The CLASS Software Architecture

Eduardo Quiñones
{eduardo.quinones@bsc.es}
Barcelona Supercomputing Center (BSC-CNS)
General Information

- **Edge and Cloud Computation:** A Highly Distributed Software Architecture for Big Data Analytics
  - Under the scope of **H2020 ICT16-2017 (RIA)** - *Big data PPP: research addressing main technology challenges of the data economy*
  - **42 months** (starting January 2018)
  - **3.900.803 €** budget
Motivation: The Importance of CLASS

1. Geographically distributed data sources and data analytics requirements, e.g., smart cities
2. The fulfillment of real-time requirements inherited from the application domain
3. Constant increment of volume, variety and velocity of data-sets

A coordination of edge and cloud resources is needed!
The Vision of CLASS

1. Significantly increase the capabilities of the data analytics
   - Integrate both responsive data-in-motion and latent data-at-rest analytics in a single complex workflow
2. Fulfill the real-time requirements
3. Use advance parallel and energy-efficiency embedded platforms at edge side

Productivity

- Programmability
- Portability/Scalability
- (Guaranteed) Performance
Main Contribution:
The CLASS Software Architecture

- Integrate technologies from different computing domains into a single development framework
  1. Powerful API for the development of advanced data-analytics methods
  2. QoS Serverless and CaaS cloud technologies
  3. Advanced orchestration methods for time-predictable workflow scheduling and deployment across the compute continuum
  4. Used of advanced embedded parallel and heterogeneous processor architectures
Smart City Use-Case

- Deployed on the **Modena Automotive Smart Area (MASA)** in the city of Modena (Italy)
  1. A **living lab urban area** with IoT connectivity and a compute continuum infrastructure
  2. **Three connected cars** equipped with sensors (cameras and LiDAR) and V2I communication

- Information exchange between the city and vehicles to enhance mobility
  1. Computation of emission of pollution in real-time
  2. Advanced Driving Assistant Systems
     - Virtual Mirror
     - Two Sources of Attention
Virtual Mirror Use-Case

Data-Analytics Methods
1. Sensor Fusion
2. Object Detection
3. Object Tracking
4. Data deduplication
5. Trajectory Prediction
6. Air pollution computation
7. Data model creation
8. Collision Detection (CD)
9. Generation of WA
10. WA alert visualization
Summary

1. CLASS aims to develop a novel **software architecture** with the following capabilities:
   - Increase data analytics capabilities by efficiently combine data-in-motion and data-at-rest analytics into **complex workflows**
   - Increase the development and deployment **productivity** of systems requiring data-analytics
   - Guarantee the **real-time properties** inherited from the domain

2. CLASS aims to apply the software architecture to develop a distributed sensing/computing infrastructure within the MASA for advanced urban mobility applications