DOS
Distributed and Operating Systems.

berlin

WaterGridSense4.0
Finales Gesamt-Projekttreffen

AP2 / AP6 / AP7

8 Dec 2021

Morgan Geldenhuys, Felix Lorenz, Benjamin Pfister, Martin Haug, Lauritz Thamsen

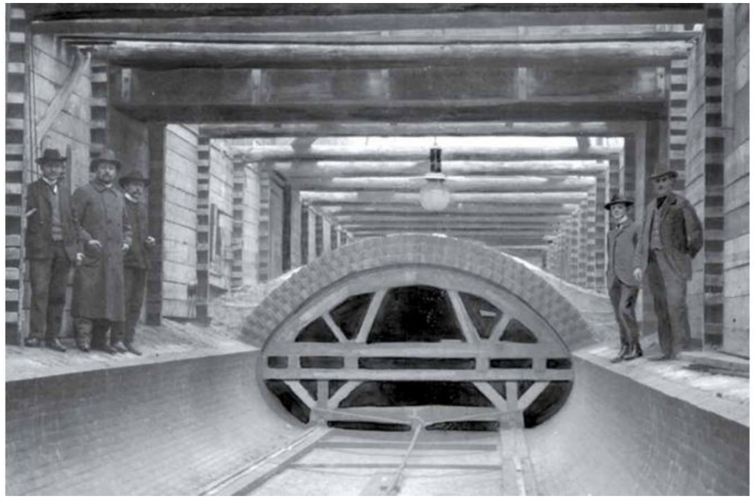
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

BETREUT VOM
PTKA
Projektträger Karlsruhe
Karlsruher Institut für Technologie

Motivation

- Water networks are critical urban infrastructure
  - Spread over wide geographical areas
  - Direct access is **very** expensive

- Monitoring can be used for
  - Detecting anomalous states
  - Optimized controls
  - Predictive maintenance



https://www.bwb.de/de/assets/downloads/Berliner-Kanalisation_2012_web.pdf

Berliner Wasserbetriebe

> 10.500 km length

245 million m^3 p.a.

152 pump stations

Leverage IoT and Cloud Computing for water network monitoring

Our WaterGridSense4.0 Project

DOS
Distributed and Operating Systems.

Team at DOS

- Morgan Geldenhuys
- Felix Lorenz
- Benjamin Pfister
- Martin Haug
- Lauritz Thamsen

WaterGridSense 4.0

Partners

Sieker Die Regenwasserexperten
HAW HAMBURG
Funke Gruppe
REMONDIS®
ACO
Berliner Wasserbetriebe
HAMBURG WASSER

Work Packages

Q4 2018

- WP1: Use Cases and Requirements
- **AP2: Overall System Architecture**
- WP3: Lab Prototype
- WP4: Field Prototype
- WP5: GIS Platform
- **AP6: Analysis Platform**
- **AP7: Anomaly Detection**
- WP9: Field / Platform Tests

Q4 2021

Requirements

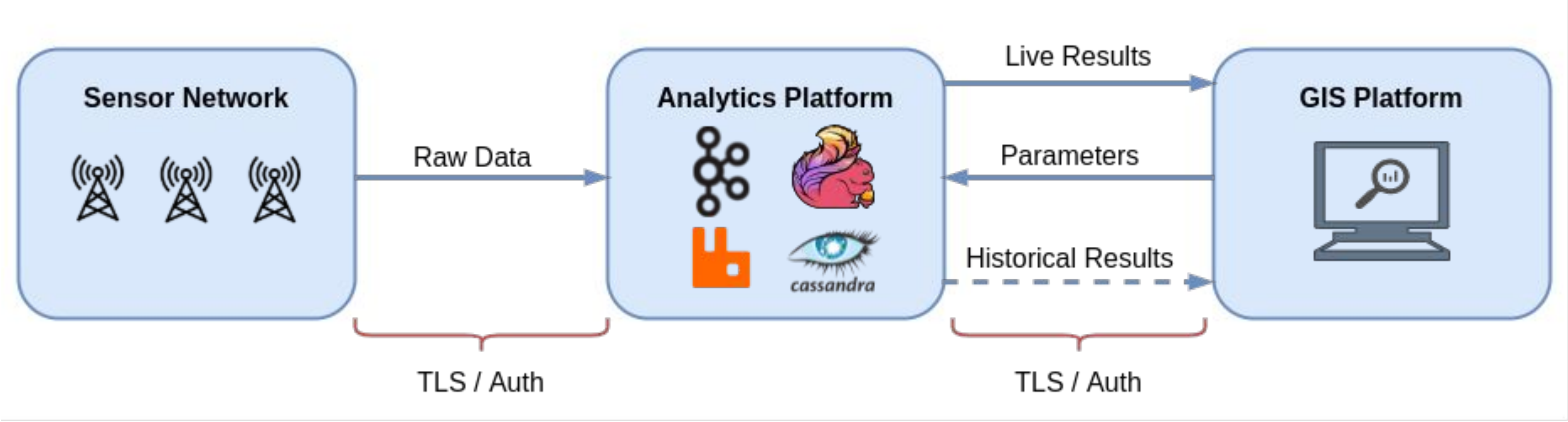
Required Characteristics of a Dependable Critical Infrastructure Monitoring Platform

