The graph & RDF benchmark reference

Full Disclosure Report of the LDBC Social Network Benchmark

An Implementation of the LDBC Social Network Benchmark's Interactive Workload over Galaxybase

May 18, 2022

General Terms

Executive Summary

Galaxybase is a proprietary graph database product developed by CreateLink. This document describes an implementation of the LDBC Social Network Benchmark's Interactive workload in Galaxybase. The implementation uses user-defined procedures written using Java 11 methods, which are compiled and loaded into the database as user-defined procedures. Thus, the benchmark implementation uses imperative queries with manually defined query evaluation plans over the data to compute the queries specified in the workload. The data schema follows the property graph data model with indices over node and edge identifiers and over properties selected by the user. This current benchmark implementation employs four additional precomputed properties to improve the operation throughput of the system, which are maintained individually during each update. The system under test and the driver communicates using remote procedure calls (RPC) over local sockets.

Declaration of Audit Success

This report contains an audited LDBC benchmark run. The results have been gathered by an independent and impartial auditor who has validated the implementation of the queries, successfully run the ACID tests associated with the claimed isolation level (serializable), and verified the overall system's configuration conformance to the description of the benchmark and its strict requirements.

DocuSigned by:
976A26F8CA5A40A...

Dr. Marton Bur
(Auditor)

5/19/2022
Date

DocuSigned by:
Gabor Szarnyas
FA33C2C8CB944A5...

Dr. Gabor Szarnyas
(Head of LDBC SNB Task Force)

5/20/2022
Date

DocuSigned by:
Yan Zhou
BBA1A13B17AE422...

Dr. Yan Zhou
(Test Sponsor Representative)

5/19/2022
Date

Table of Contents

1 System Description and Pricing Summary

1.1 Details of machines driving and running the workload

1.1.1 Machine overview

The details below were obtained from the Amazon Web Services console.

Table 1.1: Machine Type and Location

Cloud provider	Amazon Web Services
Machine region	N. Virginia
Common name of the item	r5d.12xlarge
Operating system	18.04.1-Ubuntu

1.1.2 CPU details

The details below were obtained using the command `cat /proc/cpuinfo` (Listing A.1) issued from the machine instance and the datasheet of the used CPU type.

Table 1.2: CPU details summary

Type	Intel® Xeon® Platinum 8259CL CPU @ 2.5GHz
Total number	1
Cores per CPU	24
Threads per CPU	48
CPU clock frequency	2.5GHz
Total cache size per CPU	L1 cache: 1.5MB L2 cache: 24MB L3 cache: 33MB

1.1.3 Memory details

The total size of the memory installed is 372GB, and this information was obtained using the `cat /proc/meminfo` command issued from the virtual machine instance. The type and frequency of the memory installed in the virtual machine was not obtainable.

1.1.4 Disk and storage details

Disk controller or motherboard type was not obtainable from the virtual machine instance. The storage consists of 2 x 900GB NVMe SSD in RAID0 configuration, formatted with ext4 filesystem. The storage size and type is from the Amazon Web Services website `https://aws.amazon.com/ec2/instance-types/r5/` (accessed: May 6, 2022).

The 4KB QD1 write performance was measured with the `fio` command and the output (Listing A.2) showed an average of 6421 IOPS.

1.1.5 Network details

The presented benchmark run only used a single machine, thus network details are not included here.

1.1.6 Machine pricing

The system pricing summary is included in the table below (prices are in USD).

Table 1.3: Pricing summary

Item	**Price**
r5d.12xlarge reserved instance machine in AWS (standard 3-year term)	$52,210
Permanent Galaxybase license	$155,200
Maintenance fee (3 years)	$55,872

1.1.7 System availability

The Galaxybase software used in the audited runs was made available on January 21, 2022. This version was deployed to the machine described previously in this section. The features of the system are summarized in the table below.

Table 1.4: System availability summary

System version	v3.3.0
System APIs	Java traversal API, Cypher
API used for implementing performance benchmark queries	Java traversal API
API used for additional querying during the audit (e.g., durability tests)	Cypher

2 Dataset Generation

2.1 General information

The data generation settings of the LDBC Datagen are described below.

Table 2.1: Datagen settings summary

Datagen version	v0.3.3
Output format	CsvComposite serializer
Scale factors	10, 30, 100, and 300

Scale factor 10 is used for query results validation only, while the other ones were used for performance measurements.

2.2 Datagen configurations

The Datagen configuration for SF10 is shown in Listing 2.1. The configurations for SF30, SF100 and SF300 are shown in Listings A.3–A.5.

Listing 2.1: Contents of `params-sf10.ini` used for scale factor 10

```
ldbc.snb.datagen.generator.scaleFactor:snb.interactive.10
ldbc.snb.datagen.serializer.numUpdatePartitions:32
ldbc.snb.datagen.serializer.dynamicActivitySerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.activity.CsvCompositeDynamicActivitySerializer
ldbc.snb.datagen.serializer.dynamicPersonSerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.person.CsvCompositeDynamicPersonSerializer
ldbc.snb.datagen.serializer.staticSerializer:ldbc.snb.datagen.serializer.snb.csv.staticserializer.CsvCompositeStaticSerializer
```

2.3 Data loading and data schema

Some of the output files produced by the Datagen is converted to a custom (i.e., vendor-specific) CSV representation which can be loaded into the database (see the attached `datagen.sh` script). The loading process takes a configuration file `mapping-sf{10,30,100,300}.json` which defines the files to process along with what properties are represented in the values in each column. See an excerpt for the mapping file for scale factor 30 in Listing 2.2 and the complete mapping file for scale factor 30 in Listing A.6. The data type definitions are done separately in `schema-sf{10,30,100,300}.json`, Listing 2.3 shows an excerpt for scale factor 30. While this FDR document shows example configuration files for scale factor 30, these files are similarly defined for other scale factors and are available in the attachments.

Listing 2.2: Excerpt from `mapping-sf30.json`, describing the data mapping

```
{
  "source": {
    "sourceType": "CSV",
    "encoding": "utf-8",
    "filePath": "data/sf30/post_0_0.csv",
    "delimiter": "|",
    "enclosingChar": "\"",
    "hasHeader": true
  },
  "vertexes": [
```

```
    {
      "type": "Post",
      "primaryId": "id",
      "pkColumnIndex": 0,
      "hasDuplicatePrimaryKey": false,
      "property": [
        {
          "columnIndex": 1,
          "alias": "imageFile"
        },
        {
          "columnIndex": 2,
          "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
          "alias": "creationDate"
        },
        {
          "columnIndex": 3,
          "alias": "locationIP"
        },
        {
          "columnIndex": 4,
          "alias": "browserUsed"
        },
        {
          "columnIndex": 5,
          "alias": "language"
        },
        {
          "columnIndex": 6,
          "alias": "content"
        },
        {
          "columnIndex": 7,
          "alias": "length"
        }
      ]
    }
  ]
}
```

