Analysing Query Performance on Different Data Layouts in Cloud Object Stores

Sunil Mallireddy (sunil.mallireddy@wisc.edu)
John Shawger (jshawger@wisc.edu)
Olesia Elfimova (oelfimova@wisc.edu)

1 Problem Description

In the past decade there was a rise of various massively scalable storage layers over cloud object stores, such as data lake [4], delta lake [1], and lakehouse [2] (these technologies have differences but this is not the focus of the project, accordingly we assume they are interchangeable in the rest of the paper). They are superior than traditional data warehouse architectures because of higher durability, cheaper cost, and ability to store unstructured data. SQL is a popular and familiar way to query data lakes. Active area of systems and databases research is focusing on various optimization techniques to improve performance of queries over data lakes. For example, Armbrust et al. [2] state that determining which data layout strategies are likely to be most effective for massive datasets in cloud object stores is an open question.

In this project we plan to investigate different data layout optimizations in cloud object stores and compare them in terms of SQL query performance.

2 Outline of Approach

Optimizations strategies we plan to investigate:

- **Z-ordering** Delta Lake provides Z-Ordering and claims that it reduces I/O on certain types of queries. Z-ordering combines multiple columns of data into a single scalar value, which can be queried by range. We would like to evaluate Z-ordering more closely, determine which types of queries it is most useful for, and measure overall system performance compared to other approaches.
- **Small File Compaction** Having a large number of small files in a data lake (rather than larger files optimized for analytics) can slow down performance considerably due to limitations with I/O throughput. Accordingly, Armbrust et al. suggest to consolidate small files into larger ones of size 1 GB that are optimized for read access. We plan to vary the default size value and see if there are query performance benefits.
- **Block Skipping** Skipping data reduces both I/O and CPU usage while executing a query [5] [6]. Block skipping involves maintaining metadata blocks of records, and skipping the entire block if it is outside a query. Comparing an approach like this to the above two strategies may yield different performance benefits on certain types of queries or datasets. To evaluate the effectiveness of block skipping, we would implement a version of the algorithm and incorporate it in an existing system like Delta Lake.
- **Data Clustering and Zone Maps** Zone maps provide query optimization by clustering data in the table. They are especially effective when clustered by filtering columns [7]. The fact table data can be clustered by columns of dimension tables together with zone maps, which maintain min/max value ranges of these clustering columns over zones of fact table data. In that case, the performance of snowflake queries in data warehouses can be significantly improved. We would like to explore whether this approach can be applied in cloud object stores.

Benchmarks for query performance evaluation [3]:

- TPC-H, TPC-DS, YCSB

Timeline:

- March 14 - Have system setup ready, including rent of cloud instance, installation of open source delta lake project [4], ingestion of data in a lake.
- March 28 - Implement optimization strategies in our system.
- April 11 - Conduct preliminary analysis and record query performance on different benchmarks.
- May 1 - Finalize presentation slides.
- May 8 - Finalize the report.

References

[1] M. Armbrust, T. Das, L. Sun, B. Yavuz, S. Zhu, M. Murthy, J. Torres, H. van Hovell, A. Ionescu, A. Łuszczak, M. undefinedwitakowski, M. Szafrański, X. Li, T. Ueshin, M. Mokhtar, P. Boncz, A. Ghodsi, S. Paranjpye, P. Senster, R. Xin, and M. Zaharia. Delta lake: High-performance acid table storage over cloud object stores. *Proc. VLDB Endow.*, 13(12):3411–3424, aug 2020.

[2] M. Armbrust, A. Ghodsi, R. Xin, and M. Zaharia. Lakehouse: a new generation of open platforms that unify data warehousing and advanced analytics. In *Proceedings of CIDR*, 2021.

[3] M. Barata, J. Bernardino, and P. Furtado. Survey on big data and decision support benchmarks. In H. Decker, L. Lhotská, S. Link, M. Spies, and R. R. Wagner, editors, *Database and Expert Systems Applications*, pages 174–182, Cham, 2014. Springer International Publishing.

[4] R. Hai, S. Geisler, and C. Quix. Constance: An intelligent data lake system. In *Proceedings of the 2016 International Conference on Management of Data*, SIGMOD '16, page 2097–2100, New York, NY, USA, 2016. Association for Computing Machinery.

[5] L. Sun, M. J. Franklin, S. Krishnan, and R. S. Xin. Fine-grained partitioning for aggressive data skipping. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 1115–1126, 2014.

[6] L. Sun, M. J. Franklin, J. Wang, and E. Wu. Skipping-oriented partitioning for columnar layouts. *Proceedings of the VLDB Endowment*, 10(4):421–432, 2016.

[7] M. Ziauddin, A. Witkowski, Y. J. Kim, D. Potapov, J. Lahorani, and M. Krishna. Dimensions based data clustering and zone maps. *Proceedings of the VLDB Endowment*, 10(12):1622–1633, 2017.