A Coordinated Ecosystem for HL-LHC Computing R&D

Coherence & Alignment, October 2019

Hosted by IRIS-HEP

THE CATHOLIC UNIVERSITY of AMERICA

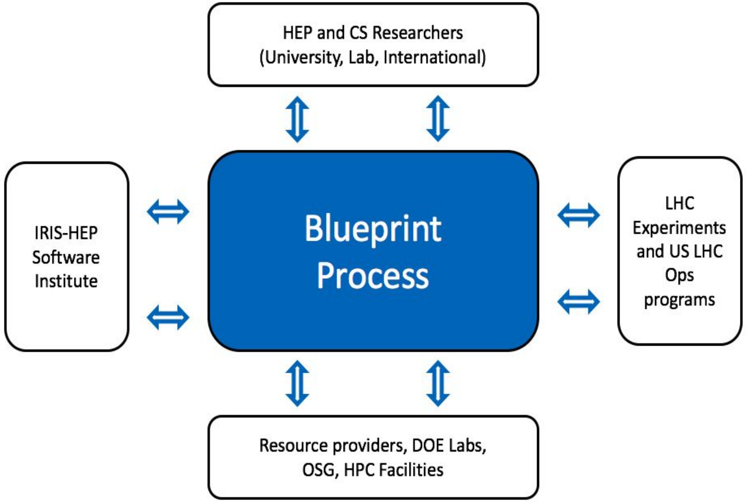
Outcomes of this mini-workshop

- Commitment to joint blueprint activities
- Commitment to joint project coordination across DOE-HEP, DOE-ASCR, NSF-EPP, and NSF-OAC funded activities.
  - Initial exploration and agreement on complementarity of DOE and NSF supported activities.
- Initial agreement on S2I2 governance

From Nov 2017 CUA Workshop

Blueprint Activity - Maintaining a Common Vision

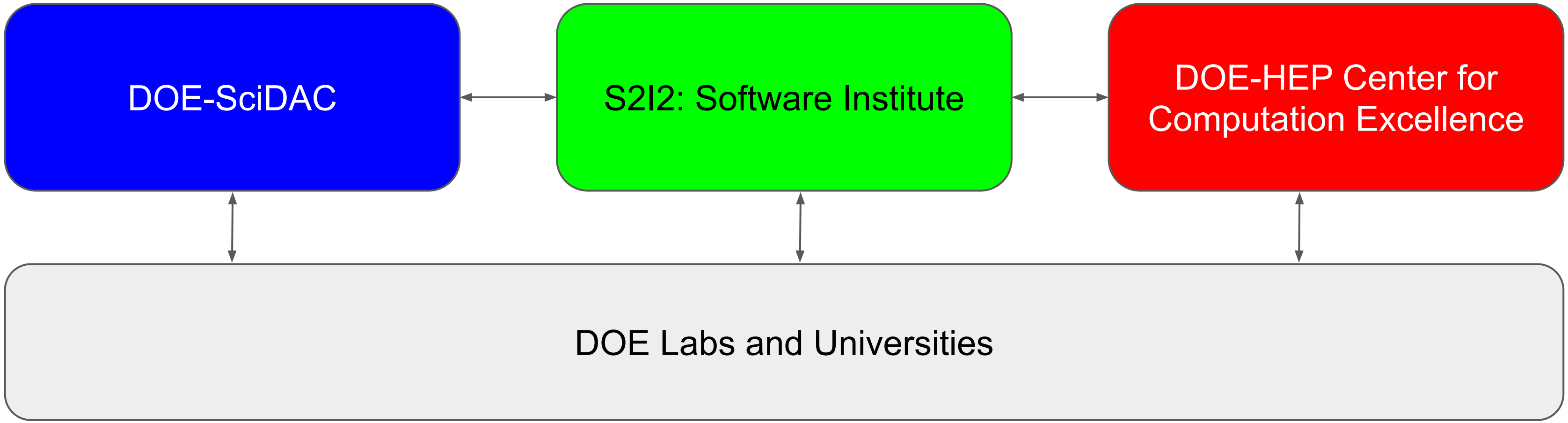
- Small **"blueprint" workshops** 3-4 times per year with key personnel and experts
- Facilitate effective collaborations by building and maintaining a common vision
- Answer specific questions within the scope of the Institute's activities or within the wider scope of HEP software & computing.
- 21 Jun - 22 Jun, 2019 - Blueprint: Analysis Systems R&D on Scalable Platforms (*NYU*)
- 10 Sep - 11 Sep, 2019 - Blueprint: Accelerated Machine Learning and Inference (*Fermilab*)



- 23 Oct - 25 Oct, 2019 - Blueprint: A Coordinated Ecosystem for HL-LHC Computing R&D (*Catholic University of America, Washington DC*)
- Others (e.g. Training) in planning

R&D Partnerships

From Nov 2017
CUA Workshop



NSF and DOE partnership, as informed by the blueprint process, will be essential to the success of the HL-LHC R&D efforts

Thank you all!

