Can Lossy Compression Benefit NVMe-based I/O?

Darren Ng[1], Duo Zhang[1], Sheng Di[2], Zhaorui Zhang[3], and Xiaoyi Lu[1]

University of California, Merced[1]
Argonne National Laboratory[2]
The Hong Kong Polytechnic University[3]

UNIVERSITY OF CALIFORNIA MERCED

PADSYS Lab

Introduction

- As scientific simulations continue to scale, modern HPC applications rapidly increases the massive amounts of data it generates, often reaching terabytes or even petabytes.
  - Scientific floating-point simulations such as CESM for climate modeling [1] and HACC for cosmology [2] can produce over 1 PB of data per run, placing immense pressure on storage system to meet both capacity and performance demands.
- NVMe-SSDs exhibit 10.93X and 34.66X faster write and read speeds than HDDs
  - However, combining them with lossy compression introduces unique challenges.
- Lossy compression improves I/O performance by about 3X on SCSI-HDDs (the first two red bars)
  - But degrades NVMe-SSD performance by factors of 4X and 8X for write and read operations, respectively (the teal bars).
- Consequently, a critical question arises:
  - *Can lossy compression truly enhance I/O performance on NVMe-SSDs, or does it introduce new performance constraints?*

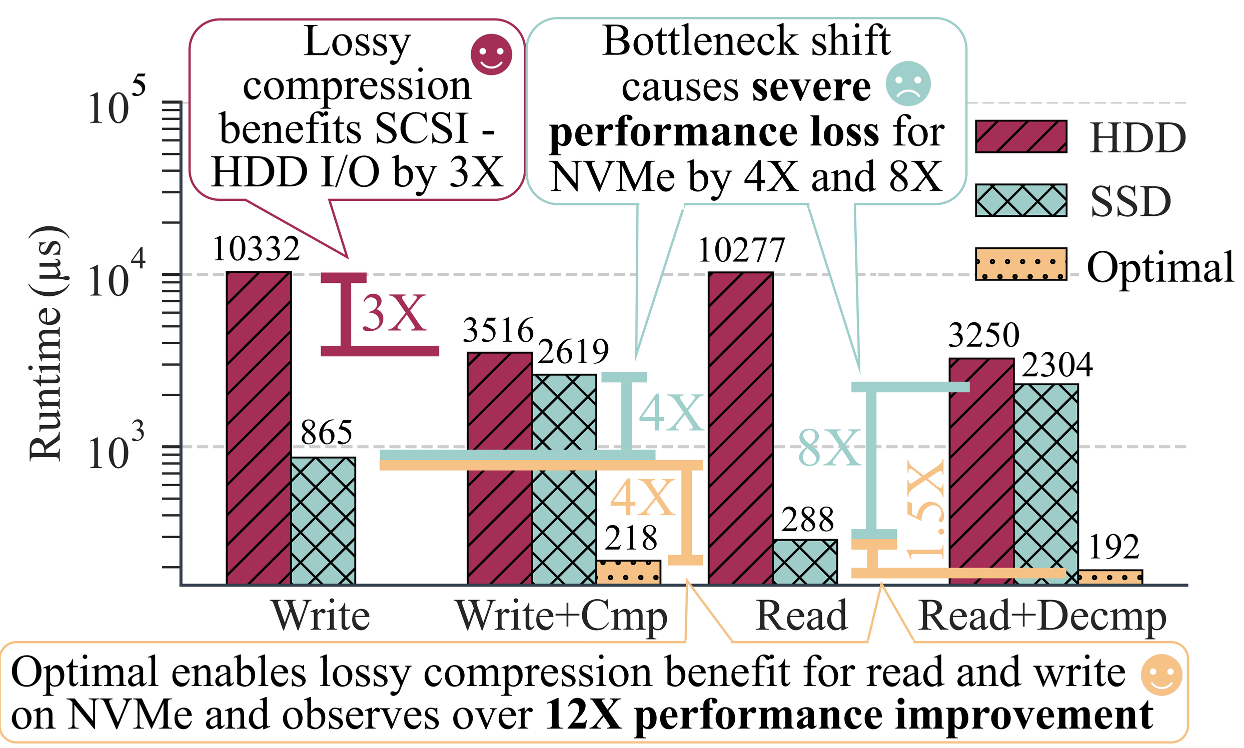

Fig. 1: Total runtime of data storage w./wo. (denoted by "+") SZx (de)compression (Cmp and Decmp) for SCSI-HDD (HDD) and NVMe-SSD (SSD).