Listing 2.3: Excerpt from `schema-sf30.json`, describing the data schema

```
{
  "version": 0,
  "graphName": "sf30",
  "vertexes": {
    "Post": {
      "type": "Post",
      "primaryId": "id",
      "property": [
    {
          "name": "creationDate",
          "type": "DATETIME"
        },
    {
          "name": "locationIP",
          "type": "STRING"
        },
```

```
        {
          "name": "browserUsed",
          "type": "STRING"
        },
        {
          "name": "content",
          "type": "STRING"
        },
        {
          "name": "length",
          "type": "INT"
        },
        {
          "name": "language",
          "type": "STRING"
        },
        {
          "name": "imageFile",
          "type": "STRING"
        }
      ]
    },
    ...
  },
  "edges": {
    ...
  }
}
```

Data loading times are shown for each scale factor in the table below. Values were measured using the difference between the time values returned by the Java built-in `System.currentTimeMillis()`. The column **CSV loading time** shows how long it took to create a graph from the input CSV files and also to build and index over *id* properties, but they do not include CSV conversion times (the CSV conversion was executed outside the measurements and was not timed). The column **Data preprocessing time** shows how much time it took to create additional precomputed properties (similar to materialized views) *personKnowsPerson.weight*, *forumHasMemberPerson.count*, *postHasCreatorPerson.creationDate* and *commentHasCreatorPerson.creationDate*. The column **Total time** contains the sum of the CSV loading and data preprocessing times.

Table 2.2: Data loading times

Scale factor	CSV loading time (s)	Data preprocessing time (s)	Total time (s)
30	1 679.45	161.76	1 841.21
100	5 335.88	573.43	5 909.31
300	16 203.40	1 666.19	17 869.59

3 Test Driver Details

3.1 Driver implementation

A test driver adaptation for the SUT was provided by the test sponsor and is available as part of the attachment package. The archive created from the version of the driver used for the audited run is included in the attachments of this report.

The SUT-specific test driver class `com.ldbc.impls.workloads.ldbc.snb.galaxybase.interactive.GalaxybaseInteractiveDb` (via multiple intermediate classes) extends the class `com.ldbc.driver.Db` provided in the LDBC SNB Interactive driver package. Internally, the `GalaxybaseInteractiveDb` relies on remote procedure calls (RPC) using local sockets to communicate with the SUT.

3.2 Benchmark configuration of driver

The driver applied time compression ratio values of

- TCR=0.0038 for scale factor 30,
- TCR=0.0163 for scale factor 100, and
- TCR=0.053 for scale factor 300.

The complete configuration files for the different scale factors are shown in Listings A.7–A.9, and are also included in the attached supplementary materials.

4 Performance Metrics

The performance metrics reported here show benchmark runs with scale factors 30, 100 and 300. The performance summary tables below highlight key performance characteristics.

Table 4.1: Summary of results for scale factor 30

Benchmark duration	Benchmark operations	Throughput	Query on-time compliance
2h 14m 35.740s	74 990 196	9 285.86 $\frac{\text{operations}}{\text{second}}$	100.00%

Table 4.2: Summary of results for scale factor 100

Benchmark duration	Benchmark operations	Throughput	Query on-time compliance
2h 05m 26.944s	63 988 107	8 501.21 $\frac{\text{operations}}{\text{second}}$	100.00%

Table 4.3: Summary of results for scale factor 300

Benchmark duration	Benchmark operations	Throughput	Query on-time compliance
2h 07m 24.645s	63 989 669	8 370.52 $\frac{\text{operations}}{\text{second}}$	100.00%

Table 4.4: Detailed performance benchmark results for scale factor 30 in microseconds

Query	Total count	Min.	Max.	Mean	P_{50}	P_{90}	P_{95}	P_{99}
Complex1	523 933	465	346 384	8 681.31	5 508	6 830	8 034	125 660
Complex2	368 170	53	174 056	1 930.08	1 484	2 409	2 822	6 496
Complex3	128 512	11 474	281 888	20 244.07	18 475	24 585	26 700	62 124
Complex4	378 397	58	229 528	2 853.48	2 316	3 824	4 389	8 970
Complex5	189 198	28 225	961 792	337 757.53	320 480	521 600	576 768	677 856
Complex6	43 109	234	145 424	3 047.59	3 212	5 481	6 125	11 987
Complex7	283 797	39	220 432	464.26	112	262	337	4 095
Complex8	1 513 588	1 570	241 512	4 273.26	3 136	6 924	8 412	13 817
Complex9	35 475	9 432	527 840	165 935.19	154 592	239 440	268 720	335 520
Complex10	368 170	1 626	280 064	38 679.43	37 138	45 850	49 478	97 048
Complex11	681 114	62	231 824	924.78	561	813	938	4 747
Complex12	309 598	127	291 392	23 991.27	22 355	31 612	35 428	69 392
Complex13	716 963	48	221 800	743.21	328	1 031	1 194	4 459
Complex14	278 006	85	291 328	3 791.86	1 061	1 957	29 770	51 572
Short1	7 302 529	31	137 800	78.90	59	79	87	188
Short2	7 302 529	29	162 496	269.85	208	402	606	1 119
Short3	7 302 529	33	157 000	562.04	306	964	2 517	4 019
Short4	7 302 328	33	200 720	95.75	73	95	105	218
Short5	7 302 328	31	148 488	83.90	69	90	98	164
Short6	7 302 328	32	200 944	87.70	73	95	103	163
Short7	7 302 328	31	138 872	114.99	94	154	177	243
Update1	4 197	137	121 984	904.11	285	498	2 214	8 328
Update2	2 562 611	39	184 808	207.85	79	104	117	1 059
Update3	3 724 097	39	185 376	206.87	79	104	117	1 076
Update4	71 805	82	141 104	262.83	137	180	243	1 195
Update5	7 724 095	43	218 200	230.37	100	148	188	1 080
Update6	939 752	70	177 536	239.61	155	220	260	753
Update7	2 755 361	81	412 048	295.88	156	203	230	1 513
Update8	273 349	58	151 272	1 244.82	542	2 598	5 052	9 410

Table 4.5: Detailed performance benchmark results for scale factor 100 in microseconds

