Snowmass: Computational Frontier
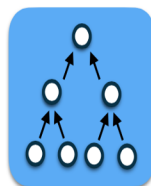
BNL Snowmass Retreat
December 17, 2021
Peter Boyle (BNL/HET)

- Slides based on material by CompF conveners, with thanks to:

Steve Gottlieb (Indiana U),
Ben Nachman (LBNL), Oliver Gutsche (FNAL; till Dec 2021),
Daniel Elvira (FNAL; From Dec 2021)

- All credits go to the authors
- All mistakes are mine!

Computational Frontier: Topical Working Groups

CompF01
Experimental
Algorithm
Parallelization

Guiseppi Cerati (FNAL), Katrin
Heitmann (ANL), Walter Hopkins (ANL)

CompF02
Theory
Calculations
& Simulation

Peter Boyle (BNL), Daniel Elvira
(FNAL), Ji Qiang (LBNL)

CompF03
Machine
Learning

Phiala Shanahan (MIT), Kazu Terao
(SLAC), Daniel Whiteson (Irvine)

CompF04
Storage and Processing
Resource Access
(Facility and Infrastructure R&D)

Wahid Bhimji (NERSC), Rob Gardner
(U. Chicago), Frank Würthwein (UCSD)

CompF05
End User
Analysis

Gavin Davis (U. Mississippi),
Peter Onyisi (U. Texas at Austin),
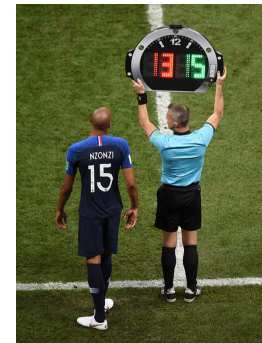Amy Roberts (UC Denver)

CompF06
Quantum
Computing

Travis Humble (ORNL), Gabriel Perdue
(FNAL), Martin Savage (U. Washington)

CompF07
Reinterpretation & Long-term
Preservation of Data and Code

Kyle Cranmer (NYU), Mike Hildreth (Notre
Dame), Matias Carrasco Kind (Illinois/NCSA)

Substitutions:

CompF2: Kevin Pedro (FNAL)
CompF4: Meifeng Lin (BNL)
CompF7: Stephen Baileuy (LBL)

Promotions:

Daniel Elvira as convener

Snowmass 2021: communication

https://snowmass21.org/computational/start

Join our Slack channels!

comp_frontier_topics

compf01-expalgos

compf02-theorycalcsim

compf03-ml

compf04-storeandprocess

compf05-useranalysis

compf06-quantum

compf07-preservation

Join our topical group meetings!

Join our email lists!

Computational Frontier: Liaisons

Energy Frontier
Daniel Elvira (FNAL)

Neutrino Frontier
Alex Himmel (FNAL)

Rare Processes & Precision
Stefan Meinel (Arizona)

Cosmic Frontier
Deborah Bard (NERSC)
Brian Yanny (FNAL)

Computational Frontier

Theory Frontier
Steven Gottlieb (Indiana)

Accelerator Science/Technology
Jean-Luc Vay (LBNL)

Instrumentation Frontier
Darin Acosta (Florida)

Community Engagement
David Bruhwiler (RadiaSoft)

Computational Frontier: Scope & Outcome

Our main time horizon should be ~10 years (HL-LHC, DUNE, LSST, etc.), but it is also useful to think about the next-to-next experiments and what R&D/funding opportunities we may need to be ready for the computing of the future.

Outcome:

- Letters of Intent: 18+116+55+25+26+34+19 (many multi-frontier)

- CompF : whitepaper abstracts by Jan 31. white paper submission to arXiv: Mar 15.

- Every topical group writes a document about their findings & points out opportunities and challenges
- The Computational Frontier writes one document consolidating all topical working groups
- Snowmass 2021 writes one document consolidating all frontiers

Amdahl's law: parallelization cannot improve the serial component of an algorithm

Multicore CPU / HTC

Independent progress

"**ant colony**" model

GPU / DOE Exascale

Coupled progress

"**marching army**" model

Only effective if all elements
do SAME thing at same time

Cross-frontier themes

- Many different needs, from different experiments or different algorithms
  - difficult to have a one-fits-all solution, even within a frontier
  - possible exceptions: accelerated FFTs in Cosmic and Neutrino frontier, real time processing (trigger/broker applications)
- Transitioning from HTC to HPC (or using both)
  - evolution of the programming model
- (Optimal) use of heterogeneous resources
  - how to keep the GPUs busy?
- Interplay between ML and traditional reco algorithms
  - switch to ML approaches vs rewriting algorithms
  - avoid separate workflows, ensure feedback between the two

Cross cutting issues

GPU architectures are for good reason the Exascale in immediate US roadmap

In principle problems:
- Not all algorithms are suitable

In practice problems
- Preexisting investment in software
- Software shared in international collaborations
- Various GPU systems have *different* programming interfaces
- Spatial / FPGA ?