Motivation

- Our study reveals that both the compression and decompression stages dominate the overall runtime in data write and read I/O processes respectively, leading us to the question:
  - Where exactly do these bottlenecks occur?
- Our analysis uncovers three primary issues
  - 1) Dataset Characteristics: Compression and decompression time depend heavily on dataset features (e.g., data type). However, the current I/O operation paradigm often overlooks this factor.
  - 2) Control Mechanism: Existing error-bounded approaches do not effectively enhance I/O performance, whereas using a fixed compression ratio simplifies control over the compressed data size and offers a pathway for future performance improvements.
  - 3) Suboptimal NVMe Configurations: Improperly tuned NVMe parameters, such as the number of threads, can lead to significant performance variability. Proper optimization of these settings ensures efficient I/O operations and maximizes overall performance.
- Our study highlights the crucial need to balance dataset-specific characteristics, computational efficiency and I/O performance to maximize overall NVMe-based storage performance.
- With this design, we aim to answer a third crucial question:
  - *How can we optimize this process to mitigate performance losses?*
- To address this, we breakdown the lossy compressed I/O pipeline and perform a parameter study on multiple lossy compressors.

References

[1] Flavio Lehner, Fortunat Joos, Christopher C Raible, Juliette Mignot, Andreas Born, Kathrin M Keller, and Thomas F Stocker. Climate and carbon cycle dynamics in a cesm simulation from 850 to 2100 ce. Earth system dynamics, 6(2):411–434, 2015.
[2] Katrin Heitmann, Thomas D Uram, Hal Finkel, Nicholas Frontiere, Salman Habib, Adrian Pope, Esteban Rangel, Joseph Hollowed, Danila Korytov, Patricia Larsen, et al. Hacc cosmological simulations: First data release. The Astrophysical Journal

Characterization and Findings

- We firstly observe the time breakdown of the entire serial pipeline.
- FIO Storage parameters:
  - block size of 4KB
  - I/O depth of 16
  - number of jobs to 1
- We find that lossy compression occupies a substantial amount of the runtime
  - NVMe read/write is visually miniscule

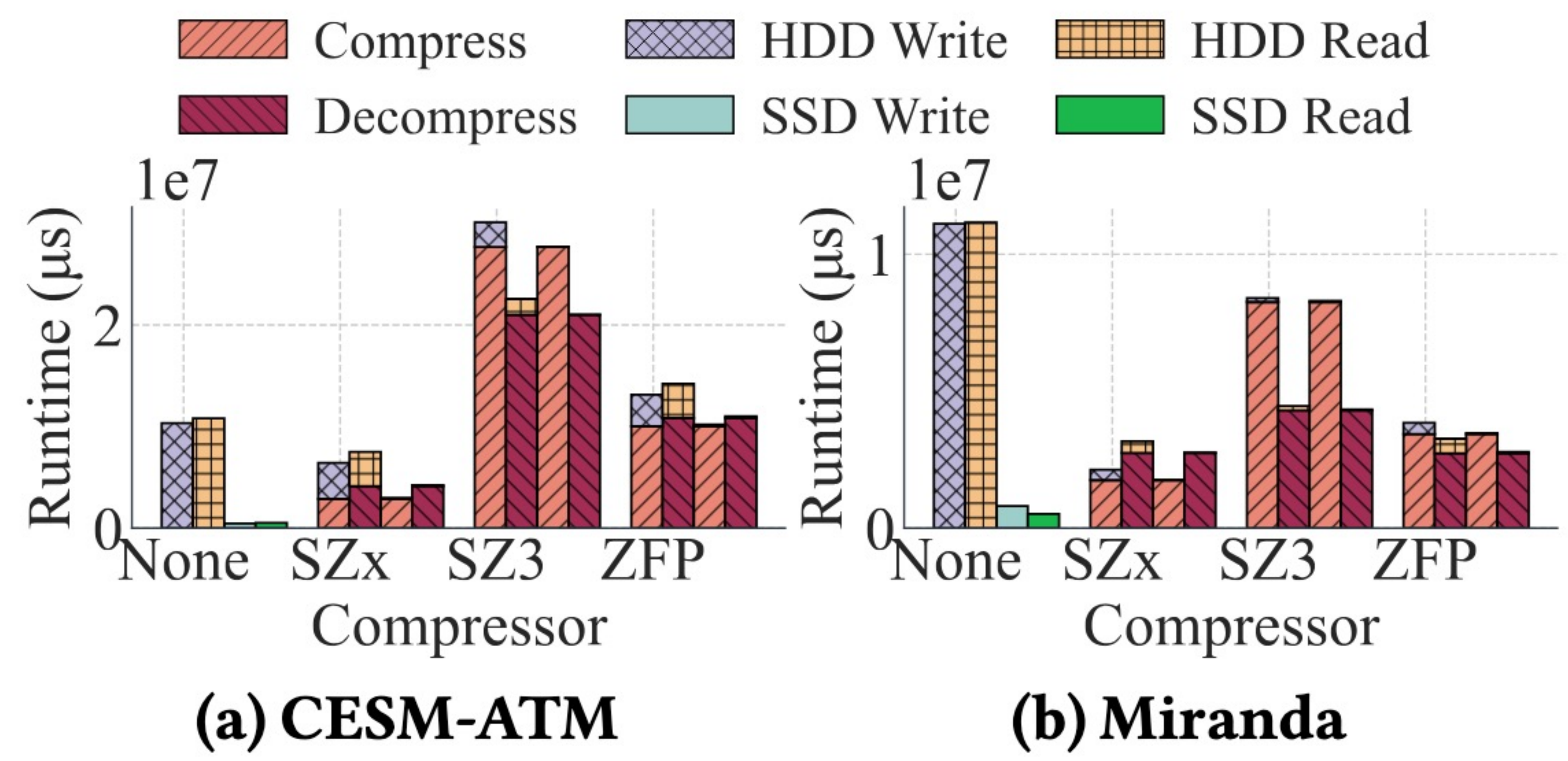

Fig. 2: Time breakdown of storing and loading various scientific datasets with HDDs and SSDs

Finding 1: Writing the dataset directly to the NVMe SSD often requires less runtime compared to utilizing lossy compression

Finding 2: The performance of each compressor is heavily influenced by the dataset's characteristics.

