

Ministry of education and science of
Ukraine
National University of Kyiv-Mohyla academy



Development of cloud-native streaming data pipeline for data
visualization and decision making

Volodymyr Materynskyi
Thesis supervisor: Dmytro Cherkasov

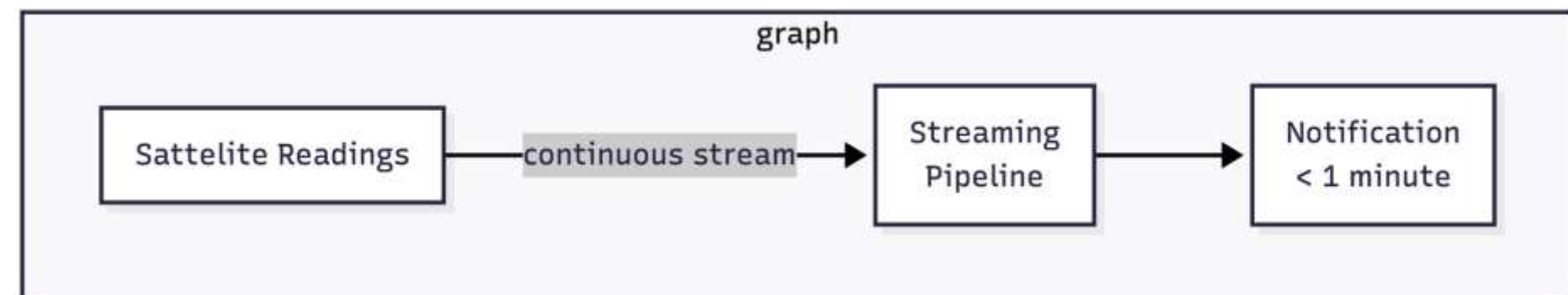
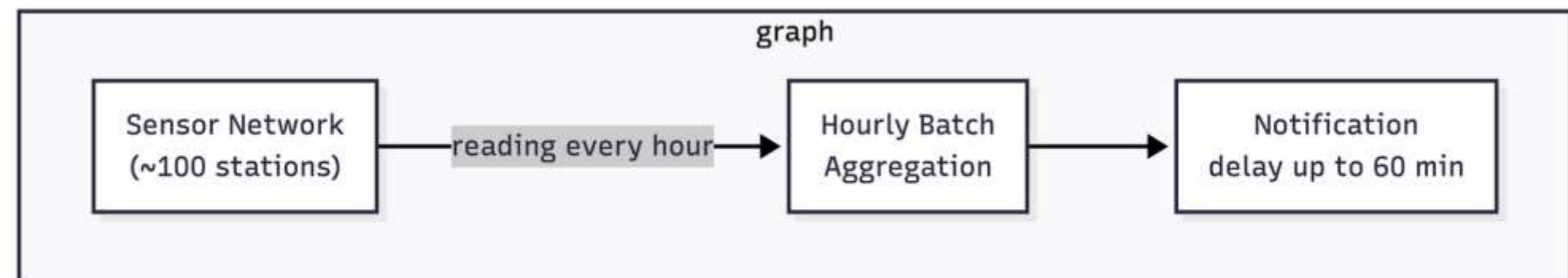
Motivation & Problem Statement



