CS223 Project – Spring 2024

In this project, you are asked to implement a transaction processing system for a distributed system. You will implement the transaction processing protocol described in the following paper:

Zhang, Yang, Russell Power, Siyuan Zhou, Yair Sovran, Marcos K. Aguilera, and Jinyang Li. "Transaction chains: achieving serializability with low latency in geo-distributed storage systems." In Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles, pp. 276-291. 2013.

You need to implement the transaction processing component of the paper (but you do not need to implement the part about the language and static analysis of transactions and automatically identifying them.) Instead, you will come up with your own transactions manually that satisfy the properties of transaction chains.

The steps you need to take for this project:

1. Read and understand the theory of transaction chopping and SC-Cycles.
2. Propose an application that you will implement. Write the table schemas and a set of transactions that correspond to the tasks and requests handled by the application. Your application must have at least three tables and at least 7 transactions. The SC-Graph of 6 out of the 7 transactions must have NO sc-cycles. However, the SC-Graph of the 7 transactions together leads to an SC-cycle.
3. Propose a partitioning strategy across at least three nodes. Try to make the partitioning strategy as efficient as possible. Discuss the efficiency goals you took into consideration. One such consideration is to try to make the first hop of transactions be closest to the users issuing the request for a transaction.
4. Implement a transaction chopping transaction processing layer to process the transactions and applications you proposed across three nodes (these can be virtual nodes in the same physical machine).
5. What will happen if we use direct edges between siblings? Is it still correct to consider directed cycles as SC-cycles? If not, how can we address this issue? Please propose a solution under the following conditions:
    a. Transactions with at most two hops
    b. Any arbitrary transactions
6. Implement parts 5.a and 5.b in your transaction processing layer. This can be integrated into your existing transaction processing layer as an additional option.
7. Write a report of at least 7 pages (acm conference style). A suggested outline is:
    - Abstract and introduction 1 to 1.5 pages.
    - Background (about transaction chopping and any other related concepts, and a related work subsection talking about research relevant to this area): ~1 page.
    - Application design, describing the application in terms of the table schemas, partitioning of the data, and transactions (and how a subset are free from sc-cycles): ~1–1.5 pages.

- System design describing how you implemented the transaction chopping protocol and directed sibling edges: ~2–3 pages.
- Evaluation, describing experimental evaluation results of the system's latency and throughput while varying different parameters (e.g., the size of the database, the number of concurrent requests): 1.5–2 pages.
- Conclusion: 0.25 pages
- References: ~0.5 to 1 page

Tentative Timeline:

- Presentations: June 6.
- Final report: June 7.