- Second, we perform a parameter study on lossy compressors
- We evaluate the distribution of different error-bounded methods
  - Absolute error bound
  - Relative error bound
  - Fixed ratio bound
- We find that fixed ratio bounding provides the best compression ratio distribution
  - This offers better control on compression output

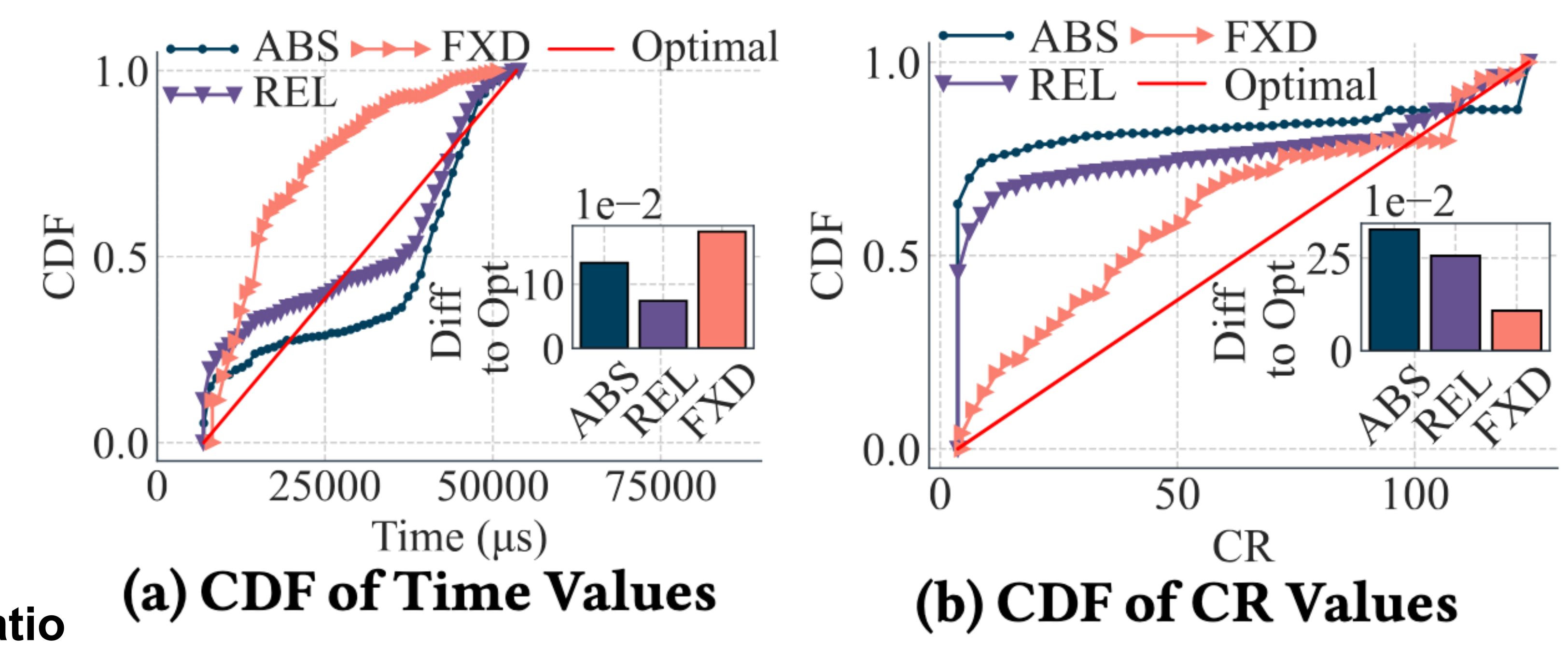

Fig. 3: CDF of Compression cost (a) and compression ratio (b) when bounding SZx by absolute error, relative error bounds and fixed compression ratio

Finding 3: Fine-grained compression ratio control through error-bounding tuning is difficult

Conclusion and Future Work

- In this poster, we examined the challenges of integrating lossy compression with high-performance NVMe-SSDs, highlighting how the computational overhead of compression can diminish its storage efficiency benefits in high-speed storage system.
- In future work, we aim to apply our insights to an informed design.

Acknowledgement

This work was supported in part by NSF research grants (#2321123, #2340982, and #2505106) and a DOE research grant DE-SC0024207. Results in this paper were obtained using the CloudLab testbed. We also thank our lab mate Seoyeong Lee for her valuable feedback.