1. The platform must be *available*, i.e. should deliver the service that is required through being scalable.
2. The platform must be *reliable*, i.e. provide fault tolerance through service redundancy and replication.
3. The platform should be *maintainable*, i.e. abstract away the complexity from users by making it relatively straightforward to use and maintain.

Reproducibility & Dissemination

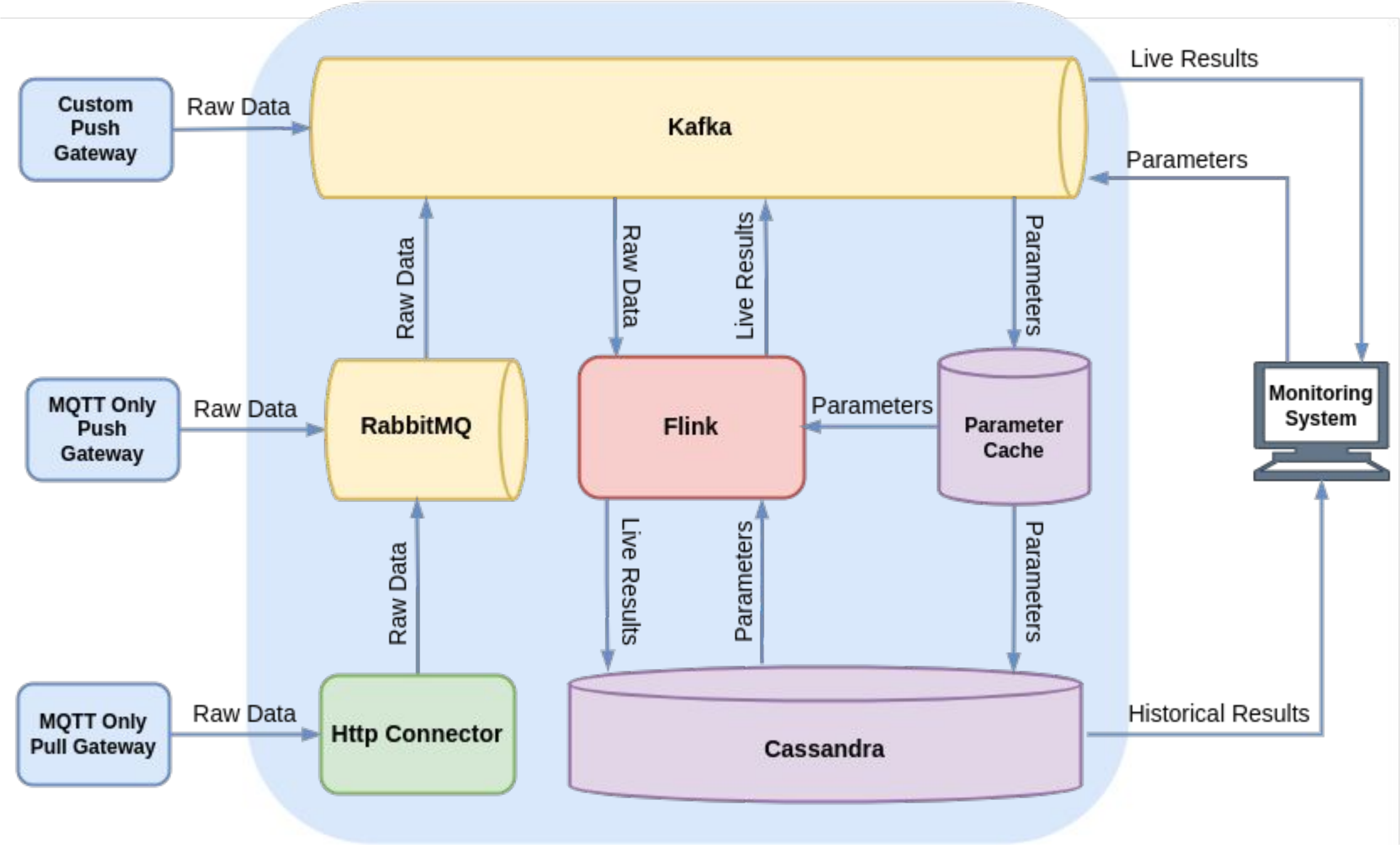
4. The platform should be easily reproducible, i.e. only use FOSS and provide deployment code.

