

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
3.2.1. SPB Primary Metrics.....	14
3.2.2. Query Execution Report.....	14
3.2.3. 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 3 (256M 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-3 256M
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 : 8859807, 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=232000000
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-SF3
endpointUpdateURL=http://52.91.73.231:7200/repositories/LDBC-SF3/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)	10:00:32
End Time (in ms)	12:59:54
Generation Time	10 762 417 ms (=2h 59m 22s)

Table 6: Data Generation Times

2.3. Bulk Loading

Data generation and bulk loading went together.

Start Time (in ms)	10:00:32
End Time (in ms)	12:59:54
Generation Time	10 762 417 ms (=2h 59m 22s)

Table 7: Bulk Loading Time

2.4. Additional setup

Lucene connector time	6m38s
Create a geospatial index	7.7s

2.5. Raw results

```
Loaded 5123 files with Creative Works in 10762417 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 : 8859807,
  dbpedia locations : 268862, geonames locations : 1285124)
.....
```

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

```
real 190m40.377s
user 21m54.144s
sys 1m41.859s
```

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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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=232000000 # (SF3
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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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	37.7325

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	4490	36251	4538

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	103 ms	84 ms	727 ms	36251
Update	115 ms	94 ms	658 ms	4490
Delete	108 ms	87 ms	265 ms	4538

Table 12: Editorial Report

3.2.3. Raw results

Seconds : 1200

2022-12-02 14:58:04 (completed query mixes : 0)

Editorial:

4 agents

36251 inserts (avg : 103 ms, min : 84 ms, max : 727 ms)

4490 updates (avg : 115 ms, min : 94 ms, max : 658 ms)

4538 deletes (avg : 108 ms, min : 87 ms, max : 265 ms)

45279 operations (36251 CW Inserts (0 errors), 4490 CW Updates (0 errors), 4538 CW Deletions (0 errors))

37.7325 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 27m32.138s
user 1m28.087s
sys 0m6.210s

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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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=232000000 # (SF3
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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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)
216.5542	31.1383

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	259865	3703	29959	3704

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	20 ms	2 ms	427 ms	21656
Q2	2 ms	1 ms	94 ms	21655
Q3	15 ms	5 ms	1277 ms	21655
Q4	6 ms	1 ms	254 ms	21655
Q5	32 ms	2 ms	597 ms	21657
Q6	43 ms	25 ms	332 ms	21655
Q7	5 ms	2 ms	118 ms	21655
Q8	28 ms	22 ms	442 ms	21657
Q9	72 ms	1 ms	601 ms	21655
Q10	38 ms	1 ms	454 ms	21654
Q11	7 ms	2 ms	766 ms	21655
Q12	6 ms	2 ms	123 ms	21656

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	125 ms	89 ms	638 ms	29959
Update	139 ms	105 ms	566 ms	3703
Delete	129 ms	86 ms	527 ms	3704

Table 12: Editorial Report

4.2.3. Raw results

Seconds : 1200

2022-12-03 01:00:58 (completed query mixes : 21652)

Editorial:

4 agents

29959 inserts (avg : 125 ms, min : 89 ms, max : 638 ms)

3703 updates (avg : 139 ms, min : 105 ms, max : 566 ms)

3704 deletes (avg : 129 ms, min : 86 ms, max : 527 ms)

37366 operations (29959 CW Inserts (0 errors), 3703 CW Updates (0 errors), 3704 CW Deletions (0 errors))

31.1383 average operations per second

Aggregation:

8 agents

21656 Q1	queries	(avg : 20 ms, min : 2 ms, max : 427 ms, 0 errors)
21655 Q2	queries	(avg : 2 ms, min : 1 ms, max : 94 ms, 0 errors)
21655 Q3	queries	(avg : 15 ms, min : 5 ms, max : 1277 ms, 0 errors)
21655 Q4	queries	(avg : 6 ms, min : 1 ms, max : 254 ms, 0 errors)
21657 Q5	queries	(avg : 32 ms, min : 2 ms, max : 597 ms, 0 errors)
21655 Q6	queries	(avg : 43 ms, min : 25 ms, max : 332 ms, 0 errors)
21655 Q7	queries	(avg : 5 ms, min : 2 ms, max : 118 ms, 0 errors)
21657 Q8	queries	(avg : 28 ms, min : 22 ms, max : 442 ms, 0 errors)
21655 Q9	queries	(avg : 72 ms, min : 1 ms, max : 601 ms, 0 errors)
21654 Q10	queries	(avg : 38 ms, min : 1 ms, max : 454 ms, 0 errors)
21655 Q11	queries	(avg : 7 ms, min : 2 ms, max : 766 ms, 0 errors)
21656 Q12	queries	(avg : 6 ms, min : 2 ms, max : 123 ms, 0 errors)

259865 total retrieval queries (0 errors)

216.5542 average queries per second

Stopping the benchmark...