First, thanks to all the participants - especially those who participated in the discussion on Thursday!

- Significant progress on understanding topical area coverage
- Overviews on Wednesday helped inform the discussions and provide context

Questions Addressed at this workshop

1. How does the ensemble of US Software R&D efforts fit together to implement the HL-LHC Software/Computing roadmap described in the Community White Paper and meet the challenges of the HL-LHC? Which areas are not covered by US R&D efforts?

2. How do the US Software R&D efforts collaborate with each other and with international efforts? How do these efforts align with and leverage national exascale, national NSF OAC priorities and trends in the broader community?

3. How should the US R&D efforts be structured and organized in order to impact planned updates (all in ~2021/2022) to the HSF Community White Paper, the software/computing part of the US Snowmass process and HL-LHC experiment-specific software/computing TDRs?

Summary of Thursday Discussions

- A summary of the Wednesday discussions will be part of the final workshop closeout report.
  - Due to time constraints, we won't be summarizing these in the closeout slides.
  - Additionally, impact beyond LHC isn't included in these slides.
  - Several items (e.g. - how LSST utilizes Jupyter to enable analysts) need follow-up, even if it's not part of HL-LHC planning.
- Thursday:
  - Iterated through each of the CWP areas to review the progress & projects since last meeting. Came up with ideas for milestones and deliverables to better align.
  - Guided discussion on how to best engage with the ECP and (proposed) CCE project.
  - Started to build timeline of projects going out to 2023.

CWP Area Overview

Where are resources deployed? Does not indicate full coverage!

- DOMA
- Physics Generators
- Visualization
- Data and Software Preservation
- Machine Learning
- Software Trigger & Event Reconstruction
- Detector Simulation
- Workflow and Resource Management
- Data-Flow Processing Framework
- Data Analysis & Interpretation
- Networking, Storage Infrastructure and Facilities
- Software Development, Deployment, Validation, Verification
- Training

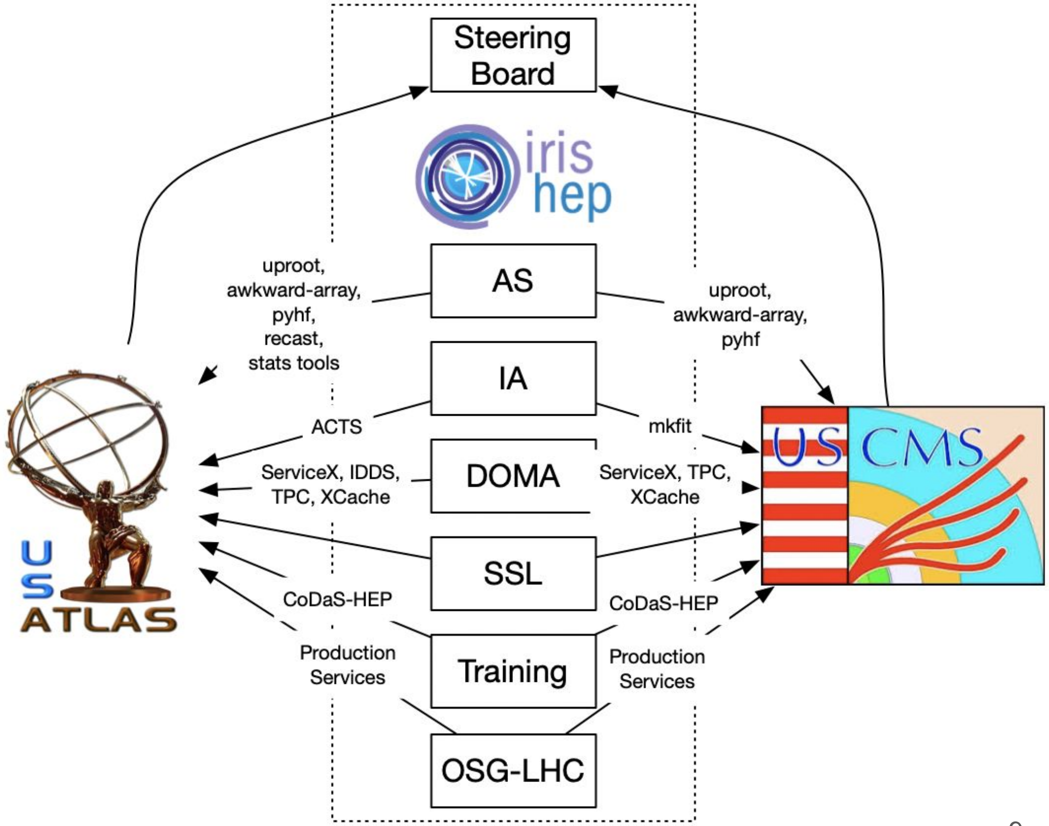
Interaction Between IRIS-HEP and Ops

In its first year, IRIS-HEP has established links across its focus areas to both LHC operations programs.