Query	Total count	Min.	Max.	Mean	P_{50}	P_{90}	P_{95}	P_{99}
Complex1	369 993	44	611 200	23 103.42	7 037	9 801	229 632	274 704
Complex2	259 996	42	303 696	1 640.68	1 188	1 750	2 096	6 405
Complex3	78 210	31 114	459 920	54 658.66	50 944	66 132	73 916	140 016
Complex4	267 217	57	312 752	2 906.34	2 262	3 982	4 702	9 809
Complex5	123 331	2 440	725 664	230 111.66	225 288	321 104	361 584	435 104
Complex6	22 165	60	179 504	7 541.73	7 990	15 299	17 312	31 876
Complex7	253 153	38	246 344	521.15	103	255	362	4 181
Complex8	1 923 965	41	313 504	538.28	132	213	286	4 220
Complex9	18 254	50	461 472	151 057.30	144 912	204 512	249 152	312 832
Complex10	240 496	47	462 048	50 097.52	48 152	59 880	66 512	126 256
Complex11	437 265	54	317 216	1 119.87	705	922	1 094	5 342
Complex12	218 632	50	368 784	30 224.56	28 266	40 128	46 346	90 180
Complex13	506 306	44	309 936	1 091.08	617	1 259	1 452	5 601
Complex14	196 323	43	350 768	8 055.02	1 332	23 917	28 993	54 058
Sort1	6 168 974	30	178 408	93.69	58	81	95	401
Sort2	6 168 974	29	148 504	235.34	179	317	439	837
Sort3	6 168 974	31	168 744	675.01	358	1 256	2 985	4 743
Sort4	6 169 406	33	195 504	110.18	71	97	113	443
Sort5	6 169 406	31	134 752	95.65	67	91	103	302
Sort6	6 169 406	32	155 128	98.49	71	95	108	295
Sort7	6 169 406	31	160 032	130.90	95	163	194	366
Update1	2 471	150	176 888	1 180.09	307	1 435	2 912	12 181
Update2	2 172 995	39	241 384	232.43	79	107	141	2 063
Update3	4 223 704	40	286 592	231.89	79	107	143	2 203
Update4	41 796	85	236 696	320.73	143	204	305	2 285
Update5	4 927 042	44	250 704	254.14	99	154	211	2 119
Update6	738 684	79	210 312	305.89	166	230	280	1 645
Update7	3 590 745	75	247 552	318.91	157	208	255	2 087
Update8	190 818	60	327 680	1 336.57	531	2 896	5 380	10 755

Table 4.6: Detailed performance benchmark results for scale factor 300 in microseconds

Query	Total count	Min.	Max.	Mean	P_{50}	P_{90}	P_{95}	P_{99}
Complex1	326 326	44	907 136	14 075.74	8 056	10 017	10 990	396 976
Complex2	229 311	44	312 144	1 277.65	888	1 455	1 759	6 927
Complex3	59 750	93 576	546 624	156 157.10	152 040	194 320	209 920	259 440
Complex4	235 679	45	344 880	3 354.60	2 701	4 820	5 492	13 930
Complex5	101 006	2 433	789 728	280 544.12	270 032	409 616	446 784	529 568
Complex6	14 629	60	158 816	18 781.95	4 879	42 130	46 532	77 072
Complex7	265 140	37	325 968	436.86	100	243	312	4 513
Complex8	2 828 157	41	367 152	446.34	128	191	220	4 739
Complex9	12 035	53	465 024	148 705.99	131 728	222 448	265 152	319 952
Complex10	192 829	45	400 528	60 427.09	58 990	78 028	86 088	127 720
Complex11	353 520	76	346 656	1 187.11	851	1 048	1 166	6 582
Complex12	192 829	42	359 680	32 632.02	30 925	46 922	53 576	88 640
Complex13	446 551	49	344 592	1 507.92	1 283	1 956	2 252	7 017
Complex14	173 152	48	384 960	10 563.06	2 058	24 738	30 599	56 864
Short1	6 817 920	31	116 340	75.85	58	78	86	154
Short2	6 817 920	30	314 576	186.69	162	237	284	437
Short3	6 817 920	31	310 048	767.01	432	1 407	3 250	5 338
Short4	6 816 037	34	312 896	94.71	74	98	108	181
Short5	6 816 037	30	325 200	83.18	70	93	101	153
Short6	6 816 037	32	144 512	86.05	73	97	105	153
Short7	6 816 037	30	120 028	118.22	98	165	200	274
Update1	1 924	142	275 728	898.52	291	480	640	9 501
Update2	1 450 882	41	311 536	252.13	82	107	120	1 242
Update3	2 235 838	42	333 632	252.16	82	107	119	1 191
Update4	32 817	92	122 848	312.20	141	183	243	1 038
Update5	4 396 694	44	354 352	272.18	100	150	190	1 131
Update6	590 149	88	345 424	347.37	171	263	413	1 681
Update7	1 960 498	84	333 792	349.77	158	209	241	2 226
Update8	172 045	62	118 504	1 294.21	447	2 805	4 936	12 065

5 Validation of the Results

The scale factor 10 data set was used for validating the correctness of the implementation over the SUT. The validation data set was created with the SNB Interactive reference implementation over Neo4j, running the Community Edition of version 4.4.1. The system with the driver configuration shown in Listing A.10 successfully returned the expected result sets for the queries of the benchmark.

6 ACID Compliance

6.1 Transaction isolation level

The benchmark was executed using the *serializable* isolation level setting of the SUT, which is more strict than the *read committed* isolation level minimally required by the SNB Interactive specification.

6.2 SNB Interactive ACID test results

The ACID test implementations were reviewed to conform to the ACID test specifications, with all specified test cases implemented. Furthermore, test execution was successful, no atomicity and isolation test failed with serializable isolation level transaction settings.

6.3 Recovery and durability

6.3.1 Recovery

Durability tests were using the regular benchmark workload with scale factor 30, and the server machine was shut down using the command `sleep 2h && sudo reboot`. This command was issued right after the start of the benchmark run. The database server process was manually started again after the crash and it was ready in 2 minutes and 43 seconds, which was not different from a regular startup time for the server.

6.3.2 Durability

From the driver log, the last update operations before the crash were obtained using the commands below.

```
$ grep LdbcUpdate1 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate1AddPerson|1651780041734|1651780041734|472|0|1349400557708
$ grep LdbcUpdate2 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate2AddPostLike|1651780043007|1651780043007|79|0|1349400892580
$ grep LdbcUpdate3 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate3AddCommentLike|1651780043007|1651780043007|57|0|1349400892474
$ grep LdbcUpdate4 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate4AddForum|1651780042988|1651780042988|137|0|1349400887474
$ grep LdbcUpdate5 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate5AddForumMembership|1651780043005|1651780043005|95|0|1349400892010
$ grep LdbcUpdate6 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate6AddPost|1651780042982|1651780042982|226|0|1349400886053
$ grep LdbcUpdate7 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate7AddComment|1651780043007|1651780043007|157|0|1349400892563
$ grep LdbcUpdate8 LDBC-SNB-results_log.csv | tail -n 1
LdbcUpdate8AddFriendship|1651780042948|1651780042948|264|0|1349400877093
```

From the logs, the last completed updates were retrieved for each update operation. The log entries include the operation name, actual and scheduled start time, the execution time, the delay between scheduled and actual start times, and the initial query start time without the scale factor multiplier (this latter one is included in the last column). Using this information, the query parameters were obtained from the initial CSV files generated by the Datagen. To check whether the graph entities in the driver log entries were persisted in the database, custom read queries were executed after database restart. The queries returned the data that was committed according to the logs, so the system passed this check. These queries are included in the `RecoveryQueries.java` attachment.

7 Supplementary Materials

The table below shows the list of supplementary materials. These materials are made available with this full disclosure report to allow reproducibility of results.

