LDBC Social Network Benchmark @ Neo Technology

Alex Averbuch[1]

[1]Neo Technology

Table of Contents

Introduction

- LDBC Social Network Benchmark
  - DATA: social network, user activity, realistic
  - SCALABLE: 1, 3, 10, 30, 100, ...
  - QUERIES: 14 reads, 8 updates, (8 additional short reads soon)
  - TOOLS
    - DATA GENERATOR: `github.com/ldbc/ldbc_snb_datagen`
    - DRIVER: `github.com/ldbc/ldbc_driver`
  - SPEC: `github.com/ldbc/ldbc_snb_docs`

Introduction

- Neo4j Implementation
  - VERSIONS: 2.0, 2.1, 2.2
  - API: Java, Cypher
  - VALIDATED: Neo4j Cypher x Neo4j Java x Sparksee
    - `github.com/ldbc-dev/ldbc_snb_interactive_validation`
  - CODE: `github.com/ldbc-dev/ldbc_snb_workload_interactive_neo4j`

Introduction

552 commits | 3 branches | 0 releases | 2 contributors

branch: master ▾ ldbc_snb_workload_interactive_neo4j / +

Merge branch '2.1-maint'

alexaverbuch authored 5 hours ago — latest commit 7ad54c5ac4

benchmark-run	Merge branch '2	a month ago
data-import	Merge branch '2.1-maint	6 hours ago
ldbc_driver @ 7fc7a11	update ldbc driver to get parser	5 hours ago
report	update ldbc_driver	9 months ago
schema	added queries 1, 3 and 4 from LDBC interactive workload	a year ago
steps @ 56760b0	have Maven demand Java 7	a month ago
.gitignore	ensure driver returns non-zero when failing so that the build can pic...	a month ago
.gitmodules	removing jdbc submodule	a month ago
README.md	restructuring so that we can build uberjars	a month ago
pom.xml	Merge branch '2.1-maint'	a month ago
queries.txt	update to use new update streams from data generator <-- fixes GCT tr...	5 days ago

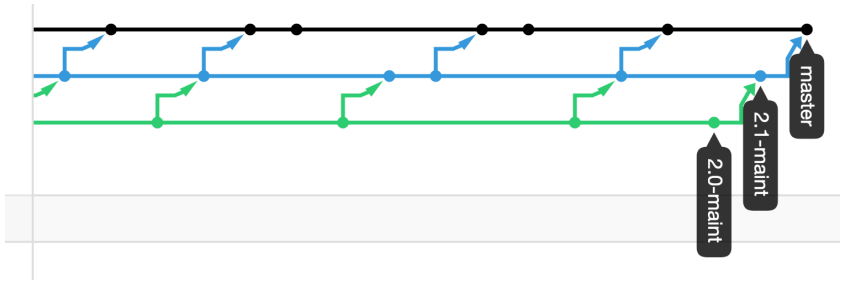


Table of Contents

Instructions

1. GENERATE DATA
   - `github.com/ldbc/ldbc_snb_datagen`

Instructions

1. GENERATE DATA
   - `github.com/ldbc/ldbc_snb_datagen`
2. IMPORT DATA
   java -jar data-import-VERSION.jar db/ csv/

Instructions

1. GENERATE DATA
   - `github.com/ldbc/ldbc_snb_datagen`
2. IMPORT DATA
   java -jar data-import-VERSION.jar db/ csv/
3. RUN BENCHMARK
   java -jar benchmark-run-VERSION.jar [args]
   - `github.com/ldbc/ldbc_driver`

