Improving performances of an embedded RDBMS with a hybrid CPU/GPU processing engine

Samuel Cremer[1,2], Michel Bagein[1], Saïd Mahmoudi[1], Pierre Manneback[1]
[1]*Computer Science Department, University of Mons, Belgium*
[2]*Computer Engineering Department, Haute Ecole en Hainaut, Mons, Belgium*
samuel.cremer@heh.be, {michel.bagein, said.mahmoudi, pierre.manneback}@umons.ac.be

Relational database management systems (RDBMS) are still widely required by numerous business applications on servers, client devices, and even embedded inside end-user applications. Boosting processing speed and energy efficiency without compromising functionalities represents a big challenge. To achieve this goal, we propose to boost an existing embedded RDBMS by making it able to use hardware architectures with high memory bandwidth like GPUs. Moreover, GPUs benefit from much more computation power with the same order of energy consumption than CPUs.

The objective of this work is to exploit specificities of modern CPU / GPU architectures to improve performances of SQLite. SQLite is an open source embedded relational database which is a key component of many applications and systems. In this study, massively parallel processing is combined with strategic data placement, closer to computing units, with aim of minimizing data transfers between CPU and GPU. Our solution, named CuDB, is an embedded in-memory relational database engine. The SQLite API remains unchanged, allowing developers to migrate easily from SQlite to CuDB even on already existing applications. Thanks to their massive data parallel architecture, GPUs can outperform CPUs when selection queries are applied on large datasets, but as a counterpart, GPUs are not suited for processing small datasets. In order to reach high efficiency in most cases, CuDB embeds a hybrid processing engine which switches dynamically between CPU and GPU processing units according to the data volume it has to process.

According to content and selectivity of queries, measured speedups are between 4 and 400, with a clear trend to reach better speedups when SQL queries are more complex and/or selectivity is sharper. In this study, performances and energy efficiency with different GPU ranges are also compared. Measured processing durations show that even an entry-level GPU can provide noticeable speedups while delivering high energy efficiency.

Keywords: In-Memory DB, embedded RDBMS, GPU, CUDA, SQLite, hybrid architecture, parallelism