Computing Continuous SPARQL Query over RDF Streams on Storm Platform

By Usman Younas and Sander Breukink

# Introduction

The amount of data on the web has massively increased over the years, and it doesn’t show any signs of stopping. Hence the demand for analyzing it in reasonable amount of time is also increasing. The Resource Description Framework (RDF) is a data model whose purpose is to form a comprehensive framework to integrate data from different fields. It is a flexible data model used in the Semantic Web (a Web of data) on which we can do querying or reasoning. In this report we discuss several ways of executing SPARQL queries on RDF data and compare their results based on several testing scenarios. For this we will use the Apache Storm Framework with different topologies. We will also run these topologies on a testing set and compare these results.

# Decomposing queries

State how we decompose the queries

# Spouts

How spouts work

# Bolts

How bolts work

# Testing scenarios

Serveral joins etc.

# Testing environment

Grid 5000 network

# Benchmark results

Results from the benchmarks

# Conclusion

Todo