- A *<u>subset</u>* of the projects & contributions where the organizations interact are highlighted below.
- The U.S. LHC Ops programs help guide the R&D activities through the steering board.

IRIS-HEP includes effort to help bring R&D projects through **integration** (SSL) and **production** (OSG-LHC).

- Even then, **requires close collaboration** with the Operations program to ensure we derive value!



Coordinating with CCE & ECP

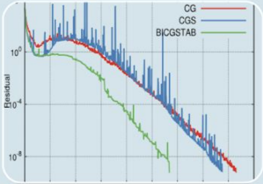
Opportunity: Software stack of ECP aims to have broader impact.

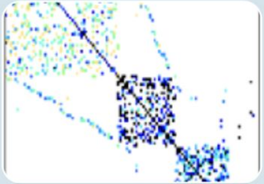
- USATLAS and USCMS determine interests in ECP products
  - Form group of people and preliminary list of areas of interest in ECP products
  - ECP organizes more extensive briefing of ECP program for USATLAS and USCMS
  - USATLAS & USCMS define list of projects to use ECP products
  - If approved, CCE should play a coordination role and should provide a home for discussions
- Explore how we can have more formal engagement between LHC & ECP
  - What is the process of asking ECP-funded projects for bigger changes that would require development effort?

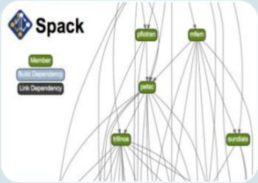
(Incidentally) Yearly allocations do not work for HEP, need programmatic allocations at significant enough scale to have an impact

  - Prototyping multi-year requests @NERSC/ERCAP, need to find solution for LCF/ALCC/INCITe
  - Consult with ECP

ECP software technologies are a fundamental underpinning in delivering on DOE's exascale mission





National Nuclear Security Administration

Programming Models & Runtimes
- **Enhance & prepare OpenMP and MPI programming models** (hybrid programming models, deep memory copies) for exascale
- Development of **performance portability tools** (e.g. Kokkos and Raja)
- Support alternate models for potential benefits and risk mitigation: PGAS (UPC++/GASNet) ,task-based models (Legion, PaRSEC)
- Libraries for deep memory hierarchy & power management

Development Tools
- Continued, multifaceted capabilities in portable, open-source LLVM compiler ecosystem to support expected ECP architectures, including support for F18
- **Performance analysis tools** that accommodate new architectures, programming models, e.g., PAPI, Tau

Math Libraries
- **Linear algebra, iterative linear solvers, direct linear solvers, integrators and nonlinear solvers optimization, FFTs, etc**
- Performance on new node architectures; extreme strong scalability
- Advanced algorithms for multi-physics, multiscale simulation and outer-loop analysis
- Increasing quality, ... math libraries

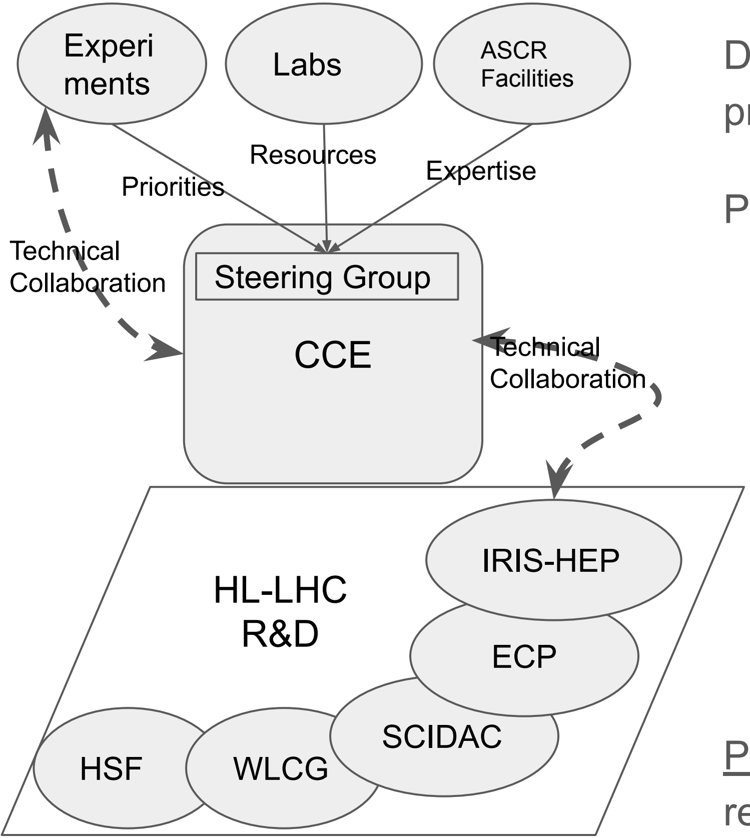
Data and Visualization
- **I/O libraries: HDF5, ADIOS, PnetCDF,**
- **I/O via the HDF5 API**
- Insightful, memory-efficient in-situ visualization and analysis – **Data reduction via scientific data compression**
- Checkpoint restart
- **Filesystem support for emerging solid**

