Algorithms for Elastic Big Data Stream Processing

Alexandre da S. Veith

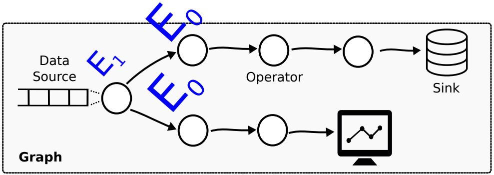
Advised by Marcos D. de Assunção, Laurent Lefèvre

alexandre.veith@ens-lyon.fr

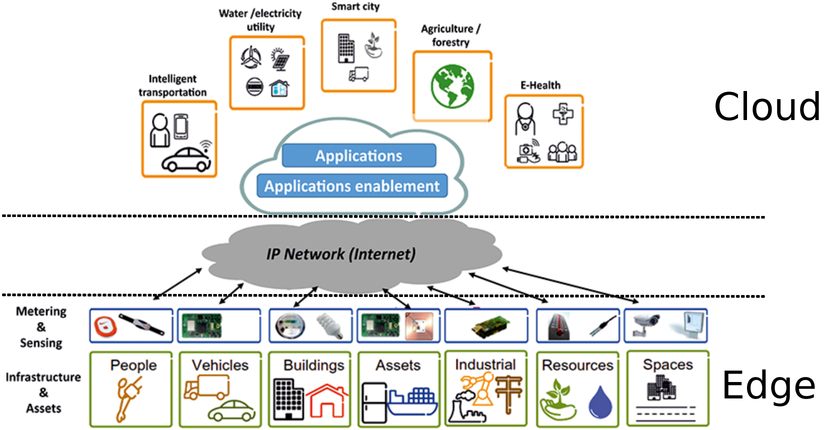
23rd April 2018

Motivation

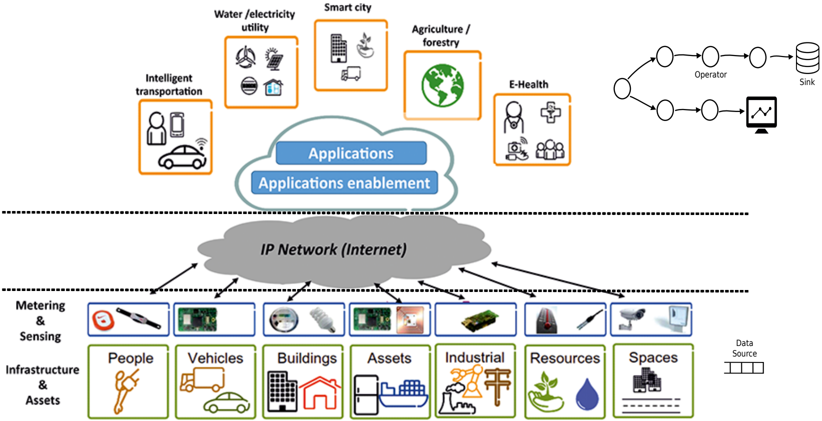
- Today's instruments and services are producing **ever-increasing** amounts of data that require quick analysis(**low end-to-end latencies**)



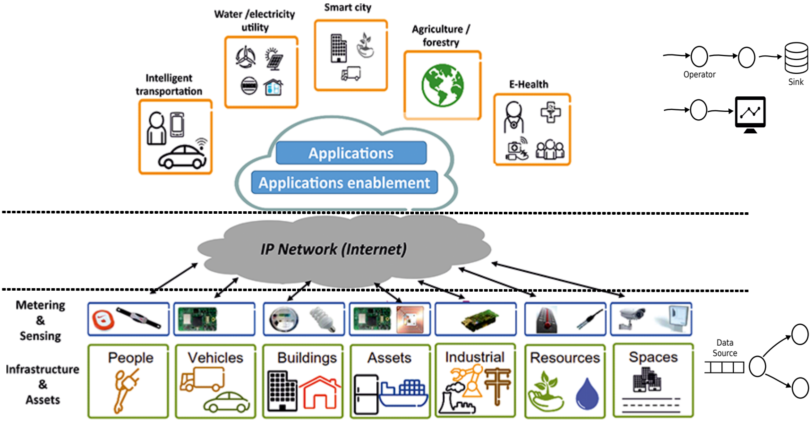
Edge Infrastructures and Data Stream Processing