AP2: Overall System Architecture



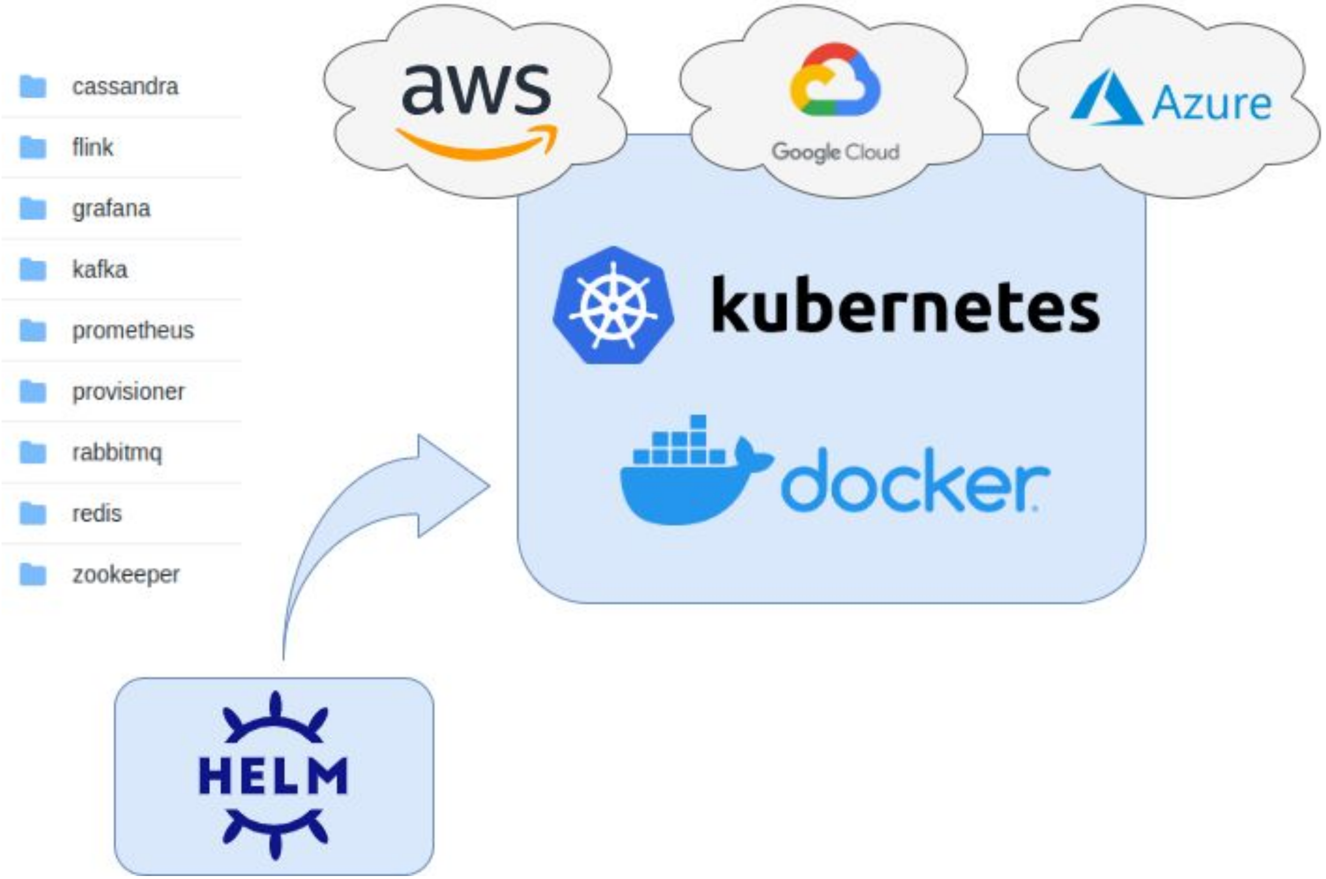
- Interfaces & Dataflows
- Systems & Configurations
- Data Formats & Enrichment
- Security & Encryption

AP6: Analysis Platform



DOS
Distributed and Operating Systems.