Table 7.1: Supplementary materials

File (or parent folder) name	Purpose (empty for parent folders)
./	
README.md	Guide to set up benchmark environment
check/	
Acid.java, acid.sh	ACID tests implementation and runner script
RecoveryQueries.java, recoveryQueries.sh	Recovery check implementation and runner script
galaxybase-bolt-driver-....jar	Dependency for accessing the database for checks
code/	
code/galaxybase-apoc-ldbc.zip	Benchmark query implementation package
code/ldbc_snb_interactive-0.3.4.zip	SUT-specific LDBC driver implementation
code/ldbc_snb_driver-0.3.4.zip	Generic LDBC driver base project snapshot
database/	
build.sh	Database initializer script
galaxybase.tar.gz	Database binaries
load.sh	Database bulk importer tool
move.sh	Data mover script from Datagen to workspace
remove.sh	Cleanup script
database/json/	
mapping-sf{1,10,30,100,300}.json	Data schema mapping files
schema-sf{1,10,30,100,300}.json	Data schema descriptor files
datagen/	
datagen.sh	Datagen runner and CSV converter script
ldbc_snb_datagen.zip	LDBC datagen snapshot
params-sf{10,30,100,300}.ini	Datagen parameters for the used scale factors
docs/	
Galaxybase LDBC (sf{acid,10,30,100,300}).pdf	Instructions to run the benchmark
implements/	
Warmup.java, warmup.sh	Supplementary property initializer program and script
params.sh	Benchmark data mover script
galaxybase-bolt-driver-....jar	Dependency for accessing the database for benchmark
implements/galaxybase/	
sf{30,100,300}-interactive-benchmark.properties	Driver configurations for performance benchmark
interactive-validate.properties	Results validation driver settings
run.sh	Benchmark runner script
result.sh	Script to collect benchmark results
implements/galaxybase/target/	
galaxybase-0.3.4-SNAPSHOT....jar	Database binary
results/	
results-sf{30,100,300}.zip	Driver output files and console logs for the selected scale factors

The attachment folder directory structure is as follows (based on the output of the `tree` command):

```
attachments
├── README.md
├── check
│   ├── Acid.java
│   ├── RecoveryQueries.java
│   ├── acid.sh
│   ├── galaxybase-bolt-driver-3.3.1-SNAPSHOT-jar-with-dependencies.jar
│   └── recoveryQueries.sh
├── code
│   ├── galaxybase-apoc-ldbc.zip
│   ├── ldbc_snb_driver-0.3.4.zip
│   └── ldbc_snb_interactive-0.3.4.zip
├── database
│   ├── build.sh
│   ├── galaxybase.tar.gz
│   ├── json
│   │   ├── mapping-sf{1,10,30,100,300}.json
│   │   └── schema-sf{1,10,30,100,300}.json
│   ├── load.sh
│   ├── move.sh
│   └── remove.sh
├── datagen
│   ├── datagen.sh
│   ├── ldbc_snb_datagen.zip
│   ├── params-sf{1,10,30,100,300}.ini
│   └── pretreatment.sh
├── docs
│   └── Galaxybase LDBC ({acid,sf10,sf30,sf100,sf300}).pdf
├── implements
│   ├── Warmup.java
│   ├── galaxybase
│   │   ├── interactive-create-validation-parameters.properties
│   │   ├── interactive-validate.properties
│   │   ├── result.sh
│   │   ├── run.sh
│   │   ├── sf{1,30,100,300}-interactive-benchmark.properties
│   │   ├── target
│   │   │   └── galaxybase-0.3.4-SNAPSHOT-jar-with-dependencies.jar
│   │   └── validation_params.csv
│   ├── galaxybase-bolt-driver-3.3.1-SNAPSHOT-jar-with-dependencies.jar
│   ├── params.sh
│   └── warmup.sh
└── results
    └── sf{30,100,300}-result.zip
```

A Appendix

A.1 CPU details

Listing A.1: Output of the `cat /proc/cpuinfo` command for one core

```
$ cat /proc/cpuinfo
processor       : 0
vendor_id       : GenuineIntel
cpu family      : 6
model           : 85
model name      : Intel(R) Xeon(R) Platinum 8259CL CPU @ 2.50GHz
stepping        : 7
microcode       : 0x500320a
cpu MHz         : 3100.015
cache size      : 36608 KB
physical id     : 0
siblings        : 48
core id         : 0
cpu cores       : 24
apicid          : 0
initial apicid  : 0
fpu             : yes
fpu_exception   : yes
cpuid level     : 13
wp              : yes
flags           : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse
    sse2 ss ht syscall nx pdpe1gb rdtscp lm constant_tsc arch_perfmon rep_good nopl xtopology nonstop_tsc cpuid
    aperfmperf tsc_known_freq pni pclmu
lqdq monitor ssse3 fma cx16 pcid sse4_1 sse4_2 x2apic movbe popcnt tsc_deadline_timer aes xsave avx f16c rdrand
    hypervisor lahf_lm abm 3dnowprefetch invpcid_single pti fsgsbase tsc_adjust bmi1 avx2 smep bmi2 erms invpcid
    mpx avx512f avx512dq rdseed adx s
map clflushopt clwb avx512cd avx512bw avx512vl xsaveopt xsavec xgetbv1 xsaves ida arat pku ospke
bugs            : cpu_meltdown spectre_v1 spectre_v2 spec_store_bypass l1tf mds swapgs itlb_multihit
bogomips        : 4999.99
clflush size    : 64
cache_alignment : 64
address sizes   : 46 bits physical, 48 bits virtual
power management:
```

A.2 IO performance

Listing A.2: Output of the `fio` command

```
$ sudo fio --rw=write --ioengine=sync --fdatasync=1 --direct=1 --directory=/mnt --size=2g --bs=4k --name=iotest
iotest: (g=0): rw=write, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=sync, iodepth=1
fio-3.1
Starting 1 process
iotest: Laying out IO file (1 file / 2048MiB)

iotest: (groupid=0, jobs=1): err= 0: pid=9615: Fri May  6 07:54:19 2022
  write: IOPS=6421, BW=25.1MiB/s (26.3MB/s)(2048MiB/81643msec)
    clat (usec): min=27, max=301, avg=47.83, stdev=10.88
     lat (usec): min=27, max=301, avg=47.97, stdev=10.94
    clat percentiles (usec):
```

```
     |  1.00th=[   31],  5.00th=[   33], 10.00th=[   34], 20.00th=[   35],
     | 30.00th=[   37], 40.00th=[   51], 50.00th=[   53], 60.00th=[   54],
     | 70.00th=[   56], 80.00th=[   57], 90.00th=[   60], 95.00th=[   62],
     | 99.00th=[   69], 99.50th=[   74], 99.90th=[   82], 99.95th=[   88],
     | 99.99th=[  133]
   bw (  KiB/s): min=24440, max=28544, per=100.00%, avg=25689.04, stdev=788.09, samples=163
   iops        : min= 6110, max= 7136, avg=6422.24, stdev=197.03, samples=163
  lat (usec)   : 50=38.59%, 100=61.39%, 250=0.03%, 500=0.01%
  cpu          : usr=3.37%, sys=16.04%, ctx=1048692, majf=0, minf=14
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwt: total=0,524288,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
  WRITE: bw=25.1MiB/s (26.3MB/s), 25.1MiB/s-25.1MiB/s (26.3MB/s-26.3MB/s), io=2048MiB (2147MB), run=81643-81643
     msec

Disk stats (read/write):
    md127: ios=0/2622450, merge=0/0, ticks=0/0, in_queue=0, util=0.00%, aggrios=0/790388, aggrmerge=0/520843,
     aggrticks=0/26905, aggrin_queue=0, aggrutil=96.08%
  nvme3n1: ios=0/790316, merge=0/520865, ticks=0/27400, in_queue=0, util=96.08%
  nvme2n1: ios=0/790460, merge=0/520821, ticks=0/26411, in_queue=0, util=95.48%
```

