An IT infrastructure for smart levee monitoring and flood decision support

Bartosz Baliś, Marek Kasztelnik, Maciej Malawski, Piotr Nowakowski, Bartosz Wilk, Maciej Pawlik, Marian Bubak

AGH University of Science and Technology
Department of Computer Science
Kraków, Poland

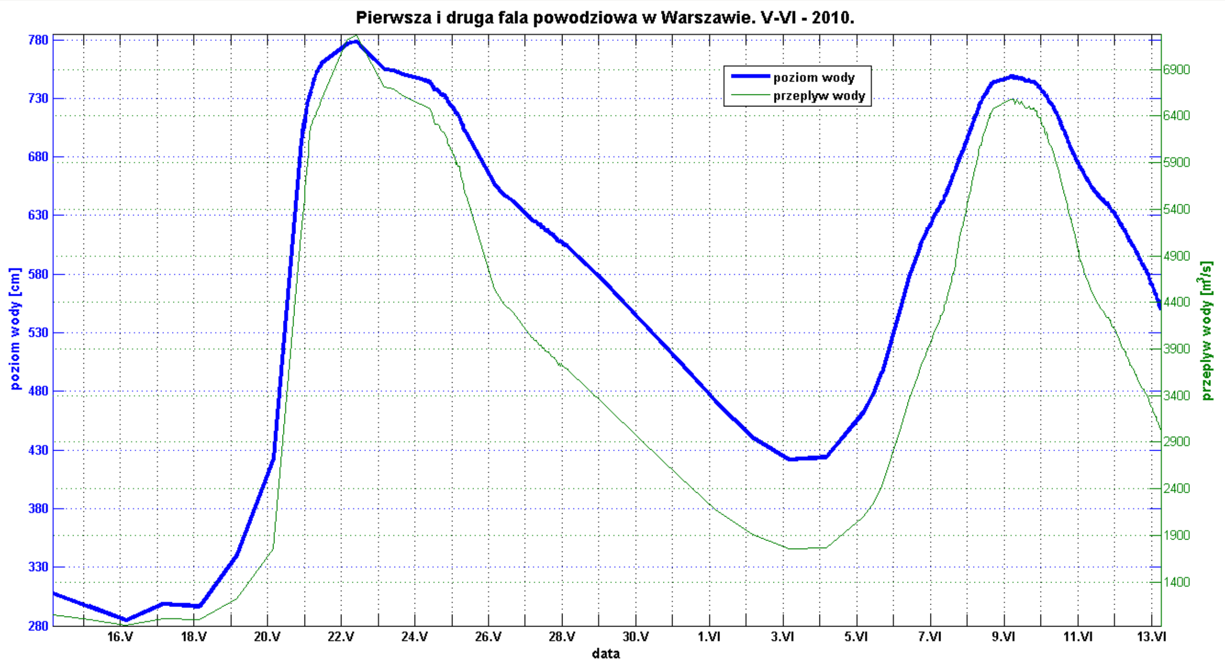
CGW ’16, Kraków, October 2016

Agenda

- Motivation, ISMOP IT infrastructure
- Data processing and resource management
- Data management
- User interfaces (live demo)
- Holistic approach to system optimization

Motivating scenario

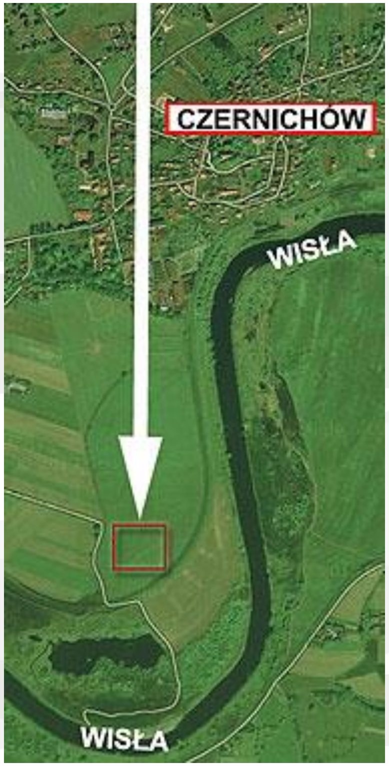
- Flood threat due to a passing water wave
- High water levels lasting up to several weeks
- Increased probability of levee failure