Virtual Deployment and Cloud Orchestration



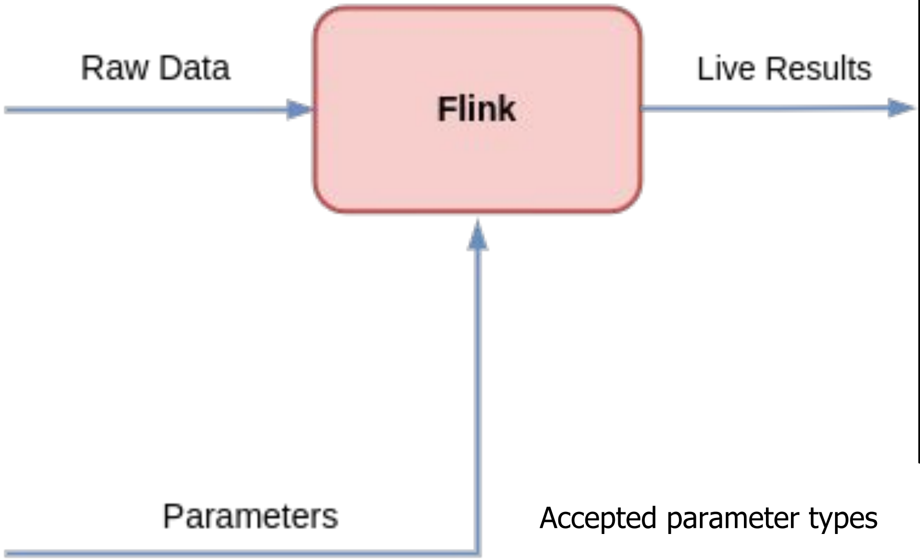
AP6: Analysis Platform

DOS
Distributed and Operating Systems.

Enrichment Job

Original SenML packet:

```
{
  "n" : "a8404177f1828cca-distance",
  "t" : 1619449203,
  "v" : 3.1415
}
```



Enrichment SenML packet:

```
{
  "n" : "a8404177f1828cca-distance",
  "t" : 1619449219,
  "v" : 1.2345,
  "g" : "53.800651,-4.064941",
  "r" : "617067112908455935",
  "ty" : "TMP36",
  "u" : "Cel",
  "c" : 28.912
}
```

Fixed format for parameter

```
{
  "n" : "a8404177f1828cca-distance",
  "t" : 1619431200,
  "p" : "geolocation",
  "v" : "52.503620,13.661910"
}
```