A.3 Datagen configuration

Listing A.3: Contents of `params-sf30.ini` used for scale factor 30

```
ldbc.snb.datagen.generator.scaleFactor:snb.interactive.30
ldbc.snb.datagen.serializer.numUpdatePartitions:32

ldbc.snb.datagen.serializer.dynamicActivitySerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.
    activity.CsvCompositeDynamicActivitySerializer
ldbc.snb.datagen.serializer.dynamicPersonSerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.person.
    CsvCompositeDynamicPersonSerializer
ldbc.snb.datagen.serializer.staticSerializer:ldbc.snb.datagen.serializer.snb.csv.staticserializer.
    CsvCompositeStaticSerializer
```

Listing A.4: Contents of `params-sf100.ini` used for scale factor 100

```
ldbc.snb.datagen.generator.scaleFactor:snb.interactive.100
ldbc.snb.datagen.serializer.numUpdatePartitions:32

ldbc.snb.datagen.serializer.dynamicActivitySerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.
    activity.CsvCompositeDynamicActivitySerializer
ldbc.snb.datagen.serializer.dynamicPersonSerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.person.
    CsvCompositeDynamicPersonSerializer
ldbc.snb.datagen.serializer.staticSerializer:ldbc.snb.datagen.serializer.snb.csv.staticserializer.
    CsvCompositeStaticSerializer
```

Listing A.5: Contents of `params-sf300.ini` used for scale factor 300

```
ldbc.snb.datagen.generator.scaleFactor:snb.interactive.300
ldbc.snb.datagen.serializer.numUpdatePartitions:32

```

```
ldbc.snb.datagen.serializer.dynamicActivitySerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.activity.CsvCompositeDynamicActivitySerializer
ldbc.snb.datagen.serializer.dynamicPersonSerializer:ldbc.snb.datagen.serializer.snb.csv.dynamicserializer.person.CsvCompositeDynamicPersonSerializer
ldbc.snb.datagen.serializer.staticSerializer:ldbc.snb.datagen.serializer.snb.csv.staticserializer.CsvCompositeStaticSerializer
```

A.4 Import configuration

Listing A.6: Content of `mapping-sf30.json` describing the data mapping

```
{
  "parallel": true,
  "splitFile": true,
  "deleteTempFile": true,
  "fastMode": true,
  "content": [
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/comment_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "vertexes": [
        {
          "type": "Comment",
          "primaryId": "id",
          "pkColumnIndex": 0,
          "hasDuplicatePrimaryKey": false,
          "property": [
            {
              "columnIndex": 1,
              "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
              "alias": "creationDate"
            },
            {
              "columnIndex": 2,
              "alias": "locationIP"
            },
            {
              "columnIndex": 3,
              "alias": "browserUsed"
            },
            {
              "columnIndex": 4,
              "alias": "content"
            },
            {
              "columnIndex": 5,
              "alias": "length"
            }
          ]
        }
```

```
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/post_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "Post",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "imageFile"
          },
          {
            "columnIndex": 2,
            "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
            "alias": "creationDate"
          },
          {
            "columnIndex": 3,
            "alias": "locationIP"
          },
          {
            "columnIndex": 4,
            "alias": "browserUsed"
          },
          {
            "columnIndex": 5,
            "alias": "language"
          },
          {
            "columnIndex": 6,
            "alias": "content"
          },
          {
            "columnIndex": 7,
            "alias": "length"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/data/company_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
```

```
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "Company",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "name"
          },
          {
            "columnIndex": 2,
            "alias": "url"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/data/university_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "University",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "name"
          },
          {
            "columnIndex": 2,
            "alias": "url"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/data/city_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
```

```
    "vertexes": [
      {
        "type": "City",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "name"
          },
          {
            "columnIndex": 2,
            "alias": "url"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/data/continent_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "Continent",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "name"
          },
          {
            "columnIndex": 2,
            "alias": "url"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/data/country_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
```

```
        "type": "Country",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "name"
          },
          {
            "columnIndex": 2,
            "alias": "url"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/forum_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "Forum",
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "title"
          },
          {
            "columnIndex": 2,
            "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
            "alias": "creationDate"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/person_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "Person",
```

```
        "primaryId": "id",
        "pkColumnIndex": 0,
        "hasDuplicatePrimaryKey": false,
        "property": [
          {
            "columnIndex": 1,
            "alias": "firstName"
          },
          {
            "columnIndex": 2,
            "alias": "lastName"
          },
          {
            "columnIndex": 3,
            "alias": "gender"
          },
          {
            "columnIndex": 4,
            "dateFormat": "yyyy-MM-dd",
            "alias": "birthday"
          },
          {
            "columnIndex": 5,
            "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
            "alias": "creationDate"
          },
          {
            "columnIndex": 6,
            "alias": "locationIP"
          },
          {
            "columnIndex": 7,
            "alias": "browserUsed"
          },
          {
            "columnIndex": 8,
            "alias": "language"
          },
          {
            "columnIndex": 9,
            "alias": "email"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/tag_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "vertexes": [
      {
        "type": "Tag",
```

```
          "primaryId": "id",
          "pkColumnIndex": 0,
          "hasDuplicatePrimaryKey": false,
          "property": [
            {
              "columnIndex": 1,
              "alias": "name"
            },
            {
              "columnIndex": 2,
              "alias": "url"
            }
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/tagclass_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "vertexes": [
        {
          "type": "TagClass",
          "primaryId": "id",
          "pkColumnIndex": 0,
          "hasDuplicatePrimaryKey": false,
          "property": [
            {
              "columnIndex": 1,
              "alias": "name"
            },
            {
              "columnIndex": 2,
              "alias": "url"
            }
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/forum_containerOf_post_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Forum_CONTAINER_OF_Post",
          "fromType": "Forum",
          "toType": "Post",
```

```
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/comment_hasCreator_person_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Comment_HAS_CREATOR_Person",
          "fromType": "Comment",
          "toType": "Person",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/post_hasCreator_person_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Post_HAS_CREATOR_Person",
          "fromType": "Post",
          "toType": "Person",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/person_hasInterest_tag_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
```

```
    "edges": [
      {
        "type": "Person_HAS_INTEREST_Tag",
        "fromType": "Person",
        "toType": "Tag",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/forum_hasMember_person_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "edges": [
      {
        "type": "Forum_HAS_MEMBER_Person",
        "fromType": "Forum",
        "toType": "Person",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
          {
            "columnIndex": 2,
            "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
            "alias": "joinDate"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/forum_hasModerator_person_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "edges": [
      {
        "type": "Forum_HAS_MODERATOR_Person",
        "fromType": "Forum",
        "toType": "Person",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
        ]
      }
    ]
```

```
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/post_hasTag_tag_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Post_HAS_TAG_Tag",
          "fromType": "Post",
          "toType": "Tag",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/forum_hasTag_tag_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Forum_HAS_TAG_Tag",
          "fromType": "Forum",
          "toType": "Tag",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/comment_hasTag_tag_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Comment_HAS_TAG_Tag",
          "fromType": "Comment",
          "toType": "Tag",
          "fromKeyColumnIndex": 0,
```

```
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/tag_hasType_tagclass_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Tag_HAS_TYPE_TagClass",
          "fromType": "Tag",
          "toType": "TagClass",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/organisation_isLocatedIn_place_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Company_IS_LOCATED_IN_Country",
          "fromType": "Company",
          "toType": "Country",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/organisation_isLocatedIn_place_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
```

```
        {
          "type": "University_IS_LOCATED_IN_City",
          "fromType": "University",
          "toType": "City",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/comment_isLocatedIn_place_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Comment_IS_LOCATED_IN_Country",
          "fromType": "Comment",
          "toType": "Country",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/post_isLocatedIn_place_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Post_IS_LOCATED_IN_Country",
          "fromType": "Post",
          "toType": "Country",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/person_isLocatedIn_place_0_0.csv",
```

```
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Person_IS_LOCATED_IN_City",
          "fromType": "Person",
          "toType": "City",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/place_isPartOf_place_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Country_IS_PART_OF_Continent",
          "fromType": "Country",
          "toType": "Continent",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/place_isPartOf_place_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "City_IS_PART_OF_Country",
          "fromType": "City",
          "toType": "Country",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
```

```
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/tagclass_isSubclassOf_tagclass_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "TagClass_IS_SUBCLASS_OF_TagClass",
          "fromType": "TagClass",
          "toType": "TagClass",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/person_knows_person_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
          "type": "Person_KNOWS_Person",
          "fromType": "Person",
          "toType": "Person",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
            {
              "columnIndex": 2,
              "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
              "alias": "creationDate"
            }
          ]
        }
      ]
    },
    {
      "source": {
        "sourceType": "CSV",
        "encoding": "utf-8",
        "filePath": "data/sf30/person_likes_post_0_0.csv",
        "delimiter": "|",
        "enclosingChar": "\"",
        "hasHeader": true
      },
      "edges": [
        {
```

```
        "type": "Person_LIKES_Post",
        "fromType": "Person",
        "toType": "Post",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
          {
            "columnIndex": 2,
            "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
            "alias": "creationDate"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/person_likes_comment_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "edges": [
      {
        "type": "Person_LIKES_Comment",
        "fromType": "Person",
        "toType": "Comment",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
          {
            "columnIndex": 2,
            "dateFormat": "yyyy-MM-dd'T'HH:mm:ss.SSS",
            "alias": "creationDate"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/comment_replyOf_post_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "edges": [
      {
        "type": "Comment_REPLY_OF_Post",
        "fromType": "Comment",
        "toType": "Post",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
```

```
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/comment_replyOf_comment_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "edges": [
      {
        "type": "Comment_REPLY_OF_Comment",
        "fromType": "Comment",
        "toType": "Comment",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/person_studyAt_organisation_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
    },
    "edges": [
      {
        "type": "Person_STUDY_AT_University",
        "fromType": "Person",
        "toType": "University",
        "fromKeyColumnIndex": 0,
        "toKeyColumnIndex": 1,
        "property": [
          {
            "columnIndex": 2,
            "alias": "classYear"
          }
        ]
      }
    ]
  },
  {
    "source": {
      "sourceType": "CSV",
      "encoding": "utf-8",
      "filePath": "data/sf30/person_workAt_organisation_0_0.csv",
      "delimiter": "|",
      "enclosingChar": "\"",
      "hasHeader": true
```

```
      },
      "edges": [
        {
          "type": "Person_WORK_AT_Company",
          "fromType": "Person",
          "toType": "Company",
          "fromKeyColumnIndex": 0,
          "toKeyColumnIndex": 1,
          "property": [
            {
              "columnIndex": 2,
              "alias": "workFrom"
            }
          ]
        }
      ]
    }
  ]
}
```

