LDBC Social Network Benchmark Task Force work charter

This charter regularizes the status of the group creating the Social Network Benchmark (SNB) as a Task Force, with respect to the more formal LDBC constitution and rules adopted since 2019 onwards. This is in line with the drive to professionalize LDBC functioning, and reflects the steady growth of audited benchmark projects, and the need for periodic Board review. The Task Force has operated, and continues to operate, as the most significant and exemplary permanent technical working body of LDBC.

Task Force Name

LDBC Social Network Benchmark Task Force

Group leaders

- Gábor Szárnyas (lead)
- David Püroja (co-lead)
- Benjamin A. Steer (co-lead)

Group members

1. Gábor Szárnyas (CWI)
2. David Püroja (independent)
3. Benjamin A. Steer (Pometry)
4. Peter Boncz (CWI) – *on a sabbatical leave until Sep 15 2023*
5. Dávid Szakállas (independent)
6. Altan Birler (TUM)
7. Zhihui Guo (Ant Group)
8. Shipeng Qi (Ant Group)

Task Force Lifetime

Permanent

Charter version date

18 February 2023

Charter versions

- 28 February 2023:   LDBC AMS-001 Revision 1
- 17 March 2023:   LDBC AMS-001 Revision 2
- 18 March 2023:   LDBC AMS-001 Revision 3

Date charter version agreed by Board

28 February 2023

Mission

The task force maintains and periodically renews the Social Network Benchmark suite's workloads. This includes the benchmark specifications as well as the software components such as the Data generator (Datagen), the benchmark drivers and the reference

implementations. The task is responsible for training auditors, conducting auditor exams and overseeing audits.

Motivation

Facilitating competition among graph(-capable) database management systems

LDBC's mission is to facilitate the adoption of graph processing technologies and to speed up technological progress in the area of graph data management. To this end, it defines benchmarks and governs their use.

The Social Network Benchmark (SNB) is currently LDBC's most widely-adopted benchmark suite. The SNB targets database management systems, including graph DBMSs and relational DBMSs with graph processing capabilities. It has two workloads: the transactional Interactive workload [Interactive paper] and analytical Business Intelligence workload [BI paper]. Both of these workloads have official audited benchmark results [Audited results].

If we look back at the history of the well-known TPC-H benchmark – which was influential in both driving innovation among RDBMSs and also inspired several design decisions in the LDBC SNB – we can see a more than 100-fold increase in throughput among its audited results over the course of 20 years (see Appendix A). With the SNB workloads, the results are promising so far (a 4x increase in the last 3 years) but we are only at the beginning of this process.

Scope of Work

The main tasks of the task force are the following:

- Maintain the SNB Interactive v1 and SNB Business Intelligence v1 workloads.
- Help the auditing processes of SNB workloads by training auditors and overseeing audits.
- Create and release the SNB Interactive v2 workload. This work is based on the MSc thesis of David Püroja, which presents an initial design and implementation of this workload [Interactive v2 MSc thesis].
- Create reference implementation(s) of the workloads using the upcoming ISO standard query languages, GQL and the SQL/PGQ extension of SQL:2023.
- Elaboration and advocacy of benchmark specifications as formal LDBC Standards, which are the central work products of LDBC.

Intended output documents

The primary output of the task force is the SNB benchmark specification [SNB specification].

For each workload, we maintain the following software components:

- Data generator
- Parameter generators
- Benchmark driver
- Reference implementations

Other intended outputs/work product

- Presentations explaining the benchmark workloads given at industry meetups, academic workshops, LDBC TUC meetings, etc.

- Auditing guidelines
- Auditor exam sheets

Mode of work and cadence

For communication, we use LDBC's Basecamp and Slack organizations. We hold conference calls on Zoom at least once per 3 months.

Changes to the specification and software code will be done via GitHub pull requests. When major changes are applied (e.g. a query specification changes), a new version of the specification will be posted on arXiv.

The majority of the task force is located in Europe (UK, Netherlands, Czech Republic, Germany). Therefore, we have regular (informal) face to face meetings for efficient collaboration.

