Reproducibility Report for ACM SIGMOD 2022 Paper: "TreeToaster: Towards an IVM-Optimized Compiler"

IMMANUEL HAFFNER, Saarland University, Germany

The evaluation was easy to reproduce and the reproduced findings confirm the evaluation done by the authors. Only acquiring the Docker image was unexpectedly difficult.

1 INTRODUCTION

This is the reproducibility report for the paper "TreeToaster: Towards an IVM-Optimized Compiler" by Darshana Balakrishnan, Carl Nuessle, Oliver Kennedy, and Lukasz Ziarek from University at Buffalo [1].

2 SUBMISSION

The authors provide a 2-page PDF document that briefly elaborates the steps required to reproduce evaluation. The evaluation software and environment are provided as a Docker image, hosted on Google Drive[1].

3 HARDWARE AND SOFTWARE ENVIRONMENT

	Paper	Repro Review
CPU	Intel Xeon	AMD Ryzen Threadripper 1900X
#cores	24	16
Base Frequency	2.5 GHz	3.8 GHz
Main Memory	192 GiB	32 GiB
Operating system	Ubuntu 16.04 LTS	ArchLinux w/ Linux 5.16.10

4 REPRODUCIBILITY EVALUATION

4.1 Process

Downloading the Docker image to our evaluation machine was unexpectedly difficult: Google Drive does not support easy `wget` or `curl` download for large files. Instead, a special client like Gdown is required. However, such clients can only download "open access" files. Hence, we had to download the Docker image to a desktop machine first and then copy it over to our evaluation machine. This is quite cumbersome since we are in a home-office due to the pandemic situation. Next time, please consider providing the Docker image on a different file hoster or host the file yourself and guard it with an `.htaccess`.

Apart from acquiring the Docker image, the overall evaluation is very straight-forward and well documented. A single script `main_script.sh` automatically downloads the required data sets, runs the benchmarks, and produces the charts presented in the paper.

4.2 Results

Overall, experiments on our machine take a little less time than on the hardware used by the authors. For Figure 1, this increases the relative amount of time spent on searching. However, the reproduced experimental results overall match the authors' results. Furthermore, our experimental

[1] https://drive.google.com/file/d/1KhU-NU24WI2dYJkRVDU6f1caoUU3Z8tC/view?usp=sharing

results appear to be more stable and show fewer outliers. This may be due to the fact that we disabled dynamic frequency scaling of the CPU, whereas the authors might have dynamic frequency scaling (Intel® TurboBoost) enabled.

REFERENCES

[1] Darshana Balakrishnan, Carl Nuessle, Oliver Kennedy, and Lukasz Ziarek. 2021. TreeToaster: Towards an IVM-Optimized Compiler. In *SIGMOD '21: International Conference on Management of Data, Virtual Event, China, June 20-25, 2021*. ACM, 155–167. https://doi.org/10.1145/3448016.3459244