

Prefetching in RocksDB/Apache Flink for streaming workloads

FREDRIK BJÖRKMAN

`fbjorkma@kth.se`

December 15, 2021

1 Thesis title

Prefetching in RocksDB/Apache Flink for streaming workloads

2 Background

This project will be carried out at KTH in collaboration with Research Institutes of Sweden (RISE) and Boston University. Modern stream processors rely on embedded key-value stores to manage state that accumulates over long-running computations and exceeds the available memory size. One of these key-value stores is RocksDB [1] which is used by many systems, such as Apache Flink [2]. An important technique in reducing the I/O wait time for disc reads is the use of prefetching data into main memory before it is used. The challenge then arise to determine what data to prefetch, since the memory accesses might not have any obvious patterns. However some unique characteristics arises around streaming state access workloads. The state access for streaming workloads tends to have a high spacial locality [3].

The goal of this project is to leverage this result by designing and implementing a novel prefetching mechanism in RocksDB that is tailored for streaming workloads, that proactively populates the cache with data that has a high probability of being accessed in the near future. To do so, we will need to continuously identify and extract key neighbourhoods from the state access traces of streaming computations. This project will be a continuation on the previous work done by Vasiliki Kalavri and John Liagouris *In support of workload-aware streaming state management* [4].

3 Research question

By designing and implementing an effective prefetch mechanism in RockDB, how will that affect the performance of state access on the streaming data.

4 Hypothesis

The expected outcome of this project is that, if an effective prefetch mechanism is in place, then this will have a positive effect on the performance.

5 Research method

This project will have a quantitative approach, where tests will be performed to measure the performance for state access on the streaming data and compare the performance with and without a prefetch mechanism tailored for streaming workloads.

6 Background of the student

I currently take two courses that are relevant: ID2222 Data mining and ID2223 Scalable machine learning. I have previously taken relevant courses are ID2201 Distributed systems, basic course and ID2203 Distributed systems, advanced course. Other previous courses that might be relevant for the project is DD2350 Algorithms, Data Structures and Complexity, IS1200 Computer Hardware Engineering and IS2202 Computer Systems Architecture (Not completed, but have gained some knowledge about caches, prefetching, RAM etc.)

7 Suggested examiner at KTH

Assistant Professor Paris Carbone (parisc@kth.se). I have been in contact with him and he has expressed an interest in being an examiner for this project.

8 Suggested supervisor at KTH

Doctoral student Sonia-Florina Horchidan (sfhor@kth.se). I have been in contact with her and she has expressed an interest in being a supervisor for this project.

9 Resources

This project as said in Section 2 will be built upon previous work done by Vasiliki Kalavri and John Liagouris *In support of workload-aware streaming state management* [4]. The project will be able to use an already build testbed built by Boston University for the performance testing on state access for streaming workloads. There are also available data sets for streaming data that can be used for testing of the performance. Both RISE and Boston University have expressed that they are able to assist with some material and support that might be needed in this project.

10 Eligibility

I have a complete bachelor. Currently I am taking my last two complete courses on my master, ID2222 Data mining and ID2223 Scalable machine learning. I have only one leftover part of a course from the master at the moment, but according to my study plan, when I begin the master thesis project all the courses will be completed.

11 Study Planning

I plan to finish all the courses before starting the master thesis project in January. That is the two courses I currently take and also an essay in ID2012 Ubiquitous Computing.

References

- [1] D. B. et al, “RocksDB,” *GitHub*, 12 2021, [Online; accessed 13. Dec. 2021]. [Online]. Available: <https://github.com/facebook/rocksdb/wiki/RocksDB-Overview>
- [2] “Apache Flink: What is Apache Flink? — Applications,” 9 2021, [Online; accessed 14. Dec. 2021]. [Online]. Available: <https://flink.apache.org/flink-applications.html>
- [3] S. Somogyi, T. Wenisch, A. Ailamaki, B. Falsafi, and A. Moshovos, “Spatial memory streaming,” in *33rd International Symposium on Computer Architecture (ISCA’06)*, vol. 34, no. 2. IEEE Computer Society, 2006. doi: 10.1109/ISCA.2006.38 pp. 252–263.
- [4] V. Kalavri and J. Liagouris, “In support of workload-aware streaming state management,” in *12th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage 20)*. USENIX Association, 7 2020. [Online]. Available: <https://www.usenix.org/conference/hotstorage20/presentation/kalavri>

Acronyms

RISE Research Institutes of Sweden