Software Ecosystem
- Develop features in Spack necessary to support all ST products in E4S, and the AD projects that adopt it
- Development of **Spack stacks for reproducible turnkey deployment** of large collections of software
- **Optimization and interoperability of containers on HPC systems**
- Regular E4S releases of the ST software stack and ... with regular integration of new ST products

NNSA ST
- Projects that have both mission role and open science role
- Major technical areas: New programming abstractions, math libraries, data and viz libraries
- Cover most ST technology areas
- Open source NNSA Software projects
- Subject to the same planning, reporting and review processes

Community has significant efforts on these, some through SciDAC

Interaction Between CCE and Ops



Discussed CCE projects; strong support for PPS as highest priority; concerns: IOS effort level.

Project Interactions(PPS, IOS, EG)

- Priorities determined by point of contacts (experiments)
- Interact regularly with experiment and IRIS-HEP technical experts.
  - **Collaboration on pilot projects vital for success; must have active participation from Ops**.
- Blueprint workshops and topical meetings to coordinate with Ops Programs, IRIS-HEP tech areas and HSF WGs such as frameworks, and DOMA

<u>Proposal</u>: to ensure coordination, IRIS-HEP & CCE have representatives on each other's governance mechanisms.

Relations/Interactions with LHCb

- IRIS-HEP Steering Committee Member is Gerhard Raven (NIKHEF)
- IRIS-HEP funding to CIncinnati & MIT for Innovative Algorithms & Analysis Systems work (IA efforts relate to Collaborative SSE efforts)
- IRIS-HEP DOMA group has initiated collaboration with LHCb with respect to data compression
- LHCb and CMS are sharing MIT Tier 2 resources [LHCb M&O award from NSF as of February, 2019].

Gaps and Opportunities:

- Limited interactions with CMS/ATLAS; no US-LHCb core computing effort.

From R&D to LHC Events

Example from history: CMS's use of threading started with an investigation analogous to CCE's (proposed) PPS.

Investigation for parallel scheduling technologies. **2012**	Selection of TBB as underlying technology **2013**	First version of multithreaded framework **2014**	1,000 algorithms converted over to threaded mode **2015**	Use of multithreading in production **2016**

Take-home: Need continuous coordination and feedback with R&D. Takes years of investment by the experiment to take successful R&D outcomes to production.

Opportunities and Prioritization

- How do we innovate event processing frameworks to best accommodate accelerator integration. Large amount of existing software makes change difficult.
- R&D to use HPC’s in HEP’s distributed computing infrastructure needs to be bolstered.
- There are some worries about HPC facility timelines and the use of Run 3 for testing.
- We need better metrics to understand how and when DOMA transitions to new facility architectures.
- Understand how to include QA for complex reconstruction & trigger algorithms.

