High Energy Physics Center for Computational Excellence

AHM Welcome/Introduction/Overview

Salman Habib
for HEP-CCE

All-Hands Meeting
December 18, 2023

HEP-CCE All Hands Meeting

December 18-19, 2023

Welcome to Argonne!

HEP-CCE Mission

HEP-CCE Primary Mission

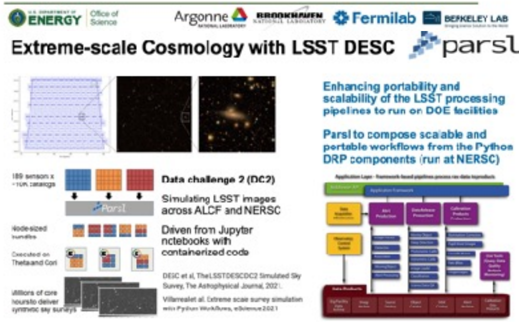
- Bring large-scale computational and data management resources to bear on pressing HEP science problems
- Focus on cross-cutting solutions leveraging ASCR and HEP expertise and resources (ESnet, LCFs, NERSC, IRI, --)
- Work closely with experiments and carry out other tasks as directed by DOE HEP
- HEP-CCE as a joint effort across the participating laboratories – key to providing the ability to integrate individual strengths and resources

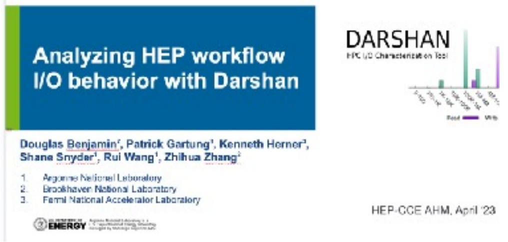
U.S. DEPARTMENT OF ENERGY | Office of Science · Argonne NATIONAL LABORATORY · Brookhaven National Laboratory · Fermilab · BERKELEY LAB Bringing Science Solutions to the World

HEP-CCE: Current Context

HEP-CCE Roles

- FY19-23; HEP-CCE funded as a joint effort – now is considered to be a base program targeting:
  - Future architectures — evolution of ASCR HPC systems and HEP priorities for future systems and engagement with the HPC environment (ASCR, cloud, international, e.g., ALCF-4, NERSC-10, OLCF-6 –)
  - Long-term engagement with ASCR facilities — LCF and NERSC projects (future HPDF and IRI connections); significant ASCR/HEP "cross-infrastructure" opportunities
  - LCFs moving to host data-intensive computing resources – separate resources and via queueing rules on exascale systems
  - HEP requirements — What do experiments need from ASCR facilities? What is possible to do on ASCR HPC systems? How do these resources fit within the HEP computing ecosystem? How can the HEP and ASCR HPC environments best engage and help each other?





HEP Experiments and the HPC Ecosystem

HEP Computing Resource Challenges

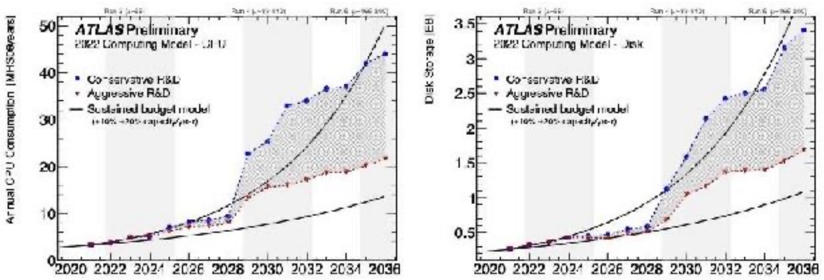
- HL-LHC example: 10X data, 10X complexity
- In 2030, LHC experiments will need:
  - > O(100) PFlops in sustained compute performance
  - O(10) Exabyte/year data throughput

