

LDBC

*The graph & RDF
benchmark reference*

Semantic Publishing Benchmark

2.0.2

Full Disclosure Report

Pjotr Scholtze

2023-01-29

Table of Contents

Preface	4
General Terms	5
1. System Description	6
1.1 Database System.....	6
1.2 Database Engine Configuration.....	6
1.3 Platform Description.....	7
1.4 Network Infrastructure Information.....	8
2. Data Generation & Loading	9
2.1 Dataset Information.....	9
2.1.1 Description.....	9
2.1.2 External Datasets.....	9
2.1.3 Dataset Characteristics.....	9
2.1.4 Data Generator Parameters.....	9
2.2 Data Generation Times.....	10
2.3. Bulk Loading.....	10
2.4. Additional setup.....	10
2.5. Raw results.....	11
3. Test case: 0 read threads, 4 write threads	12
3.1. Benchmark Test Driver.....	12
3.1.1. Basic test driver configuration details.....	12
3.1.2. Configuration Parameters for Driver Warmup.....	12
3.1.3. Configuration Parameters for Driver Execution.....	12
3.1.4. Raw test.properties.....	12
3.1.5. Test Driver Reference Data.....	13
3.2. Performance Metrics.....	14
4.2.1. SPB Primary Metrics.....	14
4.2.2. Query Execution Report.....	14
4.2.3. Raw results.....	14
4. Test case: 8 read threads, 4 write threads	16
4.1. Benchmark Test Driver.....	16
4.1.1. Basic test driver configuration details.....	16
4.1.2. Configuration Parameters for Driver Warmup.....	16
4.1.3. Configuration Parameters for Driver Execution.....	16
4.1.4. Raw test.properties.....	16
4.1.5. Test Driver Reference Data.....	17
4.2. Performance Metrics.....	18
4.2.1. SPB Primary Metrics.....	18
4.2.2. Query Execution Report.....	18
4.2.3. Raw results.....	18
5. Test case: 16 read threads, 0 write threads	20
5.1. Benchmark Test Driver.....	20
5.1.1. Basic test driver configuration details.....	20
5.1.2. Configuration Parameters for Driver Warmup.....	20
5.1.3. Configuration Parameters for Driver Execution.....	20
5.1.4. Raw test.properties.....	20
5.1.5. Test Driver Reference Data.....	21
5.2. Performance Metrics.....	22
5.2.1. SPB Primary Metrics.....	22
5.2.2. Query Execution Report.....	22

5.2.3. Raw results	22
6. Test case: 16 read threads, 4 write threads	24
6.1. Benchmark Test Driver	24
6.1.1. Basic test driver configuration details	24
6.1.2. Configuration Parameters for Driver Warmup	24
6.1.3. Configuration Parameters for Driver Execution	24
6.1.4. Raw test.properties	24
6.1.5. Test Driver Reference Data	25
6.2. Performance Metrics	26
6.2.1. SPB Primary Metrics	26
6.2.2. Query Execution Report	26
6.2.3. Raw results	26
7. Test case: 24 read threads, 0 write threads	28
7.1. Benchmark Test Driver	28
7.1.1. Basic test driver configuration details	28
7.1.2. Configuration Parameters for Driver Warmup	28
7.1.3. Configuration Parameters for Driver Execution	28
7.1.4. Raw test.properties	28
7.1.5. Test Driver Reference Data	29
7.2. Performance Metrics	30
7.2.1. SPB Primary Metrics	30
7.2.2. Query Execution Report	30
7.2.3. Raw results	30
8. Recovery	32
9. Pricing Summary	33
10. Attachment's CheckList	34

Preface

This is the Full Disclosure Report for the benchmark results produced by Semantic Publishing Benchmark v2.0 for GraphDB EE 10.1.0, Scale Factor 5 (1B Triples), using a single r6id.8xlarge AWS server.

General Terms

Test Sponsor

Test sponsor of this benchmark is Sirma AI Ad, trading as Ontotext, Bulgaria. Ontotext is a global leader in enterprise knowledge graph technology and semantic database engines. Its semantic technology is distinctive for its performance, scalability, robustness and compliance with open standards. Ontotext solutions have been applied for business critical projects in the areas of financial services (S&P), publishing (BBC, Financial Times), life sciences (Astra Zeneca), industry and infrastructure management (Johnson Controls, Siemens, Stellantis/PSA), government organizations (UK Parliament, NASA) and others. Ontotext is the developer of GraphDB semantic graph database.

