



*The graph & RDF  
benchmark reference*

***Semantic Publishing Benchmark  
Full Disclosure Report  
Irina Fundulaki  
10/06/2015***

## Table of Contents

|  |           |
|--|-----------|
| <b>Preface</b>                                     | <b>3</b>  |
| <b>General Terms</b>                               | <b>4</b>  |
| <b>1. System Description</b>                       | <b>5</b>  |
| 1.1 Database System                                | 5         |
| 1.2 Database Engine Configuration                  | 5         |
| 1.3 Platform Description                           | 5         |
| 1.4 Network Infrastructure Information             | 6         |
| <b>2. Data Generation &amp; Loading</b>            | <b>7</b>  |
| 2.1 Dataset Information                            | 7         |
| 2.1.1 Description                                  | 7         |
| 2.1.2 External Datasets                            | 7         |
| 2.1.3 Dataset Characteristics                      | 8         |
| 2.1.4 Data Generator Parameters                    | 8         |
| 2.2 Data Generation Times                          | 9         |
| 2.3 Bulk Loading                                   | 9         |
| <b>3. Benchmark Test Driver</b>                    | <b>9</b>  |
| 3.1. Basic test driver configuration details       | 9         |
| 3.2. Configuration Parameters for Driver Warmup    | 9         |
| 3.3. Configuration Parameters for Driver Execution | 10        |
| 3.3. Test Driver Reference Data                    | 10        |
| <b>4. Performance Metrics</b>                      | <b>11</b> |
| 4.1. SPB Primary Metrics                           | 11        |
| 4.2. Query Execution Report                        | 11        |
| <b>5. Recovery</b>                                 | <b>13</b> |
| <b>6. Pricing Summary</b>                          | <b>14</b> |
| <b>7. Attachment's CheckList</b>                   | <b>15</b> |

## ***Preface***

This is the Full Disclosure Report for the benchmark results produced by Semantic Publishing Benchmark v2.0 for GraphDB SE 6.3 alpha, Scale Factor SF1 (64M Triples), single CPU server, with 64GB of RAM and 2xSSD drives.

## General Terms

### Test Sponsor

Test sponsor of this benchmark is Ontotext AD, Bulgaria. It is a leading provider of core semantic technology distinctive for its performance, scale, robustness and compliance with open standards. Ontotext is unique as a developer that provides state of the art semantic technology in two distinct but complementary areas – semantic graph database engines and text analytics. Ontotext solutions have been applied for business critical projects in the areas of publishing (BBC, Financial Times, Oxford University Press), life sciences (Astra Zeneca), cultural heritage (British Museum, Getty Trust), telecommunications (Korea Telecom), government organizations (UK Parliament) and others.

Ontotext is the developer of GraphDB semantic graph database.

# 1. System Description

## 1.1 Database System

|                                      |  |
|--------------------------------------|--|
| <b>Vendor Name</b>                   | ONTOTEXT AD  |
| <b>Database Name</b>                 | GRAPHDB STANDARD EDITION   |
| <b>Version Number</b>                | 6.3 ALPHA, BUILD 4888f632  |
| <b>Database Engine Configuration</b> | SINGLE NODE; Java HotSpot(TM) 64-Bit Server VM (build 24.75-b04, mixed mode) |

Table 1: DBMS Characteristics

## 1.2 Database Engine Configuration

|  |  |
|--|--|
| <b>Cache Configuration</b>               | 18G CACHE MEMORY OF 54G TOTAL RESERVED MEMORY (JAVA -XMX54G) |
| <b>Transaction Isolation Level/Model</b> | READ COMMITTED   |
| <b>Reasoning Strategy/configuration</b>  | FORWARD-CHAINING   |
| <b>triple/quadruple indices</b>          | POS, PSO, PCSO, SPOC, PRED. LISTS                            |
| <b>indices</b>                           | GEO-SPATIAL, TEXT INDEXES (LUCENE CONNECTOR)                 |
| <b>Optimisations</b>                     | OWL:SAMEAS   |