Edge Infrastructures and Data Stream Processing



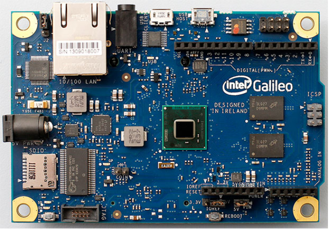
Edge Infrastructures and Data Stream Processing

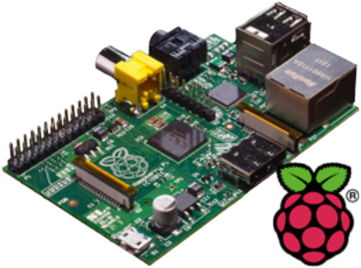


Edge Devices and Communication Constraints

Raspberry Pi 2

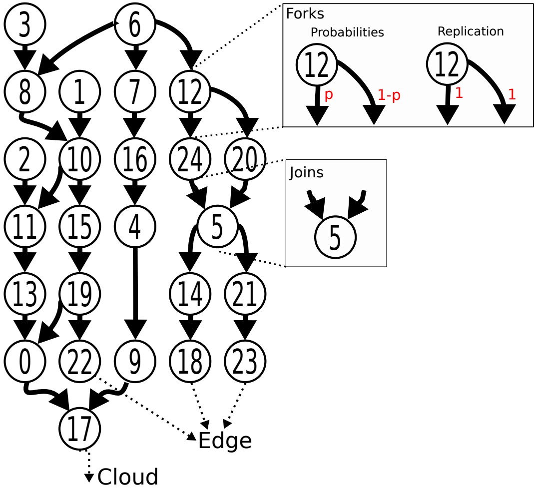
Galileo





LTE: 3G, 4G, 4.5G, and 5G
LoRaWAN
SigFOx

DSP Application Behaviors



DSP Application Trade-off

Throughput versus response time

Throughput

- Time windows to create message batches
- A decreasing in network latency can impact data transfers
- Increase the response time for a message

Response time (i.e. end-to-end application latency)

- One-at-a-time message transfer
- The response time is affected by the network latency
- Near real-time solutions

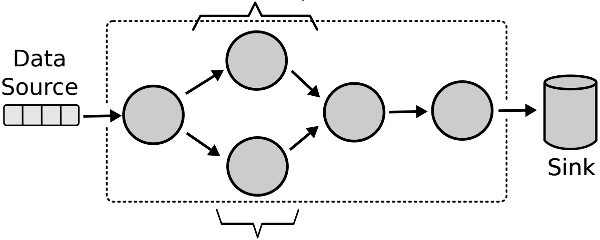
Related Work on DSP Application Placement

- Dynamic placements [Cardellini:2015, buddhika:2016]
- Static placements
  - Placement by hand [Sajjad:2016, Cheng:2015]
  - Communication is disregarded [Cheng:2015, Wu:2015, Zeng:2015, Hochreiner:2016]

de Assunção, M. D.; **da Silva Veith, Alexandre**; and Buyya, R. Distributed data stream processing and edge computing: A survey on resource elasticity and future directions Journal of Network and Computer Applications , 2018, 103, 1 – 17

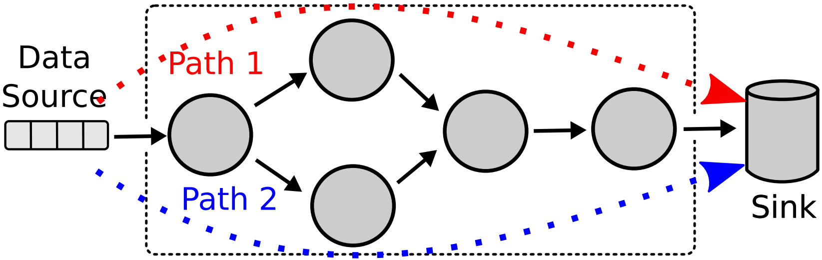
Model

Data compression/expansion factor: The ratio of the size of input events to the size of output events