HEP Long-Term Involvement in HPC

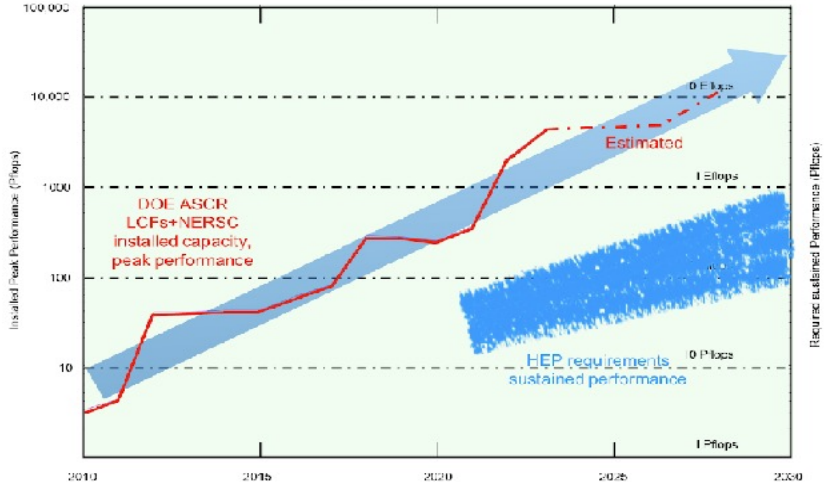
- Chosen platform for Cosmic Frontier (CMB-S4, DESI, LSST DESC, LZ, —)
- LHC experiments among top users of NERSC's Cori system (~10% of available cycles in 2021)
- DUNE has identified HPC resource utilization as a major component of its computing strategy

HPC Challenges

- Ability to run multiple HEP workflows at O(10) Pflops sustained on multiple heterogeneous exascale systems
- Match HEP I/O requirements to HPC file systems and networking infrastructure
- Address realtime to near-realtime applications and resilience challenges



ATLAS CPU and disk requirements



Evolution of ASCR and HEP computing resources

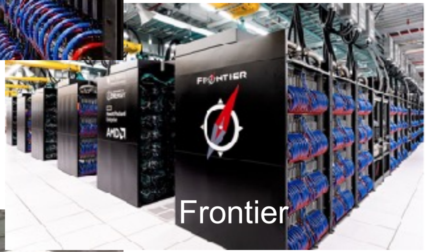
What do Experiments Need from HEP-CCE?

ASCR Liaison Role

- Track ASCR HPC system evolution, provide early access/input to ASCR facility planning and broker access to HPC testbeds
- Enable long-term allocations at ASCR facilities; effort also related to sustainability issues for programmatic DOE/SC software

Specific HEP-CCE Tasks

- Ability for HEP workflows to exploit concurrency at the node level of HPC platforms — natural event-level parallelism is not enough
- Multiple compute platforms as a portability challenge — complex HEP code needs to run on grid/cloud/HPC systems in production
- I/O needs to scale to thousands of HPC nodes — not just for data but also for software libraries, databases, configurations (~100K files)
- Ability to run complex HEP workflows robustly on HPC systems – resilience challenge
- Mixed CPU/GPU pipelines will become more common as AI/ML methods are increasingly incorporated – HPC facilities offer unique large-scale AI/ML training opportunities



ASCR HPC systems (2023)

Evolution of HEP-CCE Role

HEP-CCE Phase I

- Exploration of portability frameworks/tools and successful demonstration of running HEP applications on heterogeneous architectures
- Experiment-agnostic mimicking framework – testing structure for I/O performance across I/O packages
- I/O characterization using Darshan; improvement of Darshan for HEP applications
- HDF5 as intermediate event storage for HPC workflows
- Porting EGs from CPUs to GPUs; working with the broader EG community
- Identified need to move from applications to workflows with different requirements across the partner experiments
- Workflows for remote computing applications and improved workflow monitoring
- Data storage and data management, more generally, identified as new vectors for HEP-CCE efforts