Table 2: Database Engine Configuration

## 1.3 Platform Description

|   |  |
|---|--|
| <b>Model</b>                                | SUPERMICRO X10SRI-F, SOCKET R3, INTEL C612   |
| <b>Processors</b>                           | CPU INTEL XEON E5-1650 V3 3.5GHZ, 15MB L3 CACHE, S2011   |
| <b>Memory</b>                               | MEMORY 4 x 16GB DDR4-2133 2Rx4 ECC REG DIMM  |
| <b>No. Disks/Type/Storage Configuration</b> | 1 x HDD WESTERN DIGITAL HDD 500GB SATAIII RAID EDITION RE4 - 7200RPM 64MB CACHE – FOR THE OPERATING SYSTEM ONLY<br>2 x SSD DRIVES 400 GB SSD , SAMSUNG 845DC PRO |
| <b>Network Adapters</b>                     | I350 DUAL PORT GIGABIT LAN   |
| <b>Operating System</b>                     | 14.04 UBUNTU 64 BIT  |
| <b>File System</b>                          | EXT4, SOFTWARE RAID-0 (SSD)  |
| <b>CPU Type &amp; Count</b>                 | INTEL XEON, 1 CPU  |
| <b>No. Threads</b>                          | 12   |
| <b>No. Cores</b>                            | 6  |

|                             |                                |
|-----------------------------|--------------------------------|
| <b>Memory</b>               | 64 GB                          |
| <b>Total Disks Capacity</b> | 734 GB (SSD), 1 X 413 GB (HDD) |

**Table 3: System Configuration**

|                             |  |
|-----------------------------|--|
| <b>No. disks</b>            | 2 SSD, 1 HDD   |
| <b>Storage interfaces</b>   | SATA3  |
| <b>Storage Technologies</b> | SSD, HDD   |
| <b>RAID/HBA controller</b>  | SOFTWARE INTEL CORPORATION PATSBURG 6-PORT SATA AHCI CONTROLLER (REV 06), RAID-0 |

**Table 4: No. Disks/Type/Storage/Configuration**

**1.4 Network Infrastructure Information**

|                           |                     |
|---------------------------|---------------------|
| <b>Model</b>              | NO EXTERNAL NETWORK |
| <b>Network Switches</b>   | N/A                 |
| <b>Wiring Information</b> | N/A                 |

**Table 5: Network Infrastructure Information**

## 2. Data Generation & Loading

Generated 39 M triples, using SPB's data generator. Loaded using SPB's loading process.

### 2.1 Dataset Information

#### 2.1.1 Description

|                        |             |
|------------------------|-------------|
| Scale Factor           | SF1, 64M    |
| Reference Data Size    | 25M         |
| Data Format            | NQUADS      |
| Data Generator Version | 2.0.1369E4A |
| Time Compression ratio | N/A         |

Table 6: Dataset characteristics

#### 2.1.2 External Datasets

External datasets and ontologies can be found in the directories provided with the SPB distribution. The directories where this information is located are:

| ./DATA/DATASETS                           |  |
|---|--|
| INTERNATIONAL-FOOTBALL-COMPETITIONS-3.TTL | ENTITIES_RANKS.TTL                     |
| INTERNATIONAL-FOOTBALL-TEAMS-2.TTL        | GEONAMES_EUROPE.TTL                    |
| SCOTTISH-FOOTBALL-COMPETITIONS-1.TTL      | DBPEDIA_COMPANIES.TTL                  |
| SCOTTISH-FOOTBALL-TEAMS-2.TTL             | DBPEDIA_EVENTS.TTL                     |
| FORMULA1-COMPETITIONS-8.TTL               | ENTITIES_PREFLABELS.TTL                |
| FORMULA1-TEAMS-3.TTL                      | GEONAMES_SAMEAS_LINKS_EN_EN.TTL        |
| ENGLISH-FOOTBALL-COMPETITIONS-1.TTL       | DBPEDIA_PERSONDATA.TTL                 |
| ENGLISH-FOOTBALL-TEAMS-2.TTL              | UK-PARLIAMENT-IDENTIFIERS-PEOPLE-8.TTL |
| ./DATA/ONTOLOGIES/CORE                    |  |
| CMS.1.2.TTL                               | CREATIVEWORK.0.9.TTL                   |
| COMPANY.1.4.TTL                           | GEONAMES_ONTOLOGY_V3.1-SPB.TTL         |
| CORECONCEPTS.0.6.TTL                      | PERSON.0.2.TTL                         |
| PROVENANCE.1.1.TTL                        | SYSTEMLOGIC.TTL                        |
| TAGGING.1.0.TTL                           |  |
| ./DATA/ONTOLOGIES/DOMAIN                  |  |
| CNEWS-1.2.TTL                             | CURRICULUM-ONTOLOGY-4.TTL              |
| SPORT.2.3.TTL                             |  |