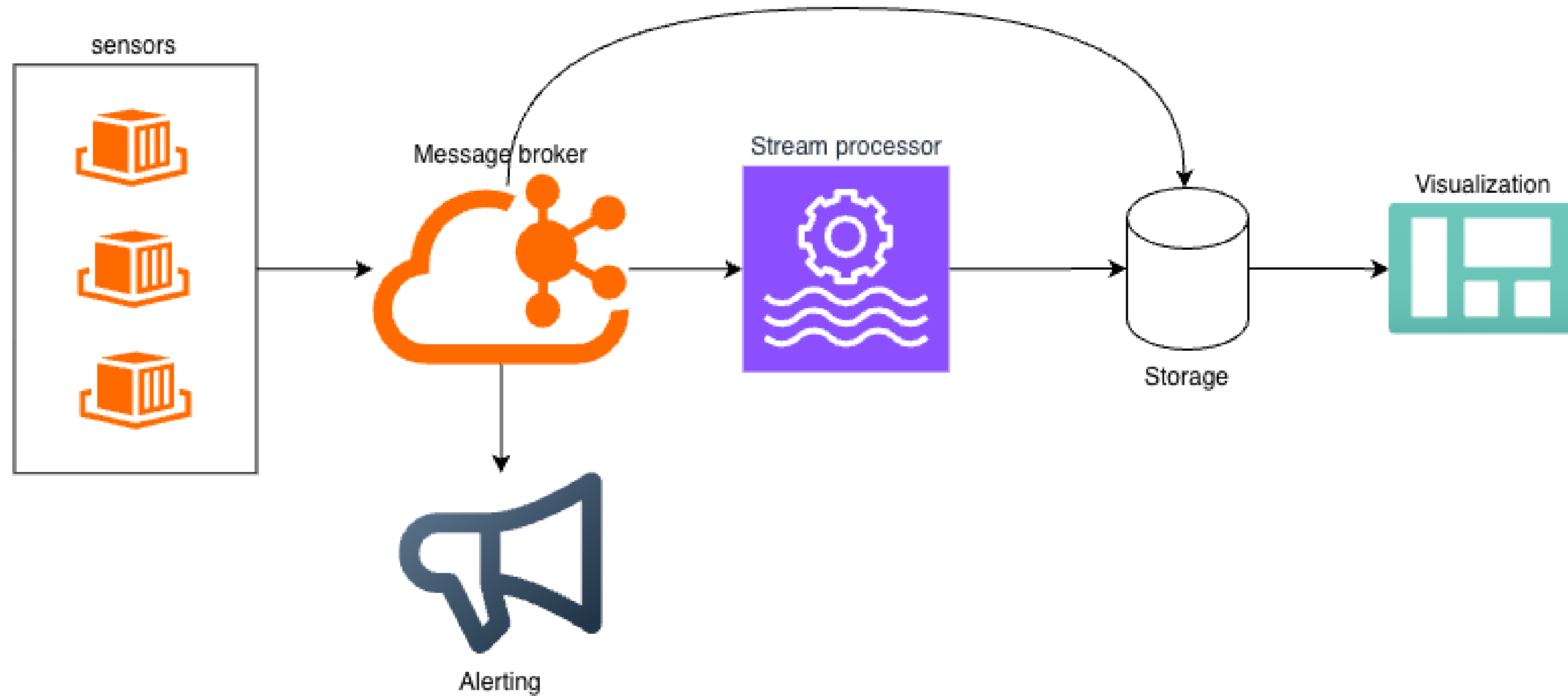
Research Objectives

- Review existing air quality monitoring solutions and technologies
- Design a real-time air quality monitoring system architecture
- Develop and implement the proposed monitoring solution
- Evaluate system performance through load testing and latency analysis