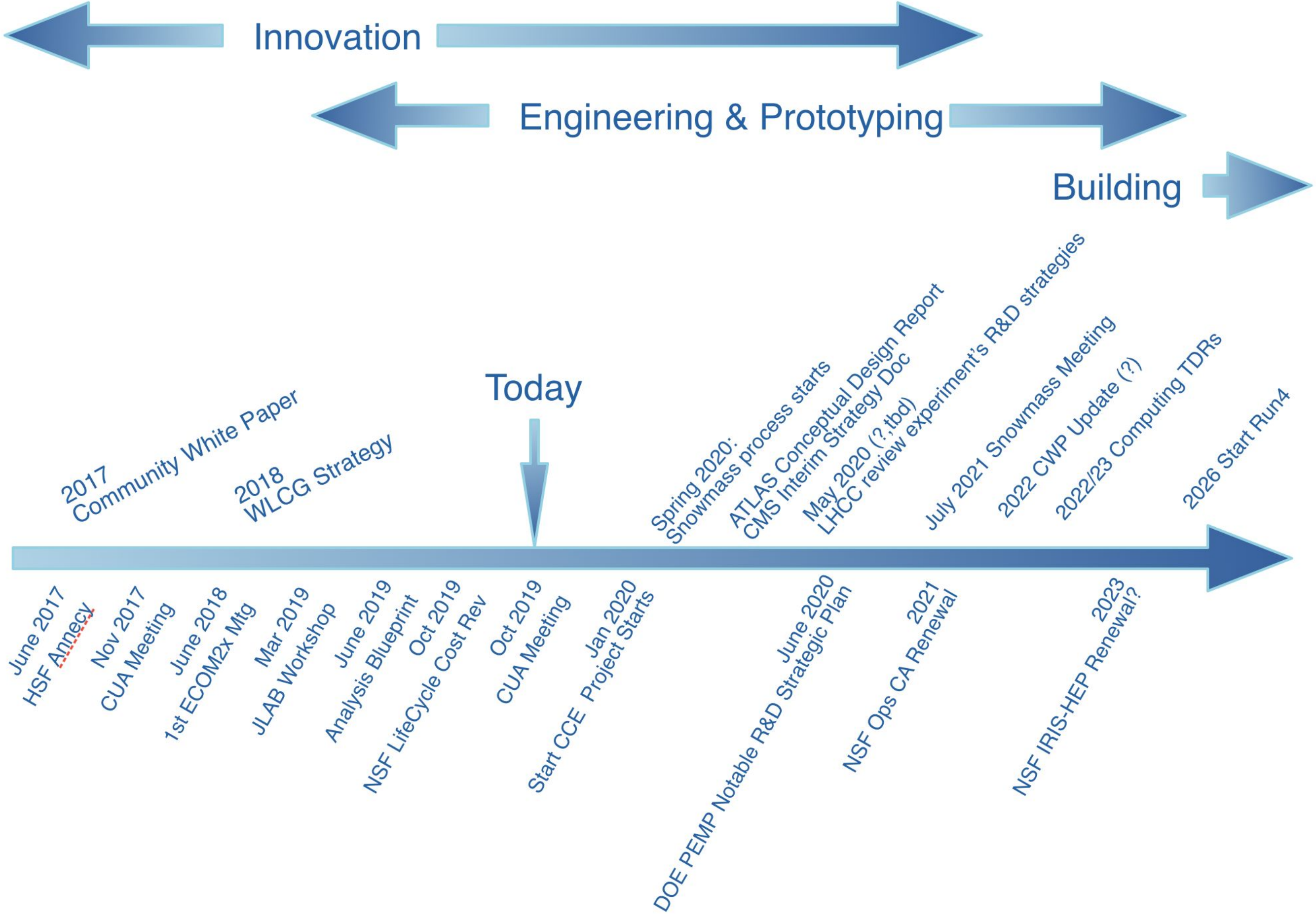
Collaboration Opportunities

- Better understand how to foster research and collaboration in building Analysis Facilities
  - Users, Interface, Researchers
- We’d like to better align plans with SciDAC-4 & CCE in DOMA.
- Close coordination with ESNet needed as computing models evolve; continuously improve understanding of usage.
- Our field needs to build wider community for generator optimization.
- Our field needs to build collaborations between Computer Scientists and Trigger/Reco experts to re-engineer algorithms
- Experiments’ schedule for choosing new low level accelerator interface technology is not matched to R&D schedule (“programming model”)

Attention & Effort Needed

- Evolution of Facilities towards the HL-LHC era
  - "**Analysis Facility**" - what specializations are needed? How do we build this out?
    - See blueprint meeting
  - Future of **U.S. storage facilities**: caching integration? Hierarchical storage approach?
    - Getting the facilities and the use cases integrated.
- Uncovered areas
  - There is an opportunity for the USA to take the lead developing the next version of GEANT optimized for accelerators.

Time Line for HL-LHC Computing R&D



Milestones & Deliverables to be scheduled

- Each of the 13 areas worked to put together a few *potential* milestones (see backup slides). <u>To be scheduled</u>:
    - CCE blueprint workshop (if funded).
    - Analysis facilities technical workshop.
    - Analysis Ecosystem (follow-on to A'dam workshop in May 2017)
    - Phone briefing for US-ATLAS / US-CMS of the relevant ECP areas.
    - Possible WLCG DOMA workshop (connected to planned Rucio in March 2020?)
    - Whitepaper on future U.S. LHC storage facility models.
    - Evolve ESNet / HEP "Blueprint" (analytics) group to work on network needs.
    - Develop improved requirements modeling for HL-LHC use of generators.
    - Whitepaper on "killer-apps" for Machine Learning in the HL-LHC context

2021?