HEP-CCE Phase 2

- Transition to mini-apps and cookbooks for interactions with experiments as distinct from work on individual applications
- Optimizing data reduction, organization, and data movement for reducing the data storage footprint
- End-to-end demos and tests for experiment workflows (including hybrid examples), testing interfaces, resilience, and other requirements (e.g., near-realtime) for successful integration with facilities
- Accelerated simulation effort to continue work on event generators and to potentially include GPU-optimized optical photon transport (important for DUNE and LZ)
- AI/ML application scale-up on HPC resources as component of hybrid workflows
- Tighter, planned coordination with HPC facilities as part of IRI and other initiatives
- Cross-cutting outreach and EDIA program with its own set of funding/resources

HEP-CCE Structure Evolution (Technical)

Current: Four Technical Teams

- Coordinated by team leads
  - Stakeholder oversight
- **PPS: Portable Parallelization Strategies**
  - Exploit massive concurrency
  - Portability requirements
- **IOS: HEP I/O and HPC Storage**
  - Accelerator-friendly data models
  - Fine-grained I/O, event batching
- **EG: Event Generators**
  - Optimizing event generators for accelerated systems
- **CW: Complex workflows**
  - Porting and running HEP workflows on HPC systems

Future: Four Technical Teams

- Coordinated by team leads
  - Stakeholder oversight
- **Optimizing Data Storage**
  - Data management, data reduction/compression
  - Optimized data delivery
- **Portable Applications to Portable Workflows**
  - HEP workflow overlay for HPC
  - Hybrid CPU/GPU application support
- **Scaling up HEP AI/ML Applications**
  - Large-scale training and hyperparameter optimization on HPC systems
- **Accelerating HEP Simulation**
  - Event generators for accelerated systems
  - Optical photon transport on GPUs
  - GPU-enabled tracking geometries

HEP-CCE Structure Evolution (Management)

Management Team
- PI, Deputy PIs (technical coordination, outreach/EDI), project coordinator*
- Technical Leads
- Institutional Council
  - PI, Deputy PIs, admin lead
  - Institutional leads

Steering Group
- PI, Deputy PIs, project coordinator
- Experiment POCs for ATLAS, CMB-S4, CMS, DUNE, Rubin/LSST DESC, LZ
- ASCR facility POCs for ALCF, NERSC, OLCF

Technical Team
- PI, Deputy PI (technical coordination), project coordinator
- Technical Leads (Data storage, Portable Applications/Workflows, Scaling up HEP ML, Accelerating HEP Simulation)
  - Task leads

*administrative support

Current HEP-CCE Organization

Leadership

- **PI/Deputy PIs:** Salman Habib, Paolo Calafiura (Technical), Kerstin Kleese Van Dam (EDI)
- **Institutional Leads:** Salman Habib (Argonne), Kerstin Kleese Van Dam (BNL), Adam Lyon (Fermilab), Peter Nugent (LBNL)
- **Experiment Reps:** Ed Moyse (ATLAS), Julian Borrill/Ted Kisner (CMB-S4), Jim Chiang (Rubin/LSST DESC), Lindsey Gray (CMS), Andrew Norman (DUNE), Maria Elena Monzani (LZ)

Technical Teams

- **PAW: Portable Applications and Workflows** (Leads: Charles Leggett, Meifeng Lin)
  - Exploit massive concurrency and heterogeneous architectures
  - Portability in computing as well as running HEP workflows on HPC systems
- **SOP: Storage OPtimization** (Leads: Rob Ross, Peter van Gemmeren)
  - Accelerator-friendly data models, data reduction, IO optimization on HPC systems
  - Data delivery/orchestration
- **SIM: HEP SIMulations for Experiments** (Leads: Taylor Childers, Tom Evans, Stefan Hoeche)
  - Optimizing event generators and detector simulation for accelerated systems
- **SMiLe: Complex workflows** (Leads: Paolo Calafiura, Walter Hopkins)
  - Enabling HEP AI/ML applications at scale on ASCR HPC resources

