Discussion of EF/CompF Submissions

Peter Onyisi

29 Mar 2022

The University of Texas at Austin

Computational Frontier

https://snowmass21.org/computational/start

Software and Computing are an integral part of the science process. High Energy Physics traditionally had the largest computing resource needs and subsequently most complex software stack in science. This is not true anymore, with many other science domains predicting equal or larger resource needs. The Computational Frontier will assess the software and computing needs of the High Energy Physics community emphasizing common needs and common solutions across the frontiers. We want to gain an overall understanding of the community's needs and discuss common solutions to them in the context of current and future solutions from the HEP community, other science disciplines and industry solutions. Our focus is to facilitate discussions amongst all frontiers and don't separate them into individual groups.

- Understanding HEP computational capabilities as something that need to be planned for in advance
  - hardware & networking infrastructure
  - software platforms
  - developer ecosystem
  - interface for physicists
  - community effects (equity, workforce development, environmental impact …)
- Important: what do we need computationally now to enable the design of future physics programs?
  - HL-LHC needs are reasonably well understood at this point

CompF Organization

Frontier Conveners

Name	Institution	email
Steve Gottlieb	Indiana University	sg[at]indiana.edu
Daniel Elvira	Fermi National Accelerator Laboratory	daniel[at]fnal.gov
Ben Nachman	Lawrence Berkeley National Laboratory	bpnachman[at]lbl.gov

- Seven topical groups
  - all with potential overlaps with other frontiers
- Liaisons with other frontiers
  - that's why I'm here
  - relatively few of us in CompF are computing professionals, most wear hats of other frontiers also
- Discussions in the usual ways (Slack, email lists; linked from main Wiki page)

Topical groups

Name
CompF1: Experimental Algorithm Parallelization
CompF2: Theoretical Calculations and Simulation
CompF3: Machine Learning
CompF4: Storage and processing resource access (Facility and Infrastructure R&D)
CompF5: End user analysis
CompF6: Quantum computing
CompF7: Reinterpretation and long-term preservation of data and code

Challenges for HEP Computing

- Access to sufficient hardware resources (storage, compute, networking)
  - distributed à la LHC Grid? Supercomputer centers? Clouds?
  - scaling from single users to full experiment production?
- Using "heterogeneous" architectures – GPUs, FPGAs, etc. (and using CPUs better)
  - frameworks for portability & flattening learning curve
- Home-grown vs. industry solutions
  - perhaps our home-grown solutions can become standards?
- Sustainable support model
  - how to maintain key home-grown packages in an environment that is biased for the novel over the established
  - but also encourage young people to work in these areas
- Avoid computing standing in the way between a physicist and their work
  - friendly ecosystems, lowered barriers, efficient workflow patterns matched to real needs
  - useful training & making documentation available
  - making the things we want people to do (e.g. data & analysis preservation) easy & the path of least resistance
  - well-defined interfaces between stakeholders – theory & experiment of different frontiers
- Where next for machine learning?
  - classification a standard part of the EF repertoire, regression in some key places (e.g. NNPDF)
- What does quantum computing have to offer?

White Paper Overlaps

- Currently 57 white papers marked as relevant to CompF
  - 11 cross-listed with EF; late submissions possibly still coming
- Topics of these white papers:
  - Theory computations
  - Surveys of tools & challenges
  - Diversity, equity, inclusion
  - Improving physics sensitivity with ML
  - Additional, incidentally marked for CompF

Coverage of White Papers

- Organizing principle in CompF: over the next decade, where should we focus computing effort (personnel, hardware resources, training, software R&D …) to enable the HEP program?
  - physics goals are the motivating factors but those details are best covered by relevant frontiers
  - CompF6 (quantum computing) and CompF3 (machine learning) have particular interest in new techniques, but those are usually somewhat independent of physics channel
- Generally very clear whether a white paper is focused on computational issues or not
  - e.g. the EF cross-listings for CompF2 (Theoretical Calculations and Simulation) are mostly about the calculations, not the computational needs/techniques/libraries
  - The HEP Software Foundation's HL-LHC computing review (2008.13636) is not cross-listed with EF! But then, it doesn't discuss physics…

Some Themes

- Computing can be quite fragmented across HEP frontiers
  - e.g. neutrino experiments do not have an equivalent of the standard HepMC format for exchanging events between generators and simulation, or standard interfaces for adding new models to generators
- Collider experiments (incl. B-factories) have pioneered a lot of capabilities
  - have lots of personnel to attack problems
  - historically encountered some problems first
- Energy Frontier experiments will not be unique in (e.g.) data size challenges in the future, but can still offer their expertise & experience
  - frameworks, data formats, workflow management software, ...

Overlapping White Papers by Topical Group

- General Interest:
  - “Software and Computing for Small HEP Experiments” (2203.07645)
  - “In Search of Excellence and Equity in Physics” (2203.10393)
- CompF1 (Experimental Algorithm Optimization & Parallelization):
  - none
- CompF2 (Theoretical Calculations and Simulation):
  - “Lattice QCD Calculations of Parton Physics” (2202.07193)
  - “Vector boson fusion at multi-TeV muon colliders” (2005.10289)
  - “The Effective Vector Boson Approximation in High-Energy Muon Collisions” (2111.02442)
  - “Detector and Beamline Simulation for Next-Generation High Energy Physics Experiments” (2203.07614)
  - “PetaVolts per meter Plasmonics” (2203.11623)
  - “Event Generators for High-Energy Physics Experiments” (2203.11110)

Overlapping White Papers by Topical Group

- CompF3 (Machine Learning):
  - "The LHC Olympics 2020: A Community Challenge for Anomaly Detection in High Energy Physics" (2101.08320)
  - "Event-based anomaly detection for new physics searches at the LHC using machine learning" (2111.12119)
  - "Improving Di-Higgs Sensitivity at Future Colliders in Hadronic Final States with Machine Learning" (2203.07353)
- CompF4 (Storage and Data Processing Resource Access):
  - none
- CompF5 (End-user Analysis):
  - none
- CompF6 (Quantum Computing):
  - none
- CompF7 (Reinterpretation and Long-Term Preservation of Data and Code):
  - none

Report Writing

- Both frontiers are beginning the process of distilling white paper submissions to reports
  - want to introduce as little noise to this as possible
- Still, want conveners on both sides to keep a broad mind and notify if they think information in a white paper might be relevant to the other
- Also take the opportunity to step back for a minute and think of the big, integrated picture

Potential Discussion Prompts

Examples of cross-cutting topics to keep in the back of our minds…

- Are we giving up physics capability because of a lack of computing capability (or the wrong kind of hardware)?
  - Higher-order accuracy SM calculations? Better handling of systematics?
- Could new techniques improve detector design & optimization?
  - e.g. differentiable programming? ML techniques to drive searches through large design space?
- Are there elements of the computing environment that hinder physicist productivity?
  - things that are not impossible to do, but people avoid just because they're difficult/time-consuming?