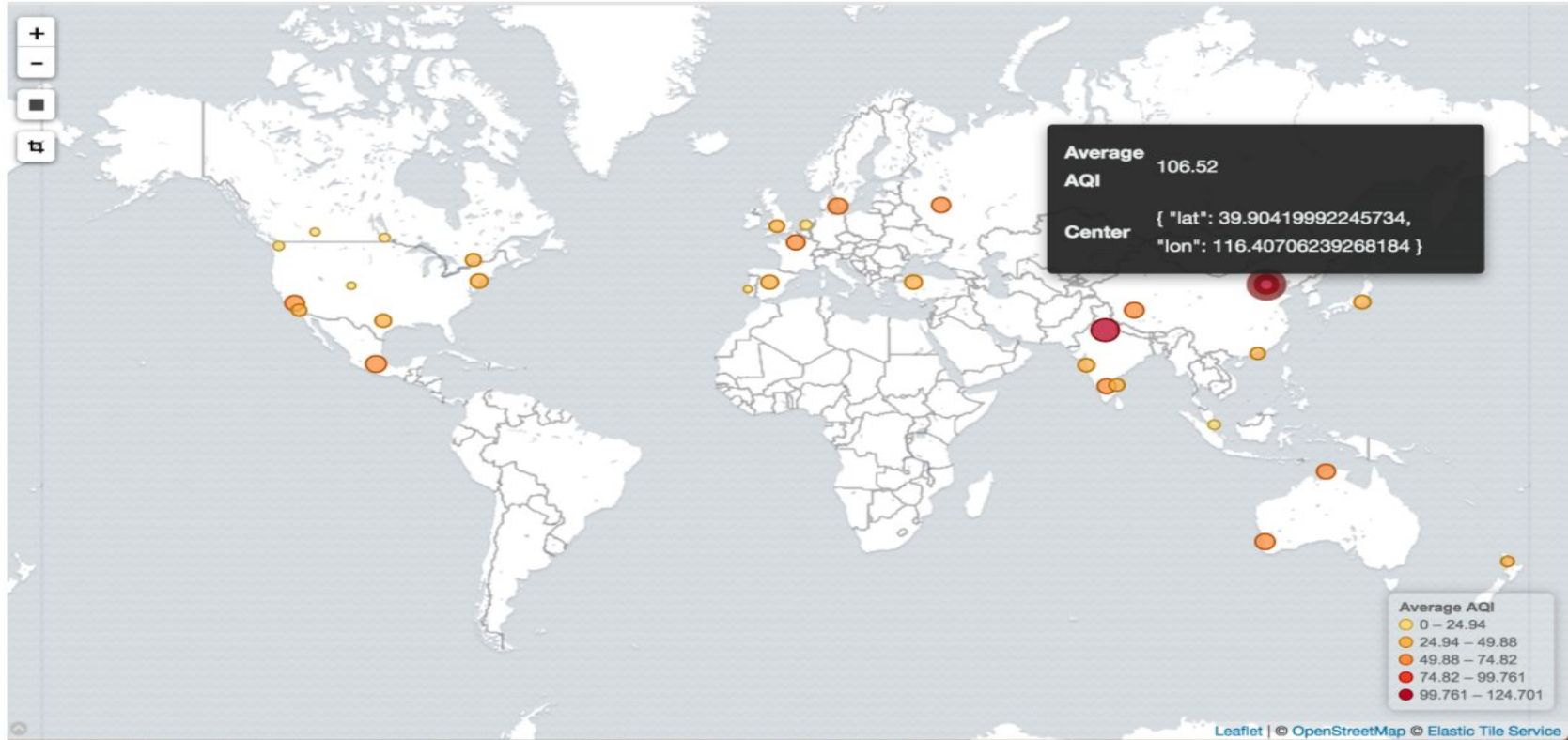


Methodologies

- 1) **Data Extraction** - Air, Weather and Traffic Data are taken from the public API and refined into JSON format suitable for us
- 2) **Ingestion Engine** - 3 topics are ingested into the kafka by partitioning.
- 3) **Load Balancing** - Self sustaining Dynamic Load Balancing
- 4) **Elasticsearch** - Storage of the data and run analytical queries
- 5) **Cloud Storage** - Provides improved availability, scalability
- 6) **Visualisation** - Running analytical queries and visualizing the data is done using Kibana

Evaluation / Results

The average Air Quality Index result for the last 7 days around the world.



Scalability and Fault Tolerance

- **Micro-service architecture**

Adding new types of messages to kafka cluster can be scaled up by creating new worker nodes for the same message and the master node takes it from then on to automatically load balance.

Future Scope

- **Load balancing framework** - A leader election protocol can be implemented to make the current centralised architecture fault tolerant.
- **Visualisation as a Service** - Port entire end to end application to cloud and run it as a real time visualisation service.