Thanks!

We hope to capture all of these items in a close-out report before the end of the year.

Backup Slides

Data Analysis Systems and Software/Data Preservation

Scope & Activities

- Covers everything from the end of the production system to the final physics paper: data query, extraction, histogramming, statistical models, and reuse and analysis.
- Analysis Facilities, Analysis Frameworks (e.g. Coffea), Statistical Models, Analysis Preservation (REANA), RECAST

Potential Milestones & Deliverables

- Blueprint meeting to understand what an analysis facility would look like
- A prototype analysis facility capable of doing a “modern” analysis with small numbers of simultaneous users
- End-to-End Data Challenge

Projects

- IRIS-HEP, SCAILFIN, Coffea, hepaccelerate, SciDAC-4, ROOT, CERN OpenLab (Spark), SWAN, REANA, RECAST

Opportunities

- Increased user-involvement as prototype stage progresses; CMS’s Spark/coffea analysis facility; connect with LCF analysis/visualization facilities

Weaknesses & Gaps

- Appropriate underlying hardware for an analysis Facility undefined, Research vs Program effort, Integration with SSL-like substrate infrastructures

Reconstruction and Trigger Algorithms

Scope & Activities

- Reconstruction and trigger algorithms are resource drivers during HL-LHC given large event rate and event complexity increases
- R&D focuses on reengineering current approaches and taking novel approaches to solve problems (typically via AI)

Potential Milestones & Deliverables

- Establish scope of accelerator reengineering effort
- Lower barriers to entry by documenting demonstrators as they are developed
- Use expert conferences/workshops to increase coherence between efforts (eg, CTD2020 in April)

Projects

- Numerous - US program has notable focus on tracking. Other areas include FPGA acceleration for trigger systems as well as algorithms for calorimetry and jet reconstruction.

Opportunities

- To further enhance collaborative R&D rather than single experiment projects
- To reduce facility costs by establishing a programming model towards modern hardware

Weaknesses & Gaps

- Some R&D faces tension between Run 3 as testing ground and HPC facility timelines
- Involving subject matter experts in reengineering to ensure long term sustainability
- CWP focus areas un(der) covered - Real-time analysis beyond LHCb, and modernizing data quality monitoring

Applications of Machine Learning

Scope & Activities

- **Usable** tools for **large-scale distributed training and optimization**
- Training methodologies that are able to detect **rare features** in high-dimensional spaces while being robust against systematic effects
- Tools to quantify **systematic effects**
- High-quality generative models satisfying **physical constraints and symmetries**

Potential Milestones & Deliverables

- Prepare a white-paper and a slide deck describing to potential collaborators state of the art in ML for HEP (Detector GANs, GNN, model-free searches) with references, curated datasets, etc. Emphasize depth of HEP expertise

Projects

- FastML: ***μ*s inference for HLT**
- The **ML fast chain**: replace expensive traditional detector simulation+reconstruction with GAN simulation+ML pattern reco. TFlop/event → GFlop/event
- **Model-free** semi/weakly/unsupervised new physics searches.

Opportunities

- Collaboration with math and CS experts towards a NSF AI institute or a DOE AI initiative call

Weaknesses & Gaps

- Not strong connections to the foundational groups in HPC ML community

Data Organization, Management and Access

Scope & Activities

- **Organization**: **Contents** of events (AOD vs xAOD vs PHYS-lite), memory layouts (CCE IOS?), data **formats** (ROOT vs HDF5 vs RNTuple), **compression**.
- **Management**: Policy-based data placement (Rucio) and alternate transfer mechanisms, database-like access (SkyHook).
- **Access**: Cache-based access; event delivery (ServiceX, IDDS).

Potential Milestones & Deliverables

- Prototype to convert NANOAOD/PhysLite into HDF5/Parquet.
- Formulate R&D topics related to columnar analysis to start discussion with ECP.
- Develop whitepaper on U.S. LHC storage facilities model for HL-LHC.

Projects

- IRIS-HEP (ServiceX, IDDS, SkyHook, columnar analysis, XCache). DIANA (compression).
- DOE-HEP (Rucio)
- U.S. ATLAS Ops (Rucio, XCache)
- U.S. CMS Ops (Rucio, XCache)
- (proposed) CCE IOS.

Opportunities

- Columnar-based data analysis promises significant improvements in data rates.
  - Essential for using accelerators; potential overlap with ECP.

Weaknesses & Gaps

- Need an agreed-upon facilities model (esp. Integrating caching).
- Better define metrics to understand when we should transition to new models (e.g., caching).
- Could use better alignment with SciDAC-4 & CCE plans.

