

LDBC



*The graph & RDF
benchmark reference*

*Semantic Publishing Benchmark
Full Disclosure Report
Irina Fundulaki
24/04/2015*

Table of Contents

Preface	3
General Terms	4
1. System Description	5
1.1 Database System	5
1.2 Database Engine Configuration	5
1.3 Platform Description	5
1.4 Network Infrastructure Information	6
2. Data Generation & Loading	7
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
3. Benchmark Test Driver	9
3.1. Basic test driver configuration details	9
3.2. Configuration Parameters for Driver Warmup	9
3.3. Configuration Parameters for Driver Execution	9
3.3. Test Driver Reference Data	10
4. Performance Metrics	10
4.1. SPB Primary Metrics	11
4.2. Query Execution Report	11
5. Recovery	13
6. Pricing Summary	14
7. Attachment's CheckList	15

Preface

This is the Full Disclosure Report for the benchmark results produced by Semantic Publishing Benchmark v2.0 for GraphDB EE 6.2beta , Scale Factor SF3 (256M 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

<i>Vendor Name</i>	ONTOTEXT AD
<i>Database Name</i>	GRAPHDB ENTERPRISE EDITION
<i>Version Number</i>	6.2 BETA, BUILD 0xee41aa6a-20150423172407
<i>Database Engine Configuration</i>	SINGLE NODE; Java HotSpot(TM) 64-bit Server VM (build 24.75-b04, mixed mode)

Table 1: DBMS Characteristics

1.2 Database Engine Configuration

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

Table 2: Database Engine Configuration

1.3 Platform Description

<i>Model</i>	SUPERMICRO X10SRI-F, SOCKET R3, INTEL C612
<i>Processors</i>	CPU INTEL XEON E5-1650 V3 3.5GHZ,15MB L3 CACHE, S2011
<i>Memory</i>	MEMORY 4 x 16GB DDR4-2133 2RX4 ECC REG DIMM
<i>No. Disks/Type/Storage Configuration</i>	1 x HDD WESTERN DIGITAL HDD 500GB SATAIII RAID EDITION RE4 - 7200RPM 64MB CACHE – FOR THE OPERATING SYSTEM ONLY 2 x SSD DRIVES 400 GB SSD , SAMSUNG 845DC PRO
<i>Network Adapters</i>	I350 DUAL PORT GIGABIT LAN
<i>Operating System</i>	14.04 UBUNTU 64 BIT
<i>File System</i>	EXT4, SOFTWARE RAID-0 (SSD)
<i>CPU Type & Count</i>	INTEL XEON, 1 CPU
<i>No. Threads</i>	12
<i>No. Cores</i>	6

Memory	64 GB
Total Disks Capacity	734 GB (SSD), 2 x 2,7 TB (HDD), 1 X 413 GB (HDD)

Table 3: System Configuration

No. disks	2 SSD, 1 HDD
Storage interfaces	SAS SATA3, LOCAL STORAGE
Storage Technologies	SOFTWARE RAID-0
RAID/HBA controller	

Table 3: No. Disks/Type/Storage/Configuration

1.4 Network Infrastructure Information

Model	NO EXTERNAL NETWORK
Network Switches	N/A
Wiring Information	N/A

Table 4: Network Infrastructure Information

2. Data Generation & Loading

Generated 231 M triples, using SPB's data generator. Loaded using SPB's loading process.

2.1 Dataset Information

2.1.1 Description

Scale Factor	SF3, 256M
Reference Data Size	25M
Data Format	NQUADS
Data Generator Version	2.0.1369E4A498737C2A18F6AC73DD3F7069C871D5D6
Time Compression ratio	N/A

Table 5: 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	CREATIVWORK.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 6: External Datasets & Ontologies (Listing)

2.1.3 Dataset Characteristics

Statistics for the datasets are produced and stored in file
./generated.256/dataset.info

NO. GENERATED TRIPLES	231,000,000
NO. CREATIVE WORKS	8,821,390
NO. TOTAL TRIPLES	482,240,341

Table 7: Dataset Characteristics

2.1.4 Data Generator Parameters

Data Generator Parameters can be found in files
./definitions.properties and
./test.256.properties.