A.5 Benchmark configuration

Listing A.7: Contents of `interactive-benchmark-sf30.properties` used for scale factor 30

```
endpoint=ldbc://127.0.0.1:62000
threadCount=48
user=admin
password=admin
graphName=sf30
queryDir=queries/
printQueryNames=false
printQueryStrings=false
printQueryResults=false


status=1
thread_count=48
name=LDBC-SNB
results_log=true
time_unit=MICROSECONDS
time_compression_ratio=0.0038
peer_identifiers=
workload_statistics=false
spinner_wait_duration=1
help=false
ignore_scheduled_start_times=false

workload=com.ldbc.driver.workloads.ldbc.snb.interactive.LdbcSnbInteractiveWorkload
db=com.ldbc.impls.workloads.ldbc.snb.galaxybase.interactive.GalaxybaseInteractiveDb
operation_count=75000000
ldbc.snb.interactive.parameters_dir=../sf30/substitution_parameters/
ldbc.snb.interactive.updates_dir=../sf30/social_network/
ldbc.snb.interactive.short_read_dissipation=0.2
ldbc.snb.interactive.update_interleave=156

warmup=19000000
```

```
33
34 ## frequency of read queries (number of update queries per one read query)
35 ldbc.snb.interactive.LdbcQuery1_freq=26
36 ldbc.snb.interactive.LdbcQuery2_freq=37
37 ldbc.snb.interactive.LdbcQuery3_freq=106
38 ldbc.snb.interactive.LdbcQuery4_freq=36
39 ldbc.snb.interactive.LdbcQuery5_freq=72
40 ldbc.snb.interactive.LdbcQuery6_freq=316
41 ldbc.snb.interactive.LdbcQuery7_freq=48
42 ldbc.snb.interactive.LdbcQuery8_freq=9
43 ldbc.snb.interactive.LdbcQuery9_freq=384
44 ldbc.snb.interactive.LdbcQuery10_freq=37
45 ldbc.snb.interactive.LdbcQuery11_freq=20
46 ldbc.snb.interactive.LdbcQuery12_freq=44
47 ldbc.snb.interactive.LdbcQuery13_freq=19
48 ldbc.snb.interactive.LdbcQuery14_freq=49
49
50 # *** For debugging purposes ***
51
52 ldbc.snb.interactive.LdbcQuery1_enable=true
53 ldbc.snb.interactive.LdbcQuery2_enable=true
54 ldbc.snb.interactive.LdbcQuery3_enable=true
55 ldbc.snb.interactive.LdbcQuery4_enable=true
56 ldbc.snb.interactive.LdbcQuery5_enable=true
57 ldbc.snb.interactive.LdbcQuery6_enable=true
58 ldbc.snb.interactive.LdbcQuery7_enable=true
59 ldbc.snb.interactive.LdbcQuery8_enable=true
60 ldbc.snb.interactive.LdbcQuery9_enable=true
61 ldbc.snb.interactive.LdbcQuery10_enable=true
62 ldbc.snb.interactive.LdbcQuery11_enable=true
63 ldbc.snb.interactive.LdbcQuery12_enable=true
64 ldbc.snb.interactive.LdbcQuery13_enable=true
65 ldbc.snb.interactive.LdbcQuery14_enable=true
66
67 ldbc.snb.interactive.LdbcShortQuery1PersonProfile_enable=true
68 ldbc.snb.interactive.LdbcShortQuery2PersonPosts_enable=true
69 ldbc.snb.interactive.LdbcShortQuery3PersonFriends_enable=true
70 ldbc.snb.interactive.LdbcShortQuery4MessageContent_enable=true
71 ldbc.snb.interactive.LdbcShortQuery5MessageCreator_enable=true
72 ldbc.snb.interactive.LdbcShortQuery6MessageForum_enable=true
73 ldbc.snb.interactive.LdbcShortQuery7MessageReplies_enable=true
74
75 ldbc.snb.interactive.LdbcUpdate1AddPerson_enable=true
76 ldbc.snb.interactive.LdbcUpdate2AddPostLike_enable=true
77 ldbc.snb.interactive.LdbcUpdate3AddCommentLike_enable=true
78 ldbc.snb.interactive.LdbcUpdate4AddForum_enable=true
79 ldbc.snb.interactive.LdbcUpdate5AddForumMembership_enable=true
80 ldbc.snb.interactive.LdbcUpdate6AddPost_enable=true
81 ldbc.snb.interactive.LdbcUpdate7AddComment_enable=true
82 ldbc.snb.interactive.LdbcUpdate8AddFriendship_enable=true
```