END OF RUN, all agents shut down...

real 27m37.114s
user 20m6.485s
sys 1m29.789s

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=0
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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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=232000000 # (SF3
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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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)
384.2900	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	461148	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	23 ms	2 ms	361 ms	38428
Q2	2 ms	1 ms	69 ms	38429
Q3	15 ms	4 ms	1763 ms	38430
Q4	6 ms	1 ms	385 ms	38429
Q5	34 ms	2 ms	714 ms	38429
Q6	42 ms	25 ms	119 ms	38430
Q7	4 ms	2 ms	76 ms	38433
Q8	31 ms	23 ms	254 ms	38428
Q9	92 ms	1 ms	454 ms	38428
Q10	44 ms	1 ms	177 ms	38429
Q11	8 ms	2 ms	1833 ms	38427
Q12	5 ms	2 ms	100 ms	38428

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

2022-11-30 14:41:10 (completed query mixes : 38420)

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

38428	Q1	queries	(avg : 23	ms, min : 2	ms, max : 361	ms, 0 errors)
38429	Q2	queries	(avg : 2	ms, min : 1	ms, max : 69	ms, 0 errors)
38430	Q3	queries	(avg : 15	ms, min : 4	ms, max : 1763	ms, 0 errors)
38429	Q4	queries	(avg : 6	ms, min : 1	ms, max : 385	ms, 0 errors)
38429	Q5	queries	(avg : 34	ms, min : 2	ms, max : 714	ms, 0 errors)
38430	Q6	queries	(avg : 42	ms, min : 25	ms, max : 119	ms, 0 errors)
38433	Q7	queries	(avg : 4	ms, min : 2	ms, max : 76	ms, 0 errors)
38428	Q8	queries	(avg : 31	ms, min : 23	ms, max : 254	ms, 0 errors)
38428	Q9	queries	(avg : 92	ms, min : 1	ms, max : 454	ms, 0 errors)
38429	Q10	queries	(avg : 44	ms, min : 1	ms, max : 177	ms, 0 errors)
38427	Q11	queries	(avg : 8	ms, min : 2	ms, max : 1833	ms, 0 errors)
38428	Q12	queries	(avg : 5	ms, min : 2	ms, max : 100	ms, 0 errors)

461148 total retrieval queries (0 errors)
384.2900 average queries per second

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

real 27m51.643s
user 32m55.250s
sys 2m19.312s

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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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=232000000 # (SF3
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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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)
335.4758	25.6592

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	402571	3019	24704	3068

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	26 ms	2 ms	436 ms	33546
Q2	3 ms	1 ms	89 ms	33548
Q3	18 ms	5 ms	1897 ms	33550
Q4	7 ms	1 ms	529 ms	33546
Q5	40 ms	2 ms	1128 ms	33548
Q6	46 ms	25 ms	558 ms	33551
Q7	5 ms	2 ms	106 ms	33546
Q8	32 ms	23 ms	580 ms	33548
Q9	117 ms	2 ms	5494 ms	33546
Q10	48 ms	1 ms	632 ms	33547
Q11	9 ms	2 ms	2276 ms	33545
Q12	7 ms	2 ms	113 ms	33550

Table 11: Aggregation Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Insert	151 ms	88 ms	848 ms	24704
Update	169 ms	121 ms	301 ms	3019
Delete	158 ms	101 ms	415 ms	3068

Table 12: Editorial Report

6.2.3. Raw results

Seconds : 1200

2022-12-21 22:27:58 (completed query mixes : 33541)

Editorial:

4 agents

24704 inserts (avg : 151 ms, min : 88 ms, max : 848 ms)

3019 updates (avg : 169 ms, min : 121 ms, max : 301 ms)