<p>TEST.256.PROPERTIES</p> <p>DATAGENERATORWORKERS=8 DATASETSIZE=231000000 CREATIVEWORKSPATH=./GENERATED.256 QUERYSUBSTITUTIONPARAMETERS=100000 ENABLECOMPRESSIONONGENERATEDDATA=FALSE</p>
<p>DEFINITIONS.PROPERTIES</p> <p>ABOUTSALLOCATIONS=0.1006, 0.2313, 0.3088, 0.2278, 0.1035, 0.028 MENTIONSALLOCATIONS=0.9477, 0.0382, 0.0093, 0.0031, 0.0012, 0.0005 ENTITYPOPULARITY=0.05, 0.95 USEPOPULARENITIES=0.3, 0.7 CREATIVEWORKTYPESALLOCATION=0.45, 0.35, 0.2 ABOUTANDMENTIONSALLOCATION=0.85, 0.15 EDITORIALOPERATIONSALLOCATION=0.8, 0.1, 0.1 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 EXPONENTIALDECAYUPPERLIMITOFCWS=5000 EXPONENTIALDECAYRATE=0.1 EXPONENTIALDECAYTHRESHOLDPERCENT=0.05 MAJOREVENTS=5 MINOREVENTS=100 SEEDYEAR=2011 DATAGENERATIONPERIODYEARS=1 CORRELATIONSAMOUNT=50 CORRELATIONSMAGNITUDE=60 CORRELATIONDURATION=0.2 CORRELATIONENTITYLIFESPAN=0.5 MINLAT = 36.5 MAXLAT = 70.0 MINLONG = -9.5 MAXLONG = 31.0</p>

Table 8: Data Generator Parameters

2.2 Data Generation Times

Start Time	16:31:09
End Time	16:46:00
Generation Time	870984 (IN MS)

Table 9: Data Generation Times

2.3. Bulk Loading

Start Time	16:46:00
End Time	22:32:00
Bulk Load Time	19654705 (IN MS)

Table 10: Bulk Loading Time

3. Benchmark Test Driver

3.1. Basic test driver configuration details

Basic Test driver configuration details can be found in ./test.256.properties file.

No. read threads	8
No. write threads	2

Table 11: Basic Test Driver Configuration Details

3.2. Configuration Parameters for Driver Warmup

Test driver configuration details for Driver Warmup can be found in ./test.256.properties file.

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

Table 12: Configuration Parameters for Driver Warmup

3.3. Configuration Parameters for Driver Execution

Test driver configuration details for Driver Execution can be found in ./test.256.properties file.

```

AGGREGATIONAGENTS=8
EDITORIALAGENTS=2
BENCHMARKRUNPERIODSECONDS=1800
QUERYTIMEOUTSECONDS=300
BENCHMARKBYQUERYMIXRUNS=0
MINUPDATERATETHRESHOLDOPS=0.0
MINUPDATERATETHRESHOLDREACHTIMEPERCENT=0
MAXUPDATERATETHRESHOLDOPS=0.0
ENDPOINTURL=HTTP://LOCALHOST:4082/OPENRDF-SESAME/REPOSITORIES/LDBC256
ENDPOINTUPDATEURL=HTTP://LOCALHOST:4082/OPENRDF-
SESAME/REPOSITORIES/LDBC256/STATEMENTS
QUERIESPATH=./DATA/SPARQL
RUNBENCHMARK=TRUE

```

Table 13: 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	

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)
29.8986	N/A	9.5017

Table 14: Primary Metrics

4.2. Query Execution Report

Duration of Bulk Load (in ms)	Duration of Measurement Window (in minutes)	No. complete analytical mixes per second	No. complete interactive mixes per second	No. complete update operations performed
19654705	30MIN	N/A	4052	17103

Table 10: 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	202	4	1785	202	4057
Q2	6	2	1257	6	4057
Q3	44	9	2462	43	4056
Q4	24	3	1729	23	4057
Q5	88	3	2301	88	4057
Q6	166	56	2021	168	4058
Q7	76	6	1225	76	4057
Q8	150	91	1617	151	4055
Q9	427	4	2819	430	4055
Q10	2155	2	13520	2153	4054
Q11	24	4	7895	25	4057
Q12	125	3	1873	125	4055

Table 11: 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 15: Analytical Mix Report

Query	Arithmetic Mean of Execution Time	Min Exec Time	Max Exec Time	90th percentile of Exec Time	Count of Executions
inserts	58	28	1738	58	13619
updates	126	69	1802	126	1709
deletes	64	35	1713	65	1775

Table 16: Update Operation Report

5. Recovery

Execution Duration	N/A
Time to Recover	N/A

Table 14: 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	
	HDD HITACHI 3TB , ULTRASTAR, 7K4000, HUS724030ALE640, 7200 RPM, SATAIII, 64MB	
	SSD DRIVES 400 GB SSD , SAMSUNG 845DC PRO	
	RAID CONTROLLER LSI LSI00419 MEGARAID 9341-4i SGL SAS 3/SATA CONTROLLER, 4 CHANNEL, 12GB/S	
	HDD WESTERN DIGITAL HDD 500GB SATAIII RAID EDITION RE4 - 7200RPM	
TOTAL		\$4252 w/o VAT

Table 15: 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 17: Checklist