Intended timescales

- June 2023: SNB Interactive v2.0 RC
- July 2023: SNB Interactive v2.0 submitted for academic publication
- October 2023: SNB Interactive v2.0 submitted for approval

Related Task Forces or Working Groups

The Financial Benchmark task force works on the LDBC Financial Benchmark. This benchmark shares several core principles with the Social Network Benchmark (e.g. scalable data set and parameter generation) but it is a standalone benchmark developed independently from the SNB.

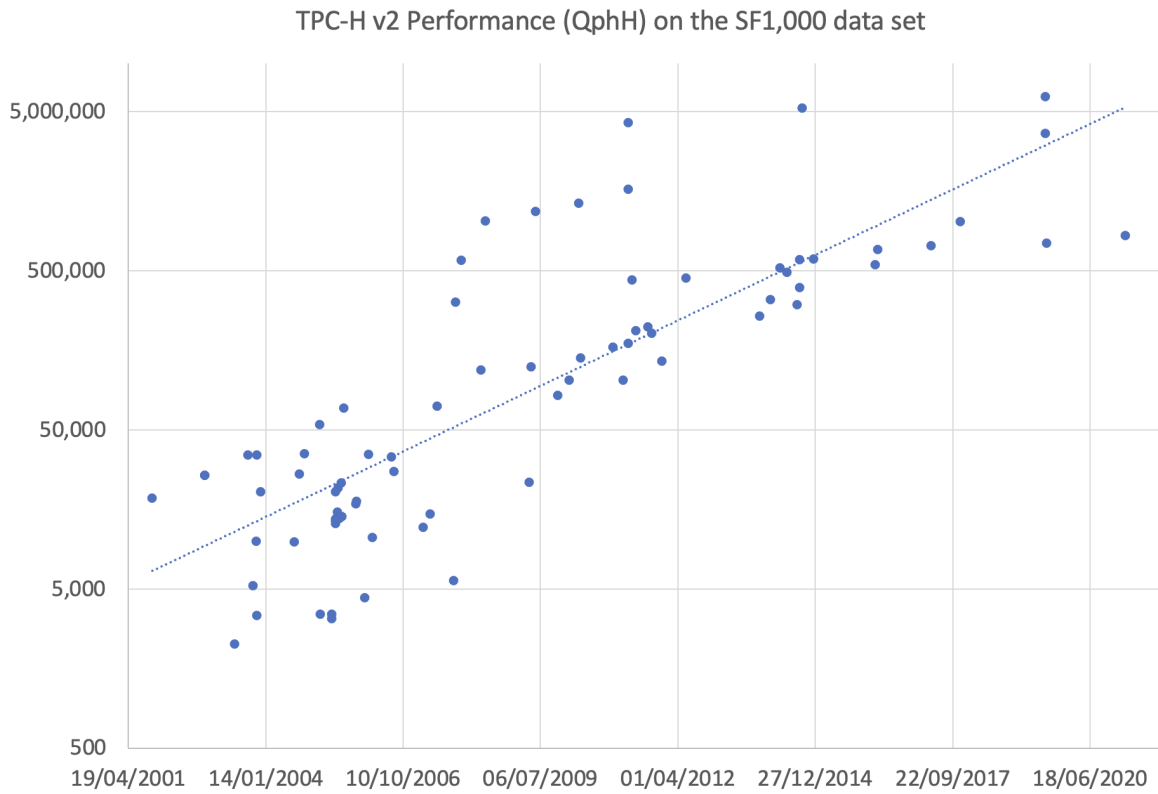## References to relevant documents, standards, etc.

[SNB specification] https://arxiv.org/abs/2001.02299

[BI paper] https://ldbcouncil.org/docs/papers/ldbc-snb-bi-vldb-2022.pdf

[Interactive paper] https://ldbcouncil.org/docs/papers/ldbc-snb-interactive-sigmod-2015.pdf

[Interactive v2 MSc thesis]
https://ldbcouncil.org/docs/papers/msc-thesis-david-puroja-snb-interactive-v2-2023.pdf

[Audited results] https://ldbcouncil.org/benchmarks/snb/

Appendix A

TPC-H audited results between 2001 and 2021

TPC-H v2 Performance (QphH) on the SF1,000 data set

SNB Interactive audited results between 2020 and 2023

SNB Interactive throughput on SF300