Semantic Publishing Benchmark 2.0.2
Full Disclosure Report
Pjotr Scholtze
2023-01-29

Table of Contents

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 3x r6id.8xlarge AWS servers in a cluster setup (another r6id.8xlarge AWS machine was used for running the benchmark).

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

No of read threads	0
No of write threads	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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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

3.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
0	11.9775

Table 9: Primary Metrics

3.2.2. 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	1396	11545	1432

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	318 ms	254 ms	1755 ms	11545
Update	392 ms	282 ms	1447 ms	1396
Delete	370 ms	256 ms	717 ms	1432

Table 12: Editorial Report

3.2.3. Raw results

```
Seconds : 1200
2022-12-07 11:34:33 (completed query mixes : 0)
        Editorial:
            4 agents

            11545 inserts (avg : 318    ms, min : 254    ms, max : 1755    ms)
            1396  updates (avg : 392    ms, min : 282    ms, max : 1447    ms)
```

```
            1432  deletes (avg : 370     ms, min : 256     ms, max : 717     ms)

            14373 operations (11545 CW Inserts (2 errors), 1396 CW Updates (0 errors), 1432 CW
Deletions (2 errors))
            11.9775 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    28m41.474s
user    1m14.469s
sys     0m3.693s
```

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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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)
77.9942	7.9983

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	93593	950	7670	978

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	58 ms	6 ms	397 ms	7799
Q2	4 ms	2 ms	189 ms	7799
Q3	63 ms	23 ms	1122 ms	7799
Q4	19 ms	5 ms	270 ms	7799
Q5	134 ms	7 ms	2525 ms	7800
Q6	50 ms	32 ms	204 ms	7800
Q7	16 ms	6 ms	188 ms	7801
Q8	109 ms	80 ms	448 ms	7799
Q9	78 ms	3 ms	420 ms	7799
Q10	142 ms	3 ms	478 ms	7799
Q11	41 ms	3 ms	1655 ms	7800
Q12	17 ms	8 ms	160 ms	7799

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	488 ms	260 ms	3327 ms	7670
Update	540 ms	289 ms	3309 ms	950
Delete	513 ms	276 ms	3331 ms	978

Table 12: Editorial Report

4.2.3. Raw results

```
Seconds : 1200
2022-12-08 10:42:14 (completed query mixes : 7796)
        Editorial:
                4 agents

                7670  inserts (avg : 488    ms, min : 260    ms, max : 3327   ms)
                950   updates (avg : 540    ms, min : 289    ms, max : 3309   ms)
```

```
            978   deletes (avg : 513     ms, min : 276     ms, max : 3331    ms)

            9598 operations (7670 CW Inserts (20 errors), 950 CW Updates (0 errors), 978 CW Deletions
(2 errors))

            7.9983 average operations per second

      Aggregation:
            8 agents

            7799  Q1   queries (avg : 58     ms, min : 6       ms, max : 397     ms, 0 errors)
            7799  Q2   queries (avg : 4      ms, min : 2       ms, max : 189     ms, 0 errors)
            7799  Q3   queries (avg : 63     ms, min : 23      ms, max : 1122    ms, 0 errors)
            7799  Q4   queries (avg : 19     ms, min : 5       ms, max : 270     ms, 0 errors)
            7800  Q5   queries (avg : 134    ms, min : 7       ms, max : 2525    ms, 0 errors)
            7800  Q6   queries (avg : 50     ms, min : 32      ms, max : 204     ms, 0 errors)
            7801  Q7   queries (avg : 16     ms, min : 6       ms, max : 188     ms, 0 errors)
            7799  Q8   queries (avg : 109    ms, min : 80      ms, max : 448     ms, 0 errors)
            7799  Q9   queries (avg : 78     ms, min : 3       ms, max : 420     ms, 0 errors)
            7799  Q10  queries (avg : 142    ms, min : 3       ms, max : 478     ms, 0 errors)
            7800  Q11  queries (avg : 41     ms, min : 3       ms, max : 1655    ms, 0 errors)
            7799  Q12  queries (avg : 17     ms, min : 8       ms, max : 160     ms, 0 errors)

            93593 total retrieval queries (0 errors)
            77.9942 average queries per second


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

real    28m16.615s
user    14m39.965s
sys     0m46.808s
```