Storage infrastructure and Facilities

Scope & Activities

- Storage at facilities and in the network
- From cold storage and its dynamic use to low latency analysis storage
- Improve processing at HPCs through object stores?
- Optimize random access for analysis vs. whole file transformation to save space?
- Understand storage hierarchy and scale implications of HL-LHC.

Potential Milestones & Deliverables

- Define APIs to dynamically interact with cold storage; more than “Is this on tape, fetch it for me”.
- Requirements review, use cases, cost analysis and fundamental definition of our storage workflows
    - Give to facilities and vendors to come up with solutions
    - Prepare DOE exascale storage round table

Projects

- IRIS-HEP SSL, OSG-LHC, SLATE, SCAILFIN
- WLCG QoS group ➜ Data carousel
- DOE exascale storage round table

Opportunities

- Can we avoid the 2nd tape copy of RAW data and rely more on transatlantic network?
- Develop a Facilities R&D program focusing on resource flexibility, "substrate" layer, multi-prem service mesh & orchestration APIs

Weaknesses & Gaps

- Currently the experiments / community define storage hierarchies themselves. Is this what we’re going to do in 5-10 years? Or are the facilities optimizing this for the experiments?

Data Transfer and networking infrastructure

Scope & Activities
- Continue to clarify the bandwidth needs and workflows ➜ ESnet concerned if the workflows are changing dramatically for HL-LHC (bulk vs. streaming)
- What are the expected deliverables from the infrastructure ➜ know by 2024/2025 so that ESnet can procure for HL-LHC
- How are sites deploying DTNs? Impact of caching/streaming? Esp. for HPC sites.

Potential Milestones & Deliverables
- Rename analytics group, change mandate to look back (understand historical network flows) and forward (planning for future flows)
- Create blueprint to define metrics & requirements and define plan for evolution/innovation

Projects
- SAND network analytics, OSG networking area, ESnet FABRIC, SENSE
- Working group: ESnet, SURFnet, NorduNet, Geant, Internet2, Canarie (Canada) ➜ transatlantic network
- WLCG working groups: Network Function Virtualization WG, Throughput WG
- PerfSonar group

Opportunities
- Help from ESnet to connect HPC centers to distributed data infrastructure of experiments
- Federated service orchestration across facility edge networks to accelerate innovation & reduce operations costs

Weaknesses & Gaps
- Updating data transfer protocols ➜ coordinate OSG with ESnet studies

Workflow and Resource management

Scope & Activities

- Workflow and workload management systems must seamlessly and optimally orchestrate among all available resources - grid, HPC, clouds
- Systems flexible enough to accommodate future workflows on all resources - eg. ML workflows, workflows specific to new architectures, etc
- Support coscheduling, offloading, edge and streaming services, etc

Potential Milestones & Deliverables

- 6 month: develop a list of potential HPC use cases for HEP.
- 6-12 months: requirements review and survey of industry solutions.
- 1 year: document on requirements, on work needed, and priorities for workflow and workload management systems.

Projects

- Data carousel - manage hot, warm, cold data
- Distributed hyperparameter scans
- Coscheduling/splitting on CPU and GPU
- Streaming data, event service

Opportunities

- Increase in available computing resources
- Reduction in disk storage usage
- Reduction in network usage
- Optimal use of resources for future workflows

Weaknesses & Gaps

- Integration of distributed resources with HPCs
- Usability of accelerators by HEP applications
- Running data intensive applications on HPCs

Event Processing Frameworks

Scope & Activities

- New computing landscape dominated by parallel processing and heterogeneity poses many questions requiring:
  a. Language support for heterogeneous computing
  b. Data models that are adapted for execution in heterogeneous environments.
  c. Scheduling tools
  d. Interfaces to other toolkits such as ML toolkits
  e. Interaction and interface between frameworks and workload management

Potential Milestones & Deliverables

- Decide on a programming model that people will use to write an algorithm on an accelerator.
- Frameworks integrate scheduling with this programming model.

