WAVES

Big Data Platform for Real-time Semantic Stream Management

WAVES Project

OUTLINE

- What is WAVES?
- Why WAVES?
- How WAVES?
- Achievements
- Contact

WHAT IS WAVES?

Massive Semantic Streams empowering Innovative Big Data Platform

What is a data stream?

- Golab & Oszu (2003): “A data stream is a real-time, continuous, ordered (implicitly by arrival time or explicitly by timestamp) sequence of items. It is impossible to control the order in which items arrive, nor is it feasible to locally store a stream in its entirety.”
- Massive volumes of data, items arrive at a high rate.

ABOUT WAVES

> *"The real innovation in smart cities will come from integrating technologies."*
>
> *Colette Malonye, European Commission's Head of Unit, Smart Cities*

Smart Cities challenges rise to a new level of complexity with every year's population growth. New megacities are being created at a dizzying pace around the world. Reaching the next level in development will require new ways of thinking that include cutting edge technologies based on the Internet Of Things, **Big Data** ecosystem and **Linked Data**.
Atos SE contributes to this paradigm shift by innovating and demonstrating new ways of using data to simplify and improve the management of cities through the development of a new project called WAVES. Rapidly growing cities threaten dangerously the availability of critical resources such as potable water. By conceiving a smart water management application dedicated to prevent leaks in the underground pipeline system, Atos SE aims at bringing majors water actors' awareness to whole new level.
WAVES project deploys an abstract level design that cover various domains where sensor networks and Linked Data are exploited such as traffic control, power consumption and health care improvement and energy optimization. Further details available at http://waves-rsp.org/

ABOUT WAVES

SOLUTION FOR MASSIVE SEMANTIC DATA STREAMS IN REAL-TIME

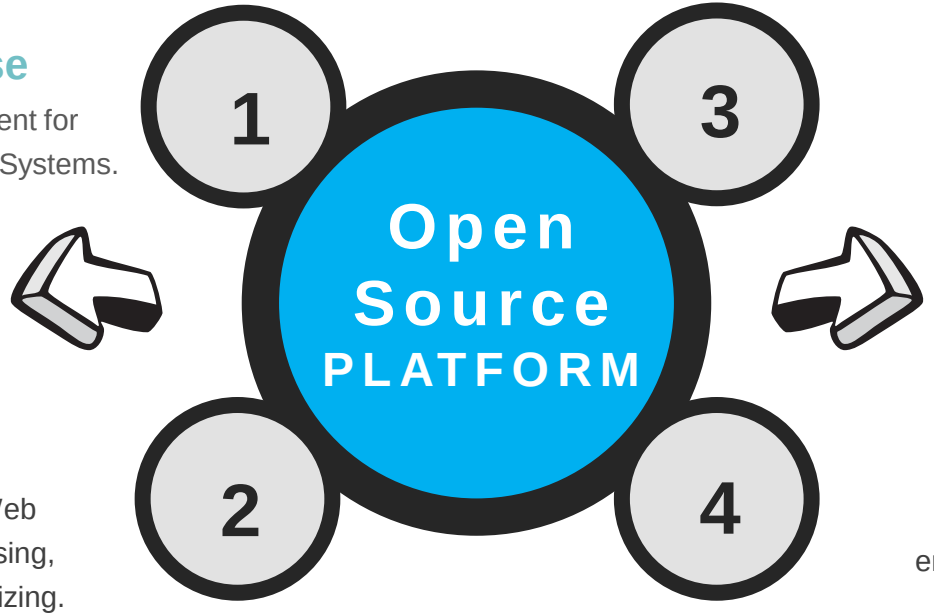
Current Use-Case

Water network management for industrial partner: Ondeo Systems.

CONTEXT

Various Tools

Big Data and Semantic Web Technologies: data cleansing, filtering, reasoning, visualizing.



Current Target

Detect anomalies in real-time in sensor networks.

GOALS

Final Objective

Design a generic inference-enabled and distributed platform for RDF stream processing.

WHY WAVES?