Instructions

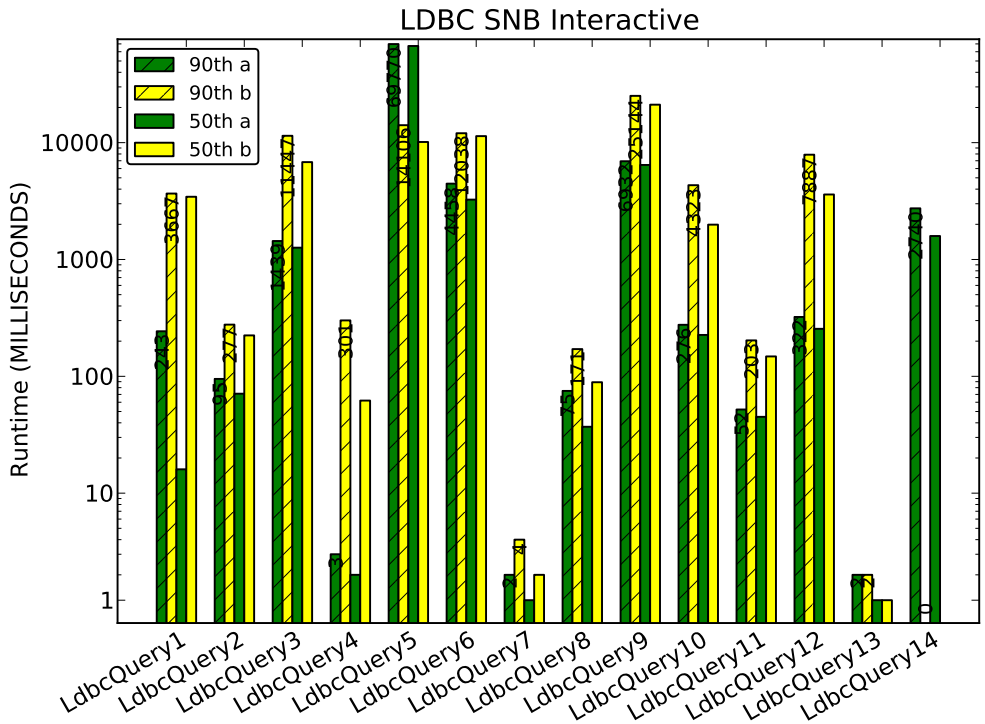
1. GENERATE DATA
   - `github.com/ldbc/ldbc_snb_datagen`
2. IMPORT DATA
   java -jar data-import-VERSION.jar db/ csv/
3. RUN BENCHMARK
   java -jar benchmark-run-VERSION.jar [args]
   - `github.com/ldbc/ldbc_driver`
4. COLLECT GIFTS
   1. results.json
   2. results_log.csv
   3. plans.json (Cypher only)

Artifacts - results.json


```
{
    "start_time":1412112651578,
    "latest_finish_time":1412113505053,
    "total_duration":853475,
    "total_count":1000,
    "unit":"MILLISECONDS",
    "all_metrics":
    [{
        "name":"com.ldbc.driver.workloads.ldbc.snb.interactive.LdbcQuery1",
        "run_time":
        {
            "count":13,
            "mean":3578.5384615384614,
            "min":2821,
            "max":5560,
            "50th_percentile":3305,
            "90th_percentile":5207,
            "95th_percentile":5207,
            "99th_percentile":5560
        }
    },{
        "name":"com.ldbc.driver.workloads.ldbc.snb.interactive.LdbcQuery10",
        "run_time":
        {
            "count":15,
            "mean":3264.4666666666667,
            "min":978,
            "max":9138,
            "50th_percentile":2438,
            "90th_percentile":5568,
            "95th_percentile":5568,
            "99th_percentile":9138
        }
    }]
}
```

Artifacts - results.json

- make_compare_charts.py
  `github.com/ldbc/ldbc_driver/plotting`



Artifacts - results_log.json

```
operation_type|scheduled_start_time_MILLISECONDS|actual_start_time_MILLISECONDS|execution_duration_MILLISECONDS|result_code
LdbcQuery6|1415833132316|1415833132535|984|0
LdbcQuery6|1415833132136|1415833132535|1015|0
LdbcQuery9|1415833132146|1415833132535|1073|0
LdbcQuery9|1415833132336|1415833132535|1126|0
LdbcQuery10|1415833132406|1415833133550|144|0
LdbcQuery3|1415833132376|1415833133520|282|0
LdbcQuery9|1415833132526|1415833133662|171|0
LdbcQuery6|1415833132496|1415833133608|226|0
LdbcQuery6|1415833132676|1415833133695|144|0
LdbcQuery10|1415833132856|1415833133840|11|0
LdbcQuery3|1415833132796|1415833133833|29|0
LdbcQuery6|1415833132856|1415833133835|28|0
LdbcQuery9|1415833132716|1415833133803|68|0
LdbcQuery14|1415833133046|1415833133863|21|0
```

