Understanding the Data-Processing Challenges in Intelligent Vehicular Systems

Stefania Costache, Vincenzo Gulisano and Marina Papatriantafilou

VEHICLE TO INFRASTRUCTURE communication

INFRASTRUCTURE TO INFRASTRUCTURE communication

VEHICLE TO VEHICLE communication

RSDs
(Road-Side Devices)

- **Data can be processed locally**
  - Lower latency
  - Higher privacy
  - Access to sensor-local data
- **Less prone to cascade failures**
- **Higher cumulative processing capacity**
- **Inherent load distribution**
- **Challenges / Limitations:**
  - Load balancing
  - Fault tolerance

CDC
(Cloud Data Center)

- **Global view of data**
- **Higher per-device processing capacity**
- **Easier load balancing**
- **Challenges / Limitations:**
  - Privacy preservation
  - High transmission latency
  - Infrastructure communication capacity

WHERE to PROCESS DATA?

HOW to PROCESS DATA?

ONE-AT-A-TIME PROCESSING

time

Process each new incoming message inmediately

- **Millisecond Latency**
- **Ease of use (semantic and syntactic transparency)**
- **Fault tolerance**
- **Deterministic processing left to the user**

MICRO-BATCH PROCESSING

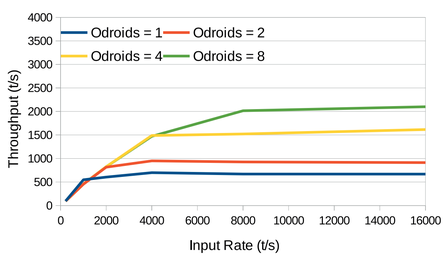
time

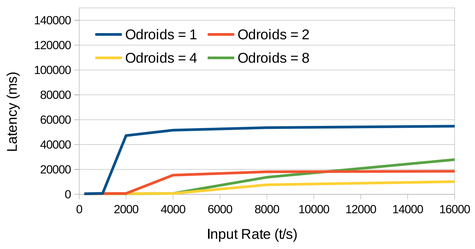
Group messages over small periods of time (postponing processing) and process as a batch

- **Deterministic Processing**
- **Ease of use (semantic and syntactic transparency)**
- **Fault tolerance**
- **Second latency**

REPRESENTATIVE RESULTS

ONE-AT-A-TIME PROCESSING



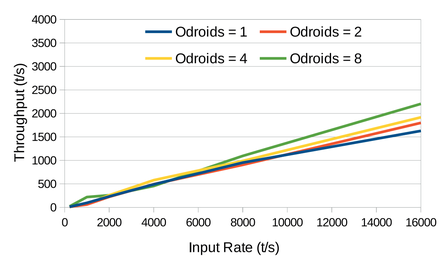


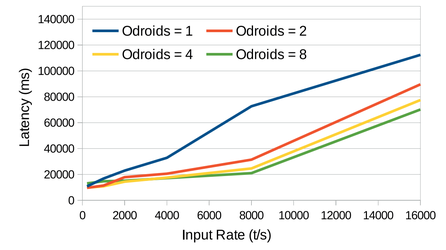
The use-case builds on the Linear Road benchmark, which simulates vehicular traffic on linear expressways that needs to be continuously processed to (i) detect possible accidents and (ii) compute tolls and notify vehicles.

To rely on devices representative of RSDs, we setup a testbed composed of 8 single-board ODROIDXU4 computers (~100$), each with 2 quad-core CPUs and 2 GB of RAM. The one-at-a-time and micro-batch processing are run using Apache Flink and Apache Spark frameworks, respectively.

The performance is assessed by means of processing throughput, measured in tuples/second (t/s), and latency, measured in milliseconds (ms). Latency is crucial when extracting information about critical situations, such as accidents, traffic jams or dangerous road conditions among others.

MICRO-BATCH PROCESSING





CHALMERS UNIVERSITY OF TECHNOLOGY

Distributed Computing and Systems
Computer Science and Engineering

The research leading to these results has been partially supported by the SAFER research center and the Transport Area of Advance at Chalmers, by the collaboration framework of Chalmers Energy Area of Advance project "SN09: EXAMINE", by the Swedish Civil Contingencies Agency (MSB) through the project "RICS", by the Swedish Foundation for Strategic Research (framework project FiC) and with support from the Swedish Energy Agency under the program Energy, IT and Design.

Stefania Costache. IBM, USA.
svcostac@us.ibm.com
(This work was done while the author was affiliated with Chalmers University of Technology, Sweden)

Vincenzo Gulisano and Marina Papatriantafilou
Computer Science and Engineering,
Chalmers University of Technology, Sweden
{vinmas,ptrianta}@chalmers.se