



Contribution ID: 344

Type: Oral

Extreme compression for Large Scale Data store

Tuesday, November 5, 2019 11:30 AM (15 minutes)

For the last 5 years Accelogic pioneered and perfected a radically new theory of numerical computing code-named “Compressive Computing”, which has an extremely profound impact on real-world computer science. 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 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.

Consider for promotion

No

Primary authors: Dr LAURET, Jerome (Brookhaven National Laboratory); Dr GONZALEZ, Juan (Accelogic, LLC); VAN BUREN, Gene (Brookhaven National Laboratory); NUNEZ, Rafael (Accelogic, LLC)

Co-authors: CANAL, Philippe (Fermi National Accelerator Lab. (US)); NAUMANN, Axel (CERN)

Presenter: VAN BUREN, Gene (Brookhaven National Laboratory)

Session Classification: Track 6 – Physics Analysis

Track Classification: Track 6 – Physics Analysis