3068 deletes (avg : 158 ms, min : 101 ms, max : 415 ms)

30791 operations (24704 CW Inserts (0 errors), 3019 CW Updates (0 errors), 3068 CW Deletions (0 errors))

25.6592 average operations per second

Aggregation:

16 agents

33546 Q1	queries	(avg : 26	ms, min : 2	ms, max : 436	ms, 0 errors)
33548 Q2	queries	(avg : 3	ms, min : 1	ms, max : 89	ms, 0 errors)
33550 Q3	queries	(avg : 18	ms, min : 5	ms, max : 1897	ms, 0 errors)
33546 Q4	queries	(avg : 7	ms, min : 1	ms, max : 529	ms, 0 errors)
33548 Q5	queries	(avg : 40	ms, min : 2	ms, max : 1128	ms, 0 errors)
33551 Q6	queries	(avg : 46	ms, min : 25	ms, max : 558	ms, 0 errors)
33546 Q7	queries	(avg : 5	ms, min : 2	ms, max : 106	ms, 0 errors)
33548 Q8	queries	(avg : 32	ms, min : 23	ms, max : 580	ms, 0 errors)
33546 Q9	queries	(avg : 117	ms, min : 2	ms, max : 5494	ms, 0 errors)
33547 Q10	queries	(avg : 48	ms, min : 1	ms, max : 632	ms, 0 errors)
33545 Q11	queries	(avg : 9	ms, min : 2	ms, max : 2276	ms, 0 errors)
33550 Q12	queries	(avg : 7	ms, min : 2	ms, max : 113	ms, 0 errors)

402571 total retrieval queries (0 errors)

335.4758 average queries per second

Stopping the benchmark...

END OF RUN, all agents shut down...

real 27m51.293s
user 33m19.509s
sys 2m14.131s

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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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=232000000 # (SF3
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-SF3
endpointUpdateURL=http://54.92.203.255:7200/repositories/LDBC-SF3/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)
413.1558	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	495787	0	0	0

Table 10: Query Execution Report

Query	Arithmetic Mean of Execution Time	Minimum Execution Time	Maximum Execution Time	Count of Executions
Q1	28 ms	2 ms	711 ms	41317
Q2	3 ms	1 ms	69 ms	41316
Q3	19 ms	5 ms	2766 ms	41316
Q4	8 ms	1 ms	610 ms	41316
Q5	41 ms	2 ms	829 ms	41318
Q6	48 ms	28 ms	148 ms	41317
Q7	6 ms	3 ms	101 ms	41315
Q8	35 ms	24 ms	392 ms	41318
Q9	167 ms	2 ms	882 ms	41312
Q10	52 ms	1 ms	197 ms	41315
Q11	11 ms	2 ms	2449 ms	41313
Q12	7 ms	2 ms	74 ms	41314

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 21:13:09 (completed query mixes : 41304)

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

41317	Q1	queries	(avg : 28	ms, min : 2	ms, max : 711	ms, 0 errors)
41316	Q2	queries	(avg : 3	ms, min : 1	ms, max : 69	ms, 0 errors)
41316	Q3	queries	(avg : 19	ms, min : 5	ms, max : 2766	ms, 0 errors)
41316	Q4	queries	(avg : 8	ms, min : 1	ms, max : 610	ms, 0 errors)
41318	Q5	queries	(avg : 41	ms, min : 2	ms, max : 829	ms, 0 errors)
41317	Q6	queries	(avg : 48	ms, min : 28	ms, max : 148	ms, 0 errors)
41315	Q7	queries	(avg : 6	ms, min : 3	ms, max : 101	ms, 0 errors)
41318	Q8	queries	(avg : 35	ms, min : 24	ms, max : 392	ms, 0 errors)
41312	Q9	queries	(avg : 167	ms, min : 2	ms, max : 882	ms, 0 errors)
41315	Q10	queries	(avg : 52	ms, min : 1	ms, max : 197	ms, 0 errors)
41313	Q11	queries	(avg : 11	ms, min : 2	ms, max : 2449	ms, 0 errors)
41314	Q12	queries	(avg : 7	ms, min : 2	ms, max : 74	ms, 0 errors)

495787 total retrieval queries (0 errors)
413.1558 average queries per second

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

real 28m21.573s
user 40m53.979s
sys 2m48.738s

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