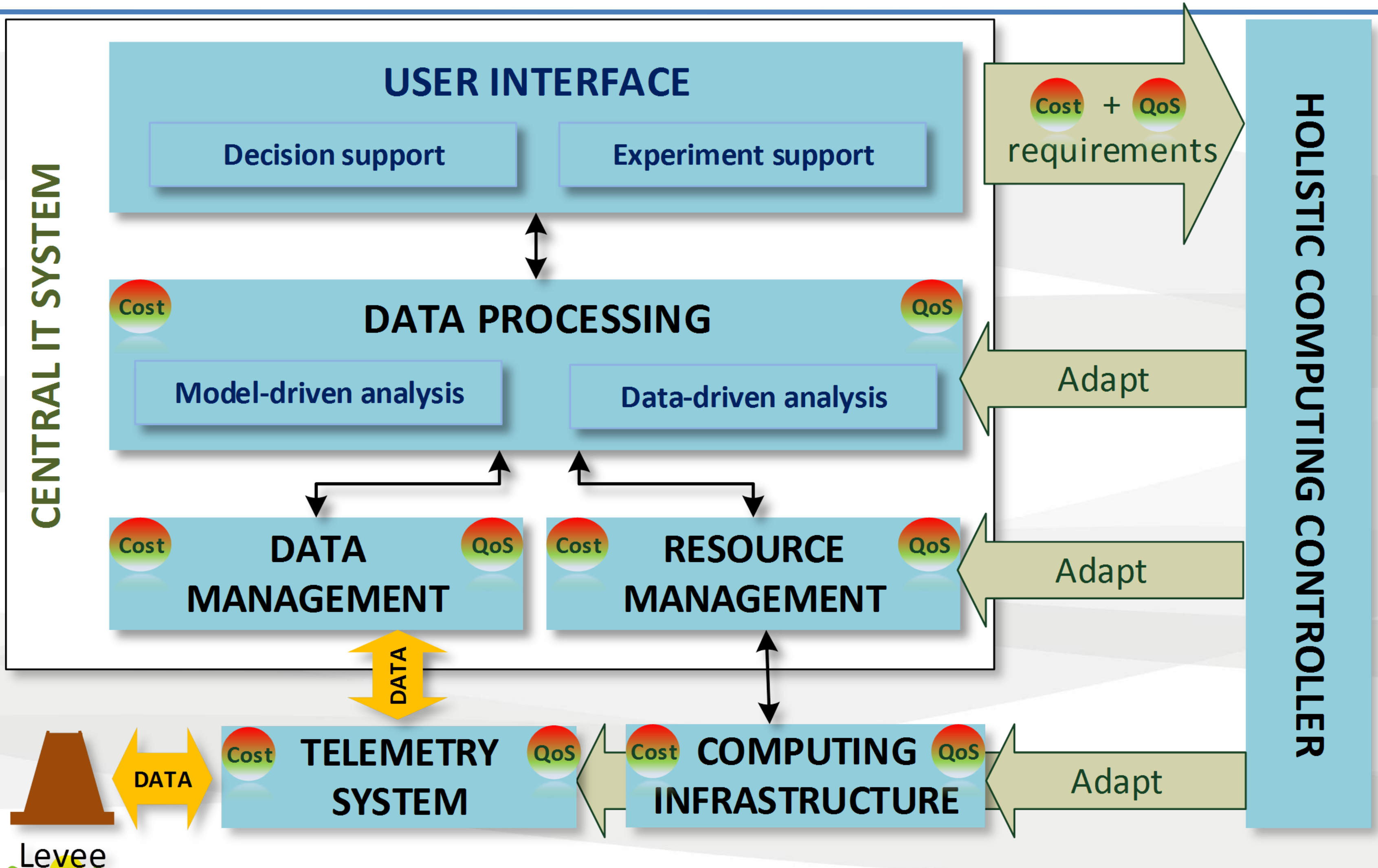
ISMOP project overview

ISMOP: a smart levee monitoring and flood decision support system

- Construction of an experimental levee
- Design of wireless sensors
- Innovative data transmission system
- Model- and data-driven modeling of levee behavior
- **Monitoring and decision support system**