1. System Description

1.1 Database System

Vendor Name	Ontotext
Database Name	GraphDB Enterprise Edition
Version No	10.1.0
Database Engine Configuration	G1 Garbage Collector with 80gb of heap memory and a few internal properties. (-XX:+UseG1GC -Xmx80g -Din-memory-literal-properties=true -Ddefault.min.distinct.threshold=100m)
Full Docker run command	docker run -d -p 7200:7200 -v /opt/temp-test-data/gdb-home:/opt/graphdb/home ontotext/graphdb:10.1.0 /opt/graphdb/dist/bin/graphdb -Dgraphdb.home=/opt/graphdb/home -XX:+UseG1GC -Xmx80g -Din-memory-literal-properties=true -Ddefault.min.distinct.threshold=100m -Dgraphdb.cluster.rpc.port=8008

Table 1: DBMS Characteristics

1.2 Database Engine Configuration

Cache Configuration	80GB Reserved memory (Java: -Xmx80g)
Transaction Isolation Level/Model	Not configurable
Reasoning Strategy/configuration	Tailormade SPB ruleset has been loaded
triple/quadruple indices	N/A
indices	GEO-SPATIAL, TEXT INDEXES (LUCENE CONNECTOR)
Optimisations	OWL:SAMEAS

Table 2: Database Engine Configuration

1.3 Platform Description

Model	AWS EC2 Instance: r6id.8xlarge
Processors	Intel Xeon Platinum 8375C, 3500 Mhz, 32 cores, L1 cache: 2560KiB, L2 cache: 40MiB, L3 cache: 54MiB
Memory	256GB, DIMM DDR4 Static column Pseudo-static Synchronous Window DRAM 3200MHz (0.3 ns)
No Disks/Type/Storage Configuration	/dev/nvme1n1= 1900GB (NVMe SSD)
Network Adapters	Elastic Network Adapter (ENA) – Amazon.com, Inc. (mtu 9001)
Operating System	Linux ip-172-31-1-44 5.15.0-1019-aws #23-Ubuntu SMP Wed Aug 17 18:33:13 UTC 2022 x86_64 x86_64 x86_64 GNU/Linux
File System	/dev/nvme1n1 ext4 /
CPU Type & Count	Intel Xeon Platinum 8375C
No vcpu Cores	32
Memory	256GB
Total Disks Capacity	1900GB
Model	AWS EC2 Instance: r6id.8xlarge

Table 3: System Configuration

No disks	1 SSD
Storage interfaces	1 x 1900GB NVMe SSD
Storage Technologies	NVMe
RAID/HBA controller	

Table 3: No Disks/Type/Storage/Configuration

1.4 Network Infrastructure Information

Model	Elastic Network Adapter (ENA) – Amazon.com, Inc.
Network Switches	AWS
Wiring Information	AWS

Table 4: Network Infrastructure Information

2. Data Generation & Loading

2.1 Dataset Information

2.1.1 Description

Scale Factor	SF-5 1B
Data Format	Direct SPARQL endpoint insert
Data Generator Version	2.0.2 (basic-querymix-graphdb)
Time Compression ratio	N/A

Table 5: Dataset characteristics

2.1.2 External Datasets

2.1.3 Dataset Characteristics

(reference data entities size : 1155229, greatest Creative Work id : 37271243, dbpedia locations : 268862, geonames locations : 1285124)

2.1.4 Data Generator Parameters

```
#Benchmark Threads Configuration
aggregationAgents=8
editorialAgents=2
dataGeneratorWorkers=4

#Benchmark Configuration Parameters
warmupPeriodSeconds=400
benchmarkRunPeriodSeconds=30
datasetSize=976000000
generatedTriplesPerFile=50000
queryTimeoutSeconds=300
querySubstitutionParameters=100000
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/repositories/LDBC-SF5/statements
ontologiesPath=./data/ontologies
referenceDatasetsPath=./data/datasets
creativeWorksPath=./generated
queriesPath=./data/sparql
validationPath=./data/validation
definitionsPath=./definitions.properties
scriptsPath= *** full path to /data/scripts ***
```

```

enableCompressionOnGeneratedData=false
creativeWorksInfo=dataset.info
generateCreativeWorksFormat=n-quads
currentRateReportPeriodSeconds=0
saveDetailedQueryLogs=true
verbose=true

#Benchmark Operational Phases
loadOntologies=true
loadReferenceDatasets=true
generateCreativeWorks=true
loadCreativeWorks=true
generateQuerySubstitutionParameters=true
validateQueryResults=false
warmUp=false
runBenchmark=false
runBenchmarkOnlineReplicationAndBackup=false
checkConformance=false

```

2.2 Data Generation Times

Data generation and bulk loading went together.

Start Time (in ms)	01:03:01
End Time (in ms)	18:04:58
Generation Time	61 317 016 ms (=17h 01m 57s)

Table 6: Data Generation Times

2.3. Bulk Loading

Data generation and bulk loading went together.