It cannot be cost effective or possible to migrate *ALL* software
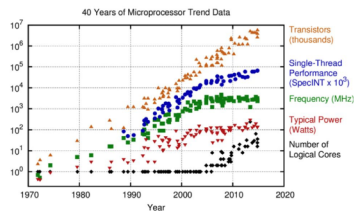- Right hybrid mix of architectures?

Significant staff effort issues
- Staff effort, Staff retention, Staff career paths, Software support
- Important: flag these issue with CompF5,7

CompF1: "Experimental Algorithm Parallelization"

Computing landscape

The computing landscape has been transforming in the last few years: end of Dennard scaling, emerging of GPUs, building of exascale machines.

This means that adiabatic improvements from past solutions may not work or may be suboptimal. This is an opportunity to re-think how we process our data, and define new solutions for a higher science throughput.

40 Years of Microprocessor Trend Data

Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2015 by K. Rupp

Relationships with other working groups

The definition of "experimental algorithms" is broad, covering the topics of other WGs.

We'll focus on the area not covered by others. It means central (i.e. not analysis specific), non-ML algorithms whose inputs are experimental data (both offline and software trigger). This may have different meaning for different physics frontiers!

Frameworks are not specifically covered in other groups, and we'd be happy to discuss implications of parallel execution for frameworks in our WG.
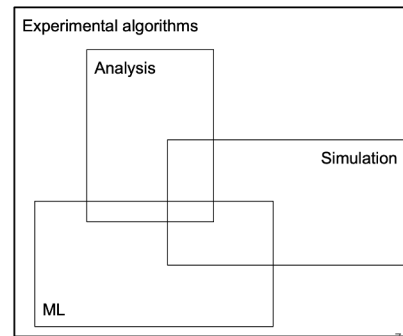
Functional areas of our working group

- Parallelization of detector reconstruction algorithms, physics object reconstruction/calibration algorithms
- Utilization of CPU, accelerator hardware and what comes next in 5-10 years
- Developing better algorithms in addition to parallelization
- Portability solutions that support the same algorithm implementation on multiple hardware architectures

CompF2: Theoretical Calculations and Simulation

Peter Boyle, Ji Qiang, Daniel Elvira (until Dec 2021), Kevin Pedro (from Jan 2022)

Six subtopics
1. Event Generators
2. Accelerator Modelling
3. Detector Modelling
4. Theoretical calculations (Perturbative)
5. Theoretical calculations (Lattice)
6. Cosmic simulations

Likely issues:
Technical:
- GPU and HPC porting of HEP software
- Perturbative software not suited to batch computing
- International computing commitments

Resource:
- Effective compute provisioning for non-parallel software
- Cannot be cost effective or possible to port ALL software
- Software support for large and small experiments (Geant, MadGraph)

Personnel
- Computational scientist staff effort, retention, career paths

- Background of preexisting community papers (HSF, USQCD) in areas

- Community points of contact nominated with baseline responsibility for white papers. Further white papers likely.

- Please refer to our questions we'd like addressed on Snowmass Wiki
- Intended to help us translate *your physics* into *computing needs*

CompF3: Machine Learning

Particle physics-specific ML

Particle physics often has unique stats challenges
- Symmetries, boundaries, limits
- Data on manifolds or subsurfaces
- Sensitivity to uncertainties
- Heterogenous data structures

Simulation and ML

Simulation very expensive
- Fast simulation with ML
- Limitations and possibilities

Interpretability/validation

What has the machine learned?
- Reverse engineering ML strategy
- Exactness proofs
- Uncertainty measures
- Data reconstruction
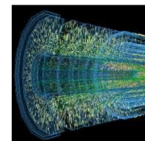
Community Tools

Standard tools and packages
- Software
- Data structures
- Adapting industry tools