The ISMOP IT infrastructure



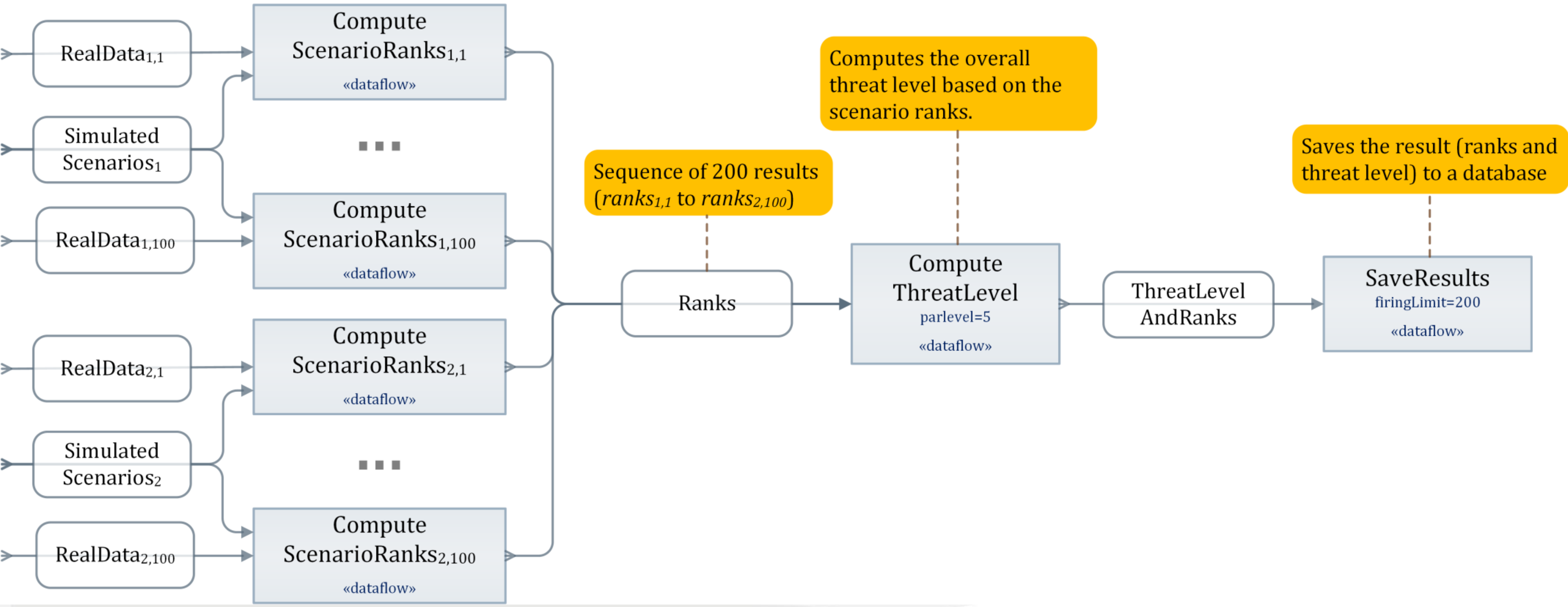
DATA PROCESSING AND RESOURCE MANAGEMENT

Challenge: urgent computing support

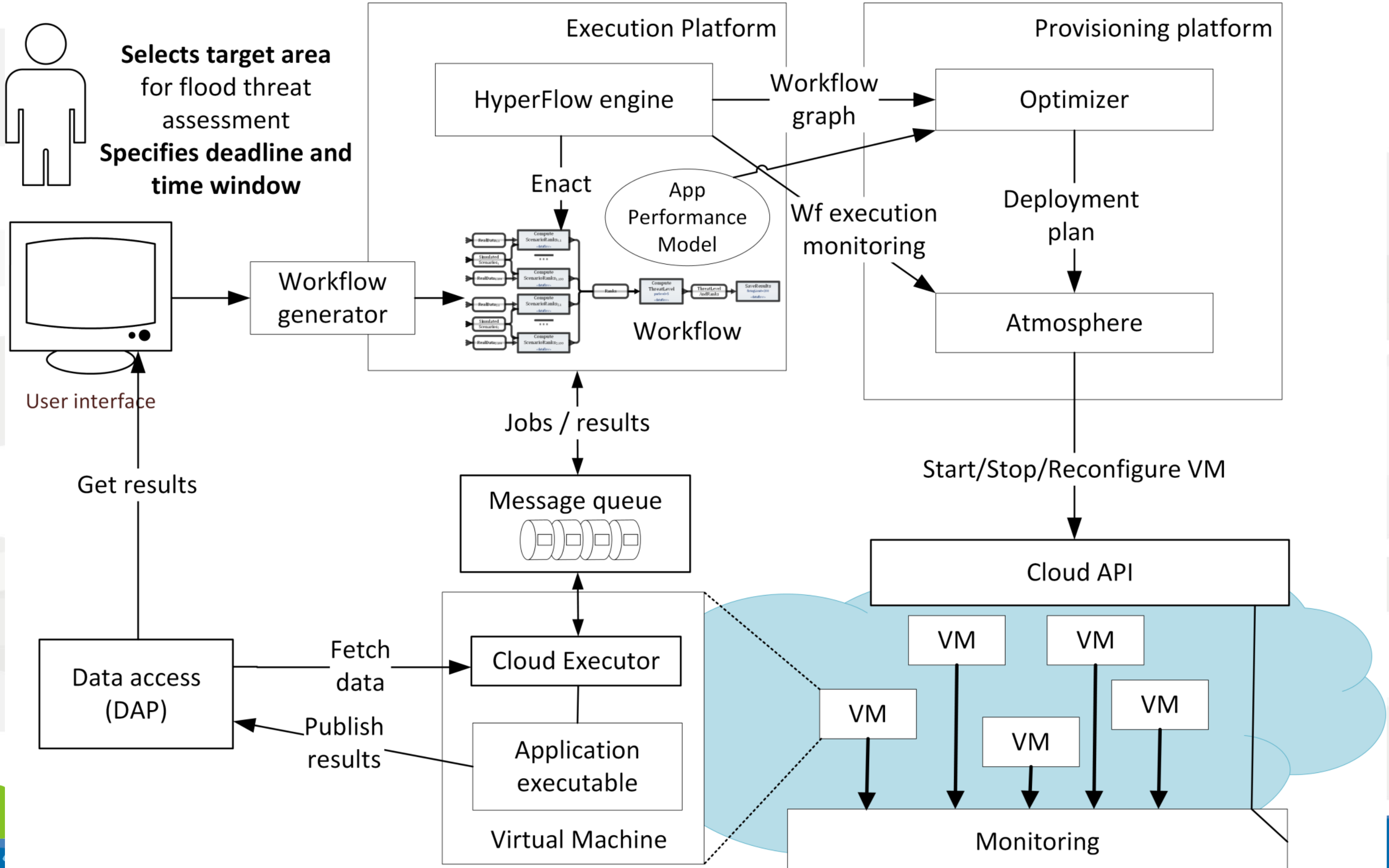
- Assess flood threat risk for large area (50+ km) of levees
- Compute results by a specified deadline

