Demo: YABench
Yet Another RDF Stream Processing Benchmark

Maxim Kolchin (ITMO University)
Peter Wetz (TU Wien)
Related Work

• 2012: LSBench [1]
  – First RSP Benchmark
  – Results: conceptual and technical differences of engines, performance shortcomings

• 2012: SRBench [2]
  – Functional evaluation based on a defined queryset
  – Results (simplified): Engines still limited

• 2013: CSRBench [3]
  – Correctness checking wrt engines’ operational semantics
  – Results: none passes all tests + discussion why queries fail
Metrics

- Precision/Recall per Window
- Window Size (expected)
- Result Size (expected/actual)
- Delay
- Performance
  - Memory Usage (% and absolute)
  - CPU Usage (%)
  - Spawned Threads
1. It determines the scope of the next window that will report based on the given window size $\alpha$, window slide $\beta$, and the required report policy.

2. Then, the scope $(t_s, t_e]$ is used to select window content from the input stream $\mathbb{S}$ where the relevant triples fall in the defined scope.

3. It computes the expected result by executing the SPARQL query $q'$ on the window content of the query engine.

4. The result of this query is compared with the next result of the continuous query $q$ and precision/recall metrics are computed.

5. The remaining metrics, i.e., delay, window size, and result size, are computed.
Workflow

Generator

C-SPARQL

CQELS

Oracle

Reporting
Let’s do some streaming ...
References


* The work done by Maxim Kolchin was financially supported by the Ministry of Education and Science of the Russian Federation (Grant #RFMEFI57514X0101).