

Extreme Compression for Large Scale Data Store

Jérôme Lauret^{1,*}, Juan Gonzalez², Gene Van Buren¹, and Rafael Nuñez²

Co-authors: Philippe Canal³ and Axel Naumann⁴

¹Brookhaven National Laboratory, P.O. Box 5000, Upton, NY 11973-5000, USA

²Accelogic, LLC, 1633 Bonaventure Blvd, Weston, FL 33326, USA

³Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510-5011, USA

⁴Conseil Européen pour la Recherche Nucléaire, Esplanade des Particules 1, P.O. Box 1211, Geneva 23, Switzerland

Abstract. For the last 5 years Accelogic pioneered and perfected a radically new theory of numerical computing codenamed "Compressive Computing", which has an extremely profound impact on real-world computer science [1]. At the core of this new theory is the discovery of one of its fundamental theorems which states that, under very general conditions, the vast majority (typically between 70% and 80%) of the bits used in modern large-scale numerical computations are absolutely irrelevant for the accuracy of the end result. This theory of Compressive Computing provides mechanisms able to identify (with high intelligence and surgical accuracy) the number of bits (i.e., the precision) that can be used to represent numbers without affecting the substance of the end results, as they are computed and vary in real time. The bottom line outcome would be to provide a state-of-the-art compression algorithm that surpasses those currently available in the ROOT framework, with the purpose of enabling substantial economic and operational gains (including speedup) for High Energy and Nuclear Physics data storage/analysis. In our initial studies, a factor of nearly x4 (3.9) compression was achieved with RHIC/STAR data where ROOT compression managed only x1.4.

In this contribution, we will present our concepts of "functionally lossless compression", have a glance at examples and achievements in other communities, present the results and outcome of our current, ongoing R&D, as well as present a high-level view of our plan to move forward with a ROOT implementation that would deliver a basic solution readily integrated into HENP applications. As a collaboration of experimental scientists, private industry, and the ROOT Team, our aim is to capitalize on the substantial success delivered by the initial effort and produce a robust technology properly packaged as an open-source tool that could be used by virtually every experiment around the world as means for improving data management and accessibility.

1 Introduction

1.1 Challenges

Existing high energy and nuclear physics (HENP) experiments are producing petabytes of analysis data per year, and the historical growth of that produced data has generally been

*e-mail: jlaret@bnl.gov

exponential. Analyses that use these data thrive on rapid (or "live") access not only to data under current production, but to prior years' accumulated data as well. However, sufficiently performant live storage and networking infrastructure to deliver on those demands are costly, justifying investigations of alternative solutions to throwing money at the infrastructure.

The problem is not a new one for this field, and solutions have of course been implemented to work within the infrastructure that's been economically available. Examples include administrative procedures such as limiting the portions of datasets of interest which are made accessible on live storage, and limiting the duration of their accessibility. But none of these administrative solutions are ideal as they inherently restrict access in one way or another, and waiting for access is not time nor money well spent.

Neither is the problem expected to simply go away as infrastructure capabilities per unit cost come down over time: experiments aim to multiply produced datasets by factors of x10-to-x100 in the coming years [2]. All of this points to the need to further explore an engineering solution: data compression.

1.2 Compression and ROOT I/O

A very popular tool in the HENP community to store and access a wide variety of data has been ROOT's TTree, which provides optional lossless compression, recognized from the beginning as a key capability [3]. For many years, ZLIB (widely available as the ubiquitous gzip compression tool) has delivered that compression, and has been generally considered to be a reasonable compromise of file size reduction and speed. The algorithm at ZLIB's core, DEFLATE, is approaching three decades old [4].

In recent years, other possible lossless compression tools have been investigated as alternatives to ZLIB for ROOT storage, with the more recent LZ4 and ZSTD showing sufficient benefits over ZLIB to merit integration [5]. While other lossless algorithms are more aggressive, they achieve only small additional gains at a high price in compression CPU time. LZMA, for example, can take nearly an order of magnitude longer than DEFLATE for only a few percent additional size reduction. These alternatives are not explored in the study presented here. Lossy compression has also been made available in ROOT in the form of the flexible Float16_t and Double32_t data types, which provide manually configurable specification of retained resolution [6].