Objective: dynamic provisioning of cloud resources to meet the deadline

Levee breach threat assessment workflow for many levee sections

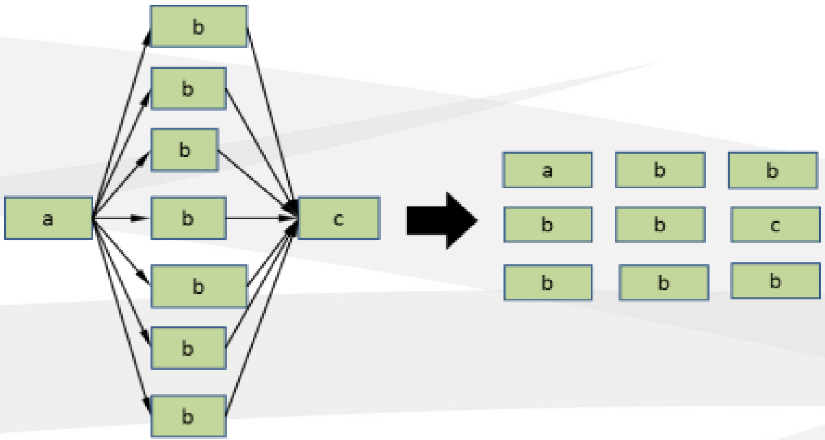


Urgent computing implementation



Resource provisioning model

- Bag-of-tasks model
  - Selection of dominating tasks
  - Uniform task runtimes



- Performance model: $T = f\,(v, d, s, \ldots)$
  - T – *total computing time*
  - v – number of VMs
  - d – time window in days
  - s – number of tasks (sections)

Resource provisioning model

$$\boldsymbol{T} = a * \frac{\boldsymbol{s} * \boldsymbol{d}}{\boldsymbol{v}} + b * \boldsymbol{v} + c \qquad (1)$$

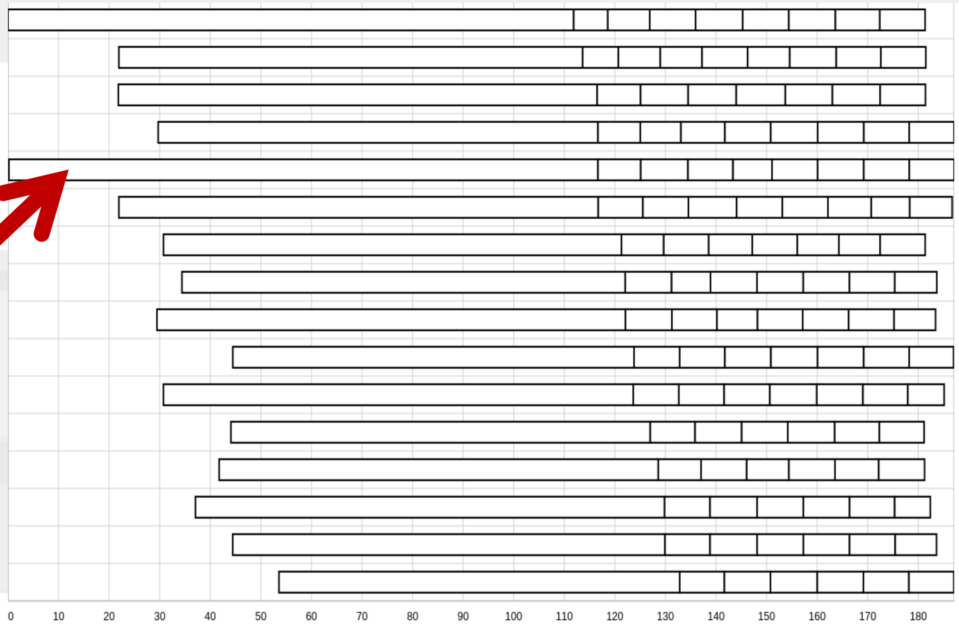
$\boldsymbol{T}$ – *total computing time*
$\boldsymbol{v}$ – number of VMs
$\boldsymbol{d}$ – time window in days
$\boldsymbol{s}$ – number of tasks (sections)