Listing A.8: Contents of `interactive-benchmark-sf100.properties` used for scale factor 100

```
1 endpoint=ldbc://127.0.0.1:62000
2 threadCount=48
3 user=admin
4 password=admin
5 graphName=sf100
```

```
queryDir=queries/
printQueryNames=false
printQueryStrings=false
printQueryResults=false
status=1
thread_count=48
name=LDBC-SNB
results_log=true
time_unit=MICROSECONDS
time_compression_ratio=0.0163
peer_identifiers=
workload_statistics=false
spinner_wait_duration=1
help=false
ignore_scheduled_start_times=false
workload=com.ldbc.driver.workloads.ldbc.snb.interactive.LdbcSnbInteractiveWorkload
db=com.ldbc.impls.workloads.ldbc.snb.galaxybase.interactive.GalaxybaseInteractiveDb
operation_count=64000000
ldbc.snb.interactive.parameters_dir=../sf100/substitution_parameters/
ldbc.snb.interactive.updates_dir=../sf100/social_network/
ldbc.snb.interactive.short_read_dissipation=0.2
ldbc.snb.interactive.update_interleave=48

warmup=16000000

## frequency of read queries (number of update queries per one read query)
ldbc.snb.interactive.LdbcQuery1_freq=26
ldbc.snb.interactive.LdbcQuery2_freq=37
ldbc.snb.interactive.LdbcQuery3_freq=123
ldbc.snb.interactive.LdbcQuery4_freq=36
ldbc.snb.interactive.LdbcQuery5_freq=78
ldbc.snb.interactive.LdbcQuery6_freq=434
ldbc.snb.interactive.LdbcQuery7_freq=38
ldbc.snb.interactive.LdbcQuery8_freq=5
ldbc.snb.interactive.LdbcQuery9_freq=527
ldbc.snb.interactive.LdbcQuery10_freq=40
ldbc.snb.interactive.LdbcQuery11_freq=22
ldbc.snb.interactive.LdbcQuery12_freq=44
ldbc.snb.interactive.LdbcQuery13_freq=19
ldbc.snb.interactive.LdbcQuery14_freq=49

# *** For debugging purposes ***

ldbc.snb.interactive.LdbcQuery1_enable=true
ldbc.snb.interactive.LdbcQuery2_enable=true
ldbc.snb.interactive.LdbcQuery3_enable=true
ldbc.snb.interactive.LdbcQuery4_enable=true
ldbc.snb.interactive.LdbcQuery5_enable=true
ldbc.snb.interactive.LdbcQuery6_enable=true
ldbc.snb.interactive.LdbcQuery7_enable=true
ldbc.snb.interactive.LdbcQuery8_enable=true
ldbc.snb.interactive.LdbcQuery9_enable=true
ldbc.snb.interactive.LdbcQuery10_enable=true
ldbc.snb.interactive.LdbcQuery11_enable=true
ldbc.snb.interactive.LdbcQuery12_enable=true
ldbc.snb.interactive.LdbcQuery13_enable=true
ldbc.snb.interactive.LdbcQuery14_enable=true
ldbc.snb.interactive.LdbcShortQuery1PersonProfile_enable=true
```

```
ldbc.snb.interactive.LdbcShortQuery2PersonPosts_enable=true
ldbc.snb.interactive.LdbcShortQuery3PersonFriends_enable=true
ldbc.snb.interactive.LdbcShortQuery4MessageContent_enable=true
ldbc.snb.interactive.LdbcShortQuery5MessageCreator_enable=true
ldbc.snb.interactive.LdbcShortQuery6MessageForum_enable=true
ldbc.snb.interactive.LdbcShortQuery7MessageReplies_enable=true

ldbc.snb.interactive.LdbcUpdate1AddPerson_enable=true
ldbc.snb.interactive.LdbcUpdate2AddPostLike_enable=true
ldbc.snb.interactive.LdbcUpdate3AddCommentLike_enable=true
ldbc.snb.interactive.LdbcUpdate4AddForum_enable=true
ldbc.snb.interactive.LdbcUpdate5AddForumMembership_enable=true
ldbc.snb.interactive.LdbcUpdate6AddPost_enable=true
ldbc.snb.interactive.LdbcUpdate7AddComment_enable=true
ldbc.snb.interactive.LdbcUpdate8AddFriendship_enable=true
```