Artifacts - plans.json


```
+ Neo4jQuery11EmbeddedCypher: {…},
+ Neo4jQuery12EmbeddedCypher: {…},
- Neo4jQuery13EmbeddedCypher: {
    - plan: {
        - root: {
              operatorType: "Projection",
              DbHits: "0",
              Rows: "1",
              version: "CYPHER 2.2-cost",
              KeyNames: "pathLength",
              EstimatedRows: "1",
            - children: [
                - {
                      operatorType: "Apply",
                      Rows: "1",
                      DbHits: "0",
                      EstimatedRows: "1",
                    - children: [
                        - {
                              operatorType: "CartesianProduct",
                              Rows: "1",
                              DbHits: "0",
                              EstimatedRows: "1",
                            - children: [
                                - {
                                      operatorType: "NodeIndexSeek",
                                      Index: ":Person(id)",
                                      Rows: "1",
                                      DbHits: "2",
                                      EstimatedRows: "1",
                                      children: [ ]
                                  },
                                - {
                                      operatorType: "NodeIndexSeek",
                                      Index: ":Person(id)",
                                      Rows: "1",
                                      DbHits: "2",
                                      EstimatedRows: "1",
                                      children: [ ]
                                  }
                              ]
                          },
                        - {
                              operatorType: "Optional",
                              Rows: "1",
                              DbHits: "0",
                              EstimatedRows: "9919224270400",
                            - children: [
                                - {
                                      operatorType: "ShortestPath",
                                      Rows: "0",
                                      DbHits: "0",
                                      EstimatedRows: "1",
                                      children: [ ]
                                  }
                              ]
                          }
                      ]
                  }
              ]
          }
      }
},
```