Parameters a, b, c to be determined experimentally

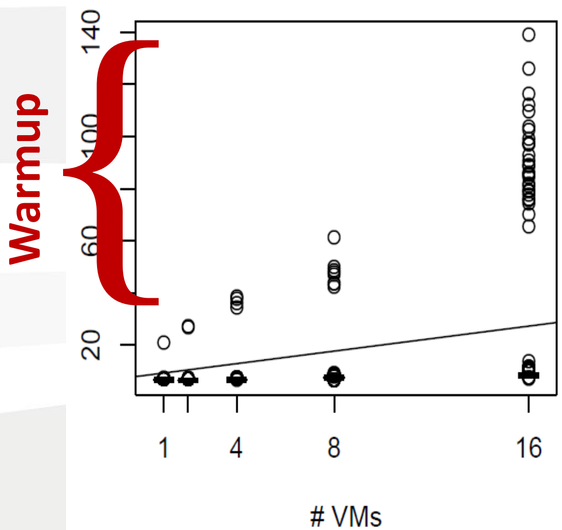
Solve eq. (1) to compute $\boldsymbol{v}$ (# of VMs) given $\boldsymbol{T}$ (deadline)

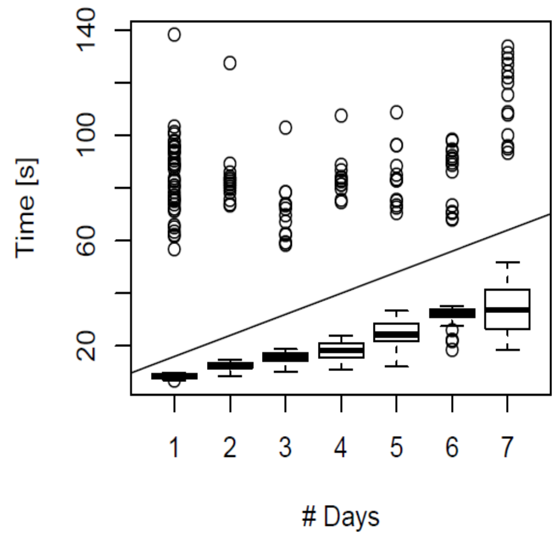
Experiments

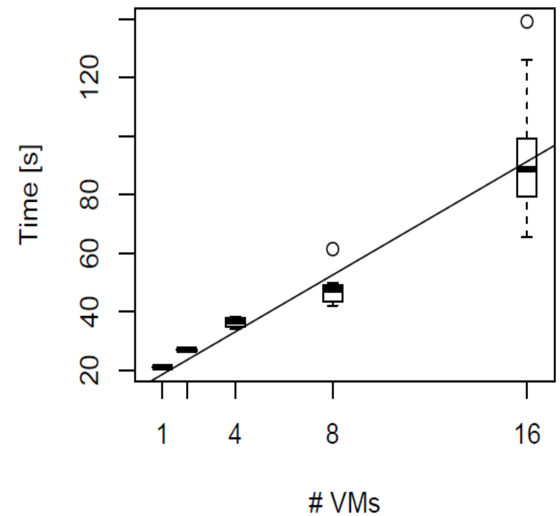
- Setup: private cloud infrastructure
  - a node with 8 cores (Xeon E5-2650)
  - virtual machines (1VCPU, 512MB RAM)
  - data for simulated scenarios (244MB total) on local disks
- Test runs:
  - 1-1024 sections
  - 1-16 VMs
  - 1-7 days time window
- Warmup tasks:



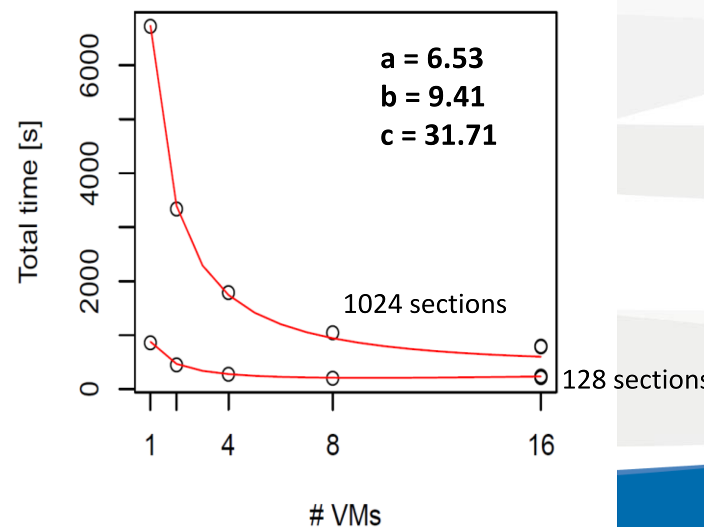
Analysis of results







- Warmup tasks clearly separated as outliers
- Linear functions
- Parameters a, b, c determined using non-linear fit
- The model fits well to the data