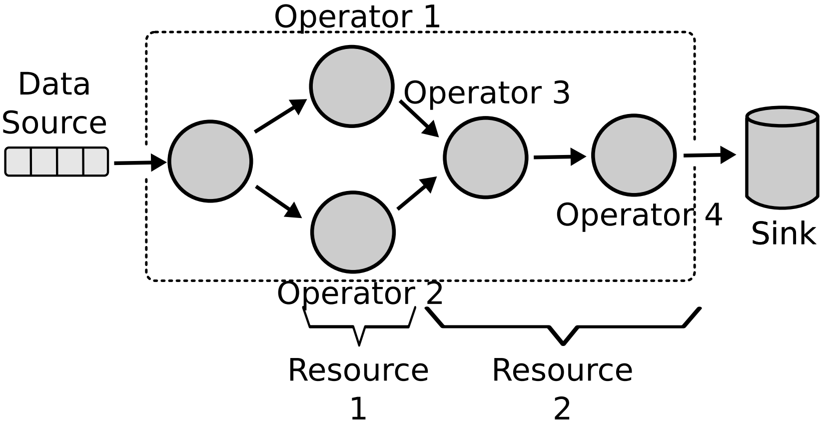


Selectivity: The ratio of number of input events to output events

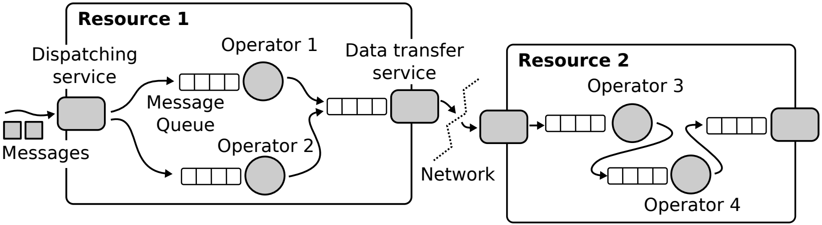
Model



Model


Operator 1
Data
Source
Operator 3
Operator 4
Sink
Operator 2
Resource
1
Resource
2

Model



Two queues: **Computation** and **Communication**

Problem Statement / Methodology

Minimize metrics such as **end-to-end application latency and energy consumption** by placing operators onto **cloud** and **edge resources**

Physical infrastructure capabilities

- CPU and memory
- Network latencies and bandwidth

Application requirements

- Selectivity
- Data compression rate
- CPU and Memory
- Data sources and sinks localization

Solution by approximation to achieve a **good result**

Evaluation (simulation and real-world) through **comparison** to the **state-of-the-art**

Formalization

End-to-end application latency

$$L_{\boxed{p_i}} = \sum_{\substack{o \in p_i \\ r \in \mathcal{R}}} mo_{\langle o,r \rangle} \times stime_{\langle o,r \rangle} + \sum_{r' \in \mathcal{R}} ms_{\langle o \rightarrow o+1, r \leftrightarrow r' \rangle} \times ctime_{\langle o,r \rangle \langle o+1, r' \rangle}$$

Paths

Formalization

End-to-end application latency

$$L_{\boxed{p_i}} = \sum_{\substack{o \in p_i \\ r \in \mathcal{R}}} mo_{\langle o,r \rangle} \times stime_{\langle o,r \rangle} + \sum_{r' \in \mathcal{R}} ms_{\langle o \rightarrow o+1, r \leftrightarrow r' \rangle} \times ctime_{\langle o,r \rangle \langle o+1, r' \rangle}$$

Paths

$$L_{p_i} = \sum_{\substack{o \in p_i \\ r \in \mathcal{R}}} \boxed{mo_{\langle o,r \rangle}} \times stime_{\langle o,r \rangle} + \sum_{r' \in \mathcal{R}} \boxed{ms_{\langle o \rightarrow o+1, r \leftrightarrow r' \rangle}} \times ctime_{\langle o,r \rangle \langle o+1, r' \rangle}$$

Mapping Operator

Mapping Stream

Formalization

End-to-end application latency

$$L \boxed{p_i} = \sum_{\substack{o \in p_i \\ r \in \mathcal{R}}} mo_{\langle o,r \rangle} \times stime_{\langle o,r \rangle} + \sum_{r' \in \mathcal{R}} ms_{\langle o \rightarrow o+1, r \leftrightarrow r' \rangle} \times ctime_{\langle o,r \rangle \langle o+1, r' \rangle}$$

Paths

$$L_{p_i} = \sum_{\substack{o \in p_i \\ r \in \mathcal{R}}} \boxed{mo_{\langle o,r \rangle}} \times stime_{\langle o,r \rangle} + \sum_{r' \in \mathcal{R}} \boxed{ms_{\langle o \rightarrow o+1, r \leftrightarrow r' \rangle}} \times ctime_{\langle o,r \rangle \langle o+1, r' \rangle}$$

