Reproducibility Report for ACM SIGMOD 2021 Paper: "COMPASS: Online Sketch-Based Query Optimization for In-Memory Databases"

DMYTRO BOGATOV, Boston University, United States

This reproducibility work is thorough and large-scale. The environment, code, and scripts are packaged in Docker images. Instructions detail how to pull the images, set them up, run the many experiments and parse the results. Reproducibility is, for the most part, automatic, with a small portion of manual labor to generate the plots.

1 INTRODUCTION

This is a reproducibility report for the paper by Izenov, Datta, Rusu, and Shin [1]. To summarize, two large figures (with subfigures, [1, Figures 6 and 7]) have been reproduced accurately enough. Due to the large scale of original experiments, the attempt was made to run only a random portion of similar runs. The results were manually checked to be similar to precomputed ones, and the plots were created from a mix of precomputed and reproduced data.

It should also be noted that judging from the reproducibility instructions and supplied scripts, it should be possible to reproduce the rest of the figures. It would just require somewhat more manual steps, time, and effort.

2 SUBMISSION

Reproducibility submission consists of detailed instructions on how to rerun the experiments. The instructions reference three large Docker images, which together contain the code, competing database engines (PostgreSQL, MonetDB, and MapD), scripts, and precomputed results. For the sake of reproducibility, the Docker images and Github repository have been cloned to replicated public repositories.

The submission contains:

- Docker image: `yizenov/reproduced_compass:latest`
  - Digest at the time of reproducibility review: `sha256:6448b6e0...09e2e2e9`
  - Replicated at: `dbogatov/reproducibility-reviews:yizenov--reproduced_compass`
  - **The image contains the COMPASS code and experiments scripts**
- Docker image: `adatta2/pg_docker:latest`
  - Digest at the time of reproducibility review: `sha256:6ca1eb0c...3ad484ca`
  - Replicated at: `dbogatov/reproducibility-reviews:adatta2--pg_docker`
  - **The image contains the PostgreSQL engine and some parsing scripts**
- Docker image: `adatta2/monetdb_docker:latest`
  - Digest at the time of reproducibility review: `sha256:a437356a...efc7dad0`
  - Replicated at: `dbogatov/reproducibility-reviews:adatta2--monetdb_docker`
  - **The image contains the MonetDB engine**
- A PDF with detailed instructions is publicly available in the GitHub repository.
- GitHub repository is referenced from the original paper
  - Original: `yizenov/compass_query_optimizer`
  - Replicated at: `repro-reviews/compass_query_optimizer`

3 HARDWARE AND SOFTWARE ENVIRONMENT

Table 1 describes the resources the original paper, and the reproducibility reviewer used to run the experiments.

Table 1. Hardware & Software environment

	Original paper [1]	This reproducibility report
Platform	Ubuntu 16.04	Debian 10, Docker 20.10.17
Machine	Unknown	GCP `n1-standard-16`
CPU model	Intel Xeon E5-2660	Intel (variable)
CPU cores	28 physical	8 physical
RAM	256 GB	60 GB
GPU	NVIDIA Tesla K80	NVIDIA Tesla T4

4 REPRODUCIBILITY EVALUATION

4.1 Process

The reproducibility process consists of pulling the Docker images, compiling COMPASS in one of the images, running the collection of experiments over the supplied SQL files, and parsing the results. A single such experiment is run via a convenient bash script. On top of the input and output directory, the run can be parametrized with using GPU or CPU and a choice of optimizer type. After running the experiments over all supplied SQL workloads, the Python scripts parse the results and produce a CSV file, which can be used to plot the results with supplied OpenOffice spreadsheet files (i.e., open-source MS Excel).

In this report, I did not run all thousands of experiments due to machine capacity and cost[1] constraints (the scripts require a GPU-enabled machine with at least 40 GB memory and at least 100 GB storage). Instead, I have run a random sample of experiments and checked that the results I have got are sufficiently close to the precomputed results supplied in the image. I was then able to run the results analysis scripts on a mix of reproduced and precomputed results.

The work was reproduced with minimal extra input from the authors. One needs some basic skills with Docker and bash to reproduce the experiments.

4.2 Results

See reproduced artifacts in Appendix A.