But the time has come to ask whether, and if so how, we can change the game. What else can be achieved with lossless compression? Can lossy compression, an obvious path to even greater reduction, satisfy the community needs?

2 Physics-Aware Compression

2.1 Basic community requirements

Here we outline some requirements for a broad spectrum of users within the HENP community that a compression tool should fulfill:

- *Data integrity*: no loss of physics!
- *Automatic high compression*: target factors (up to x4-x9) that far outperform currently used methods, without manual tuning
- *Negligible additional CPU overhead*: decompression speed competitive with that of ZLIB (historically used in ROOT) at its maximal compression (i.e. 'gzip -9'), and potentially significantly faster
- *Breadth of applicability*: suitable for any HENP experiment

- *Portability and non-obsolescence*: platform agnostic (e.g. HPC)
- *Open source approach*: no cost to the community
- *User empowerment*: anyone can use and control

These requirements do not exclude the prospect that awareness about features of the data can be leveraged in compression. Data type selection is a typical example of how such awareness has been used, but there is expertise not well known in the HENP community on how to go beyond that simple approach.

2.2 Accelogic strategies: target "zibbits" and redundancies

Accelogic¹, a team of proven experts in Compressive Computing, has pioneered fast and robust technologies to discover un-useful data fragments [1]. Straightforward examples of dispensable bits are those below the finite resolution inherent in the data before arriving at the compression algorithm, or the finite resolution needs of some calculations. To elucidate the latter, consider a quantity which may only ever be added to a second quantity always several orders of magnitude larger, rendering the finest resolution bits of the first quantity inconsequential. Accelogic's technologies bring to the table automatic identification of such zero-information bearing bits ("zibbits") for truncation. And the ability to fold in physics knowledge about the data enables further elimination of hidden redundancies.

Accelogic's techniques are already revolutionizing data compression for other patrons, such as NASA and its Cart3D software used to conduct simulations of air vehicles via computational fluid dynamics. Injection of Compressive Computing algorithms to smartly remove useless floating-point bits has recently accelerated the software's numerical communication routines by factors larger than x3 [7].

The possibility to incorporate Compressive Computing techniques in the ecosystem of HENP computing has arisen. This project with ROOT I/O aims at consolidating the theory and application of Compressive Computing to boost storage capacity of floating-point numbers, thus extending the scope of applications of the technology. In the next section, we present initial test results from this ongoing project.

3 Preliminary Results

3.1 Single variable example: particle identification (dE/dx $n\sigma$ distribution)

Energy loss (dE/dx) in the STAR TPC is used in particle identification by examining the number of standard deviations ($n\sigma$) observed signals are away from expectations for any species. We found that lossless ZLIB compression of 32-bit floating-point numbers ("floats"), which we label as our reference, provided a factor of x1.6 reduction over uncompressed data for this single variable. Accepting some lossy-ness (evident in Fig. 1) by using Float16_t [nbits=16, the number of persistent bits during I/O] plus ZLIB compression improved the reduction to x2.0. However, using the histogram as feedback for tuning the aggressiveness, Accelogic's "zibbit compression" on the floats raised the bar to x3.2 reduction without showing apparent losses here (no visible difference between the zibbit-compressed and reference histograms).

¹See [online] <http://accelogic.com> for more about Accelogic, LLC.