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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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)
180.5783	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	216694	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	48 ms	5 ms	415 ms	18056
Q2	3 ms	2 ms	134 ms	18057
Q3	56 ms	21 ms	1479 ms	18058
Q4	16 ms	3 ms	354 ms	18058
Q5	122 ms	6 ms	2638 ms	18060
Q6	46 ms	31 ms	241 ms	18057
Q7	12 ms	4 ms	208 ms	18057
Q8	103 ms	82 ms	655 ms	18059
Q9	73 ms	3 ms	427 ms	18058
Q10	138 ms	3 ms	789 ms	18057
Q11	18 ms	3 ms	1511 ms	18060
Q12	14 ms	7 ms	175 ms	18057

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
2022-12-07 16:09:50 (completed query mixes : 18051)
        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

    18056 Q1   queries (avg : 48    ms, min : 5    ms, max : 415    ms, 0 errors)
    18057 Q2   queries (avg : 3     ms, min : 2    ms, max : 134    ms, 0 errors)
    18058 Q3   queries (avg : 56    ms, min : 21   ms, max : 1479   ms, 0 errors)
    18058 Q4   queries (avg : 16    ms, min : 3    ms, max : 354    ms, 0 errors)
    18060 Q5   queries (avg : 122   ms, min : 6    ms, max : 2638   ms, 0 errors)
    18057 Q6   queries (avg : 46    ms, min : 31   ms, max : 241    ms, 0 errors)
    18057 Q7   queries (avg : 12    ms, min : 4    ms, max : 208    ms, 0 errors)
    18059 Q8   queries (avg : 103   ms, min : 82   ms, max : 655    ms, 0 errors)
    18058 Q9   queries (avg : 73    ms, min : 3    ms, max : 427    ms, 0 errors)
    18057 Q10  queries (avg : 138   ms, min : 3    ms, max : 789    ms, 0 errors)
    18060 Q11  queries (avg : 18    ms, min : 3    ms, max : 1511   ms, 0 errors)
    18057 Q12  queries (avg : 14    ms, min : 7    ms, max : 175    ms, 0 errors)

    216694 total retrieval queries (0 errors)
    180.5783 average queries per second


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

real    28m31.502s
user    31m55.211s
sys     2m14.455s
```

6. Test case: 32 read threads, 0 write threads

6.1. Benchmark Test Driver

6.1.1. Basic test driver configuration details

No of read threads	32
No of write threads	0

Table 14: Basic Test Driver Configuration Details

6.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=32
queryTimeoutSeconds=300
warmUp=true
```

6.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=32
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
```

6.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=32
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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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)
304.7250	0

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	365670	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	58 ms	6 ms	859 ms	30470
Q2	4 ms	2 ms	826 ms	30471
Q3	67 ms	22 ms	2160 ms	30477
Q4	19 ms	4 ms	803 ms	30471
Q5	143 ms	6 ms	2751 ms	30475
Q6	55 ms	28 ms	871 ms	30471
Q7	14 ms	5 ms	785 ms	30475
Q8	123 ms	87 ms	1265 ms	30472
Q9	93 ms	3 ms	1069 ms	30473
Q10	168 ms	3 ms	1103 ms	30473
Q11	23 ms	3 ms	1883 ms	30472
Q12	17 ms	8 ms	635 ms	30470

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

6.2.3. Raw results

```
Seconds : 1200
2022-12-07 18:10:19 (completed query mixes : 30459)
        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:
        32 agents

        30470 Q1   queries (avg : 58    ms, min : 6      ms, max : 859    ms, 0 errors)
        30471 Q2   queries (avg : 4     ms, min : 2      ms, max : 826    ms, 1 errors)
        30477 Q3   queries (avg : 67    ms, min : 22     ms, max : 2160   ms, 0 errors)
        30471 Q4   queries (avg : 19    ms, min : 4      ms, max : 803    ms, 0 errors)
        30475 Q5   queries (avg : 143   ms, min : 6      ms, max : 2751   ms, 0 errors)
        30471 Q6   queries (avg : 55    ms, min : 28     ms, max : 871    ms, 1 errors)
        30475 Q7   queries (avg : 14    ms, min : 5      ms, max : 785    ms, 0 errors)
        30472 Q8   queries (avg : 123   ms, min : 87     ms, max : 1265   ms, 0 errors)
        30473 Q9   queries (avg : 93    ms, min : 3      ms, max : 1069   ms, 1 errors)
        30473 Q10  queries (avg : 168   ms, min : 3      ms, max : 1103   ms, 0 errors)
        30472 Q11  queries (avg : 23    ms, min : 3      ms, max : 1883   ms, 0 errors)
        30470 Q12  queries (avg : 17    ms, min : 8      ms, max : 635    ms, 0 errors)

        365670 total retrieval queries (3 errors)
        304.7250 average queries per second