Table 7: External Datasets & Ontologies (Listing)

### 2.1.3 Dataset Characteristics

Statistics for the datasets are produced and stored in file `./generated.64/dataset.info`

|                              |           |
|------------------------------|-----------|
| <b>NO. GENERATED TRIPLES</b> | 39000000  |
| <b>NO. CREATIVE WORKS</b>    | 1489701   |
| <b>NO. TOTAL TRIPLES</b>     | 133112836 |

**Table 8: Dataset Characteristics**

### 2.1.4 Data Generator Parameters

Data Generator Parameters can be found in files `definitions.properties` and `test.64.properties`.

|  |
|--|
| <b>TEST.64.PROPERTIES</b>  |
| DATAGENERATORWORKERS=8<br>DATASETSIZE= 39000000<br>CREATIVEWORKSPATH= ./GENERATED.64<br>QUERYSUBSTITUTIONPARAMETERS=100000<br>ENABLECOMPRESSIONONGENERATEDDATA=FALSE   |
| <b>DEFINITIONS.PROPERTIES</b>  |
| ABOUTSALLOCATIONS=0.1006, 0.2313, 0.3088, 0.2278, 0.1035, 0.028<br>MENTIONSALLOCATIONS=0.9477, 0.0382, 0.0093, 0.0031, 0.0012, 0.0005<br>ENTITYPOPULARITY=0.05, 0.95<br>USEPOPULARENITIES=0.3, 0.7<br>CREATIVEWORKTYPESALLOCATION=0.45, 0.35, 0.2<br>ABOUTANDMENTIONSALLOCATION=0.85, 0.15<br>EDITORIALOPERATIONSALLOCATION=0.8, 0.1, 0.1<br>AGGREGATIONOPERATIONSALLOCATION=0.087, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083, 0.083<br>EXPONENTIALDECAYUPPERLIMITOFCWS=5000<br>EXPONENTIALDECAYRATE=0.1<br>EXPONENTIALDECAYTHRESHOLDPERCENT=0.05<br>MAJOREVENTS=5<br>MINOREVENTS=100<br>SEEDYEAR=2011<br>DATAGENERATIONPERIODYEARS=1<br>CORRELATIONSAMOUNT=50<br>CORRELATIONSMAGNITUDE=60<br>CORRELATIONDURATION=0.2<br>CORRELATIONENTITYLIFESPAN=0.5<br>MINLAT = 36.5<br>MAXLAT = 70.0<br>MINLONG = -9.5<br>MAXLONG = 31.0 |

**Table 9: Data Generator Parameters**



## 2.2 Data Generation Times

|                          |          |
|--------------------------|----------|
| Start Time (in hh:mm:ss) | 10:06:44 |
| End Time (in hh:mm:ss)   | 10:09:89 |
| Generation Time (in s)   | 153.324  |

Table 10: Data Generation Times

## 2.3. Bulk Loading

|                          |          |
|--------------------------|----------|
| Start Time (in hh:mm:ss) | 10:09:89 |
| End Time (in hh:mm:ss)   | 10:56:00 |
| Bulk Load Time (in s)    | 2766.285 |

Table 11: Bulk Loading Time

## 3. Benchmark Test Driver

### 3.1. Basic test driver configuration details

Basic Test driver configuration details can be found in test.64.properties file.

|                   |   |
|-------------------|---|
| No. read threads  | 8 |
| No. write threads | 2 |

Table 32: Basic Test Driver Configuration Details

### 3.2. Configuration Parameters for Driver Warmup

Test driver configuration details for Driver Warmup can be found in ./test.64.properties file.

|                         |
|-------------------------|
| WARMUPPERIODSECONDS=600 |
| AGGREGATIONAGENTS=8     |
| QUERYTIMEOUTSECONDS=300 |
| WARMUP=TRUE             |