Addressing an environmental issue on a global scale

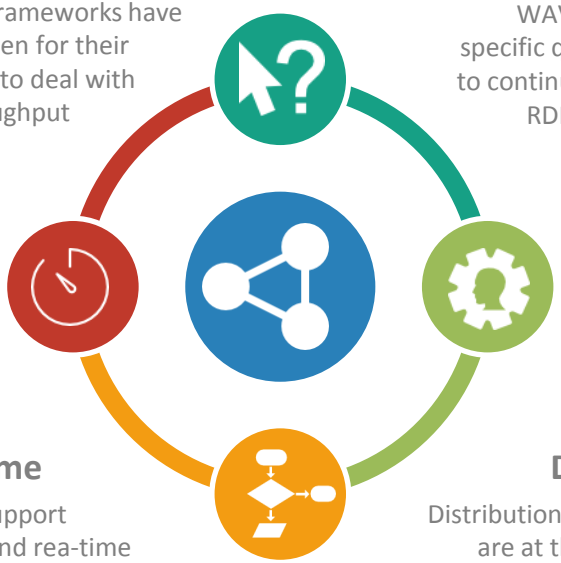
RATIONALE

Big Data

Big Data frameworks have been chosen for their capability to deal with high throughput

SPARQL Continuous Query

WAVES supports a specific query language to continuously process RDF data streams



Real-Time

WAVES support parallel and rea-time query answering over RDF data streams

Distribution

Distribution and scalability are at the heart of the architectural design to increase throughput

Multi-Purpose Stream Processing

Facing the new challenges of increasingly highly connected IOTs, WAVES is designed to analyze and act on real-time streaming data using continuous queries. These queries are executed in parallel in a distributed framework and support CEP-based operators over time-annotated RDF triples.

Reasoning Engine over RDF Data

As a stream processing platform, WAVES aims at handling high volumes of semantic data in real time with a scalable, distributed and fault tolerant architecture. This enables analysis of data in motion and anomaly detection supported by inference rules and reasoning capabilities,

APPLICATIONS DOMAIN

Website logs

Network monitoring

Financial services

Weather forecasting

circular economy

eCommerce

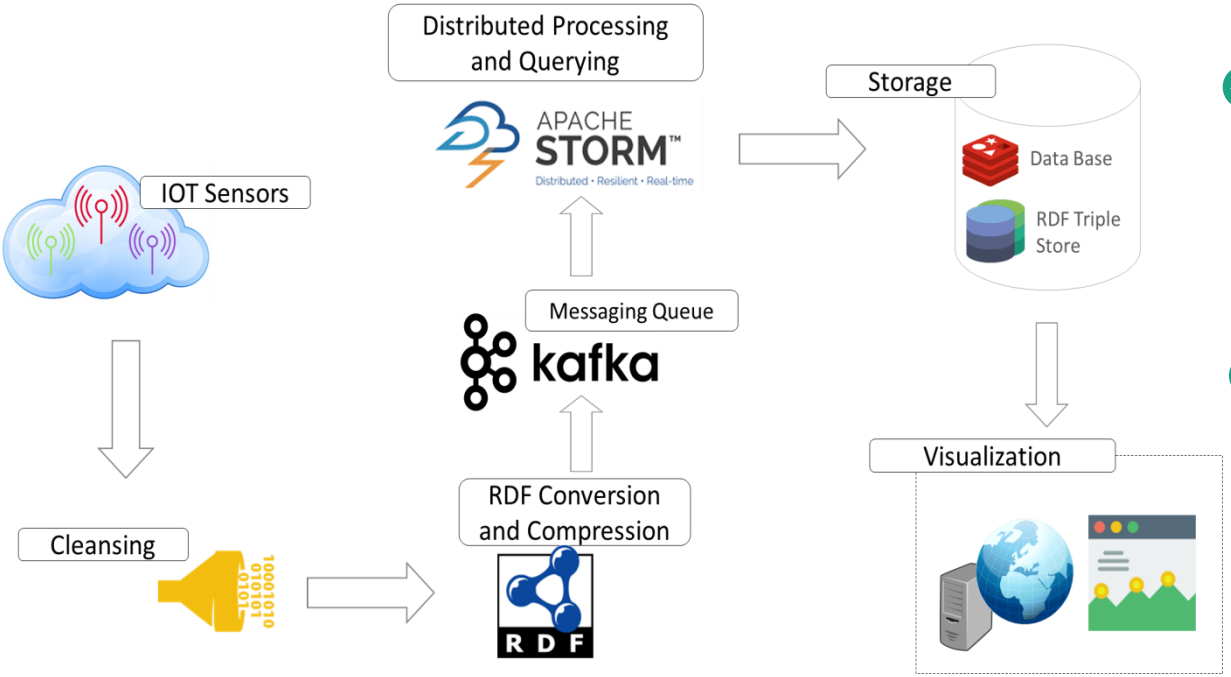
Traffic control

Power consumption

HOW WAVES?