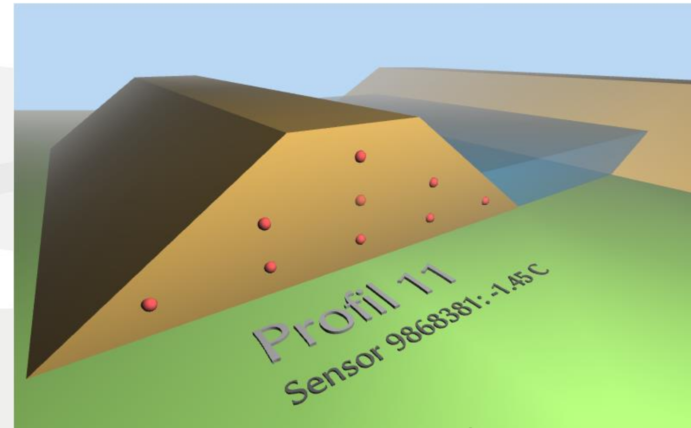


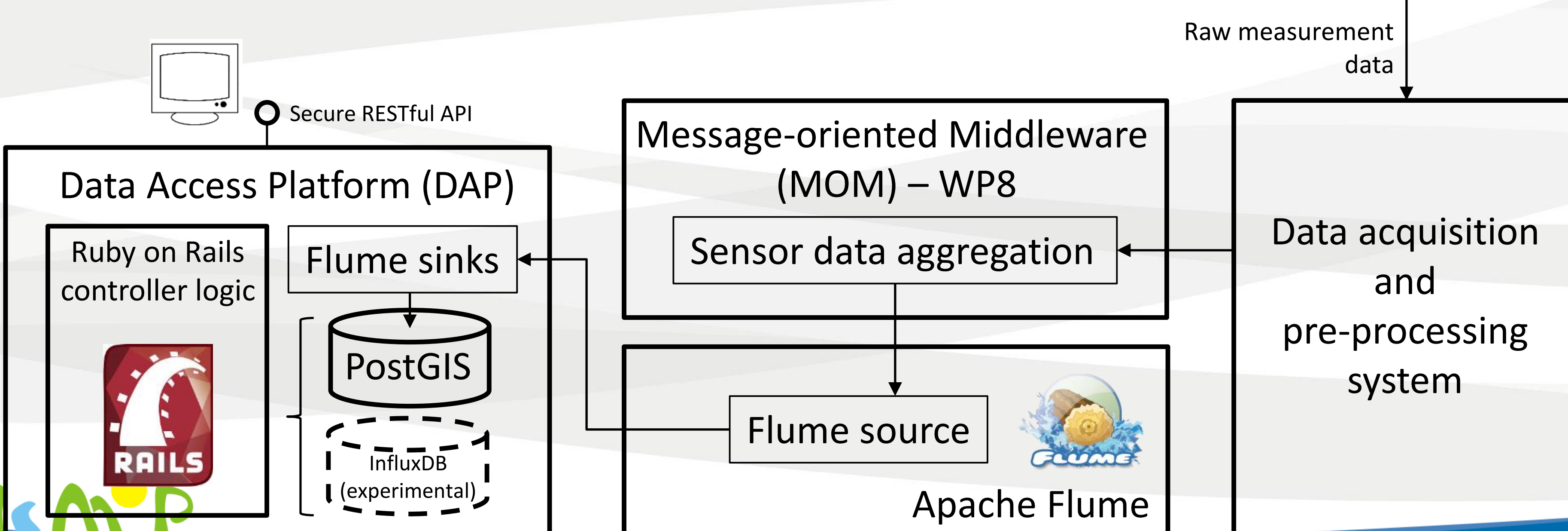
DATA MANAGEMENT

Reliable data storage

Data Access Platform (DAP):

- Ensures fault-tolerant collection of sensor data
- Provides redundant, heterogeneous backend data storage
- Implements real-time query capabilities for ISMOP UIs and simulations