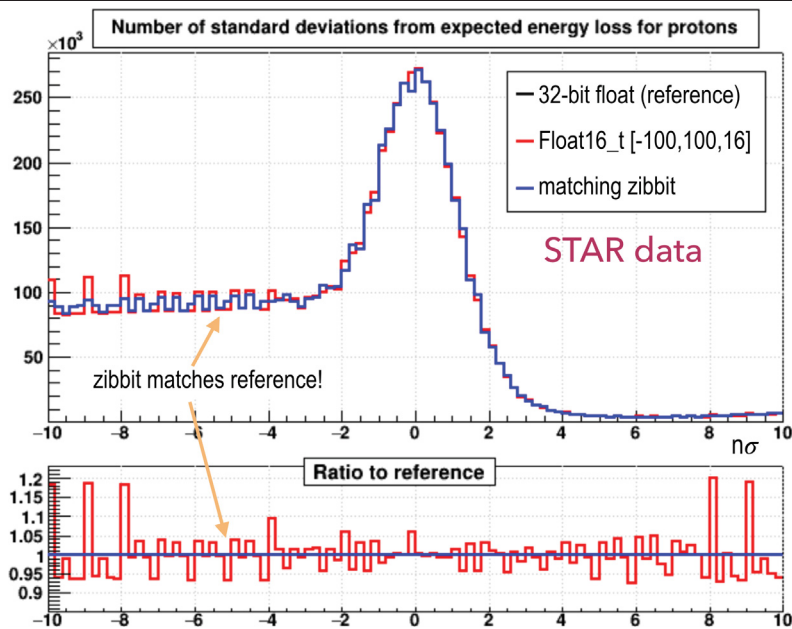


Figure 1. Example of compression impacts on the distribution of a single variable ($dE/dx \ n\sigma$) using three different compression strategies. Note that the black reference histogram is obscured by the blue matching zibbit histogram, reflected in the flat ratio of the two in the lower plot.

3.2 Multi-variable example: invariant mass reconstruction

For reconstruction of invariant masses from decay daughters with the STAR experiment at RHIC, analyses typically involve numerous stored quantities. Some are used in cuts (e.g. track projections, PID) for daughter track candidate selections, and some are used in the mass calculation itself (e.g. momenta components). Not all of these variables compress equally! For this study, we recorded the size of the full data files under different compression scenarios. Again using floats + ZLIB compression for reference, there was a reduction of x1.4 over an uncompressed file. We found that "zibbit compression" could be optimally tuned to achieve x1.9 reduction while still matching the reference histogram in reconstructing the invariant mass, as shown in Fig. 2. Moving into lossy compression with the Accellogic technology showed modification of the histograms, but no modification of the extracted physics (impacts are below the level of noise/uncertainty already present in the data), at a factor of x3.9 in file size reduction, graphed in Fig. 3. Even more aggressive lossy compression of course eventually degrades physics.

4 Plans

There are numerous steps remaining to complete this phase of the effort. These will include moving forward with auto-tuning, arming it with the intelligence to be suitable for broad applications. Optimizing the speed of the compression and decompression are important as well. We can also consider implementation and testing of further advanced Accellogic compression techniques.

A critical set of tasks for the HENP community will be the additional work involving collaboration with the ROOT Team developers on this project. That work must focus on

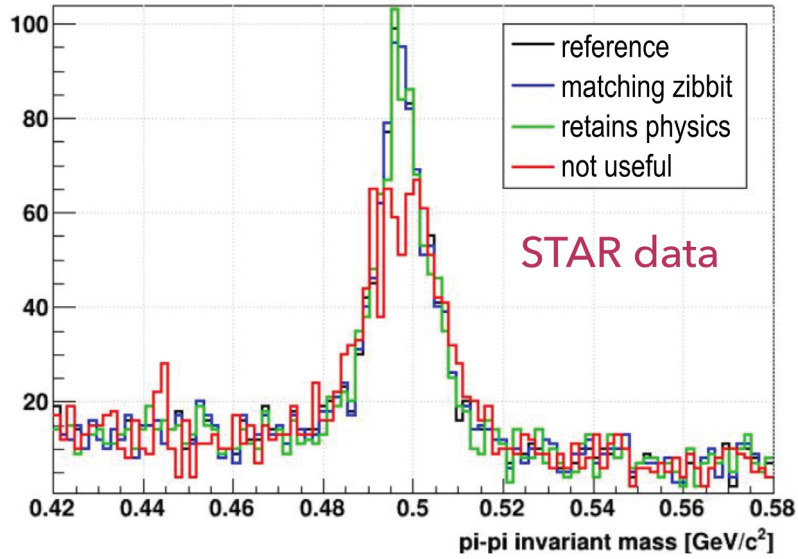


Figure 2. Example of compression impacts on a distribution dependent on many stored variables (invariant mass reconstructed from decay daughters) using three different degrees of aggressiveness for zibbit compression. The black reference histogram is generally obscured by the blue matching zibbit histogram. Zibbit compression aggressiveness can be increased to a level that still usefully retains the physics encapsulated in the histogram despite noticeably modifying the data (green histogram), or eventually to a level that results in loss of physics (red histogram).