Combining Big Data and Semantic Web technologies

SYSTEM ARCHITECTURE



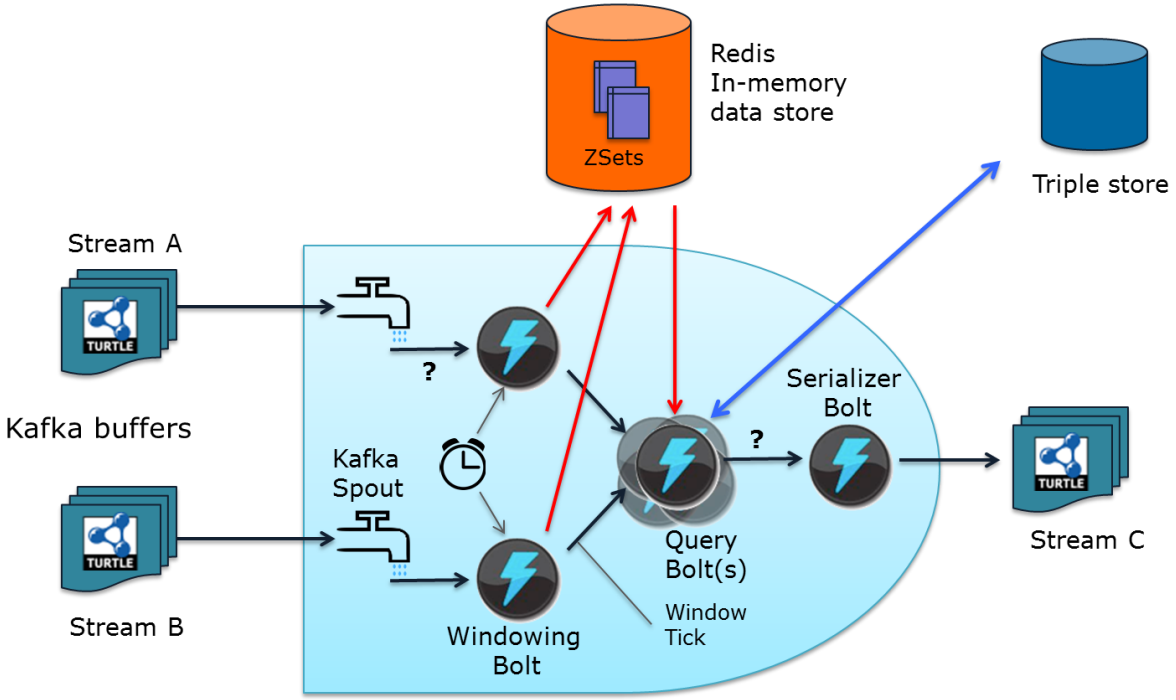
Big Data systems

WAVES architecture relies heavily on three robust components with a solid reputation within the Big Data community: Apache Storm, Kafka, and Redis.

Linked Data Principles

Waves converts sensor data to semantic streaming data based on popular ontologies such as SSN and QUDT. It supports SPARQL queries simultaneously over streaming and static data.

SYSTEM ARCHITECTURE



Distribution

The distribution in WAVES is enabled by the so-called Storm topologies. In each topology, there are at least a Kafka spout, a windowing bolt, a step bolt and a query bolt. A topology is a software unit that consumes data from Kafka and executes continuous SPARQL queries.

Modularity

The architecture is generic and multipurpose in order to handle several use cases. It contains pluggable modules, where the module is a self-contained unit in charge of executing some tasks. All modules depend on the core framework.