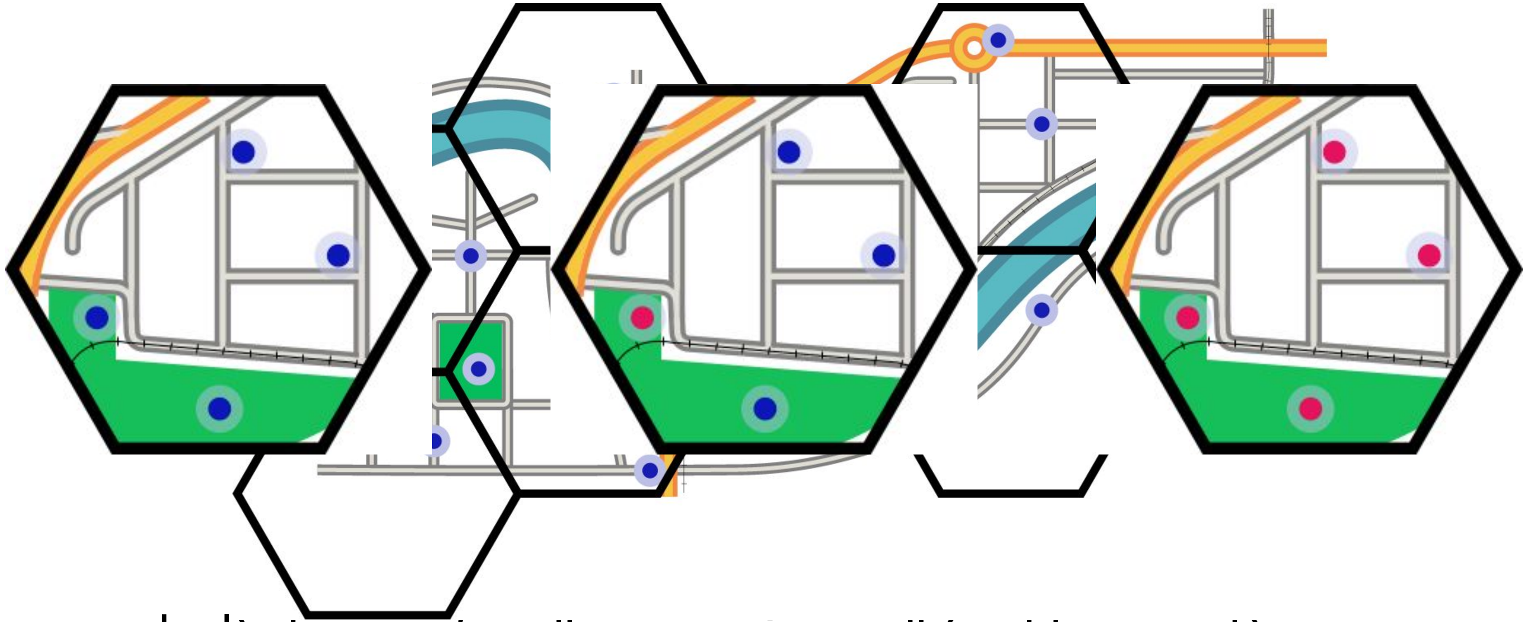
Accepted parameter types

- Geolocation - e.g. "48.8718204,2.7799161"
- Type – e.g. "TMP36"
- Unit – e.g. "Cel"
- Conversion – e.g. "[(-1, -1.59), (0, -0.80), (1, 3.76)]"

AP7: Anomaly Detection

Predictive Maintenance (Neighborhoods)

- Make use of Complex Event Processing (CEP)



This detects:

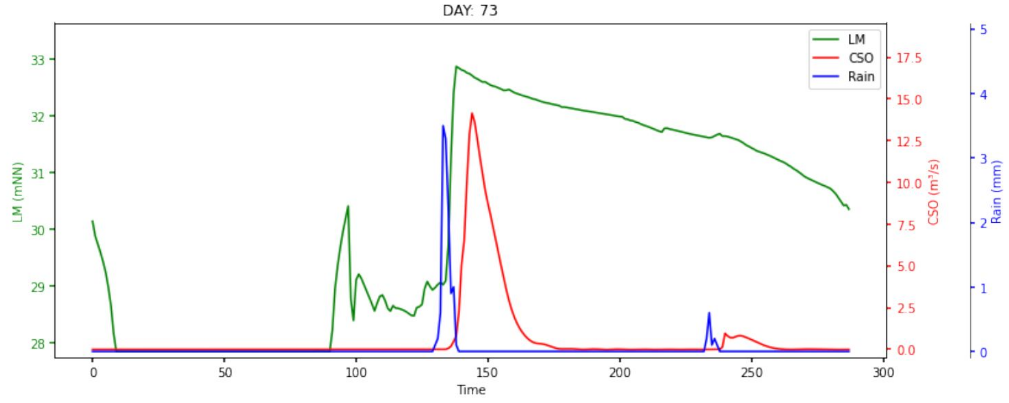
- Rapid (non-gradual) changes for all sensors in a cell (sudden event)
- Anomalies of one sensor (Overflows and Clogging)

AP7: Anomaly Detection

CSO Prediction

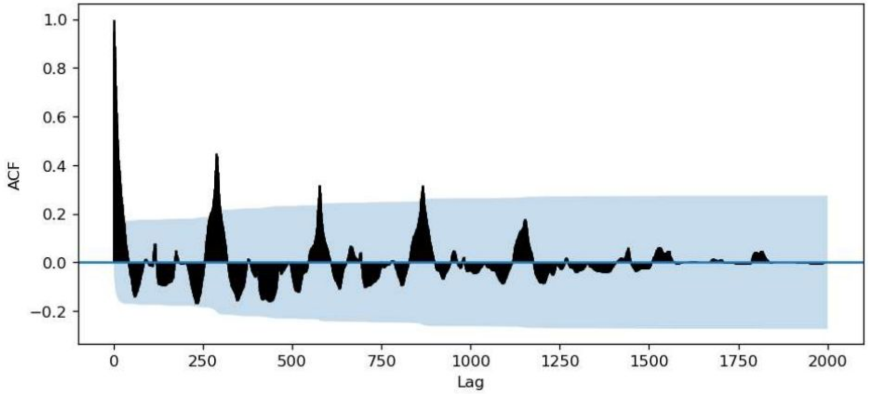
- Two B.Sc. thesis evaluated the use case of detecting CSOs on BWB data
  - A: Analysis of autocorrelation. Comparison of ARIMA, Holt-Winters, fbprophet.
  - B: Use Machine Learning (LSTM Autoencoder)