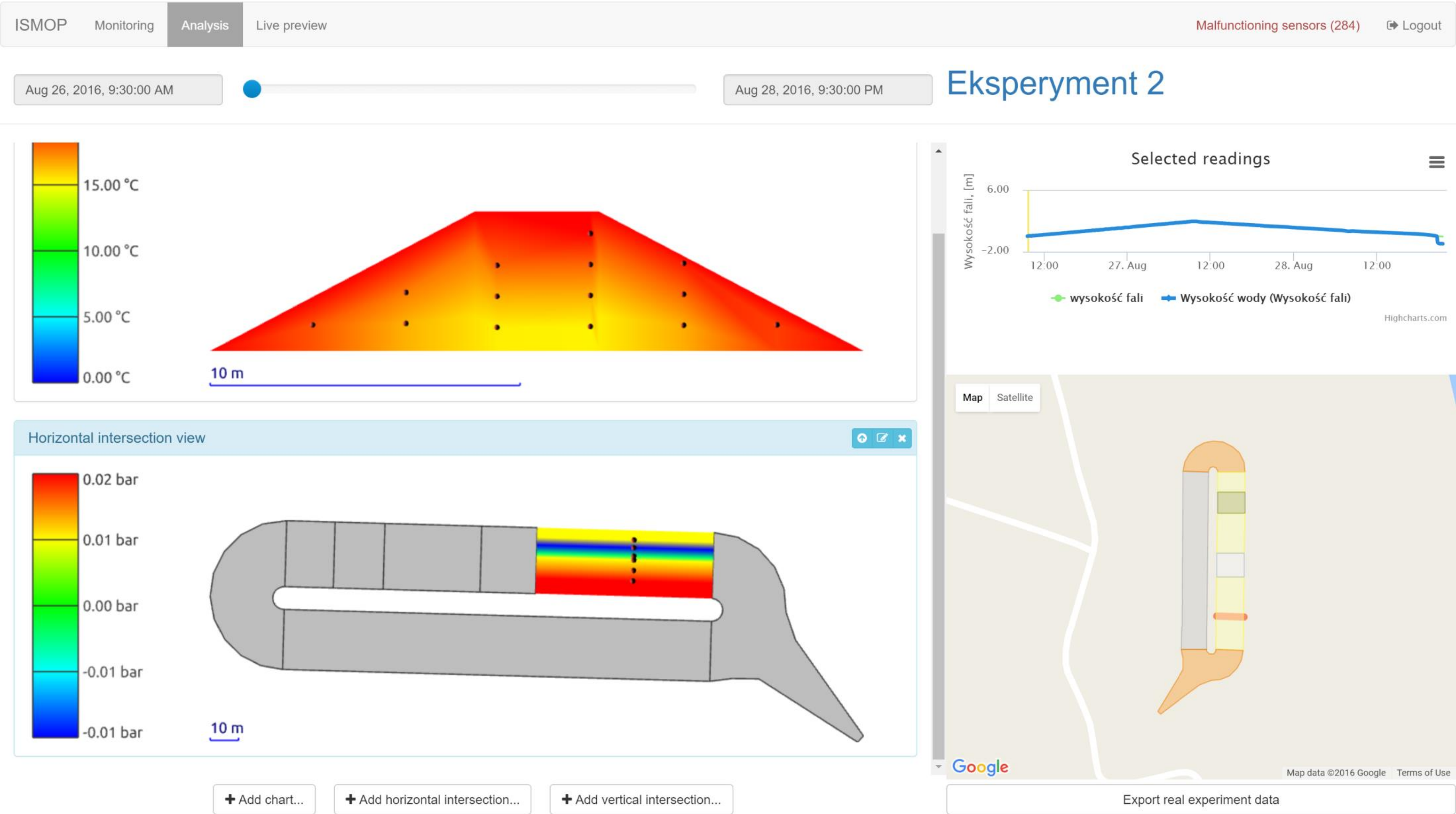
Health monitoring and data quality

- Monitoring of sensor health and reactions outages/malfunctions: more practical to implement at the data management layer
  - End-to-end principle
  - Availability of metadata
  - Access to historic data
- Data quality
  - Sanity checks: are sensor readings correct?
  - Quality metrics (accuracy, validity, completeness, consistency)
  - E.g.: Do we have all readings? Do values fall within expected range?

USER INTERFACE

User interface functions

- Experiment monitoring
- Data visualization and retrieval
- Decision support
  - Data analysis results
  - Model-based prediction of levee behavior


ISMOP
Monitoring
Analysis
Live preview
Malfunctioning sensors (284)
Logout
Aug 26, 2016, 9:30:00 AM
Aug 28, 2016, 9:30:00 PM
Eksperyment 2
15.00 °C
10.00 °C
5.00 °C
0.00 °C
10 m
Selected readings
Wysokość fali, [m]
6.00
-2.00
12:00
27. Aug
12:00
28. Aug
12:00
wysokość fali
Wysokość wody (Wysokość fali)
Highcharts.com
Horizontal intersection view
0.02 bar
0.01 bar
0.00 bar
-0.01 bar
-0.01 bar
10 m
Map
Satellite
Google
Map data ©2016 Google
Terms of Use
+ Add chart...
+ Add horizontal intersection...
+ Add vertical intersection...
Export real experiment data

HOLISTIC APPROACH TO SYSTEM OPTIMIZATION

B. Baliś, R. Brzoza-Woch, M. Bubak, M. Kasztelnik, B. Kwolek, P. Nawrocki, P. Nowakowski, T. Szydło, K. Zieliński

Problem formulation

- **Goal**: optimize the IT infrastructure in terms of
  - **QoS parameters** (measurement interval, performance, timeliness, etc.)
  - **Cost of operation**
  - **Energy consumption**
- **Problem**:
  - These objectives are contradictory
  - Their priorities depend on the mode of operation (e.g. cost important in the normal mode, perfomance and timeliness in the urgent mode)

Optimization of IT infrastructure

Multi objective optimization problem:

- **Objective functions**: non functional properties of the system
- **Decision variables**: configurations of all subsystems of the IT infrastructure