Table 43: Configuration Parameters for Driver Warmup

venelin 16/6/2015 15:43

Formatted: Normal,Normal Text

|  |  |
|--|--|
|  |  |
|--|--|

### 3.3. Configuration Parameters for Driver Execution

Test driver configuration details for Driver Execution can be found in test.64.properties file.

```

AGGREGATIONAGENTS=8
EDITORIALAGENTS=2
BENCHMARKRUNPERIODSECONDS=1200
QUERYTIMEOUTSECONDS=300
BENCHMARKBYQUERYMIXRUNS=0
MINUPDATERATETHRESHOLDOPS=0.0
MINUPDATERATETHRESHOLDREACHTIMEPERCENT=0
MAXUPDATERATETHRESHOLDOPS=0.0
ENDPOINTURL=HTTP://LOCALHOST:4082/OPENRDF-SESAME/REPOSITORIES/LDBC64
ENDPOINTUPDATEURL=HTTP://LOCALHOST:4082/OPENRDF-
SESAME/REPOSITORIES/LDBC64/STATEMENTS
QUERIESPATH=./DATA/SPARQL
RUNBENCHMARK=TRUE
    
```

**Table 54: Configuration Parameters for Driver Execution**

### 3.3. Test Driver Reference Data

The reference data for the test driver are located at:

| ./DATA/DATASETS                           |  |
|---|--|
| INTERNATIONAL-FOOTBALL-COMPETITIONS-3.TTL | ENTITIES_RANKS.TTL                     |
| INTERNATIONAL-FOOTBALL-TEAMS-2.TTL        | GEONAMES_EUROPE.TTL                    |
| SCOTTISH-FOOTBALL-COMPETITIONS-1.TTL      | DBPEDIA_COMPANIES.TTL                  |
| SCOTTISH-FOOTBALL-TEAMS-2.TTL             | DBPEDIA_EVENTS.TTL                     |
| FORMULA1-COMPETITIONS-8.TTL               | ENTITIES_PREFLABELS.TTL                |
| FORMULA1-TEAMS-3.TTL                      | GEONAMES_SAMEAS_LINKS_EN_EN.TTL        |
| ENGLISH-FOOTBALL-COMPETITIONS-1.TTL       | DBPEDIA_PERSONDATA.TTL                 |
| ENGLISH-FOOTBALL-TEAMS-2.TTL              | UK-PARLIAMENT-IDENTIFIERS-PEOPLE-8.TTL |
| ./DATA/ONTOLOGIES/CORE                    |  |
| CMS.1.2.TTL                               | CREATIVEWORK.0.9.TTL                   |
| COMPANY.1.4.TTL                           | GEONAMES_ONTOLOGY_V3.1-SPB.TTL         |
| CORECONCEPTS.0.6.TTL                      | PERSON.0.2.TTL                         |
| PROVENANCE.1.1.TTL                        | SYSTEMLOGIC.TTL                        |
| TAGGING.1.0.TTL                           |  |
| ./DATA/ONTOLOGIES/DOMAIN                  |  |
| CNEWS-1.2.TTL                             | CURRICULUM-ONTOLOGY-4.TTL              |
| SPORT.2.3.TTL                             |  |

**Table 65: Reference Data and Ontologies**

## 4. Performance Metrics

### 4.1. SPB Primary Metrics

| Query Rate Interactive mix (queries/second) | Query Rate Analytical Mix (queries per second) | Update Rate (Operations per second) |
|---|--|-------------------------------------|
| 142.7588                                    | N/A  | 10.6725                             |

Table 16: Primary Metrics

### 4.2. Query Execution Report

| Duration of Bulk Load (in s) | Duration of Measurement Window (in minutes) | No. complete analytical mixes | No. complete interactive mixes | No. complete update operations performed |
|------------------------------|---|-------------------------------|--------------------------------|--|
| 2766.285                     | 20MIN                                       | N/A                           | 13903                          | 12807                                    |

Table 17: Query Execution Report

| Query | Arithmetic Mean of Execution Time | Min Exec Time | Max Exec Time | 90th percentile of Avg Exec Time | Count of Executions |
|-------|-----------------------------------|---------------|---------------|----------------------------------|---------------------|
| Q1    | 137                               | 2             | 2857          | 137                              | 13906               |
| Q2    | 5                                 | 1             | 857           | 5                                | 13908               |
| Q3    | 9                                 | 3             | 825           | 9                                | 13907               |
| Q4    | 5                                 | 1             | 815           | 5                                | 13906               |
| Q5    | 14                                | 1             | 860           | 13                               | 13908               |
| Q6    | 88                                | 48            | 2731          | 88                               | 13905               |
| Q7    | 9                                 | 2             | 767           | 9                                | 13908               |
| Q8    | 80                                | 47            | 2709          | 80                               | 13907               |
| Q9    | 187                               | 2             | 3335          | 187                              | 13906               |
| Q10   | 98                                | 1             | 3088          | 98                               | 13908               |
| Q11   | 9                                 | 3             | 1025          | 9                                | 13908               |
| Q12   | 25                                | 1             | 884           | 25                               | 13908               |