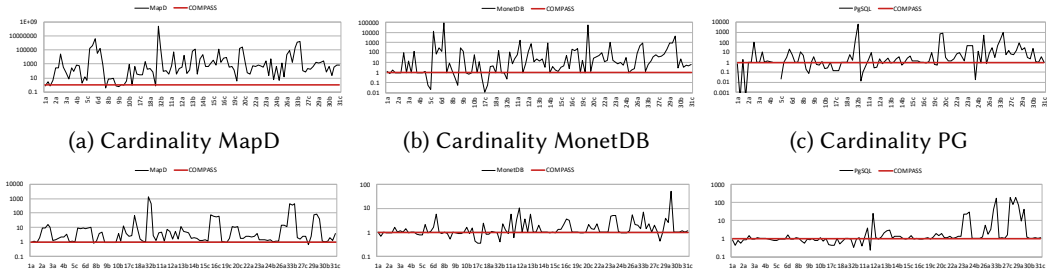
[1, Figures 6 and 7] have been reproduced (see Figs. 1 and 2). The reproduced figures lack DBMS A experiments, which is warranted given that it is an unnamed proprietary engine. Values for COMPASS, MapD, MonetDB, and PostgreSQL are expectedly different from the originally reported, but the correlations remain, and conclusions are valid (mostly the fact that COMPASS outperforms the competing solutions).

REFERENCES

[1] Yesdaulet Izenov, Asoke Datta, Florin Rusu, and Jun Hyung Shin. 2021. COMPASS: Online Sketch-Based Query Optimization for In-Memory Databases. In *Proceedings of the 2021 International Conference on Management of Data (SIGMOD '21)*. Association for Computing Machinery, 804–816. https://doi.org/10.1145/3448016.3452840

[1] I have used a GCP VM as the repro machine, and it costs about $820 per month without sustained use discount.

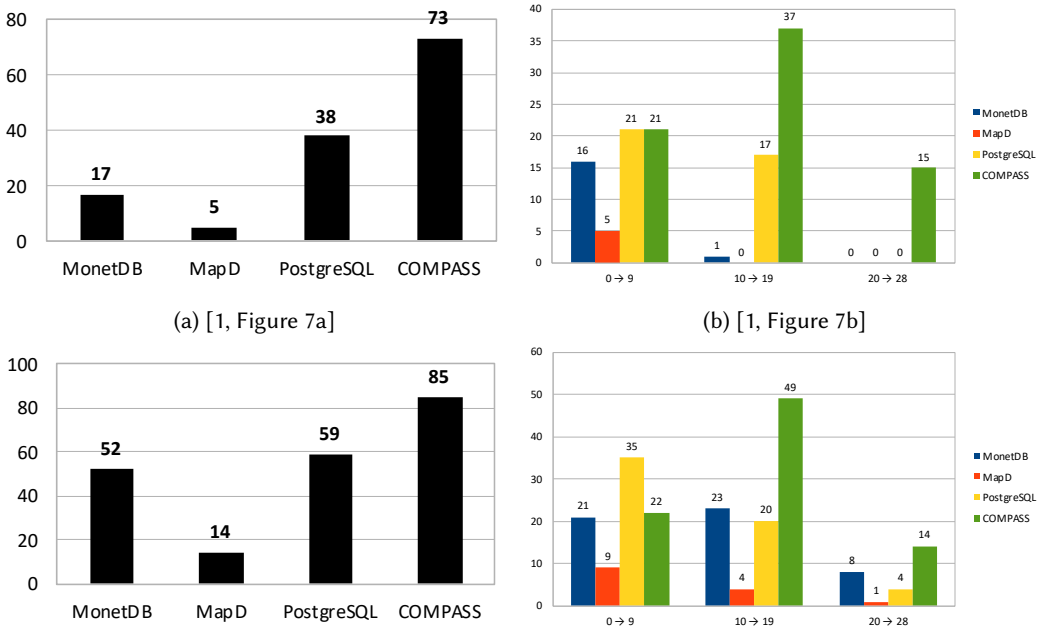
A REPRODUCED ARTIFACTS



(a) Cardinality MapD (b) Cardinality MonetDB (c) Cardinality PG

(d) Runtime MapD (e) Runtime MonetDB (f) Runtime PG

Fig. 1. [1, Figure 6]



(a) [1, Figure 7a] (b) [1, Figure 7b]

(c) [1, Figure 7c] (d) [1, Figure 7d]

Fig. 2. [1, Figure 7]