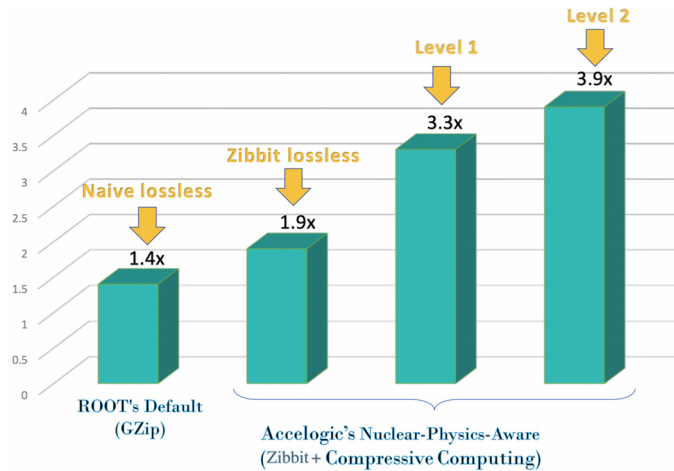


Figure 3. Comparison of file-wise size reduction achieved by various data compressions. The impacts of the "Zibbit lossless" and "Level 2" compressions vs. the reference historical ROOT default are demonstrated in Fig. 2.

fulfilling the requirements on straightforward integration into ROOT I/O, ensuring plug & play portability, and provisioning for user access and control. The end game is an open source release for transparent use with ROOT.

5 Summary

Improved data compression will undoubtedly bring considerable benefits to the HENP community. Addressing some limitations of storage and networking resources and budgets is only one side of the story. Progressing towards live and complete dataset access enables data exploration (in other words, science!), following the goals of "any data, any time!" We have demonstrated significant improvements in file size reduction over standard compression on real-word data for lossless, and physics-preserving lossy scenarios. And chances are good for further improvements! We are excited to be partnering with the ROOT team to bring this promising and highly valuable technology to the HENP community.

6 Acknowledgement and Disclaimer

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Award Number DE-SC0018521. This paper was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

References

- [1] J. Gonzalez, S. Fonseca, and R. Nunez. "Arrangements for communicating data in a computing system using multiple processor", U.S. Patent No. 10,305,980 (2019); J. Gonzalez, S. Fonseca, and R. Nunez. "Arrangements for storing more data in faster memory when using a hierarchical memory structure", U.S. Patent No. 10,114,554 (2018)
- [2] A. Klimentov, *et al.*, "BigData and computing challenges in high energy and nuclear physics", *Jour. of Inst.* **12**, proceedings of INSTR17 (2017), [online] <https://iopscience.iop.org/article/10.1088/1748-0221/12/06/C06044>
C. Alton, *et al.*, CERN-RRB-2018-085 (2018), [online] <https://cds.cern.ch/record/2638096/files/CERN-RRB-2018-085.pdf>
- [3] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework", *Nucl. Inst. and Meth. in Phys. Res.* **A389** p. 81 (1997)
- [4] P. Deutsch and J-L. Gailly, "ZLIB Compressed Data Format Specification version 3.3", RFC 1950, DOI 10.17487/RFC1950 (1996)
- [5] O. Shadura and B. P. Bockelman, "ROOT I/O compression algorithms and their performance impact within Run 3" (2019), [online] <https://arxiv.org/abs/1906.04624>
- [6] See, for example, [online] <https://root.cern.ch/root/html/tutorials/io/double32.C.html>
- [7] NASA internal report on Cart3D performance with Compressive Computing, not publicly available at this time