

# Query-driven compaction in LSM-trees

Shubham Kaushik  
Boston University  
USA

\*\*\*\*  
Boston University  
USA

\*\*\*\*  
Boston University  
USA

## ABSTRACT

In modern data systems, particularly for write-intensive workloads, log-structured merge (LSM) trees have become a most widely used technique. However, current LSM-tree designs do not capitalize on the sort-merge operations performed for the range queries, which leads to carry out almost same amount of work for each range query, even though they are same or overlapping. LSM-trees excel in write-heavy workloads due to their concept of out-of-place updates, which also leave invalid keys across multiple levels.

In this paper, we have presented a query-driven compaction strategy which removes the invalid (logically deleted) keys from the LSM-tree and flush back the valid keys which are filtered by sort-merge performed during range query. Performance experiments with the help of custom implementation in one of the popular LSM-based data store, Rocksdb.

### ACM Reference Format:

Shubham Kaushik, \*\*\*\*, and \*\*\*\*. 2023. Query-driven compaction in LSM-trees. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

- 1 INTRODUCTION
- 2 LSM BACKGROUND
- 3 PROBLEM
- 4 SOLUTION NAME
- 5 EVALUATION
- 6 RELATED WORK
- 7 CONCLUSION

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*Conference'17, July 2017, Washington, DC, USA*

© 2023 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>