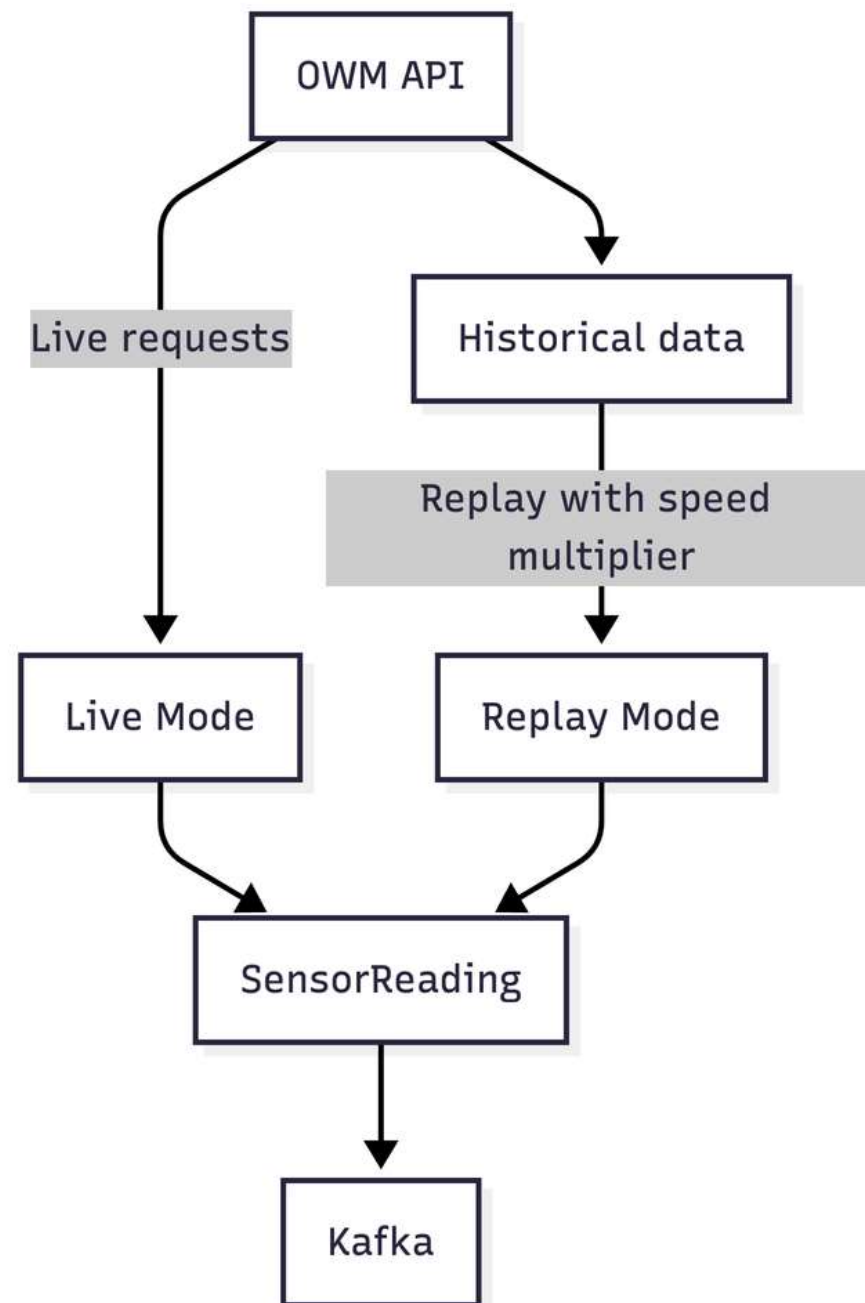
Architectural selection



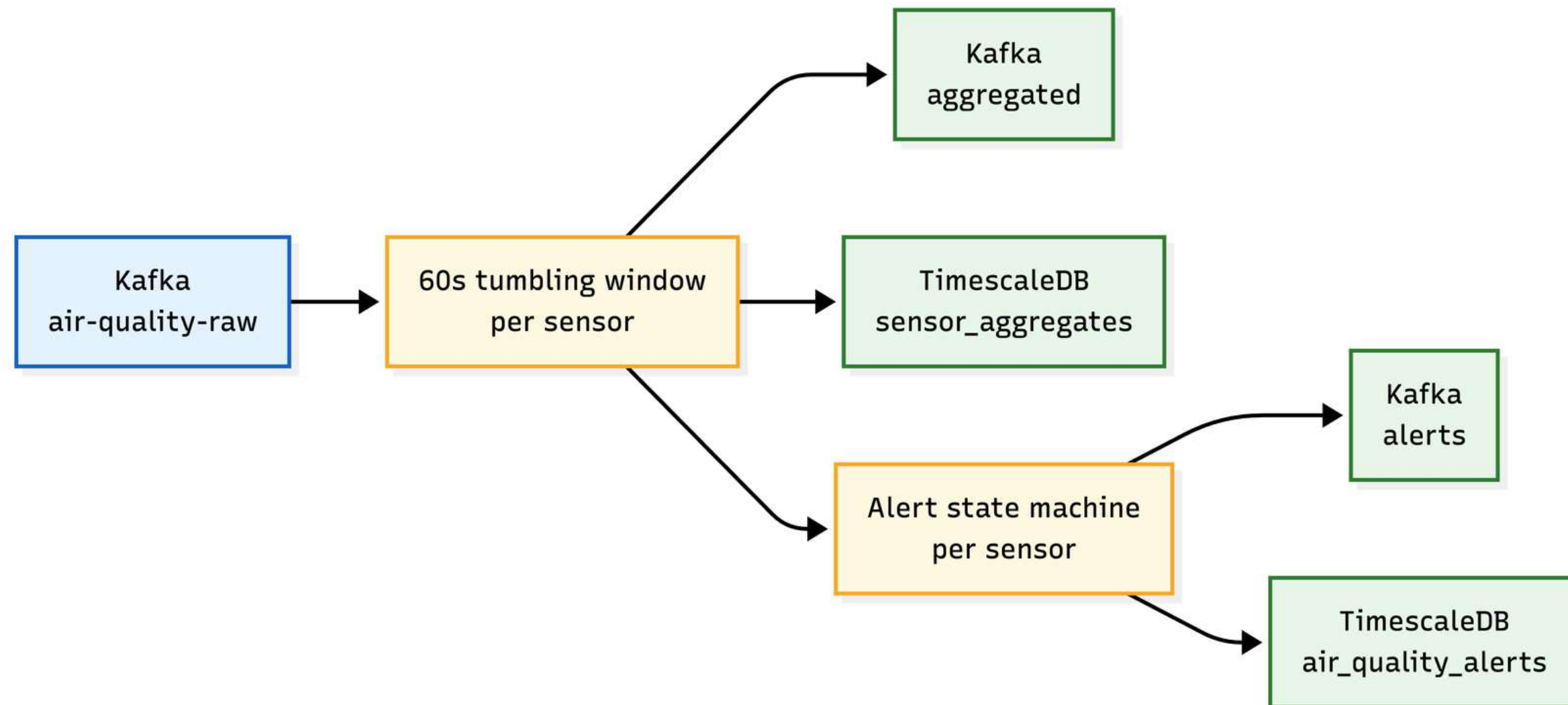
System Architecture



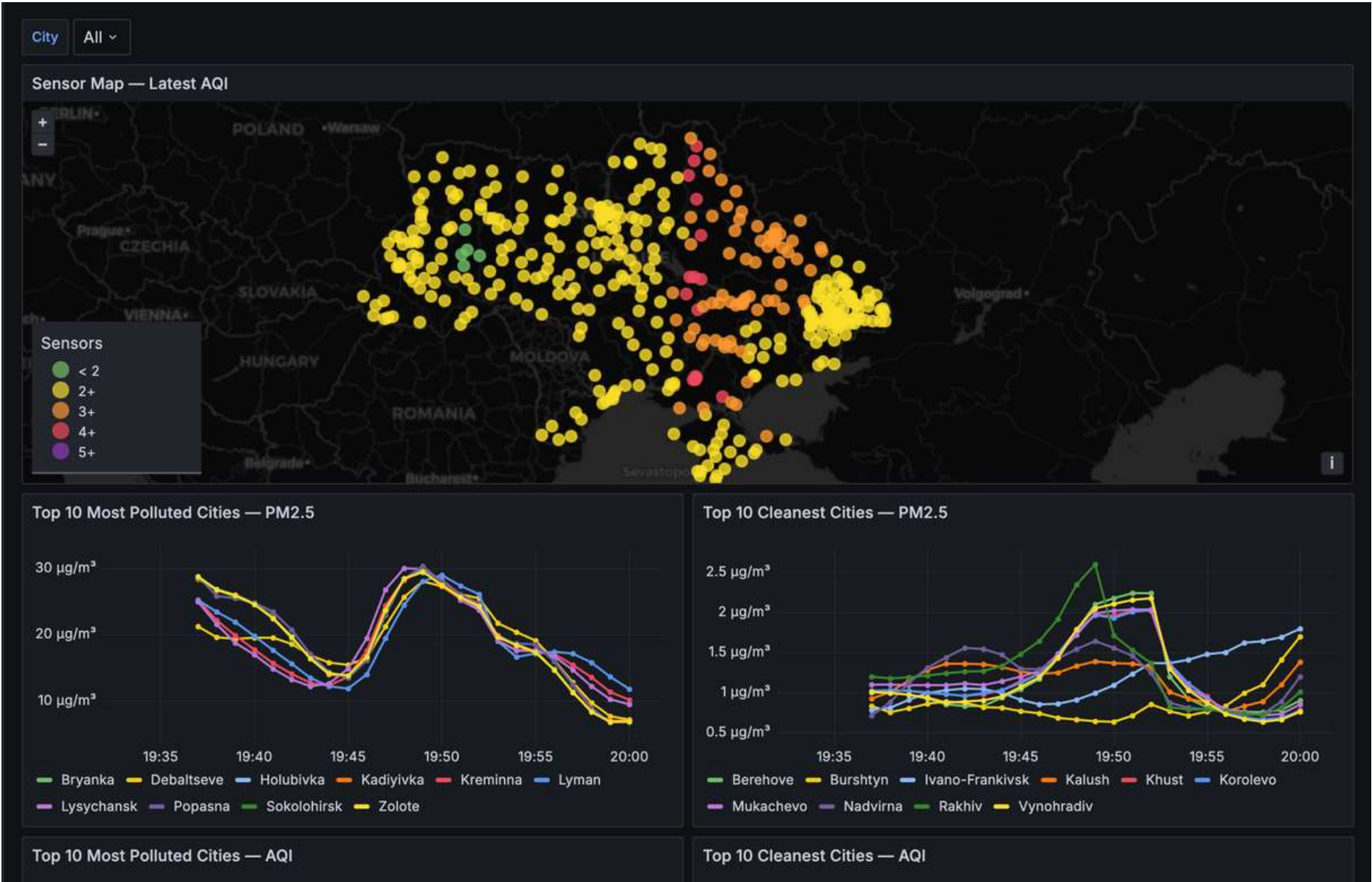
Data generation



Streaming processing



Visualization and alerting



←

🗑

⬇

🔄

From

alerts@airquality.local

Subject

[AIR QUALITY ALERT] Korosten — Fair (PM2.5: 14.7 µg/m³)

To

volodymyr2126@gmail.com

Plain text

Source

MIME

Air Quality Degradation Detected

