NEXT-GENERATION TECHNOLOGY FOR THE EXTREMELY EFFICIENT STORAGE, DISTRIBUTION, AND PROCESSING OF NUCLEAR PHYSICS DATA

Principal Investigator: Juan Gonzalez
Program Manager: Michelle Shinn

DOE SBIR Award No. DE-SC0018521
August 18, 2021

Accelogic, LLC

- Founded in 2005
- Our main business:
  - Very aggressive data compression for improved performance in scientific computing
  - Numerical integrity is preserved
- We opened a new field of science: ***Compressive Computing***
- Key principles:
  - Theorem: for each computed number, it is possible to remove certain number of bits in its floating-point representation, with ZERO LOSS of information
  - We call these bits “zibbits” (zero-information-bearing bits)
  - NOTEWORTHY: About 80% of the bits used in HPC are typically zibbits
- Key enablers (already successful in the field of HPC):
  - Solid theoretical foundations (theory is already consolidated and very rich/elegant)
  - Technologies that identify and remove zibbits at very high speeds
  - Technologies that effectively convert the “zibbit removals” into “software speedups”
- This project:
  - Expand the theory of Compressive Computing to methods that **aggressively compress NUCLEAR PHYSICS DATA in real life** (solve the storage problem, and make it work in real life for the HENP community!)
  - In collaboration with BNL and FNAL (subcontracts in place). MIT to help with CMS testing.

Project's goals

- 4x-9x storage compression factors for NP facilities
  - Unique and already proven technology (Compressive Computing)
  - Shareware-like IP (Intellectual Property) licensing of all required technologies to the HENP community
- Broadness to the general community
  - All experiments
  - Seamless integration with ROOT; ROOT "upgrade" transparent to users
  - ROOT team engaged
  - Testing to be done initially with STAR and CMS
- Software to be released
  - No cost to community
  - Open source
  - Shareware licensing
  - Ability for the community to further adapt and evolve

Milestones

- “Essentials” Prototype by Month 12
  - Done
- “Fully-Featured” Prototype by end of project at TRL 7 (“full-scale prototypical system demonstrated in relevant environment”)
  - In progress
- 2 scientific publications by end of project
  - In progress
  - 1 conference paper DONE
- No-cost extension until May 2022
  - Health issues of PI
  - COVID and other factors

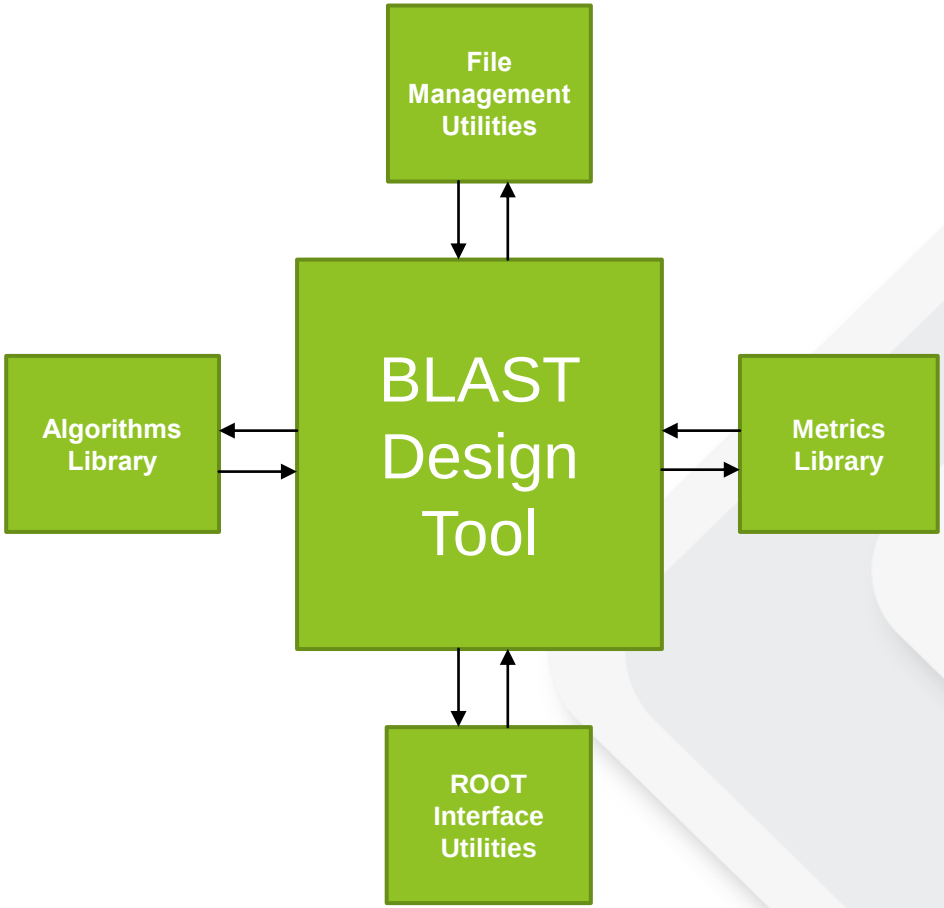
Work performed and achievements: Four main thrusts