Table of Contents

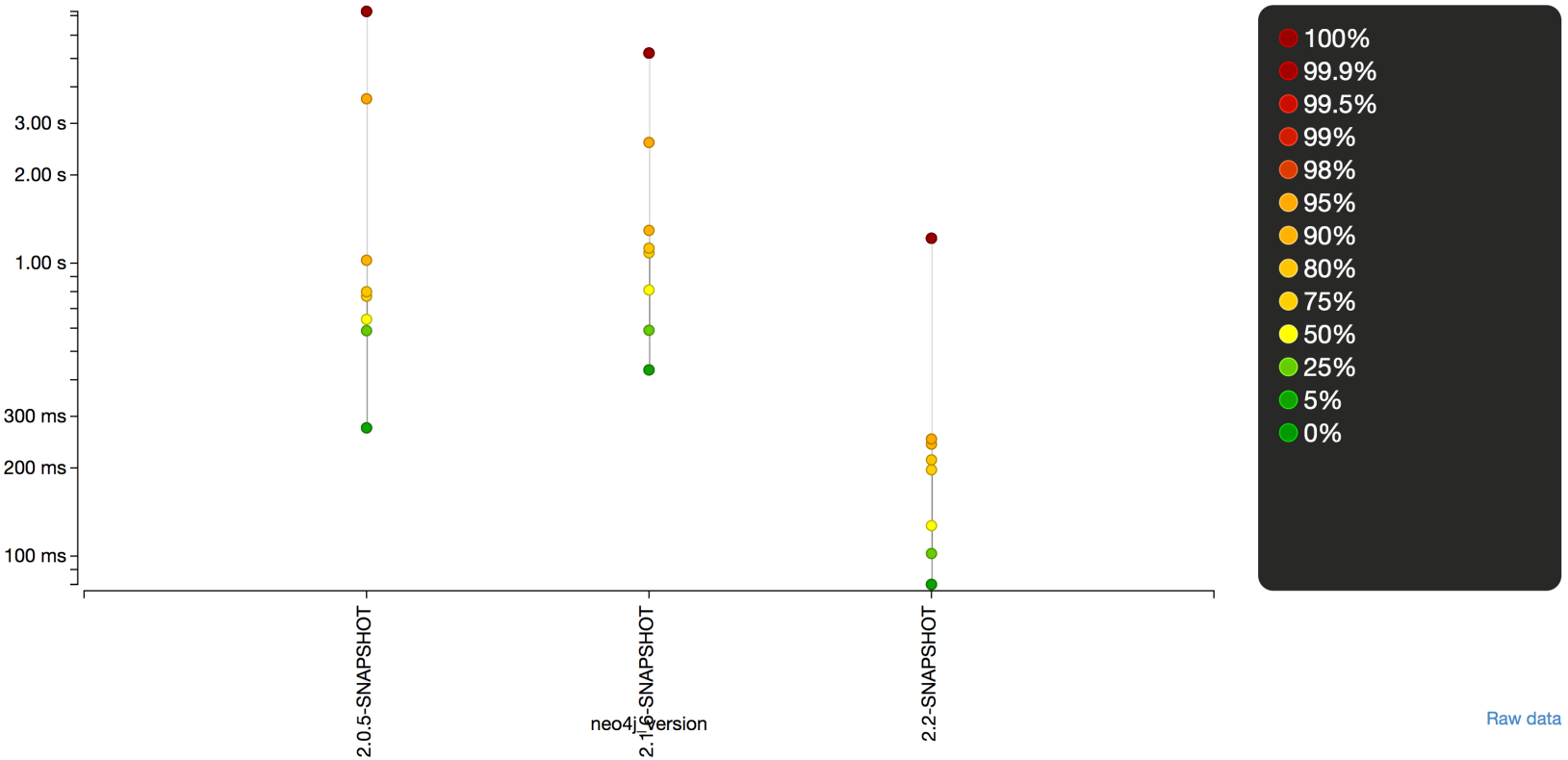
LDBC SNB Builds

- build types
  - SCALE_FACTOR: 1, 3, 10
  - NEO4J_VERSION: 2.0, 2.1, 2.2
  - API: Cypher, Java
  - WORKLOAD: Read-Only, Write-Only, Read-Write