Start Time (in ms)	01:03:01
End Time (in ms)	18:04:58
Generation Time	61 317 016 ms (=17h 01m 57s)

Table 7: Bulk Loading Time

2.4. Additional setup

Lucene connector time	28m57s
Create a geospatial index	4.5s

2.5. Raw results

```
Loaded 21542 files with Creative Works in 61317016 milliseconds
Warning : Possible wrong configuration for property 'scriptsPath' (test.properties)...
Generating query parameters
Analyzing existing reference knowledge in database, it may take a while...
  (reference data entities size : 1155229, greatest Creative Work id : 37271243,
  dbpedia locations : 268862, geonames locations : 1285124)
.....
```

```
END OF RUN, all agents shut down...
```

```
real 1056m27.178s
user  89m27.799s
sys   7m59.414s
```

3. Test case: 0 read threads, 4 write threads

3.1. Benchmark Test Driver

3.1.1. Basic test driver configuration details

<i>No of read threads</i>	0
<i>No of write threads</i>	4

Table 8: Basic Test Driver Configuration Details

3.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=0
queryTimeoutSeconds=300
warmUp=true
```

3.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=0
editorialAgents=4
benchmarkRunPeriodSeconds=1200
queryTimeoutSeconds=300
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
queriesPath=./data/sparql
runBenchmark=true
```

3.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=0
editorialAgents=4
dataGeneratorWorkers=4

#Benchmark Configuration Parameters
warmupPeriodSeconds=400
benchmarkRunPeriodSeconds=1200
datasetSize=976000000 # (SF5)
generatedTriplesPerFile=50000
queryTimeoutSeconds=300
querySubstitutionParameters=100000
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
#endpointURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10
```

```
#endpointUpdateURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10/statements
ontologiesPath=./data/ontologies
referenceDatasetsPath=./data/datasets
creativeWorksPath=./generated
queriesPath=./data/sparql
validationPath=./data/validation
definitionsPath=./definitions.properties
scriptsPath= *** full path to /data/scripts ***
enableCompressionOnGeneratedData=false
creativeWorksInfo=dataset.info
generateCreativeWorksFormat=n-quads
currentRateReportPeriodSeconds=0
saveDetailedQueryLogs=false
verbose=true
```

```
#Benchmark Operational Phases
loadOntologies=false
loadReferenceDatasets=false
generateCreativeWorks=false
loadCreativeWorks=false
generateQuerySubstitutionParameters=true
validateQueryResults=false
warmUp=true
runBenchmark=true
runBenchmarkOnlineReplicationAndBackup=false
checkConformance=false
```

3.1.5. Test Driver Reference Data

No modifications made.

3.2. Performance Metrics

4.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
0	17.4625

Table 9: Primary Metrics

4.2.2. Query Execution Report

Duration of Bulk Load	No of retrieval queries	No of update queries	No of insert queries	No of delete queries
13 333 637 ms	0	2078	16766	2111

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	0 ms	0 ms	0 ms	0
Q2	0 ms	0 ms	0 ms	0
Q3	0 ms	0 ms	0 ms	0
Q4	0 ms	0 ms	0 ms	0
Q5	0 ms	0 ms	0 ms	0
Q6	0 ms	0 ms	0 ms	0
Q7	0 ms	0 ms	0 ms	0
Q8	0 ms	0 ms	0 ms	0
Q9	0 ms	0 ms	0 ms	0
Q10	0 ms	0 ms	0 ms	0
Q11	0 ms	0 ms	0 ms	0
Q12	0 ms	0 ms	0 ms	0

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	220 ms	181 ms	952 ms	16766
Update	254 ms	193 ms	622 ms	2078
Delete	245 ms	188 ms	408 ms	2111

Table 12: Editorial Report

4.2.3. Raw results

Seconds : 1200

2022-12-05 11:34:12 (completed query mixes : 0)

Editorial:

4 agents

16766 inserts (avg : 220 ms, min : 181 ms, max : 952 ms)

2078 updates (avg : 254 ms, min : 193 ms, max : 622 ms)

2111 deletes (avg : 245 ms, min : 188 ms, max : 408 ms)

20955 operations (16766 CW Inserts (0 errors), 2078 CW Updates (0 errors), 2111 CW Deletions (0 errors))

17.4625 average operations per second

Aggregation:

0 agents

0	Q1	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q2	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q3	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q4	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q5	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q6	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q7	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q8	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q9	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q10	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q11	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)
0	Q12	queries	(avg : 0	ms, min : 0	ms, max : 0	ms, 0 errors)

0 total retrieval queries (0 errors)

0.0000 average queries per second

Stopping the benchmark...

END OF RUN, all agents shut down...

real 27m48.771s
user 1m9.748s
sys 0m4.993s

4. Test case: 8 read threads, 4 write threads

4.1. Benchmark Test Driver

4.1.1. Basic test driver configuration details

No of read threads	8
No of write threads	4

Table 10: Basic Test Driver Configuration Details

4.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=8
queryTimeoutSeconds=300
warmUp=true
```