Mapping Operator

Mapping Stream

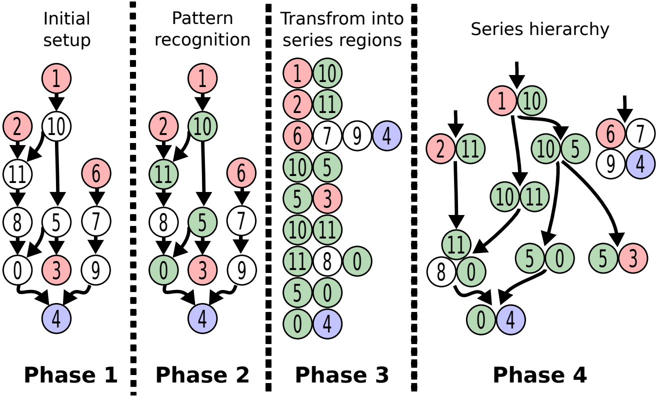
Computation time

$$stime_{\langle o,r \rangle} = \frac{1}{\mu_{\langle o,r \rangle} - \lambda_o^{in}}$$

Communication time

$$ctime_{\langle o,r \rangle \langle o+1, r' \rangle} = \frac{1}{\left(\frac{bdw_{r \leftrightarrow r'}}{\varsigma_o^{out}} \right) - \lambda_o^{out}} + l_{r \leftrightarrow r'}$$

Finding Application Patterns



Key behaviors: forks, joins, data sources and sinks

Decompose the application graph following the parallel regions

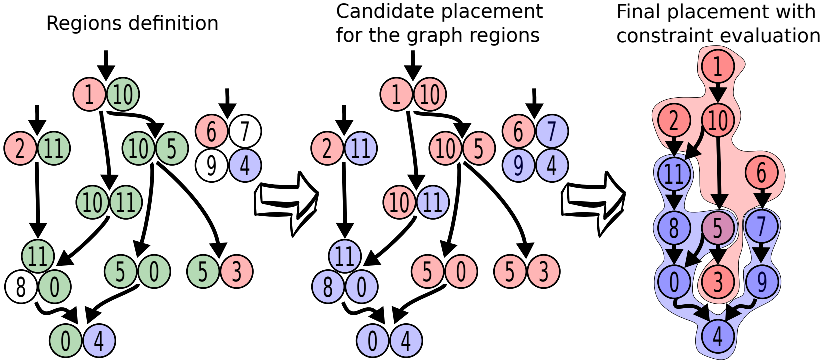
Response Time Rate (RTR) Strategy

- Response Time Rate for computational resource based on the end-to-end application latency
- Sequentially estimate the operator response time following the upstream(s) and downstream(s) connections
- Evaluate memory, cpu, and bandwidth constraints

$$L_{p_i} = \sum_{\substack{o \in p_i \\ r \in \mathcal{R}}} mo_{\langle o,r \rangle} \times stime_{\langle o,r \rangle} + \sum_{r' \in \mathcal{R}} ms_{\langle o \to o+1, r \leftrightarrow r' \rangle} \times ctime_{\langle o,r \rangle \langle o+1, r' \rangle}$$

Response Time Rate with Region Patterns (RTR+RP) Strategy

- Split the application graph following the pathways
- Calculate the Response Time Rate only to the edge side



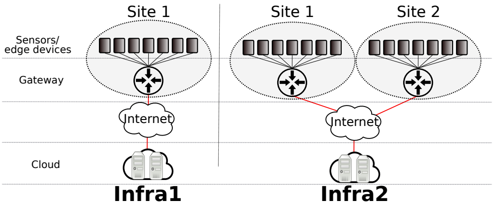
Expected Contributions

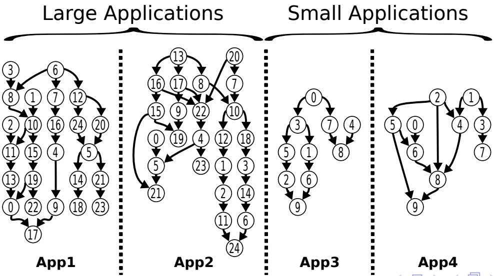
To split the DSP application graph following the paths between data sources and sinks

To evaluate the operator placement considering characteristics and requirements of applications and resources

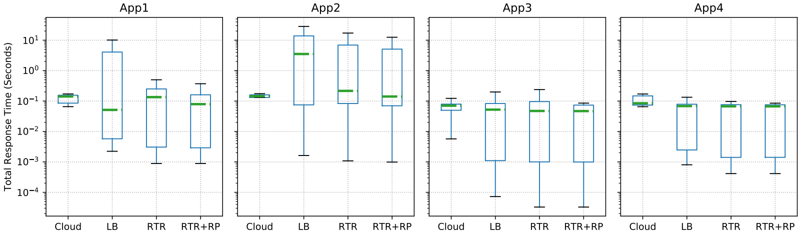
- Latency-aware
- Heterogeneity
- Capability-oriented decisions