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

real    28m45.784s
user    54m16.535s
sys     4m27.862s
```

7. Test case: 64 read threads, 0 write threads

7.1. Benchmark Test Driver

7.1.1. Basic test driver configuration details

No of read threads	64
No of write threads	0

Table 16: Basic Test Driver Configuration Details

7.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=64
queryTimeoutSeconds=300
warmUp=true
```

7.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=64
editorialAgents=0
benchmarkRunPeriodSeconds=1200
queryTimeoutSeconds=300
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
queriesPath=./data/sparql
runBenchmark=true
```

7.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=32
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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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)
408.6792	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	490415	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	95 ms	6 ms	1038 ms	40867
Q2	9 ms	2 ms	815 ms	40866
Q3	98 ms	24 ms	3359 ms	40869
Q4	31 ms	4 ms	1099 ms	40866
Q5	200 ms	6 ms	4074 ms	40870
Q6	77 ms	32 ms	1112 ms	40864
Q7	26 ms	4 ms	1036 ms	40866
Q8	179 ms	88 ms	2664 ms	40867
Q9	169 ms	3 ms	1596 ms	40867
Q10	230 ms	3 ms	1373 ms	40866
Q11	41 ms	3 ms	14427 ms	40872
Q12	31 ms	8 ms	911 ms	40875

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
2022-12-07 18:58:54 (completed query mixes : 40841)
        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:
        64 agents

        40867 Q1   queries (avg : 95    ms, min : 6     ms, max : 1038    ms, 1 errors)
        40866 Q2   queries (avg : 9     ms, min : 2     ms, max : 815     ms, 0 errors)
        40869 Q3   queries (avg : 98    ms, min : 24    ms, max : 3359    ms, 0 errors)
        40866 Q4   queries (avg : 31    ms, min : 4     ms, max : 1099    ms, 0 errors)
        40870 Q5   queries (avg : 200   ms, min : 6     ms, max : 4074    ms, 1 errors)
        40864 Q6   queries (avg : 77    ms, min : 32    ms, max : 1112    ms, 0 errors)
        40866 Q7   queries (avg : 26    ms, min : 4     ms, max : 1036    ms, 0 errors)
        40867 Q8   queries (avg : 179   ms, min : 88    ms, max : 2664    ms, 0 errors)
        40867 Q9   queries (avg : 169   ms, min : 3     ms, max : 1596    ms, 1 errors)
        40866 Q10  queries (avg : 230   ms, min : 3     ms, max : 1373    ms, 1 errors)
        40872 Q11  queries (avg : 41    ms, min : 3     ms, max : 14427   ms, 0 errors)
        40875 Q12  queries (avg : 31    ms, min : 8     ms, max : 911     ms, 0 errors)


        490415 total retrieval queries (4 errors)
        408.6792 average queries per second


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


real    29m53.565s
user    82m53.145s
sys     8m17.837s
```

8. 64 read threads, 4 write threads

8.1. Benchmark Test Driver

8.1.1. Basic test driver configuration details

No of read threads	64
No of write threads	4

Table 18: Basic Test Driver Configuration Details

8.1.2. Configuration Parameters for Driver Warmup

```
warmupPeriodSeconds=400
aggregationAgents=64
queryTimeoutSeconds=300
warmUp=true
```

8.1.3. Configuration Parameters for Driver Execution

```
aggregationAgents=64
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
```

8.1.4. Raw test.properties

```
#Benchmark Threads Configuration
aggregationAgents=64
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://52.91.73.231:7200/repositories/LDBC-SF5
endpointUpdateURL=http://52.91.73.231:7200/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
```

8.1.5. Test Driver Reference Data
No modifications made.

8.2. Performance Metrics

8.2.1. SPB Primary Metrics

Retrieval Query Rate (queries/second)	Update Rate (Operations per second)
372.5558	4.0375

Table 19: Primary Metrics

8.2.2. 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	447067	478	3867	500

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	105 ms	7 ms	1155 ms	37255
Q2	11 ms	2 ms	657 ms	37253
Q3	110 ms	30 ms	2976 ms	37256
Q4	36 ms	4 ms	796 ms	37256
Q5	223 ms	8 ms	4426 ms	37255
Q6	85 ms	33 ms	804 ms	37259
Q7	30 ms	5 ms	721 ms	37257
Q8	198 ms	89 ms	2958 ms	37252
Q9	174 ms	3 ms	1691 ms	37251
Q10	252 ms	3 ms	1391 ms	37257
Q11	44 ms	3 ms	13944 ms	37259
Q12	32 ms	7 ms	398 ms	37257