4.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=8
editorialAgents=4
benchmarkRunPeriodSeconds=1200
queryTimeoutSeconds=300
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
queriesPath=./data/sparql
runBenchmark=true
```

4.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=8
editorialAgents=4
dataGeneratorWorkers=4

#Benchmark Configuration Parameters
warmupPeriodSeconds=400
benchmarkRunPeriodSeconds=1200
datasetSize=976000000 # (SF5)
generatedTriplesPerFile=50000
queryTimeoutSeconds=300
querySubstitutionParameters=100000
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
#endpointURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10
```



```
#endpointUpdateURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10/statements
ontologiesPath=./data/ontologies
referenceDatasetsPath=./data/datasets
creativeWorksPath=./generated
queriesPath=./data/sparql
validationPath=./data/validation
definitionsPath=./definitions.properties
scriptsPath= *** full path to /data/scripts ***
enableCompressionOnGeneratedData=false
creativeWorksInfo=dataset.info
generateCreativeWorksFormat=n-quads
currentRateReportPeriodSeconds=0
saveDetailedQueryLogs=false
verbose=true
```

```
#Benchmark Operational Phases
loadOntologies=false
loadReferenceDatasets=false
generateCreativeWorks=false
loadCreativeWorks=false
generateQuerySubstitutionParameters=true
validateQueryResults=false
warmUp=true
runBenchmark=true
runBenchmarkOnlineReplicationAndBackup=false
checkConformance=false
```

4.1.5. Test Driver Reference Data

No modifications made.

4.2. Performance Metrics

4.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
69.0275	13.1450

Table 11: Primary Metrics

4.2.2. Query Execution Report

Duration of Bulk Load	No of retrieval queries	No of update queries	No of insert queries	No of delete queries
13 333 637 ms	82833	1566	12599	1609

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	54 ms	6 ms	575 ms	6901
Q2	3 ms	2 ms	134 ms	6904
Q3	78 ms	33 ms	1665 ms	6903
Q4	20 ms	4 ms	305 ms	6902
Q5	163 ms	7 ms	2862 ms	6901
Q6	60 ms	36 ms	233 ms	6904
Q7	10 ms	5 ms	142 ms	6902
Q8	128 ms	89 ms	573 ms	6903
Q9	101 ms	3 ms	569 ms	6904
Q10	180 ms	3 ms	691 ms	6902
Q11	49 ms	3 ms	1672 ms	6903
Q12	16 ms	8 ms	199 ms	6904

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	294 ms	83 ms	916 ms	12599
Update	337 ms	226 ms	656 ms	1566
Delete	323 ms	217 ms	857 ms	1609

Table 12: Editorial Report

4.2.3. Raw results

Seconds : 1200

2022-12-05 18:13:57 (completed query mixes : 6899)

Editorial:

4 agents

12599 inserts (avg : 294 ms, min : 83 ms, max : 916 ms)

1566 updates (avg : 337 ms, min : 226 ms, max : 656 ms)

1609 deletes (avg : 323 ms, min : 217 ms, max : 857 ms)

15774 operations (12599 CW Inserts (0 errors), 1566 CW Updates (0 errors), 1609 CW Deletions (0 errors))

13.1450 average operations per second

Aggregation:

8 agents

6901	Q1	queries	(avg : 54	ms, min : 6	ms, max : 575	ms, 0 errors)
6904	Q2	queries	(avg : 3	ms, min : 2	ms, max : 134	ms, 0 errors)
6903	Q3	queries	(avg : 78	ms, min : 33	ms, max : 1665	ms, 0 errors)
6902	Q4	queries	(avg : 20	ms, min : 4	ms, max : 305	ms, 0 errors)
6901	Q5	queries	(avg : 163	ms, min : 7	ms, max : 2862	ms, 0 errors)
6904	Q6	queries	(avg : 60	ms, min : 36	ms, max : 233	ms, 0 errors)
6902	Q7	queries	(avg : 10	ms, min : 5	ms, max : 142	ms, 0 errors)
6903	Q8	queries	(avg : 128	ms, min : 89	ms, max : 573	ms, 0 errors)
6904	Q9	queries	(avg : 101	ms, min : 3	ms, max : 569	ms, 0 errors)
6902	Q10	queries	(avg : 180	ms, min : 3	ms, max : 691	ms, 0 errors)
6903	Q11	queries	(avg : 49	ms, min : 3	ms, max : 1672	ms, 0 errors)
6904	Q12	queries	(avg : 16	ms, min : 8	ms, max : 199	ms, 0 errors)

82833 total retrieval queries (0 errors)

69.0275 average queries per second

Stopping the benchmark...