Experimental Setup



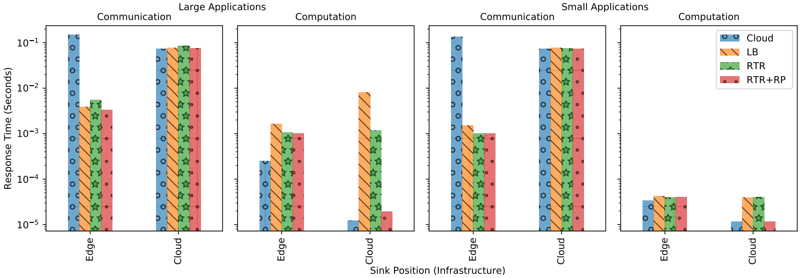


Results on Response Time



Small Applications: Improve 26% (cloud-only) and 5% (LB)
Large Applications: Improve .21% (cloud-only) and 92% (LB)
RTR+RP outperforms up to **54%** cloud-only

Results on Response Time



- 97.75% better for sinks in edge
- 1.5% worst for sinks in cloud.

Summary of Our Contributions

- A model and the DSP placement problem formalization
- Two strategies to improve the response time
- A performance comparison using a simulated environment

Publications

Book Chapter

- Dias de Assunção, Marcos; **da Silva Veith, Alexandre** (2018). Apache Spark. Encyclopedia of Big Data Technologies, 2018. DOI: 10.1007/978-3-319-63962-837-1.

International Journal

- de Assunção, M. D.; **da Silva Veith, Alexandre**; and Buyya, R. Distributed data stream processing and edge computing: A survey on resource elasticity and future directions Journal of Network and Computer Applications , 2018, 103, 1 – 17. Core: A

National Conference

- **da Silva Veith, Alexandre**; Dias de Assunção, Marcos; Lefèvre, Laurent. (2017). Assessing the Impact of Network Bandwidth and Operator Placement on Data Stream Processing for Edge Computing Environments. Conférence d'informatique en Parallélisme, Architecture et Système.

International Conferences

- **(Submitted)** C. S. Anjos, Julio; Matteussi, Kassiano; R. R. De Souza Jr, Paulo; **da Silva Veith, Alexandre**; Fedak, Gilles; Luis Victoria Barbosa, Jorge; R. Geyer, Claudio. (2018). Enabling Strategies for Big Data Analytics in Hybrid Infrastructures. International Conference on High Performance Computing and Simulation;
- **(Planning) da Silva Veith, Alexandre**; Dias de Assunção, Marcos; Lefèvre, Laurent. (2018). Latency-Aware Strategies for Placing Data Stream Processing Applications onto Edge Computing Infrastructure. International Conference on Service Oriented Computing. Core: A

Lab day-life and Training

Lab day-life

- Frequent meetings with advisors;
- AVALON chair and meeting organizer.

Training

During the first year, I completed the following training requirements:

Complementary Scientific - 64h

RSD/ASF winter school - 24h;

Parallel and Distributed Programming - 40h;

Insertion Training - 79h

FLE (University of Lyon 1) - 39h

FLE (CPU) - 40h

International Collaborations

- **Rutgers University** (Period of stay 11/28/2017-12/9/2017)
  - Fair evaluation and analysis process of stream applications
  - Apply hubs in the edge side to control the number of transferred messages
  - Setting up a platform to launch experiments
- **University Carlos III of Madrid (UC3M)** (Period of stay 1/29/2018-2/12/2018):
  - IoT-oriented scenario
  - Profiling and instrumenting applications
  - Setting up a platform to launch experiments

Frameworks Expertise

- Apache Kafka
- Mosquitto
- Apache Flink
- Apache Storm
- Apache Spark
- Apache Edgent
- OMNeT++
- G5K
- CPLEX

Perspectives

Short Term

- To conclude "Latency-Aware Strategies for Placing Data Stream Processing Applications onto Edge Computing Infrastructure" for ICSOC;
- To improve the model with: partitions and stateful operators; and reconfiguration phase;
- To write papers with Rutgers and UC3M;

Long Term

- To apply our model to a real-world evaluation using Grid'5000;
- To submit a journal paper: April 2019;
- The period of writing thesis: April-September 2019;
- PhD defense: October 2019.