Approachable Error-Bounded Lossy Compression

Robert Underwood †, Jon Calhoun †, Amy Apon †, Franck Cappello ‡
†Clemson University ‡ Argonne National Laboratory

Introduction

As data continues to grow, HPC applications need a way to make data smaller to make it easier to work with. Compression can be the answer.

	Lossless	Lossy	EBLC
Example	ZIP	JPEG	SZ, ZFP, MGARD
High	X	+	+
Ease of Use	+	X	~
Accuracy	+	X	+

My Dissertation: provide a uniform interface to compression and develop tools to allow application scientists to understand and apply Error Bounded Lossy Compression (EBLC)

LibPressio: Standard Interfaces for Compression and Analysis

Problem: Compressors and analysis codes don't agree on a common interface leading to a divserity of interfaces
Solution: LibPressio -- a library to abstract away the differences
Key Insights
A common interface needs:
- a common data abstraction
- standard settings semantics

Meta Compressors boost productivity:
- Provide services atop compressors (i.e. pre-processors, auto tuners, parallel runtimes)

Future Work:
- Support for better accelerator sharing
- Support for asynchrony/streams
- Better support for sparse problems

Automated Configuration of Compressors

FRaZ: Fixed Ratio Compression

Question: Can we tune compression using a control loop to bound compression ratio?
Approach: Use LibPressio compressor interface and derivative free optimization techniques to formulate compression configuration as an optimization problem
Key Findings:
- Binary Search doesn't work because the relationship is not monotonic (Fig. 1).
- Tuning takes only ≈2x longer than an oracle when an answer existed
- Some targets are faster to find because more error bounds meet the target (Fig. 2).

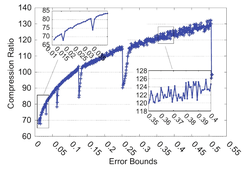

Figure 1: Error Bound vs Comprssion Ratio for SZ on QCLoud is Non-Monotonic enfeibling binary search[1]

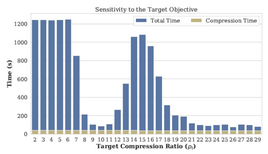

Figure 2: Runtime for Various Different Targets on Hurricane [1]

LibPressio-Opt: Bound User Metrics

Can we extend FRaZ to bound simple user metrics generally?
Approach: Convert FRaZ to a meta-compressor
Use LibPressio metrics interface:
- Allow inter-iteration early termination
- Add new "max" and "min" modes

Key Findings:
LibPressio-Opt has less overhead than:
- FRaZ - due to inter-iteration termination
- MGARD-QOI - due to less pre-compution

We can even emulate QOI mode with less overhead

More Complex Apps: HACC (On Going)

Can we bound complex metrics such as power spectra that use multiple buffers as in HACC?
Background:
- HACC –ECP astrophysics particle application
- No compressors bounds spectral error [2]

Approach:
- Implement the spectra as a metric
- Explore metrics to compare spectra (i.e. log-spectral distance, itakura-saito distace)
- Adapt LibPressio-Opt for multiple buffers
- Solve maximum compression ratio such that differences between spectra are acceptable

Future Work
- True Multi-Objective Compression
- Improve the search algorithm

Better compression task scheduling

Understanding the Effects of Compression

What are trade-offs for compressing training and testing data to save storage space?
Approach: Compare compression ratio vs quality trade-offs for many apps and benchmarks using LibPressio
Key Findings:
- SZ and ZFP outperforms common methods even for imbalanced problems (Fig. 3)
- In some cases, EBLC has better results than no compression due to reduced over-fitting

Future Work:
- Effect of EBLC for data transfers between CPU/ accelerator

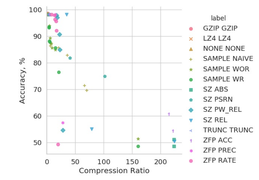

Figure 3: Accuracy vs Compression Ratio Trade-Offs for Candle-NT3. SZ and ZFP outperform sampling methods [3]

Conclusion

Error Bounded Lossy Compression has the potential to be transformative in a wide variety of domains if application scientists have the tools that they need to interface with new compressors and experiment, understand the effects of lossy compression on their data, and have tooling to support them along the way. This work provides a strong foundation to enable further research in lossy compression and enable improvements to scientific applications.

Questions: robertu@g.clemson.edu

Ciriculum-Vita

Get LibPressio

Watch This Talk

Funding Acknowledgements

This research was supported by the Exascale Computing Project (ECP),Project Number: 17-SC-20-SC, a collaborative effort of two DOE organiza-tions - the Office of Science and the National Nuclear Security Administration,responsible for the planning and preparation of a capable exascale ecosystem, including software, applications, hardware, advanced system engineering andearly testbed platforms, to support the nation's exascale computing imperative.The material was supported by the U.S. Department of Energy, Office ofScience, under contract DE-AC02-06CH11357, and supported by the NationalScience Foundation under Grant No. 1619253 and 1910197.We acknowledge the computing resources provided on Bebop, which isoperated by the Laboratory Computing Resource Center at Argonne NationalLaboratory.This material is also based upon work supported by the U.S. Department ofEnergy, Office of Science, Office of Workforce Development for Teachers andScientists, Office of Science Graduate Student Research (SCGSR) program.The SCGSR program is administered by the Oak Ridge Institute for Scienceand Education (ORISE) for the DOE. ORISE is managed by ORAU undercontract number DE-SC0014664. All opinions expressed in this paper are theauthors and do not necessarily reflect the policies and views of DOE, ORAU,or ORISE

References

[1] R. Underwood, S. Di, J. C. Calhoun and F. Cappello, "FRaZ: A Generic High-Fidelity Fixed-Ratio Lossy Compression Framework for Scientific Floating-point Data," 2020 (IPDPS)
[2] Habib, Salman, et al. "HACC: Simulating sky surveys on state-of-the-art supercomputing architectures." New Astronomy 42 (2016): 49-65.
[3] R. Underwood et al. "Machine Learning and AI with Error Bounded Lossy Compression" In Preparation