Isolated approach: each subsystem is optimized separately

Holistic appoach: the entire system (all subsystems) is optimized

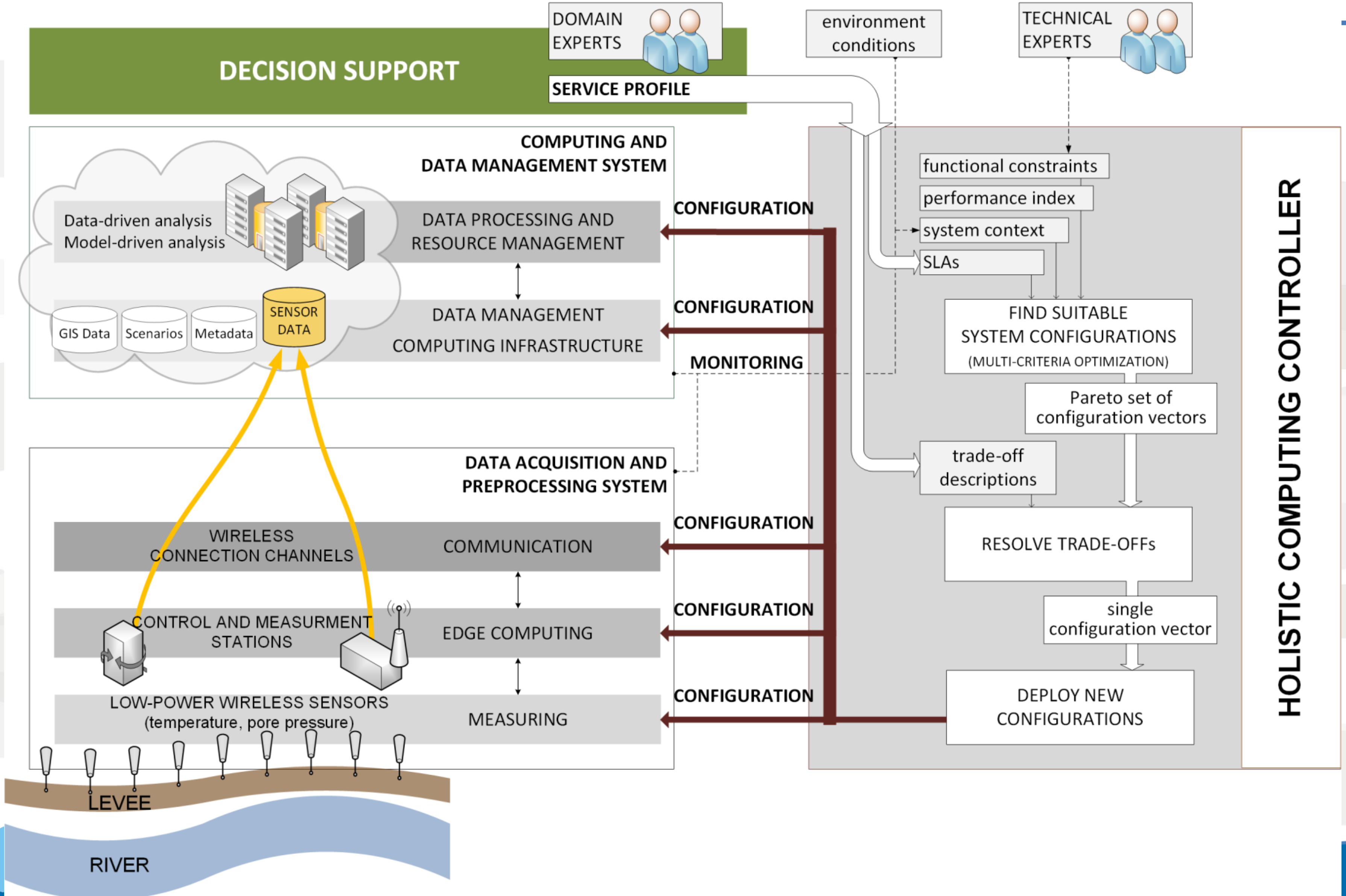
Objective functions (examples)

- **Operating cost** (OPC): expenses required to maintain operation of the system
- **Energy Efficiency** (EE): a function indicating how energy efficient the system is
- **Data measurement interval** (DMI): time interval at which sensor parameters are captured
- **Data processing interval** (DPI): time interval at which data analyses are performed
- **Data processing time** (DPT): maximum time to complete a single data analysis

Example configurable properties

Subsystem	Configurable properties
Data processing and resource orchestration	***Processing interval***: 5,15,60,720,1440 min ***Scheduling algorithm***: time-optimized, cost-optimized
Computing infrastructure	***VM allocation policy***: aggressive, conservative
Communication	***Transmission protocol order***: (Xbee-SMS-GPRS), (GPRS-SMS-Xbee)
Edge computing	***Encryption***: on, off ***Data aggregation***: off, low, high
Measurement	***Measurement interval***: 1, 5, 15, 60, 720, 1440 min ***Accuracy***: high, low

Holistic computing controller



Conclusion

- A modern IT infrastructure for environmental monitoring and decision support needs to address non-trivial technical requirements and provide demanding quality of service
- ISMOP proposed IT infrastructure is build on the IoT-Cloud technology stack
- Provides urgent computing capabilities
- Holistic approach to system management optimizes configuration of the entire IT infrastructure towards various optimization objectives

http://www.ismop.edu.pl

NEOSENTIO SWECO AGH Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie INFORMATYKA KGIS KHgi

Acknowledgements

This work is partially supported by
the ISMOP project (PBS1/B9/18/2013);
and by the AGH grant no. 11.11.230.124.

AGH University of Science and Technology

The National Centre for Research and Development