Project Coordinator:
Samantha Tezak

Evolution of ASCR Facility Roles

ASCR Facility Plans

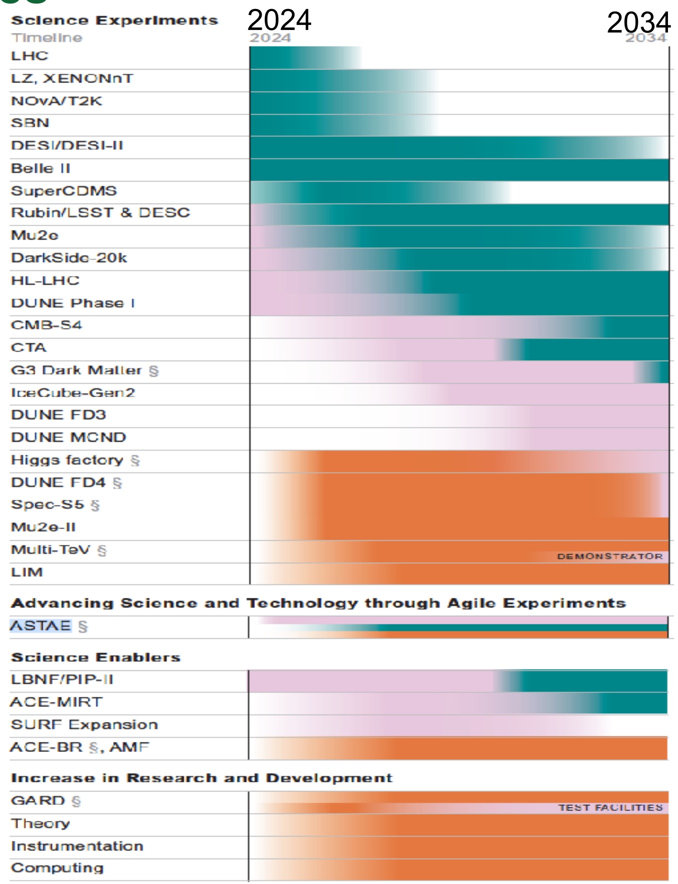
- Replacements for current systems already in planning stages – NERSC-10 in 2026+, OLCF-6 in 2027+, ALCF-4 in 2028+
- All require significant improvements in compute performance, ability to handle data-intensive applications and AI/ML workloads, all more or less within current power envelopes
- All acknowledge IRI as a scenario within which the system should operate
- Given the timescales, specific architecture choices are somewhat uncertain but CPU/GPU mixes are likely
- AI/ML will be the key driver for hardware evolution

HEP-CCE

HEP Experiment Timelines

HEP-CCE Planning

- Track experiment timelines as well as architecture/software/facility/integration roadmap
- Current choice of HEP-CCE experiments aims to provide coverage as well as aim at best use of ASCR facilities; new experiments can be added in the longer term, partly depending on P5 recommendations
- Need to work with experiments that are in planning, construction, and execution phases
- Experiment needs, priorities and readiness also factor into work with the facilities as well as with broader ASCR software efforts (including ASCR R&D)
- HEP-CCE Steering Group consisting of experiment and facility POCs will be a very important component of planning going forward
- Key HEP-CCE focus will be on pilot “end to end” demo/proof projects (aside from individual challenges) to address portability, infrastructure, and resilience challenges
- Note recommendations for computing in the P5 report are positive, but, in general, the needed scope of effort is much larger than the resource-limited outlay; CPAD-like group is being set up



Schedule

December 18

- Morning Session 1 (Plenary): Intro/overview talks (9am-10.30am)
- Morning Session 2 (Plenary): Input from experiments and stakeholders (11am-12.50pm)
- Working Lunch (get lunches and go to parallel sessions)
- Parallel Sessions: First-round planning (1pm-4pm)
- Report Out and Overnight Action Items (Plenary): Draft Plans (4pm-5.15pm)

December 19

- Morning Session 1 (Plenary): Phase 1 Wrapup Reports (9am-11am)
- Morning Session 2 (Plenary): Updated Draft Plans (11.30am-1.30pm)
- Working Lunch (1.30-2pm)
- Summary/Conclusions: Organization/Actionables (2pm-4pm)