Parquet File Writer for FPGA-accelerated Cloud Storage Layer

Jonas Dann, Gustavo Alonso
Systems Group @ ETHZ

Fall Semester 2024

1 Introduction

The volume of data handled by database systems starkly increases as the data produced by humanity grows exponentially. Current database systems, primarily built for general-purpose CPUs, face challenges with processing these increasing data volumes because performance scaling of CPUs is reaching its physical limitations. To address this, CPU vendors have turned to optimizations like instruction pipelining, coarse-grained multi-core and hardware thread parallelism, and vectorization that are themselves stagnating.

Field programmable gate arrays (FPGAs) may be used to alleviate these performance challenges and have been successfully used to accelerate machine learning and data processing workloads. Their efficiency comes from exploiting on-chip pipeline parallelism and data parallelism, leading to highly parallel multiple instructions, multiple data (MIMD) architectures. Among other workloads, the potential for acceleration of disaggregated memory and storage architectures with FPGAs has been shown (e.g., Farview [3] and AWS Redshift AQUA [2]).

Especially in the storage layer of databases, Apache Parquet [5] has become increasingly popular as a compressed file format. The goal of this project is to explore the potential of FPGAs for storage layer acceleration with a Parquet writer. The writer will compress an internal binary representation of database tables (e.g., Apache Arrow [4]) to Parquet files. In the long term, this enables us to build a distributed, FPGA-based storage layer for modern cloud databases that enables pushing down operators (e.g., filtering and repartitioning) towards storage.

2 Prerequisites

While there is flexibility in how the project is executed, we expect the following background and skills to be important for success:

Programming languages: System Verilog, C/C++, Vivado HLS (optional)

Background knowledge: FPGA programming, database systems

Tools: Vivado, git, shell

3 Work plan

The work consists of the following units:

1. Literature review on the data formats Apache Parquet [5] and Apache Arrow [4] and Coyote [1] (the underlying FPGA shell).
2. Design and implementation of a writer for Apache Parquet that transforms a binary database table format (e.g., Apache Arrow) into Parquet files on FPGAs.
3. Benchmark and analyze performance of the Parquet writer to CPU-based Parquet writers (if applicable, GPU-based).
4. Optional: Connect the writer to a Parquet reader to enable operator push-down into the storage layer.

4 Deliverables

The project should result in the following deliverables:

1. Master thesis. This should contain at least the following:
   - Introduction / motivation of the problem
   - Analysis of related work
   - Description of concepts behind design
   - Discussion of the implementation
   - Experimental evaluation
2. Complete source code for:
   - Prototype implementation
   - Benchmarking / data analysis / validation scripts

   The source code should also be accompanied with enough documentation to allow complete reproduction of the experimental results.
3. Presentation of the results and demonstration of functionality

5 Grading

The minimum requirements for a grade of 5.0 are as follows:

- Successful completion of work items 1–3.
- Production of deliverables 1–3 to a satisfactory standard.

The grade will be reduced if these goals are not achieved, except in the case of extreme extenuating circumstances (such as an unforeseeable and unsolvable technical barrier to completing the work, accompanied by an acceptable alternative work item).

A grade of 5.50 or higher will be awarded for the completion of the work to an unusually high quality (i.e., automation of obtained results, scripts to reproduce the work), the addition of extra work (i.e., integrate the design into a larger system), or a particularly novel approach to the problem (i.e., results can be published).

6 Contact

If you are interested in this thesis, please send an email with your transcript of records and CV to Jonas Dann (jonas.dann@inf.ethz.ch) or Prof. Gustavo Alonso (alonso@inf.ethz.ch).

References

[1] Dario Korolija, Timothy Roscoe, and Gustavo Alonso. "Do OS abstractions make sense on FPGAs?" In: *14th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2020, Virtual Event, November 4-6, 2020*. USENIX Association, 2020, pp. 991–1010.

[2] Nikos Armenatzoglou et al. "Amazon Redshift Re-invented". In: *SIGMOD '22: International Conference on Management of Data, Philadelphia, PA, USA, June 12 - 17, 2022*. ACM, 2022, pp. 2205–2217.

[3] Dario Korolija et al. "Farview: Disaggregated Memory with Operator Offloading for Database Engines". In: *12th Conference on Innovative Data Systems Research, CIDR 2022, Chaminade, CA, USA, January 9-12, 2022*. www.cidrdb.org, 2022.

[4] Apache Software Foundation. *Arrow Columnar Format*. `https://arrow.apache.org/docs/format/Columnar.html#arrow-columnar-format`. visited 03/2024. 2024.

[5] Apache Software Foundation. *Parquet Documentation*. `https://parquet.apache.org/docs/`. visited 04/2024. 2024.