Table 18: Interactive Mix Report

| Query  | Arithmetic Mean of Execution Time | Min Exec Time | Max Exec Time | 90th percentile of Exec Time | Count of Executions |
|--------|-----------------------------------|---------------|---------------|------------------------------|---------------------|
| Q1-Q13 | N/A                               | N/A           | N/A           | N/A                          | N/A                 |

Table 19: Analytical Mix Report

|  |  |
|--|--|
|  |  |
|--|--|

| Query   | Arithmetic Mean of Execution Time | Min Exec Time | Max Exec Time | 90th percentile of Exec Time | Count of Executions |
|---------|-----------------------------------|---------------|---------------|------------------------------|---------------------|
| inserts | 37                                | 17            | 2670          | 37                           | 10171               |
| updates | 79                                | 39            | 979           | 78                           | 1350                |
| deletes | 42                                | 20            | 1035          | 42                           | 1286                |

*Table 20: Update Operation Report*

## 5. Recovery

|                    |     |
|--------------------|-----|
| Execution Duration | N/A |
| Time to Recover    | N/A |

*Table 21: Recovery*

|  |  |
|--|--|
|  |  |
|--|--|

## 6. Pricing Summary

|              | <i>Item</i>   | <i>Price</i> |
|--------------|---|--------------|
|              | SERVER CASE SUPERMICRO CSE-825TQ-R740LPB, 2X740W REDUNDANT 80 PLUS PLATINUM   |              |
|              | MOTHERBOARD SUPERMICRO X10SRI-F, SOCKET R3, INTEL C612, UP TO 256GB DDR4, 1 PCI-E 3.0 X16, 2 PCI- 1                             |              |
|              | CPU INTEL XEON E5-1650 v3 (6 CORES) 3.5GHZ,15MB L3 CACHE, S2011   |              |
|              | COOLER SNK-P0048AP4, SUPERMICRO ,2U ACTIVE CPU HS FOR 2U ,SQUARE&NARROW   |              |
|              | MEMORY 16GB DDR4-2133 2RX4 ECC REG DIMM   |              |
|              | SSD DRIVES 400 GB SSD , SAMSUNG 845DC PRO   |              |
|              | HDD WESTERN DIGITAL HDD 500GB SATAIII RAID EDITION RE4 - 7200RPM  |              |
| <b>TOTAL</b> |   | 3 223 EUR    |
|              | GRAPHDB STANDARD EDITION, SINGLE NODE LICENSE (7 200 EUR), 3 YEARS OF MAINTENANCE (5400 EUR), 24x7 SUPPORT PACKAGE (19 500 EUR) | EUR          |
|              |   |              |
| <b>TOTAL</b> |   | 35 323 EUR   |

Table 22: Pricing Information

## 7. Attachment's CheckList

|          |   |  |
|----------|---|--|
| CLAUSE 1 | RDF DATABASE LOAD SCRIPTS IF ANY  | RDF_DATABASE_SCRIPTS.ZIP               |
| CLAUSE 2 | RDF DATABASE CONFIGURATION OF MORE DETAILS ARE AVAILABLE  | RDF_DATABASE_CONFIGURATION.ZIP         |
| CLAUSE 3 | OPERATING SYSTEM AND HARDWARE SETTINGS  | OPERATING_SYSTEM_HARDWARE_SETTINGS.ZIP |
| CLAUSE 4 | BENCHMARK TEST DRIVER CONFIGURATION AND ANY MODIFIED QUERY TEMPLATES AND (OR) ONTOLOGIES / REFERENCE DATASETS | BENCHMARK_TEST_DRIVER_CONFIG.ZIP       |
| CLAUSE 5 | ALL OF THE LOG FILES GENERATED BY THE BENCHMARK DRIVER  | LOG_FILES.ZIP                          |

**Table 23: Checklist**