Technologies & Challenges

RDF
Data

REAL-TIME
Processing

Distributed
Environement

OPENNESS
& SECURITY

TECHNICAL CHALLENGES

- Massive semantic streams
- On-the-fly computation
- Robust and secure architecture

HUMAN CHALLENGES

- Consortium of 5 members: 1 start-up, 2 large companies and 2 academics
- Management for various profiles

ACHIEVEMENTS

Exposing Realizations and Current Advancement Stage

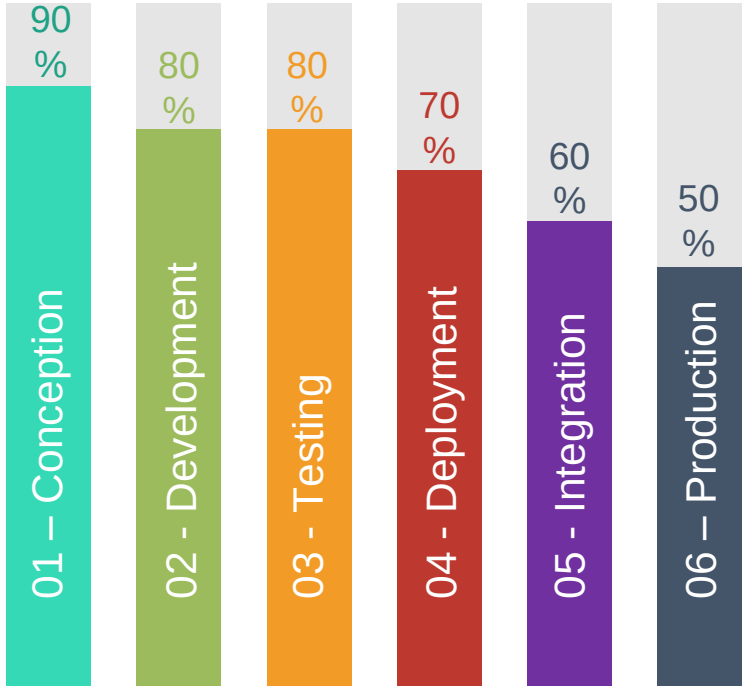
STATUS & ROAD MAP

Evolution Analysis

80%

Over the recent years, WAVES project went through several steps from the first submission for financial support to the final integration and production stages.

The organization of the project around a consortium of different members (i.e. 1 start-up, 2 large companies and 2 academics) allowed the team members to bring innovation and diversity for solving complex problems.



FULFILMENTS & REALIZATIONS

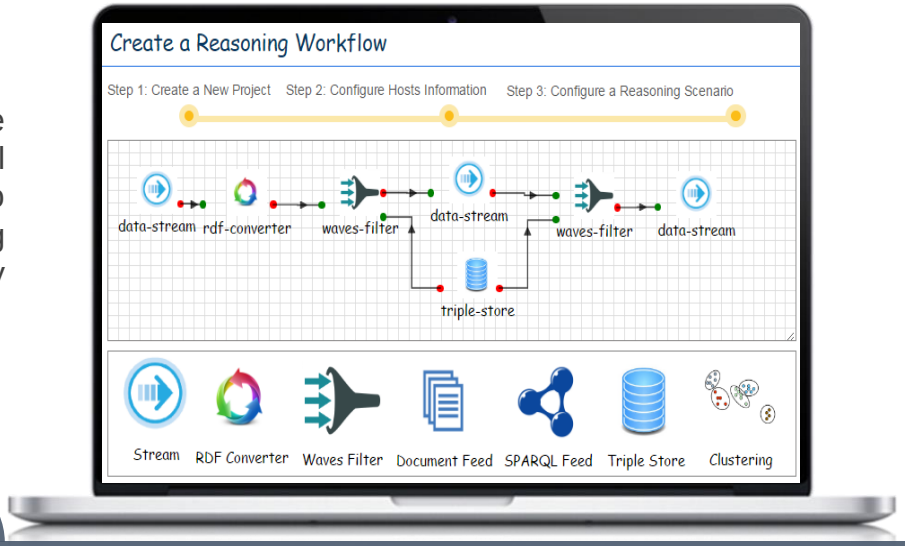
01 New API

WAVES is open-source and provides a new JAVA API for developers available at

http://waves-rsp.org/api/

02 Friendly UI

WAVES relies on a simple and effective graphical interface to allow users to configure the reasoning workflow and interact easily with the application.



03 High Performance

Compared to other engines, WAVES reaches a higher level of accuracy and fast processing under an important input load.

Cleansing

Eliminating outliers and dealing with absent values

Semantization

Converting sensor measures to semantic data

Compression

Downsizing the amount of data to reduce network overhead

Querying

Distributing queries over several machines for fast processing

Visualizing

Exposing the results of queries in a smart and attractive interface

Feel free to contact us!

We are friendly and social

80 Quai Voltaire, 95870
Bezons, France

@AtosFR AtosFR contact@waves-rsp.org