Qualitative inspection



→ CSO shortly after water level starts to rise.
→ Rain measurements appear to provide early signal

Auto-correlation

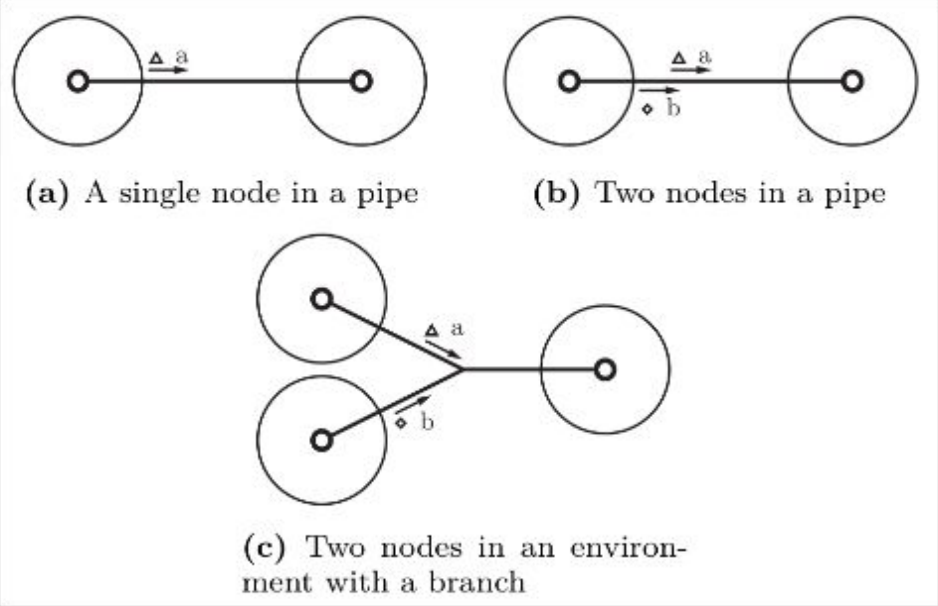


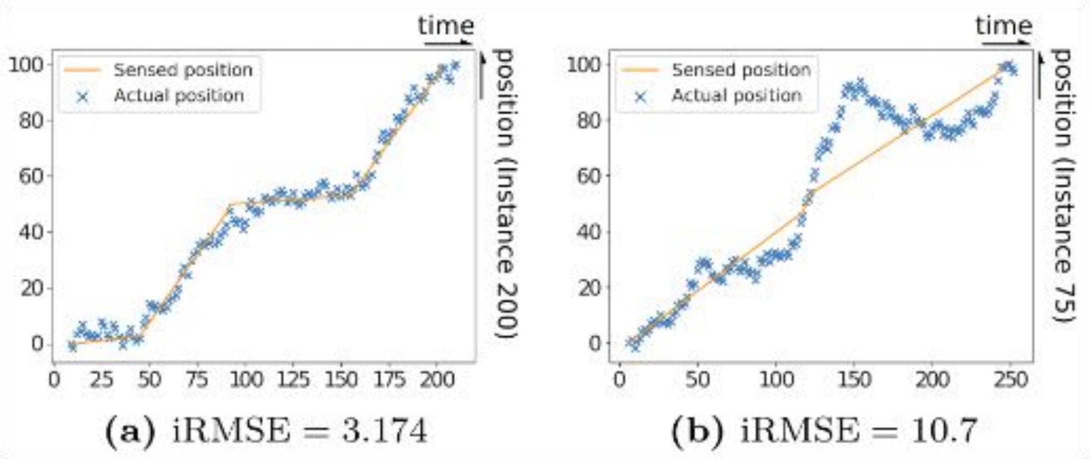
→ Measurements have auto-correlation at multiple lags.
→ Statistical methods *could* work.