END OF RUN, all agents shut down...

real 28m0.396s
user 15m8.525s
sys 1m12.269s

5. Test case: 16 read threads, 0 write threads

5.1. Benchmark Test Driver

5.1.1. Basic test driver configuration details

No of read threads	16
No of write threads	0

Table 12: Basic Test Driver Configuration Details

5.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=16
queryTimeoutSeconds=300
warmUp=true
```

5.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=16
editorialAgents=0
benchmarkRunPeriodSeconds=1200
queryTimeoutSeconds=300
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
queriesPath=./data/sparql
runBenchmark=true
```

5.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=16
editorialAgents=0
dataGeneratorWorkers=4

#Benchmark Configuration Parameters
warmupPeriodSeconds=400
benchmarkRunPeriodSeconds=1200
datasetSize=976000000 # (SF5)
generatedTriplesPerFile=50000
queryTimeoutSeconds=300
querySubstitutionParameters=100000
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
#endpointURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10
```

```
#endpointUpdateURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10/statements
ontologiesPath=./data/ontologies
referenceDatasetsPath=./data/datasets
creativeWorksPath=./generated
queriesPath=./data/sparql
validationPath=./data/validation
definitionsPath=./definitions.properties
scriptsPath= *** full path to /data/scripts ***
enableCompressionOnGeneratedData=false
creativeWorksInfo=dataset.info
generateCreativeWorksFormat=n-quads
currentRateReportPeriodSeconds=0
saveDetailedQueryLogs=false
verbose=true
```

```
#Benchmark Operational Phases
loadOntologies=false
loadReferenceDatasets=false
generateCreativeWorks=false
loadCreativeWorks=false
generateQuerySubstitutionParameters=true
validateQueryResults=false
warmUp=true
runBenchmark=true
runBenchmarkOnlineReplicationAndBackup=false
checkConformance=false
```

5.1.5. Test Driver Reference Data

No modifications made.

5.2. Performance Metrics

5.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
130.2000	0

Table 13: Primary Metrics

5.2.2. Query Execution Report

Duration of Bulk Load	No of retrieval queries	No of update queries	No of insert queries	No of delete queries
13 333 637 ms	156240	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	54 ms	6 ms	522 ms	13020
Q2	3 ms	1 ms	151 ms	13021
Q3	77 ms	29 ms	2231 ms	13020
Q4	19 ms	4 ms	545 ms	13019
Q5	166 ms	6 ms	3114 ms	13016
Q6	60 ms	34 ms	283 ms	13019
Q7	9 ms	5 ms	161 ms	13020
Q8	143 ms	92 ms	443 ms	13020
Q9	117 ms	3 ms	563 ms	13021
Q10	205 ms	2 ms	707 ms	13022
Q11	27 ms	3 ms	2169 ms	13020
Q12	15 ms	7 ms	161 ms	13022

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	0 ms	0 ms	0 ms	0
Update	0 ms	0 ms	0 ms	0
Delete	0 ms	0 ms	0 ms	0

Table 12: Editorial Report

5.2.3. Raw results

Seconds : 1200

2022-12-05 15:52:29 (completed query mixes : 13012)

Editorial:

0 agents

0 inserts (avg : 0 ms, min : 0 ms, max : 0 ms)

0 updates (avg : 0 ms, min : 0 ms, max : 0 ms)

0 deletes (avg : 0 ms, min : 0 ms, max : 0 ms)

0 operations (0 CW Inserts (0 errors), 0 CW Updates (0 errors), 0 CW Deletions (0 errors))
0.0000 average operations per second

Aggregation:

16 agents

13020	Q1	queries	(avg : 54	ms, min : 6	ms, max : 522	ms, 0 errors)
13021	Q2	queries	(avg : 3	ms, min : 1	ms, max : 151	ms, 0 errors)
13020	Q3	queries	(avg : 77	ms, min : 29	ms, max : 2231	ms, 0 errors)
13019	Q4	queries	(avg : 19	ms, min : 4	ms, max : 545	ms, 0 errors)
13016	Q5	queries	(avg : 166	ms, min : 6	ms, max : 3114	ms, 0 errors)
13019	Q6	queries	(avg : 60	ms, min : 34	ms, max : 283	ms, 0 errors)
13020	Q7	queries	(avg : 9	ms, min : 5	ms, max : 161	ms, 0 errors)
13020	Q8	queries	(avg : 143	ms, min : 92	ms, max : 443	ms, 0 errors)
13021	Q9	queries	(avg : 117	ms, min : 3	ms, max : 563	ms, 0 errors)
13022	Q10	queries	(avg : 205	ms, min : 2	ms, max : 707	ms, 0 errors)
13020	Q11	queries	(avg : 27	ms, min : 3	ms, max : 2169	ms, 0 errors)
13022	Q12	queries	(avg : 15	ms, min : 7	ms, max : 161	ms, 0 errors)

156240 total retrieval queries (0 errors)
130.2000 average queries per second

Stopping the benchmark...
END OF RUN, all agents shut down...

real 28m14.018s
user 25m1.795s
sys 2m1.805s

6. Test case: 16 read threads, 4 write threads

6.1. Benchmark Test Driver

6.1.1. Basic test driver configuration details

No of read threads	16
No of write threads	4

Table 14: Basic Test Driver Configuration Details

6.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=16
queryTimeoutSeconds=300
warmUp=true
```