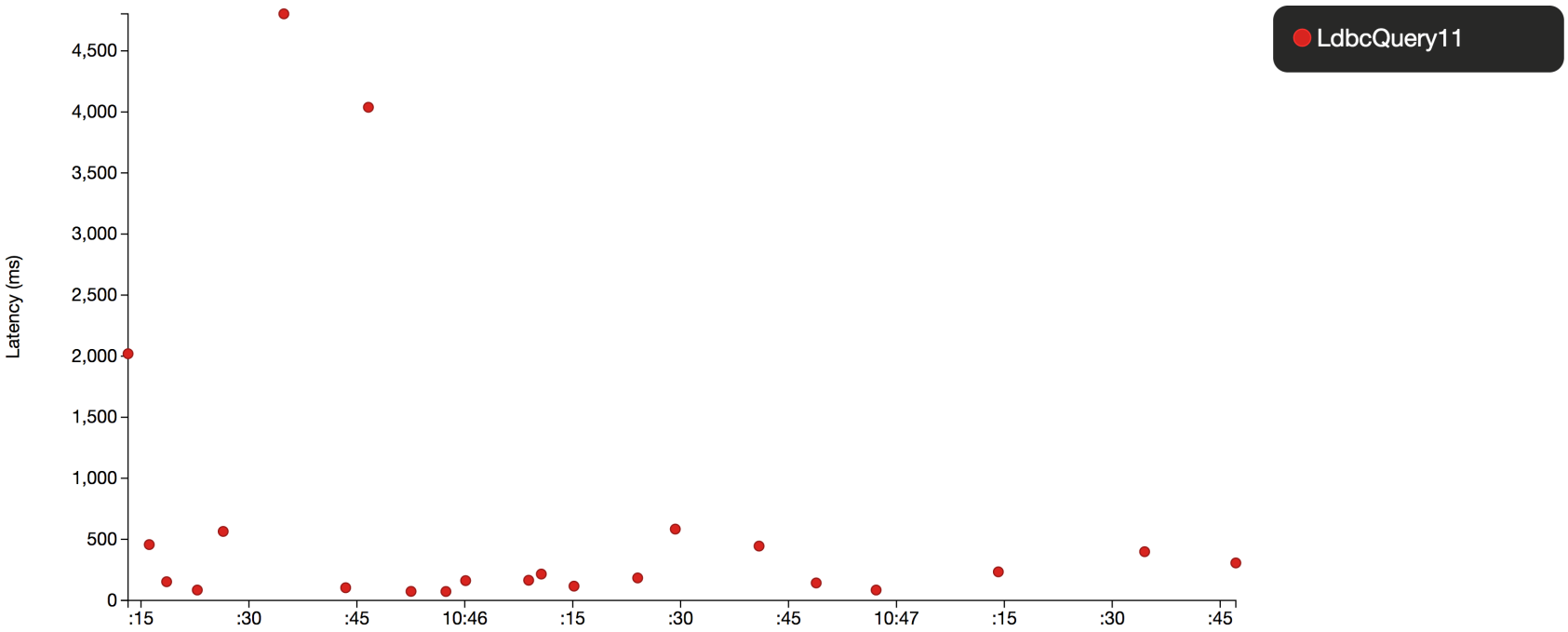
LDBC SNB Builds

- build types
  - SCALE_FACTOR: 1, 3, 10
  - NEO4J_VERSION: 2.0, 2.1, 2.2
  - API: Cypher, Java
  - WORKLOAD: Read-Only, Write-Only, Read-Write
- build process
  1. RETRIEVE CSV FILES
  2. IMPORT CSV FILES
  3. RUN BENCHMARK
  4. STORE RESULTS

Plotting - Compare Neo4j Versions (Query 11)



Plotting - Individual Query Individual Run (Query 11)



Plotting - Many Queries Individual Run

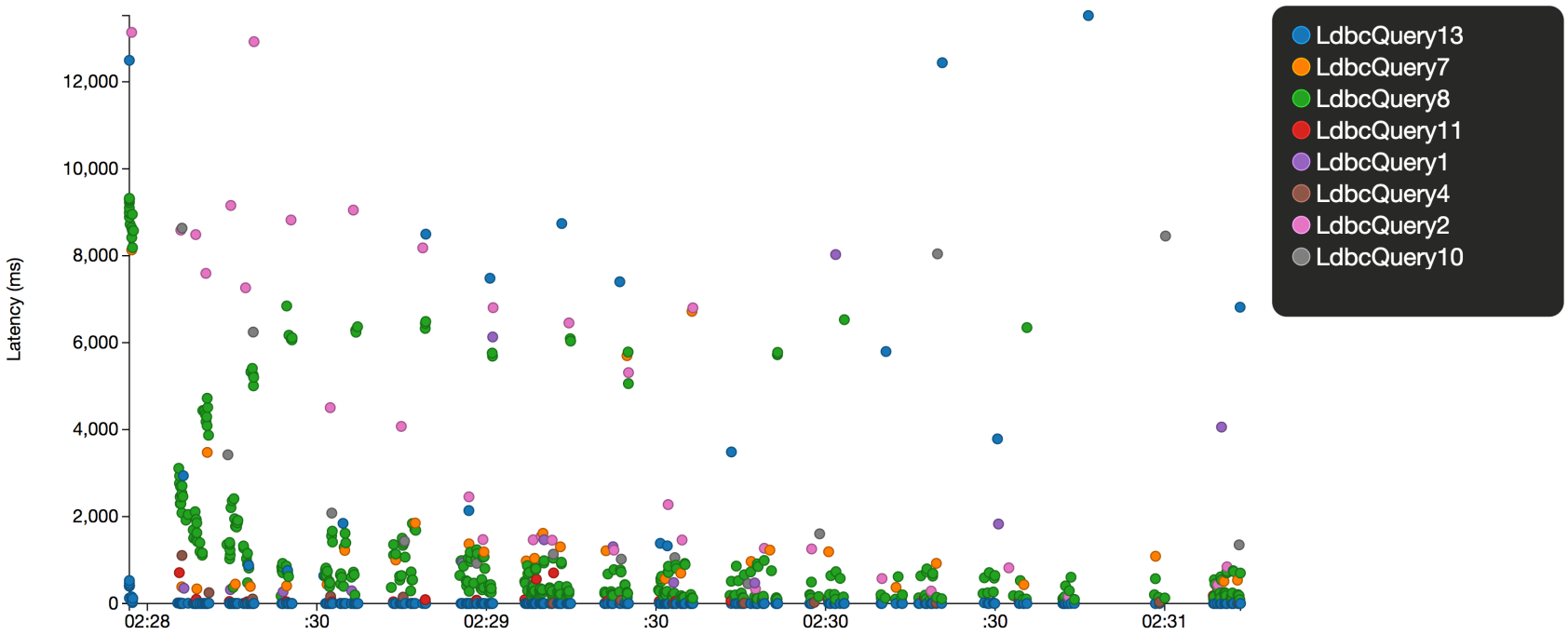


Table of Contents

Feedback

1. warm-up support
2. more plotting scripts
3. revisit query mix - ‘slow’/‘fast’ queries different between vendors

Questions