Resource needs

What computing resources are needed?
- Real-time ML
- Clusters with GPUs, FPGA, etc
- Cloud processing
- Operations intelligence

Education

What do physicists need to know?
- Physics-ML specific courses?
- Outreach to community (ie CS/ML)
- Curation of open data
- Ethics and safety of AI
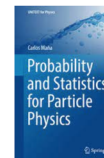
CompF4: Storage and processing resource access (Facility and Infrastructure R&D)

CompF4: Functional areas

- Provide access to data for large scale central workflows
- Provide access to data for end user analysis

- Hierarchical storage
  - Access to long-term high-latency storage (tape …)
  - Access to low latency storage (disk, ssd, nvme …)

- Access to
  - CPU resources ➜ GRID, HPC, Cloud
  - Accelerator resources ➜ GRID, HPC, Cloud
  - Specialized AI hardware

- Interconnecting everything through Network

CompF4: Mandate

- What are the workflows related to storing and accessing data of the stakeholders?
- What different storage solutions/technologies are used, will be used?
- What is the technology evolution of storage solutions/technologies? What R&D is needed?
- What are the storage and access needs of the stakeholders?
- What are the storage resource needs of the stakeholders?
- What is the role of the network in these workflows?
- How are the solutions used by the community embedded/derived from solutions from industry/other science domains
- Recruit community members to represent the physicist/analysis perspective.

Future CompF4 activity

- During fall 2020 collect:
  - Further background material / reports
  - Wide community exemplars of resource and access needs
  - Questionnaire for resource needs and research topics

- ~Oct-Nov: Focussed workshop to further gather and synthesise these requirements

CompF5: End user analysis

Mandate

Consider:

- Types of resources needed for analysis facilities
- Use of accelerators
- Analysis libraries
- Data storage formats & dataset bookkeeping
- Programming languages
- Software for collaboration: version control, messaging
- "Real-time" analysis
- Long-term reproducibility and preservation

While also considering:

- Sustainability, both technical and human
  - Documentation
  - Training
  - Long-term software support & development
  - Integration with the broader ecosystem
  - Broad applicability to the field
  - Hardware facility evolution
- Interaction with other scientific fields & industry
  - Role of proprietary technology
  - Potential contributions to computing outside HEP
- User experience
  - Ease of use & setup
  - Scalability of technologies
  - Required training and broad applicability of training to other domains

Working Group Goals

- Produce a document that identifies impediments to end user analysis and potential ways to address such issues
  - With broad scope, e.g. fragmentation of knowledge across platforms, or lack of documentation as an equity concern
  - Informed by feedback from the broad user community
- Bear in mind that the detailed landscape will certainly change in the next 5-10 years, but hopefully requirements change more slowly
- Highlight potentially transformative avenues for R&D efforts
  - Including identifying gaps not covered by current work
- LOIs/white papers are encouraged to take a "big picture" stance
  - What core issues are being addressed?
  - It's OK to have a LOI that just identifies problems without having specific solutions in mind

End user analysis survey

Still collecting responses! Fill it out at https://forms.gle/rzvtNEGxhoXYAfKj7.

We'd love to see more

- Dark matter community voices
- Nuclear physics community voices
- Theory voices
- Experimental and Test Facility voices
- Early-career voices
- YOUR VOICE, if you haven't already filled out the survey!

Very early results will be shown at the start of the parallel session

CompF6: Quantum computing
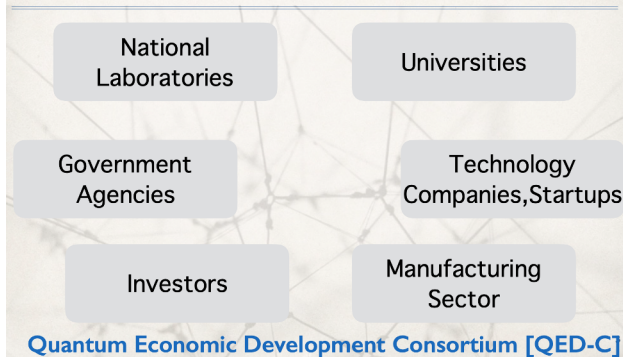
Why Quantum Computing / Quantum Information Science

✤ Quantum information science (QIS) is a major area of research emphasis in the DOE and for the nation at large.

✤ HEP has historically had a number of important roles in this field and we continue to play an important role in theoretical developments.