Projects

1. CMSSW(ART) effort at FNAL
2. LBNL effort on Athena
3. LBNL effort on Ray (UCB/Rise Lab)
4. Potential CCE effort on portability libraries (FNAL, LBNL)
5. CERN effort (Attila, SFT effort unclear)
6. Gaudi - traditionally CERN SFT - not clear if there is the required effort here
7. RAL/Edinburgh effort on Athena

Opportunities

- CCE effort can be kernel, Framework developers, can drive effort on portability libraries
- We have 4 people on C++ standards committee, our work could drive things more broadly than HEP - get vendors to adopt a single standard for using accelerators

Weaknesses & Gaps

- R&D Developers feel overly constrained by weight of existing frameworks and software
- Early choice forced by schedule constrains the field before sufficient R&D is done

Physics Generators

Scope & Activities

- Improving data and computing scalability of current generators on current HPCs
- Redesigning underlying algorithm for performance on both CPUs and GPUs
- Includes existing LO and NLO processes
- The need for NNLO processes should be better justified by the physics community
- We are working with ATLAS/CMS to share generated events

Potential Milestones & Deliverables

- Better requirements modeling:
  - current usage, estimate needed improvements
  - LO vs NLO vs NNLO (from physics groups)
- Platform independent algorithm development (theorists)
- In support of shared event generation for LHC experiments: require common defined input configs (from physics groups), data storage, formatting, and indexing.

Projects

- SciDAC4 (FNAL & ANL)
- CCE (FNAL,ANL,LBNL,BNL)
- USATLAS (ANL)

Opportunities

- Our community is driving ATLAS and CMS to begin using the same generated events.
- We must begin setting up the shared infrastructure to enable the experiment's production system to consume them
- Pool effort from NSF, DOE, and IRIS-HEP to form Joint US-LHC EvGen team

Weaknesses & Gaps

- Generator authors are not concentrated in the US
- Authors (physics theorists) motivated by physics publications and are not explicitly part of the LHC community.
- No European FTE towards generator optimization (e.g., accelerators).

Simulation

Scope & Activities

- Geant
  - Geant4 inefficient on modern hardware
  - GeantV effort recently rampped down
  - New Effort for GPU-enabled simulation
- Fast simulation
  - GAN-based approaches not yet matching parameterized models

Potential Milestones & Deliverables

- A newly-developing pilot effort for GPU-enabled simulation (Celeritas)
  - Needs to meet needs of CMS and Atlas in time for Run4
    - Aggressive R&D pilot project
- Fast simulations need to reach needed accuracy for Run4

Projects

- ECP-based pilot for Geant (FNAL, OLCF, ANL)
  - Aggressive timeline to demonstrate viability
  - Builds on elements of of GeantV
- CERN is resetting its priorities

Opportunities

- Neutron transport from OLCF (Shift) has a demonstrated solution in a similar problem space
  - Can be used as a basis for GeantX investigations
- New AI initiatives could be used for fast simulation

Weaknesses & Gaps

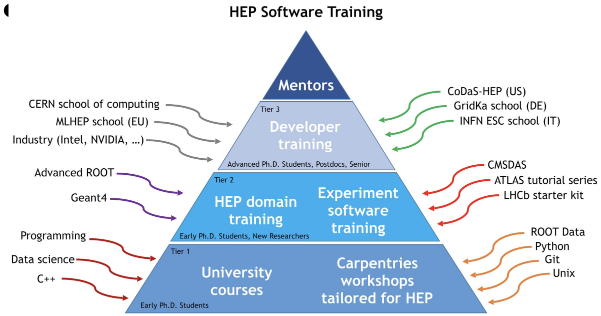
- Interactions with CERN and community on Geant are complex
  - CERN seems willing to let the US take the lead on Geant for GPUs
- Geant4 needs to be fully supported until an alternative is ready
- Developing a new version of Geant in time for Run 4 will take a significant effort

Visualization

Scope & Activities ● USCMS effort to modernize ROOT’s event display infrastructure that leverages industry tools ● USATLAS work to abstract geometry	**Projects** ● ROOT EVE ● USATLAS ● USCMS **Opportunities** ● CMS is working on geometry and so is ATLAS
Potential Milestones & Deliverables ● Sustainable geometry infrastructure capable of describing Phase2 detectors	**Weaknesses & Gaps** ● USLHC experiments have chosen different directions

Training

Scope & Activities



Potential Milestones & Deliverables

- Blueprint, best practices
- Standard curriculum - basic carpentries - shell/github/python/plotting
- Training - C++/ROOT/Geant
- Suggestion - Integrate with trainings that XSede, OSG, ECP, the LCFs provide
- Suggestion - Increase frequency of CoDaS-HEP type schools (US based)

Projects

- IRIS-HEP, FIRST-HEP
- Synergy - HSF

Opportunities

- A lot - learning more by teaching, transmit knowledge expertise to next generation, share across experiments/fields, contribute at all levels of training pyramid
- Involve even more female and under-represented participation via outreach, hackathons for broader impact
- Share/initiate training experience with neutrino/nuclear physics communities

Weaknesses & Gaps

- Challenge of finding teachers/facilitators
- Different level of engagement by experiments
- Optimal time to host training - clash with experiment specific and field specific events - collab meetings/conferences