CAS CS 561: Data Systems Architectures
Data-intensive Systems and Computing Lab
Department of Computer Science
College of Arts and Sciences, Boston University
http://bu-disc.github.io/CS561/

CS561 Spring 2024 - Research Project

Title: Quantifying *Write Amplification of LSM-tree-based Key-Value Stores with Solid-State Disks*

Background:
Log-structured merge tree (*i.e.,* LSM-tree)-based key–value stores are very popular because of their good write and read performance. NAND Flash-based Solid-state disks (*i.e.,* SSDs) have become a popular storage device alternative to hard disk drives (*i.e.,* HDDs) because of their high performance and low power consumption. LSM-tree KV stores with SSDs are deployed in many large-scale storage systems, aiming to achieve high cloud performance. However, the 'log-on-log' problem [1] can cause severe write amplification both in LSM-trees and NAND SSDs. The former, which is attributed to compaction operations in LSM-tree-based KV stores, is a burden on I/O bandwidth between the host and the device. The latter, which results from out-place updates in NAND Flash memory, blocks user I/O requests between the host and NAND Flash memory, thereby degrading the SSD performance. In this project, we first attempt to quantify the cascaded write amplification of LSM-trees running on SSD. Through this analysis, we will then attempt to derive new data placement policies for modern open-channel SSDs [2] or ZNS SSDs [3, 4] so that we can reduce write amplification, consequently improving the overall performance of LSM trees on SSDs.

Objective:
(a) Study the basics of LSM tree and SSD characteristics.
(b) Empirically quantify the write amplification of LSM-trees running on SSD. We can either develop a simulation framework or quantify it directly on RocksDB.
(c) Go over the principles of ZNS SSD or Open-Channel SSD. Research how these new devices' can be exploited to minimize the write amplification of LSM on SSD.

Responsible Mentor: *Tarikul Islam Papon*

References:
[1] Yang, J., Plasson, N., Gillis, G., Talagala, N., & Sundararaman, S. (2014). Don't stack your log on my log. In *2nd Workshop on Interactions of NVM/Flash with Operating Systems and Workloads ({INFLOW} 14)*.
[2] Wang, P., Sun, G., Jiang, S., Ouyang, J., Lin, S., Zhang, C., & Cong, J. (2014, April). An efficient design and implementation of LSM-tree based key-value store on open-channel SSD. In *Proceedings of the Ninth European Conference on Computer Systems* (pp. 1-14).
[3] Purandare, D., Wilcox, P., Litz, H., & Finkelstein, S. (2022, January). Append is near: Log-based data management on zns ssds. In *12th Annual Conference on Innovative Data Systems Research (CIDR'22)*.

[4] Lee, H. R., Lee, C. G., Lee, S., & Kim, Y. (2022, June). Compaction-aware zone allocation for LSM based key-value store on ZNS SSDs. In Proceedings of the 14th ACM Workshop on Hot Topics in Storage and File Systems (pp. 93-99).