=====

City: Korosten

Sensor: ua_704901

Coordinates: 50.951199999999998, 28.638590000000002

Severity: Fair

Trend: STABLE

Measurements

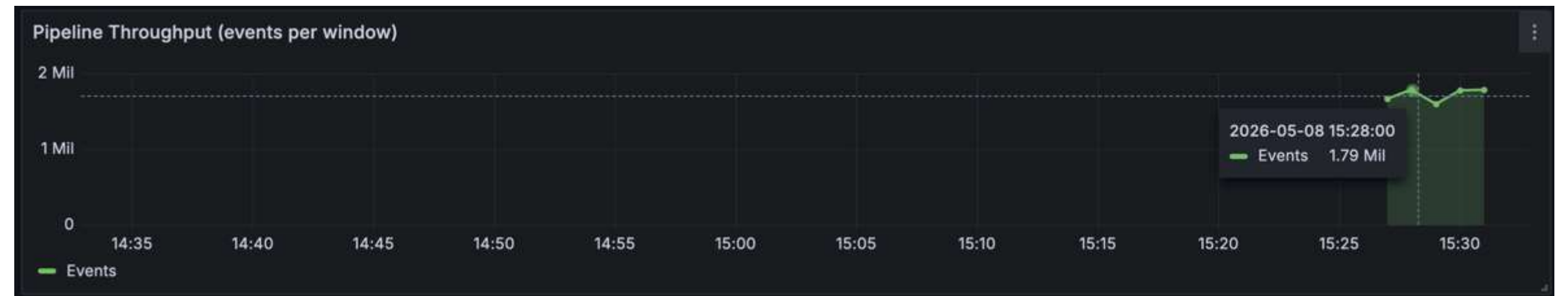
Peak PM2.5: 14.7 µg/m³ (safe threshold: 25 µg/m³)

Average AQI: 2 / 5 (European Air Quality Index)

Consecutive elevated windows: 10 (~10 minute(s) sustained)

Air Quality Monitoring Pipeline | automated alert

Evaluation and results



Future work

- NASA FIRMS satellite fire/smoke ingestion
- AWS cloud deployment (Terraform already partially designed)
- Cross-city anomaly correlation
- ML-based anomaly detection replacing Z-score

Conclusion

- Conducted a review of existing air quality monitoring solutions
- Designed and implemented a reproducible real-time streaming pipeline
- Implemented onset/recovery alerting
- Achieved sub-minute end-to-end latency at ~414k events/min

