Enabling the Physics Goals – Computational Frontier

V. Daniel Elvira (Fermilab), Steven Gottlieb (Indiana University), Ben Nachman (LBNL) based (to the best of our ability) on all the contributors to the Computational Frontier Snowmass effort

Important disclaimer

The content of these slides do not fully reflect discussions and comments received during the CSS

- Pre-CSS versions of the TG and overall CompF reports are available at:
  https://snowmass21.org/computational (under draft reports)
- We are receiving comments from multiple sources, please share yours here: https://docs.google.com/document/d/1BNl1OAFWlavwmLSavdGCSZ7HaBbag215MZaKV3CldlA
- It will take a few days to incorporate comments in a consistent way

A few notes about the Computational Frontier

Timeframe

- The Snowmass process is mostly about the long-term, the post 2035 period
- However, the near-term program, before 2035, is not a given, and the Snowmass process must make a strong case for it
- The Computational Frontier received advice to focus on the upcoming 10-15 years
  - Secure funding to bridge the needs/resources gap of near-future programs
  - Rapidly evolving technologies make predictions beyond 2035 very inaccurate
- The near-term program should also address the needs of feasibility studies and R&D for future facilities – not too much in any TG report at the moment

Distinctive features of computing

- Computing technology and software paradigms change on a much faster scale than the life of particle physics experiments
- Many S&C efforts are neither funded nor managed as projects, unlike facilities and experimental devices

P5 recommendation on computing - 2014

https://www.usparticlephysics.org/

Recommendation 29 (on page 21):

Strengthen the global cooperation among laboratories and universities to address computing and scientific software needs, and provide efficient training in next-generation hardware and data-science software relevant to particle physics. Investigate models for the development and maintenance of major software within and across research areas, including long-term data and software preservation.

Computing as a P5 science driver - 2014

The 2014 P5 report has two paragraphs on computing as a science driver

- HEP played a leading role in high-throughput distributed/grid computing, online data processing, high performance computing and networking, large-scale storage, www
- Successfully managed software development and operations on a global scale
- Computing in HEP continues to evolve based on needs and opportunities
  - High performance computing and novel algorithms for realistic beams
  - Volume and complexity of physics data from LHC experiments stresses computing infrastructure and expertise
  - Cosmology programs to extend data needs as vast new surveys and high-throughput instruments come online
  - Theory computations to increase in importance as higher fidelity modeling is required

Computing as a P5 science driver - 2022

Computing is essential to all experiments and many theoretical studies

- Data volumes, detector complexity, precision required in calculations and simulation will continue to grow in near- and far-future experiments and surveys
- Size and complexity of the software and computing is commensurate with that of the experimental instruments; projects may also need software-detector codesign
- Software and Computing (S&C) is a global endeavor, efforts should be coordinated with worldwide partners

S&C technologies are changing the face of high energy physics

– Trend towards computing hardware heterogeneity and specialization, and increased use of high-performance computing facilities
– AI/ML not on the horizon in 2013, now widespread in every HEP area
– Quantum computing is entering the stage with potential impact on quantum many-body systems, event generators, data analysis, etc.

Main recommendation

We recommend the creation of a standing **Coordinating Panel for Software and Computing (CPSC)** under the auspices of DPF mirroring the panel for advanced detectors (CPAD) established in 2012.

Promote, coordinate, and assist the HEP community on Software and Computing, working with scientific collaborations, grassroots organizations, institutes and centers, community leaders, and funding agencies on the evolving HEP Software and Computing needs of experimental, observational, and theoretical aspects of the HEP programs. The scope should include research, development, maintenance, and user support.

In addition we have identified critical challenges that are limiting the physics output of the US HEP community, including:

Critical challenges and recommendations

1. **Long-term development, maintenance, and user support of essential software packages cutting across project or discipline boundaries is largely unsupported.**
   - Grants typically fund ground-breaking Research and Development (R&D) or development of new software, but not modernization, maintenance and user support of existing tools.
   - Examples include:
     - Event generators and simulation tools like Geant4 that do not belong to a particular facility, experiment, or survey.
     - Frameworks associated with one or more experiments.
     - Data and software preservation for utilization after an experiment has ended.

Recommendation: **the US HEP community should take a leading role in long-term development, maintenance, and user support of essential software packages** with targeted investment.

Critical challenges and recommendations

2. Research and development (R&D) for software and computing cutting across project or discipline boundaries receive insufficient support.

- Computational HEP is a vehicle for cross-cutting R&D.
  Supporting research in this area at a variety of scales would be broadly impactful
- Examples include S&C for theoretical calculations/generators, cosmological, accelerator and detector modeling, machine learning methodology and hardware ecosystems, algorithms and packages across experiment boundaries
  - USQCD provides a successful example of shared, community driven, open source software development supported by DOE and NSF
  - The HEP Software Foundation (HSF) is a successful example of an experimental community organization to facilitate cooperation in common efforts internationally

Recommendation: through existing and reshaped funding programs, **cross-cutting R&D efforts should be supported from proof of concept to prototyping and (if successful) to production ready deliverables**

Critical challenges and recommendations

3. **Scarcity of personnel and expertise jeopardizes the ability for full and optimal use of heterogeneous and high performance computing (HPC) resources**
   - Experiment, cloud, and flagship computer facilities use many different hardware platforms, making it difficult and expensive to port software for optimal use of resources
   - Many algorithms are not suitable for parallelization and will require traditional CPU-based computing for the foreseeable future

Recommendation: **support computing professionals/researchers, and physicists for code re-engineering and adaptation** to use heterogeneous resources effectively. To serve the needs of inherently serial algorithms, **traditional CPU-based hardware should coexist with heterogeneous resources**

Critical challenges and recommendations

4. **Investment in training and career paths within HEP for S&C researchers is insufficient**
    - Sustainable efforts in HEP computation require continual recruitment and training of the HEP workforce.
    - Successful attempts to meet training needs have been done through individual experiments, the NSF/DOE institutes, the HEP Software Foundation, and through growing numbers of university courses
    - Faculty/staff positions in S&C for HEP are scarce and expertise shortfall endemic.

Recommendation: **documentation and training efforts at multiple levels should be encouraged and extended.** To enhance their career opportunities, **bridge positions and awards should be created for HEP scientists and engineers working on software and computing**

Snowmass process preceded by European Strategy Report

Unsurprisingly, the recommendations of the Computational Frontier resonate with the European Strategy, spelled out in their recent report:

https://europeanstrategy.cern/european-strategy-for-particle-physics

Both exploratory research and theoretical research with direct impact on experiments should be supported, including recognition for the activity of providing and developing computational tools.

(...)

Large-scale data-intensive software and computing infrastructures are an essential ingredient to particle physics research programmes. The community faces major challenges in this area, notably with a view to the HL-LHC. As a result, the software and computing models used in particle physics research must evolve to meet the future needs of the field. The community must vigorously pursue common, coordinated R&D efforts in collaboration with other fields of science and industry, to develop software and computing infrastructures that exploit recent advances in information technology and data science. Further development of internal policies on open data and data preservation should be encouraged, and an adequate level of resources invested in their implementation.

(...)

Particle physics, with its fundamental questions and technological innovations, attracts bright young minds. Their education and training are crucial for the needs of the field and of society at large. For early-career researchers to thrive, the particle physics community should place strong emphasis on their supervision and training. Additional measures should be taken in large collaborations to increase the recognition of individuals developing and maintaining experiments, computing and software.