An Outlook on Graph Stream Processing Benchmarks
Who I Am

• Assistant Professor @ University of Tartu

• Maitre de Conférence @ INSA Lyon since 01/10/2021

• Maintainer of RSP4J library
What is Stream Processing

Sort out all the colours in the streams
How many boxes, grouped by color, there are in the last minute?
What About Graph?
Streaming Graphs
Streaming Graphs
Streaming Graphs
Dynamic Graphs
Dynamic Graphs
Dynamic Graphs
Dynamic Graphs
Dynamic Graphs
Dynamic Graphs
Graph Streams
RDF Streams
RDF Stream Processing

State-of-the-art
RSP State-of-the-art: Timeline

- SKB (2008)
- C-SPARQL (2010)
- MorphStream (2012)
- INSTANS (2013)
- CQELS (2012)
- CSRbench (2013)
- LSBench (2011)
- SRBench (2012)
- Citybench (2015)
- StreamWatDiv (2016)
- RSP4j (2021)
- Strider (2018)
LSBench
Le Phuoc et al.

• Simple: Social Media (Synthetic)
• Relevant: Static + Streaming
• Portable: not for the code
• Scalable: Generator
• KPI: Max Throughput & Memory
SRBench
Y Zhang et al.

- Simple: Weather/Sensor
- Relevant: Real Data
- Portable: data only + query
- Scalable
- KPI: Query Coverage
(C)SRBench
Daniele Dell’Aglio et al.

- Simple: Weather/Sensor
- Relevant: Real Data
- Portable: data only + query
- Scalable (a bit): parametric query
- KPI: Query Coverage & Correctness
Citybench
MI Ali et al.

- Simple: Smart-City/sensor
- Relevant: Real Data, static and streaming
- Portable: generation is hard-coded
- Scalable: query generator
StreamWatDiv
L Gao et al.

- Simple: User Activities
- Relevant: static + streaming
- Portable: test driver
- Scalable: Generator
- KPI: throughput & latency
- https://dsg.uwaterloo.ca/watdiv/stream-watdiv
DEBS Gran Challenge 2017

V Gulisano et al.

- Simple: User Activities
- Relevant: real data
- Portable: RDF
- Scalable
- KPI: throughput
The goal of a domain specific benchmark is to foster technological progress by guaranteeing a fair assessment.

Benchmark

Approach Y

Measure

Measure

2m
Benchmarking databases

Benefits of benchmarking

Ying Zhang, Peter Boncz

© Jim Gray, 2005
Conclusion

Edges through time

- Fast prototyping of Graph Stream Processing engines
- Potentially a test driver for benchmarks
- ESWC 2021 Best Resource Paper award!