✤ We are becoming more important on the hardware side and have important contributions to make to QIS.

✤ Furthermore, there are a number of deeply interesting science questions quantum technologies enable us to ask in HEP.

✤ In short - we can play an important role in this endeavor and our science will benefit.

Toward a Quantum Ecosystem

National Laboratories

Universities

Government Agencies

Technology Companies, Startups

Investors

Manufacturing Sector

Quantum Economic Development Consortium [QED-C]

## Areas for Snowmass

Thank you are presen and contri

• This is a new area for Snowmass, HEP long-range planning

• Looking to identify and assess all "quantum" areas impacting and advancing HEP during the next decade

✤ Quantum networks
✤ Data analysis in HEP with quantum computers
✤ NISQ-era quantum devices for HEP
✤ Simulation of quantum field theories
✤ Quantum simulation and hardware co-design
✤ Tensor networks
✤ Quantum information, error correction and holography
✤ QFTs on AdS
✤ Search strategies for new particles using SRF cavities
✤ Quantum computing for event generators
✤ Quantum algorithms for quantum sensing
✤ Algorithm development for beyond NISQ-era devices

Quantum Comp. and Tech. for HEP

Summary

• QIS : an emerging and disruptive impact on HEP

• HEP anticipated to impact and be impacted

• Simulation, Communication, Sensing, Data

• Close collaboration with theory, HPC and experiment

• Close collaboration between Labs, Universities, Tech companies

• Close collaboration with other domain sciences, QIS

• New to Snowmass, next decade expected to be transformative
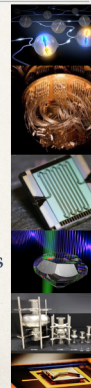
CompF7: Reinterpretation and long-term preservation of data and code

Functional (Focus) Areas:

- Public data
  - (comes in many forms … HepData, public likelihoods, CERN OpenData, data for education/outreach)
  - Tools for generating annotated public data and software
  - Tools for sharing data and software
- Not-yet-public Data
  - Tools for generating annotated "private" data and software
- Tools for combining results across experiments and frontiers
- Tools for archiving and re-running analyses (RECAST/REANA, …)
  - Internal-to-experiment and external "public" preservation

- ❖ Obvious overlap with all physics groups, as well as other computational areas
- ❖ Will try to join/convene as many joint sessions as possible moving forward

Group Mandate, Activities, Questions:

- Define the stakeholders and consumers of the data and software
  - What are the needs/requirements of the stakeholders?
    - (probably most difficult question to answer)
- What resources are needed?
  - e.g. long-term storage with external access, infrastructure for preserving executable code, etc.
  - metadata infrastructure
- What technologies are available or will be available, what is the technology evolution of these tools?
  - To be discussed in common with CompF5: End User Analysis:
    - version control
    - Containers/VMs
    - proprietary software/licenses

Group Mandate, Activities, Questions (cont.):

- How are/will the stakeholders use these technologies?
- What are the workflows that are used to combine results across experiments and frontiers?
- What tools are used/needed by the stakeholders to combine results across experiments and frontiers?
- What will the technological evolution of these tools look like?
- How are other science domains handling this topic?
- What are other science domains using, what is industry using?

Overall Goals:

- Raise awareness/visibility of preservation issues across frontiers
- Communicate current efforts/technologies to other groups/frontiers
- Mediate incorporation of these concepts and objectives into *all* reports and guidelines (where appropriate)
- Production of general guidelines (aspirations?) for preservation of scientific results

14

- CompF5 and CompF7 represent skills, human resources and preservation of investment

- Important to engage to make a powerful statement of need.

Timeline:

To allow topical group conveners sufficient time to consider white papers, we ask all who wish to contribute to submit (via email to topical group conveners) at least a title and abstract by the end of January, 2022. White papers whose title and abstract are submitted later are not guaranteed full consideration. The general Snowmass deadline also applies.

Time Schedule

- January 31, 2022: White Paper submission to arXiv (preferred) or at least title and abstract to topical group conveners
- March 15, 2022: Official deadline for white paper submission to arXiv
- May 31, 2022: Preliminary reports by the Topical Groups
- June 30, 2022: Preliminary reports by the Frontiers
- July, 2022: Snowmass Community Summer Study (CSS) at UW-Seattle
- September 30, 2022: All final reports by TGs and Frontiers
- October 31, 2022: Snowmass Book and the on-line archive documents