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	966 ms	302 ms	3772 ms	3867
Update	1048 ms	365 ms	2902 ms	478
Delete	1033 ms	336 ms	3386 ms	500

Table 12: Editorial Report

8.2.3. Raw results

```
Seconds : 1200
2022-12-22 09:44:30 (completed query mixes : 37224)
        Editorial:
            4 agents

            3867  inserts (avg : 966    ms, min : 302    ms, max : 3772    ms)
            478   updates (avg : 1048   ms, min : 365    ms, max : 2902    ms)
```

```
            500   deletes (avg : 1033    ms, min : 336     ms, max : 3386    ms)

            4845 operations (3867 CW Inserts (14 errors), 478 CW Updates (0 errors), 500 CW Deletions
(1 errors))

            4.0375 average operations per second

       Aggregation:
            64 agents

            37255 Q1   queries (avg : 105    ms, min : 7      ms, max : 1155    ms, 0 errors)
            37253 Q2   queries (avg : 11     ms, min : 2      ms, max : 657     ms, 0 errors)
            37256 Q3   queries (avg : 110    ms, min : 30     ms, max : 2976    ms, 0 errors)
            37256 Q4   queries (avg : 36     ms, min : 4      ms, max : 796     ms, 0 errors)
            37255 Q5   queries (avg : 223    ms, min : 8      ms, max : 4426    ms, 0 errors)
            37259 Q6   queries (avg : 85     ms, min : 33     ms, max : 804     ms, 0 errors)
            37257 Q7   queries (avg : 30     ms, min : 5      ms, max : 721     ms, 0 errors)
            37252 Q8   queries (avg : 198    ms, min : 89     ms, max : 2958    ms, 0 errors)
            37251 Q9   queries (avg : 174    ms, min : 3      ms, max : 1691    ms, 0 errors)
            37257 Q10  queries (avg : 252    ms, min : 3      ms, max : 1391    ms, 0 errors)
            37259 Q11  queries (avg : 44     ms, min : 3      ms, max : 13944   ms, 0 errors)
            37257 Q12  queries (avg : 32     ms, min : 7      ms, max : 398     ms, 0 errors)

            447067 total retrieval queries (0 errors)
            372.5558 average queries per second

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

real    28m55.272s
user    81m12.265s
sys     7m48.296s
```

9. Recovery

Execution Duration	N/A
Time to Recover	N/A

Table 14: Recovery

10. 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 three r6id.8xlarge Amazon machines, costing a fixed $82,422.36 for all scenarios for three years (3x $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:

```
CPU_cores_license = number_of_read_threads + (number_of_machines * number_of_write_threads)
```

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

	Item	Price
3X	AMAZON R6ID.8XLARGE	$82,422.36
1X	GRAPHDB LICENSE (FOR 12 CPU CORES) (EE)	$90,000
0X	MAINTENANCE	$0
TOTAL		$172,422

Table 20 Pricing information: 0 read threads, 4 write threads

	Item	Price
3X	AMAZON R6ID.8XLARGE	$82,422.36
1X	GRAPHDB LICENSE (FOR 16 CPU CORES) (EE)	$120,000
0X	MAINTENANCE	$0
TOTAL		$202,422

Table 21 Pricing information: 16 read threads, 0 write threads

	Item	Price
3X	AMAZON R6ID.8XLARGE	$82,422.36
1X	GRAPHDB LICENSE (FOR 32 CPU CORES) (EE)	$240,000
0X	MAINTENANCE	$0
TOTAL		$322,422

Table 22 Pricing information: 32 read threads, 0 write threads

	Item	Price
3X	AMAZON R6ID.8XLARGE	$82,422.36
1X	GRAPHDB LICENSE (FOR 64 CPU CORES) (EE)	$480,000
0X	MAINTENANCE	$0
TOTAL		$562,422

Table 23 Pricing information: 64 read threads, 0 write threads

	Item	Price
3X	AMAZON R6ID.8XLARGE	$82,422.36
1X	GRAPHDB LICENSE (FOR 20 CPU CORES) (EE)	$150,000
0X	MAINTENANCE	$0
TOTAL		$232,422

Table 24 Pricing information: 8 read threads, 4 write threads

	Item	Price
3X	AMAZON R6ID.8XLARGE	$82,422.36
1X	GRAPHDB LICENSE (FOR 76 CPU CORES) (EE)	$570,000
0X	MAINTENANCE	$0
TOTAL		$652,422

Table 25 Pricing information: 64 read threads, 4 write threads

11. 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 26: Checklist