6.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=16
editorialAgents=4
benchmarkRunPeriodSeconds=1200
queryTimeoutSeconds=300
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
queriesPath=./data/sparql
runBenchmark=true
```

6.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=16
editorialAgents=4
dataGeneratorWorkers=4

#Benchmark Configuration Parameters
warmupPeriodSeconds=400
benchmarkRunPeriodSeconds=1200
datasetSize=976000000 # (SF5)
generatedTriplesPerFile=50000
queryTimeoutSeconds=300
querySubstitutionParameters=100000
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
#endpointURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10
```



```
#endpointUpdateURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10/statements
ontologiesPath=./data/ontologies
referenceDatasetsPath=./data/datasets
creativeWorksPath=./generated
queriesPath=./data/sparql
validationPath=./data/validation
definitionsPath=./definitions.properties
scriptsPath= *** full path to /data/scripts ***
enableCompressionOnGeneratedData=false
creativeWorksInfo=dataset.info
generateCreativeWorksFormat=n-quads
currentRateReportPeriodSeconds=0
saveDetailedQueryLogs=false
verbose=true
```

```
#Benchmark Operational Phases
loadOntologies=false
loadReferenceDatasets=false
generateCreativeWorks=false
loadCreativeWorks=false
generateQuerySubstitutionParameters=true
validateQueryResults=false
warmUp=true
runBenchmark=true
runBenchmarkOnlineReplicationAndBackup=false
checkConformance=false
```

6.1.5. Test Driver Reference Data

No modifications made.

6.2. Performance Metrics

6.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
105.7608	10.4492

Table 15: Primary Metrics

6.2.2. Query Execution Report

Duration of Bulk Load	No of retrieval queries	No of update queries	No of insert queries	No of delete queries
13 333 637 ms	126913	1224	10077	1238

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	61 ms	8 ms	575 ms	10577
Q2	4 ms	2 ms	153 ms	10577
Q3	95 ms	36 ms	2470 ms	10575
Q4	25 ms	4 ms	570 ms	10576
Q5	214 ms	7 ms	7194 ms	10575
Q6	67 ms	35 ms	268 ms	10576
Q7	12 ms	5 ms	226 ms	10577
Q8	164 ms	88 ms	462 ms	10576
Q9	184 ms	3 ms	5956 ms	10577
Q10	240 ms	3 ms	1594 ms	10573
Q11	33 ms	3 ms	2781 ms	10579
Q12	20 ms	8 ms	204 ms	10575

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	370 ms	82 ms	1348 ms	10077
Update	422 ms	240 ms	738 ms	1224
Delete	409 ms	239 ms	895 ms	1238

Table 12: Editorial Report

6.2.3. Raw results

Seconds : 1200

2022-12-21 23:32:21 (completed query mixes : 10569)

Editorial:

4 agents

10077 inserts (avg : 370 ms, min : 82 ms, max : 1348 ms)

1224 updates (avg : 422 ms, min : 240 ms, max : 738 ms)

1238 deletes (avg : 409 ms, min : 239 ms, max : 895 ms)

12539 operations (10077 CW Inserts (1 errors), 1224 CW Updates (0 errors), 1238 CW Deletions (0 errors))

10.4492 average operations per second

Aggregation:

16 agents

10577 Q1	queries	(avg : 61	ms, min : 8	ms, max : 575	ms, 0 errors)
10577 Q2	queries	(avg : 4	ms, min : 2	ms, max : 153	ms, 0 errors)
10575 Q3	queries	(avg : 95	ms, min : 36	ms, max : 2470	ms, 0 errors)
10576 Q4	queries	(avg : 25	ms, min : 4	ms, max : 570	ms, 0 errors)
10575 Q5	queries	(avg : 214	ms, min : 7	ms, max : 7194	ms, 0 errors)
10576 Q6	queries	(avg : 67	ms, min : 35	ms, max : 268	ms, 0 errors)
10577 Q7	queries	(avg : 12	ms, min : 5	ms, max : 226	ms, 0 errors)
10576 Q8	queries	(avg : 164	ms, min : 88	ms, max : 462	ms, 0 errors)
10577 Q9	queries	(avg : 184	ms, min : 3	ms, max : 5956	ms, 0 errors)
10573 Q10	queries	(avg : 240	ms, min : 3	ms, max : 1594	ms, 0 errors)
10579 Q11	queries	(avg : 33	ms, min : 3	ms, max : 2781	ms, 0 errors)
10575 Q12	queries	(avg : 20	ms, min : 8	ms, max : 204	ms, 0 errors)

126913 total retrieval queries (0 errors)

105.7608 average queries per second

Stopping the benchmark...

END OF RUN, all agents shut down...

real 28m8.504s
user 22m35.107s
sys 1m49.773s

7. Test case: 24 read threads, 0 write threads

7.1. Benchmark Test Driver

7.1.1. Basic test driver configuration details

No of read threads	24
No of write threads	0

Table 16: Basic Test Driver Configuration Details

7.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=24
queryTimeoutSeconds=300
warmUp=true
```

7.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=24
editorialAgents=0
benchmarkRunPeriodSeconds=1200
queryTimeoutSeconds=300
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
queriesPath=./data/sparql
runBenchmark=true
```

7.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=24
editorialAgents=0
dataGeneratorWorkers=4

#Benchmark Configuration Parameters
warmupPeriodSeconds=400
benchmarkRunPeriodSeconds=1200
datasetSize=976000000 # (SF5)
generatedTriplesPerFile=50000
queryTimeoutSeconds=300
querySubstitutionParameters=100000
benchmarkByQueryMixRuns=0
minUpdateRateThresholdOps=0.0
minUpdateRateThresholdReachTimePercent=0
maxUpdateRateThresholdOps=0.0
endpointURL=http://54.92.203.255:7200/repositories/LDBC-SF5
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF5/statements
#endpointURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10
```

```
#endpointUpdateURL=http://localhost:8080/openrdf-sesame/repositories/ldbc10/statements
ontologiesPath=./data/ontologies
referenceDatasetsPath=./data/datasets
creativeWorksPath=./generated
queriesPath=./data/sparql
validationPath=./data/validation
definitionsPath=./definitions.properties
scriptsPath= *** full path to /data/scripts ***
enableCompressionOnGeneratedData=false
creativeWorksInfo=dataset.info
generateCreativeWorksFormat=n-quads
currentRateReportPeriodSeconds=0
saveDetailedQueryLogs=false
verbose=true
```