AP7: Anomaly Detection

GRAL

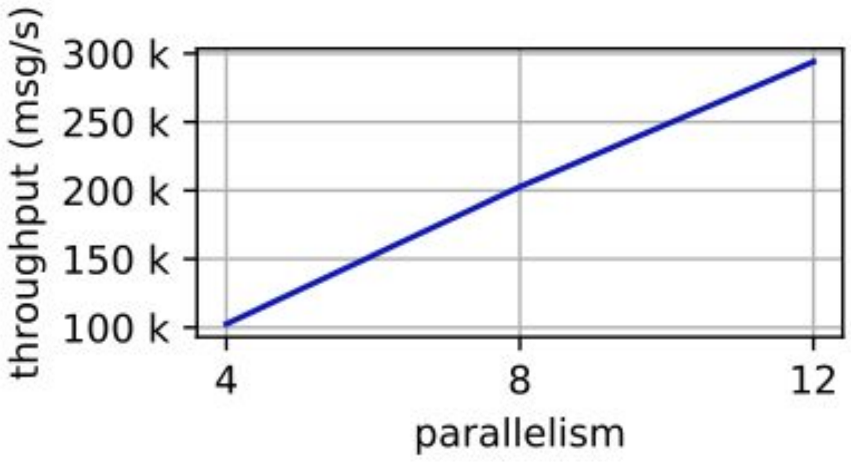
- Approach for localization of floating sensors



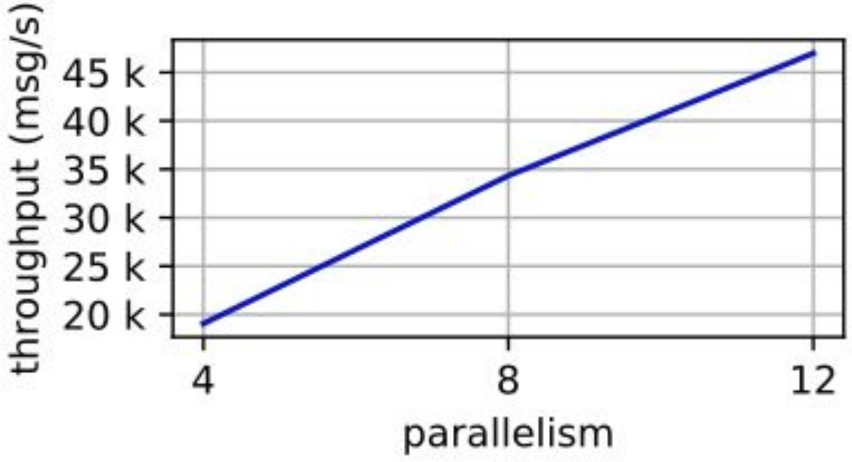


Scalability Experiment

- Evaluate scalability by measuring the maximal throughput rates over 3 different cluster sizes.



Scalability evaluation of Data Enrichment Job



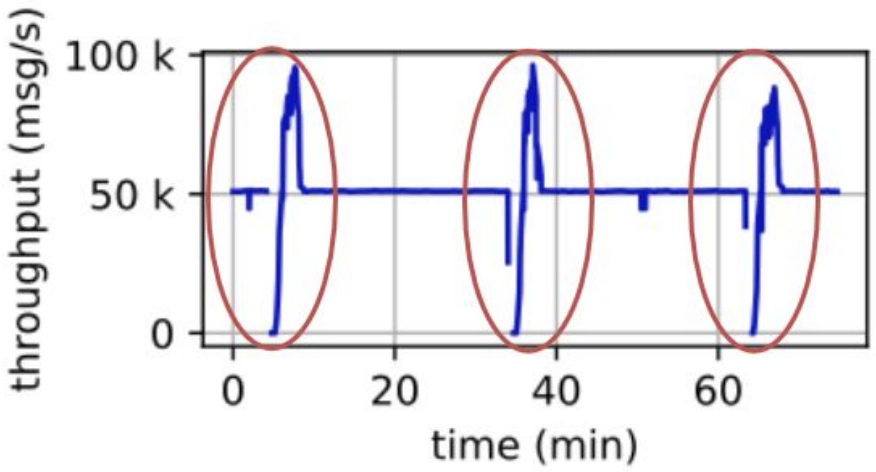
Scalability evaluation of Neighbourhood Analytics Job