Listing A.9: Contents of `interactive-benchmark-sf300.properties` used for scale factor 300

```
endpoint=ldbc://127.0.0.1:62000
threadCount=48
user=admin
password=admin
graphName=sf300
queryDir=queries/
printQueryNames=false
printQueryStrings=false
printQueryResults=false

status=1
thread_count=48
name=LDBC-SNB
results_log=true
time_unit=MICROSECONDS
time_compression_ratio=0.053
peer_identifiers=
workload_statistics=false
spinner_wait_duration=1
help=false
ignore_scheduled_start_times=false

workload=com.ldbc.driver.workloads.ldbc.snb.interactive.LdbcSnbInteractiveWorkload
db=com.ldbc.impls.workloads.ldbc.snb.galaxybase.interactive.GalaxybaseInteractiveDb
operation_count=64000000
ldbc.snb.interactive.parameters_dir=../sf300/substitution_parameters/
ldbc.snb.interactive.updates_dir=../sf300/social_network/
ldbc.snb.interactive.short_read_dissipation=0.2
ldbc.snb.interactive.update_interleave=17
warmup=16000000
## frequency of read queries (number of update queries per one read query)
ldbc.snb.interactive.LdbcQuery1_freq=26
ldbc.snb.interactive.LdbcQuery2_freq=37
ldbc.snb.interactive.LdbcQuery3_freq=142
ldbc.snb.interactive.LdbcQuery4_freq=36
ldbc.snb.interactive.LdbcQuery5_freq=84
ldbc.snb.interactive.LdbcQuery6_freq=580
ldbc.snb.interactive.LdbcQuery7_freq=32
ldbc.snb.interactive.LdbcQuery8_freq=3
ldbc.snb.interactive.LdbcQuery9_freq=705
```

```
ldbc.snb.interactive.LdbcQuery10_freq=44
ldbc.snb.interactive.LdbcQuery11_freq=24
ldbc.snb.interactive.LdbcQuery12_freq=44
ldbc.snb.interactive.LdbcQuery13_freq=19
ldbc.snb.interactive.LdbcQuery14_freq=49

# *** For debugging purposes ***

ldbc.snb.interactive.LdbcQuery1_enable=true
ldbc.snb.interactive.LdbcQuery2_enable=true
ldbc.snb.interactive.LdbcQuery3_enable=true
ldbc.snb.interactive.LdbcQuery4_enable=true
ldbc.snb.interactive.LdbcQuery5_enable=true
ldbc.snb.interactive.LdbcQuery6_enable=true
ldbc.snb.interactive.LdbcQuery7_enable=true
ldbc.snb.interactive.LdbcQuery8_enable=true
ldbc.snb.interactive.LdbcQuery9_enable=true
ldbc.snb.interactive.LdbcQuery10_enable=true
ldbc.snb.interactive.LdbcQuery11_enable=true
ldbc.snb.interactive.LdbcQuery12_enable=true
ldbc.snb.interactive.LdbcQuery13_enable=true
ldbc.snb.interactive.LdbcQuery14_enable=true

ldbc.snb.interactive.LdbcShortQuery1PersonProfile_enable=true
ldbc.snb.interactive.LdbcShortQuery2PersonPosts_enable=true
ldbc.snb.interactive.LdbcShortQuery3PersonFriends_enable=true
ldbc.snb.interactive.LdbcShortQuery4MessageContent_enable=true
ldbc.snb.interactive.LdbcShortQuery5MessageCreator_enable=true
ldbc.snb.interactive.LdbcShortQuery6MessageForum_enable=true
ldbc.snb.interactive.LdbcShortQuery7MessageReplies_enable=true

ldbc.snb.interactive.LdbcUpdate1AddPerson_enable=true
ldbc.snb.interactive.LdbcUpdate2AddPostLike_enable=true
ldbc.snb.interactive.LdbcUpdate3AddCommentLike_enable=true
ldbc.snb.interactive.LdbcUpdate4AddForum_enable=true
ldbc.snb.interactive.LdbcUpdate5AddForumMembership_enable=true
ldbc.snb.interactive.LdbcUpdate6AddPost_enable=true
ldbc.snb.interactive.LdbcUpdate7AddComment_enable=true
ldbc.snb.interactive.LdbcUpdate8AddFriendship_enable=true
```

A.6 Validation configuration

Listing A.10: The contents of `interactive-validate.properties`

```
endpoint=ldbc://127.0.0.1:62000
threadCount=1
user=admin
password=admin
graphName=sf10
queryDir=queries/
printQueryNames=false
printQueryStrings=false
printQueryResults=false

status=1
thread_count=1
```

```
name=LDBC-SNB
results_log=true
time_unit=MICROSECONDS
time_compression_ratio=0.001
peer_identifiers=
workload_statistics=false
spinner_wait_duration=1
help=false
ignore_scheduled_start_times=true

workload=com.ldbc.driver.workloads.ldbc.snb.interactive.LdbcSnbInteractiveWorkload
db=com.ldbc.impls.workloads.ldbc.snb.galaxybase.interactive.GalaxybaseInteractiveDb
operation_count=10000

validate_workload=true
validate_database=validation_params.csv
ldbc.snb.interactive.parameters_dir=../sf10/substitution_parameters/
ldbc.snb.interactive.short_read_dissipation=0.2
ldbc.snb.interactive.update_interleave=466

## frequency of read queries (number of update queries per one read query)
ldbc.snb.interactive.LdbcQuery1_freq=1
ldbc.snb.interactive.LdbcQuery2_freq=1
ldbc.snb.interactive.LdbcQuery3_freq=1
ldbc.snb.interactive.LdbcQuery4_freq=1
ldbc.snb.interactive.LdbcQuery5_freq=1
ldbc.snb.interactive.LdbcQuery6_freq=1
ldbc.snb.interactive.LdbcQuery7_freq=1
ldbc.snb.interactive.LdbcQuery8_freq=1
ldbc.snb.interactive.LdbcQuery9_freq=1
ldbc.snb.interactive.LdbcQuery10_freq=1
ldbc.snb.interactive.LdbcQuery11_freq=1
ldbc.snb.interactive.LdbcQuery12_freq=1
ldbc.snb.interactive.LdbcQuery13_freq=1
ldbc.snb.interactive.LdbcQuery14_freq=1

# *** For debugging purposes ***

ldbc.snb.interactive.LdbcQuery1_enable=true
ldbc.snb.interactive.LdbcQuery2_enable=true
ldbc.snb.interactive.LdbcQuery3_enable=true
ldbc.snb.interactive.LdbcQuery4_enable=true
ldbc.snb.interactive.LdbcQuery5_enable=true
ldbc.snb.interactive.LdbcQuery6_enable=true
ldbc.snb.interactive.LdbcQuery7_enable=true
ldbc.snb.interactive.LdbcQuery8_enable=true
ldbc.snb.interactive.LdbcQuery9_enable=true
ldbc.snb.interactive.LdbcQuery10_enable=true
ldbc.snb.interactive.LdbcQuery11_enable=true
ldbc.snb.interactive.LdbcQuery12_enable=true
ldbc.snb.interactive.LdbcQuery13_enable=true
ldbc.snb.interactive.LdbcQuery14_enable=true

ldbc.snb.interactive.LdbcShortQuery1PersonProfile_enable=true
ldbc.snb.interactive.LdbcShortQuery2PersonPosts_enable=true
ldbc.snb.interactive.LdbcShortQuery3PersonFriends_enable=true
ldbc.snb.interactive.LdbcShortQuery4MessageContent_enable=true
ldbc.snb.interactive.LdbcShortQuery5MessageCreator_enable=true
```

```
ldbc.snb.interactive.LdbcShortQuery6MessageForum_enable=true
ldbc.snb.interactive.LdbcShortQuery7MessageReplies_enable=true

ldbc.snb.interactive.LdbcUpdate1AddPerson_enable=true
ldbc.snb.interactive.LdbcUpdate2AddPostLike_enable=true
ldbc.snb.interactive.LdbcUpdate3AddCommentLike_enable=true
ldbc.snb.interactive.LdbcUpdate4AddForum_enable=true
ldbc.snb.interactive.LdbcUpdate5AddForumMembership_enable=true
ldbc.snb.interactive.LdbcUpdate6AddPost_enable=true
ldbc.snb.interactive.LdbcUpdate7AddComment_enable=true
ldbc.snb.interactive.LdbcUpdate8AddFriendship_enable=true
```