```
#Benchmark Operational Phases
loadOntologies=false
loadReferenceDatasets=false
generateCreativeWorks=false
loadCreativeWorks=false
generateQuerySubstitutionParameters=true
validateQueryResults=false
warmUp=true
runBenchmark=true
runBenchmarkOnlineReplicationAndBackup=false
checkConformance=false
```

7.1.5. Test Driver Reference Data

No modifications made.

7.2. Performance Metrics

7.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
158.1042	0

Table 17: Primary Metrics

7.2.2. Query Execution Report

Duration of Bulk Load	No of retrieval queries	No of update queries	No of insert queries	No of delete queries
13 333 637 ms	189725	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	63 ms	6 ms	518 ms	15813
Q2	4 ms	1 ms	155 ms	15812
Q3	95 ms	33 ms	2530 ms	15813
Q4	24 ms	4 ms	610 ms	15812
Q5	204 ms	6 ms	3903 ms	15808
Q6	70 ms	34 ms	286 ms	15807
Q7	12 ms	5 ms	178 ms	15813
Q8	176 ms	94 ms	584 ms	15808
Q9	160 ms	2 ms	746 ms	15806
Q10	228 ms	3 ms	800 ms	15810
Q11	32 ms	2 ms	2607 ms	15810
Q12	20 ms	7 ms	188 ms	15813

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	0 ms	0 ms	0 ms	0
Update	0 ms	0 ms	0 ms	0
Delete	0 ms	0 ms	0 ms	0

Table 12: Editorial Report

7.2.3. Raw results

Seconds : 1200

2022-12-21 21:55:42 (completed query mixes : 15799)

Editorial:

0 agents

0 inserts (avg : 0 ms, min : 0 ms, max : 0 ms)

0 updates (avg : 0 ms, min : 0 ms, max : 0 ms)

0 deletes (avg : 0 ms, min : 0 ms, max : 0 ms)

0 operations (0 CW Inserts (0 errors), 0 CW Updates (0 errors), 0 CW Deletions (0 errors))
0.0000 average operations per second

Aggregation:

24 agents

15813	Q1	queries	(avg : 63	ms, min : 6	ms, max : 518	ms, 0 errors)
15812	Q2	queries	(avg : 4	ms, min : 1	ms, max : 155	ms, 0 errors)
15813	Q3	queries	(avg : 95	ms, min : 33	ms, max : 2530	ms, 0 errors)
15812	Q4	queries	(avg : 24	ms, min : 4	ms, max : 610	ms, 0 errors)
15808	Q5	queries	(avg : 204	ms, min : 6	ms, max : 3903	ms, 0 errors)
15807	Q6	queries	(avg : 70	ms, min : 34	ms, max : 286	ms, 0 errors)
15813	Q7	queries	(avg : 12	ms, min : 5	ms, max : 178	ms, 0 errors)
15808	Q8	queries	(avg : 176	ms, min : 94	ms, max : 584	ms, 0 errors)
15806	Q9	queries	(avg : 160	ms, min : 2	ms, max : 746	ms, 0 errors)
15810	Q10	queries	(avg : 228	ms, min : 3	ms, max : 800	ms, 0 errors)
15810	Q11	queries	(avg : 32	ms, min : 2	ms, max : 2607	ms, 0 errors)
15813	Q12	queries	(avg : 20	ms, min : 7	ms, max : 188	ms, 0 errors)

189725 total retrieval queries (0 errors)
158.1042 average queries per second

Stopping the benchmark...
END OF RUN, all agents shut down...

real 28m50.099s
user 31m34.261s
sys 2m44.701s

8. Recovery

<i>Execution Duration</i>	N/A
<i>Time to Recover</i>	N/A

Table 14: Recovery

9. Pricing Summary

See <https://calculator.aws/#/createCalculator/EC2> for the current details on the r6id.8xlarge machine(s). Tables 18-22 describe pricing information for three years of usage. Using the same r6id.8xlarge Amazon machine, costing a fixed \$27,474.12 for all scenarios for three years (\$763.17 per month, standard saving plan, no upfront). The license cost is variable based on the number of cores used, not the number of cores available. Each core costs \$2500 per year. The maintenance cost is \$0 since this is already covered in the license cost. The number of cores required for a license is calculated as follows:

$$\text{CPU_cores_license} = \text{number_of_read_threads} + (\text{number_of_machines} * \text{number_of_write_threads})$$

This is because the writing threads are not distributed through the cluster, but the reading threads are.

	<i>Item</i>	<i>Price</i>
1X	AMAZON R6ID.8XLARGE	\$27,474.12
1X	GRAPHDB LICENSE (FOR 4 CPU CORES) (EE)	\$30,000
0X	MAINTENANCE	\$0
TOTAL		\$57,474

Table 18 Pricing information: 0 read threads, 4 write threads

	<i>Item</i>	<i>Price</i>
1X	AMAZON R6ID.8XLARGE	\$27,474.12
1X	GRAPHDB LICENSE (FOR 16 CPU CORES) (EE)	\$120,000
0X	MAINTENANCE	\$0
TOTAL		\$147,474

Table 19 Pricing information: 16 read threads, 0 write threads

	<i>Item</i>	<i>Price</i>
1X	AMAZON R6ID.8XLARGE	\$27,474.12
1X	GRAPHDB LICENSE (FOR 12 CPU CORES) (EE)	\$90,000
0X	MAINTENANCE	\$0
TOTAL		\$117,474

Table 20 Pricing information: 8 read threads, 4 write threads

	<i>Item</i>	<i>Price</i>
1X	AMAZON R6ID.8XLARGE	\$27,474.12
1X	GRAPHDB LICENSE (FOR 20 CPU CORES) (EE)	\$150,000
0X	MAINTENANCE	\$0
TOTAL		\$177,474

Table 21 Pricing information: 16 read threads, 4 write threads

	<i>Item</i>	<i>Price</i>
1X	AMAZON R6ID.8XLARGE	\$27,474.12
1X	GRAPHDB LICENSE (FOR 24 CPU CORES) (EE)	\$180,000
0X	MAINTENANCE	\$0
TOTAL		\$207,474

Table 22 Pricing information: 24 read threads, 0 write threads

10. Attachment's CheckList

CLAUSE 1	RDF Database load scripts if any
CLAUSE 2	RDF Database configuration of more details are available
CLAUSE 3	Operating System and Hardware Settings
CLAUSE 4	Benchmark test driver configuration and any modified query templates and (or) ontologies / reference datasets
CLAUSE 5	All of the log files generated by the benchmark driver

Table 23: Checklist