- Scaleout behaviour was approximately linear, indicating no bottlenecks

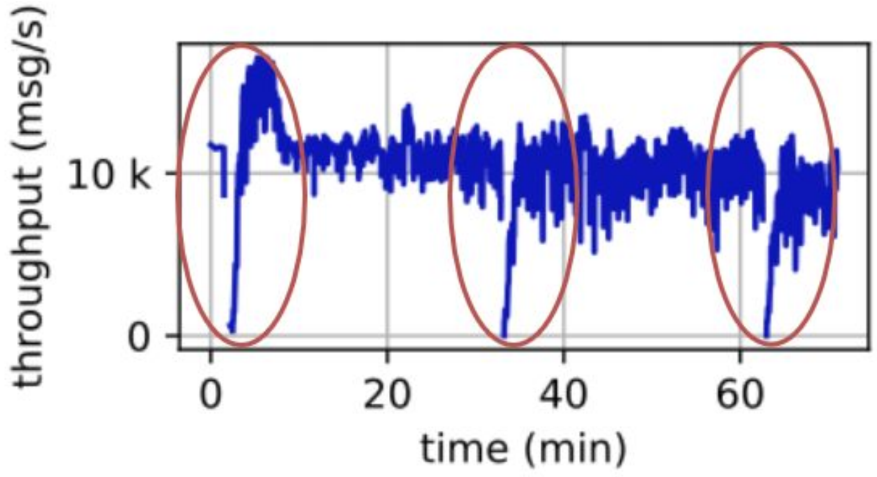
System Evaluation

Reliability Experiment

- 3 failures injected into random worker node for each running job



Fault tolerance evaluation of Data Enrichment Job



Fault tolerance evaluation of Neighbourhood Analytics Job

- Average recovery times were 194 and 243 seconds respectively

DOS
Distributed and
Operating Systems.

Publications

- Lorenz, F. and Geldenhuys, M. and Sommer, H. and Jakobs, F. and Lüring, C. and Skwarek, V. and Behnke, I. and Thamsen, L. (2020). **A Scalable and Dependable Data Analytics Platform for Water Infrastructure Monitoring**. IEEE International Conference on Big Data (pp. 3488 - 3493). IEEE.
- Lorenz, F. and Thamsen, L. and Wilke, A. and Behnke, I. and Waldmüller-Littke, J. and Komarov, I. and Kao, O. and Paeschke, M. F**ingerprinting Analog IoT Sensors for Secret-Free Authentication**. Conference on Computer Communications and Networks (ICCCN). IEEE. 2020.
- Geldenhuys, M. and Will, J. and Pfister, B. and Haug, M. and Scharmann, A. and Thamsen, L. (2021). **Dependable IoT Data Stream Processing for Monitoring and Control of Urban Infrastructures**. IEEE International Conference on Cloud Engineering (IC2E). IEEE.
- Haug, M. and Lorenz, F. and Thamsen, L. (2021). **GRAL: Localization of Floating Wireless Sensors in Pipe Networks.** IEEE International Conference on Cloud Engineering (IC2E). IEEE.

Open access papers
available on ARXIV

Source code available at github:
https://github.com/dos-group/water-analytics-cluster
https://github.com/dos-group/water-analytics-enrichment

Teaching

Presentations in several courses at TU Berlin

- As an application example in bachelor and master Computer Science courses
- Presentation of project ideas and results in the European teaching network *ide3a* (https://ide3a.net/)
    - 1st and 2nd winter schools on **Smart Sensing** (Winter 2020 and 2021)
    - 1st winter school on **Smart Cities** (Spring 2021)
    - Lecture series on **Critical Infrastructures and Digitalisation** (Summer 2021)
- Supervision of several theses related to the project

international alliance for digital e-learning, e-mobility and e-research in academia

Research / Teaching

Theses

- Martin Haug. 2019. **GRAL: Localization of Floating Wireless Sensors in Pipe Networks**.
- Viktoria Bill. 2020. **Using Statistical Models to Predict Combined Sewer Overflows Based on Distributed Sensor Data**.
- Tim Stahl. 2020. **Using IoT Devices and Hidden Markov Models to Estimate Clogging of Street inlets**.
- Kyra Kerz. 2020. **Autoencoder-Based Anomaly Detection for Combined Overflow Monitoring**.
- Marcin Ozimirski. 2020. **Sink-directed Routing for Wake-Up Sensor Mesh Networks.**
- Benjamin Pfister. 2021. **Fine-grained Failure Injection in Distributed Stream Processing Environments**.
- Ricardo Romanowski. 2021. **Praktische Evaluation einer neuen Methode zur Authentifikation analoger IoT Sensoren**.

Summary

- Established the requirements of a dependable critical infrastructure monitoring platform.
- Presented the overall system architecture as agreed upon in the Interface Control Document.
- Described the Analysis Platform, the systems, and the flow of data within the pipeline.
- Explained the use of virtual deployment and cloud orchestration for maintainability.
- Highlighted the methods used for data enrichment and anomaly detection.
- Presented the results of experiments for scalability and reliability.
- Showed the outcomes with regards to research and teaching to promote/increase the body of knowledge around this topic.