1. Algorithms and securing compression performance (Lead: Accelogic)
   - ***Status:*** *Substantial progress already made, maximally optimal algorithms discovered, theoretical elements established, preparation of key patents and tech documents is in progress*
2. "Hook-up" with NP experiments (Lead: BNL)
   - ***Status:*** *Requirements in place, first prototype for invariant mass in progress*
3. Implementation/testing/validation work
   - ***Status:*** *First prototype successfully completed and tested, prototypes for further refinements in progress, CMS testing to begin soon*
4. ROOT integration (Lead: FNAL)
   - ***Status:*** *Work is currently in early stages*

Work performed and achievements: Product vision

- BLAST ("Bit Layered Adaptive Self Tuner")
  - Will pack all the technologies into a shareware product
  - Visible member of the "compressia" product-line-to-be (compressia_data)
  - Release will include perpetual royalty-free licenses of all patents for the HE/NP community (including all previous proprietary Compressive Computing patents as needed for project success)
- Commercialization plan has matured
  - Patents will be an important protagonist product

Work performed and achievements: BLAST architecture (evolving)



Work performed and achievements: Algorithms and securing compression performance

- Basic library of algorithms implemented and tested in new architecture
- Advanced classes of new algorithms discovered for:
  - Improved compression factors
  - Improved speed
  - Improved adaptability (self morphing optimally to any given data, autotuning capability)
- Strong theory discovered and its consolidation being pursued
  - Several optimality properties demonstrated for the new algorithms
  - Malleability: user can choose between optimal speed, optimal compression factor, or in between
  - Discovery of essential mechanism that links the physics with the problem of *optimal joint quantization/compression*
- Preparation of patent applications and tech documents is in progress

Work performed and achievements: Implementation/testing/validation work

- "Essentials" prototype done
  - Validated with RHIC/STAR (outside ROOT)

- To come:
  - Validation with LHC/CMS (outside ROOT)
  - Incorporation within ROOT and massive testing for both RHIC/STAR and LHC/CMS
  - Reach-out to other experiments for further validation and potential opportunities
    - After alpha version is integrated with ROOT
    - Timing of the reach-out is important: too early could distract us, too late could hamper impact or quality (right after ROOT integration seems optimal)
  - "Fully featured" prototype with all aspects integrated

Work performed and achievements: ROOT integration

- Work is currently in early stages
- Technical approaches, including key parameters and functional requirements from both sides (i.e., BLAST and ROOT) have been discussed with FNAL

Conclusions

- Project on the way to securing 4x-9x compression factors with industrial quality robustness; paradigm for maximal broadness and impact on whole NP community via ROOT also going well
- Major theory and algorithms discovered with seemingly optimal performance for NP-like problems; maximally fast algorithms also discovered
- First software production milestone already successfully achieved
- Product and commercialization concepts have reached maturation under the BLAST name (Bit Layered Adaptive Self Tuner)
- One scientific paper already written and presented
- Contracts with the two National Labs already in place, but they took longer than expected (COVID was the biggest factor)

If interested, please contact us!

- Possibility of including more experiments in the testing efforts, after ROOT integration
- Write to us:

Juan Gonzalez
juan.gonzalez@accelogic.com