

## 1 MOTIVATION & OBJECTIVES

In most parts of the world, fast growing urbanization has faced several challenges throughout the last decades. Achieving a sustainable, environment-friendly and highly operating cities for a better quality of life requires innovative solutions brought by cutting edge technology.

WAVES project deploys water management solutions through real-time infrastructure to achieve an advanced level of comprehensive awareness.

**REAL TIME PROCESSING**

**SEMANTIC WEB**

**INTERNET OF THINGS**   **HIGH THROUGHPUT**   **DISTRIBUTED ENVIRONMENT**

## 3 ARCHITECTURE: ABSTRACT LEVEL

WAVES is a coherent and consistent platform created by combining several components of different natures linked together in an efficient fashion. Since WAVES is deeply related to massive semantic streams in a real-time environment, the architecture needs to be generic and multipurpose in order to handle several use cases.

Time Reference   Unique object Ids

Sliding Windows   RDF stores

As Waves aims to distribute stream processing among a cluster of machines, it need a mechanism to distribute stream events to the processing nodes. Several high throughput middlewares are used in the context of big data streaming.

DATA SOURCES → DATA CLEANING (Filter 1, 2, 3) → DATA CONVERSION (Semantic Conversion) → DISTRIBUTION SYSTEM (Node 1, 2, 3, ..., N) → QUERY ENGINE (Query 1, 2, 3)

Additional components: DATA COMPRESSION (A, B, C), QUEUE MESSAGING, VISUALIZATION INTERFACE (Summary, Alerts, Prediction), RDF TRIPLE STORE.

## PARTNERS

**AVERAGE WATER LOSS IN 2015**

**30%**

## 2 CHALLENGES:

WAVES platform will have the ability to collect data streams from various sources in order to create new insight for users accessed through innovative information following several processing steps such as cleaning, pre-processing, filtering, summary and visualization, Building such system provides multiple challenges on various levels:

**01 OPENNESS**   **02 SCALABILITY**   **03 CONCURRENCY**   **04 SECURITY**

## 4 RESULTS: CURRENT PERFORMANCES

❖ RDF has quite a verbose data structure, therefore being able to reduce the size of each transfer is of a critical importance.

**Precision**   **Recall**

Legend: No Compression (Blue), GZIP (Green), RDSC (Light Green)

- WAVES results compared to C-SPARQL are shown with regard to Precision and Recall for a query of medium complexity on 4s window and 1s step.

## 5 ARCHITECTURE: IMPLEMENTATION

WAVES arghitecture relies heavily on three robust components with a solid reputation within the Big Data Community:

- ✓ STORM
- ✓ KAFKA
- ✓ REDIS

PROCESSING: APACHE STORM™

MESSAGING QUEUE: kafka

DATA BASES: MESSAGING, SESAME TRIPLE STORE

REST WEB SERVICE

CLEANING: K-MEANS, DBSCAN, OPTICS, ROCK

RDF CONVERSION

## ACKNOWLEDGEMENTS

- WAVES is an FUI project funded by two public partners: BPI France and Cap Digital.
- The project's duration is 3 years and involves 5 different partners anf a global team of 17 members.

## CONTACT DETAILS

**Gabriel KEPEKLIAN**  
gabriel.kepeklian@atos.net

**Badre BELABESS**  
badre.belabess@atos.net