High Energy Physics
Center for Computational Excellence
Overview

Meifeng Lin

Slide by Salman Habib

Placing HEP-CCE in Context

HEP-CCE

HEP-CCE Primary Mission

- Bring large-scale computational resources to bear on pressing HEP science problems
- Focus on enabling cross-cutting solutions leveraging ASCR and HEP expertise and resources

HEP-CCE Roles

- ASCR-HEP connection (liaison and cross-cutting technical R&D)
  - Future architectures — evolution of ASCR High Performance Computing (HPC) systems and HEP priorities for future systems and engagement with the HPC environment (ASCR, cloud, international, —)
  - Long-term engagement with ASCR facilities — Leadership Computing Facility (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 (connected to IRI); DOE’s FASST initiative for frontier AI4S
  - 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?

Slide by Salman Habib

HEP Experiments and the HPC Ecosystem

HEP-CCE

HEP Computing Resource Challenges

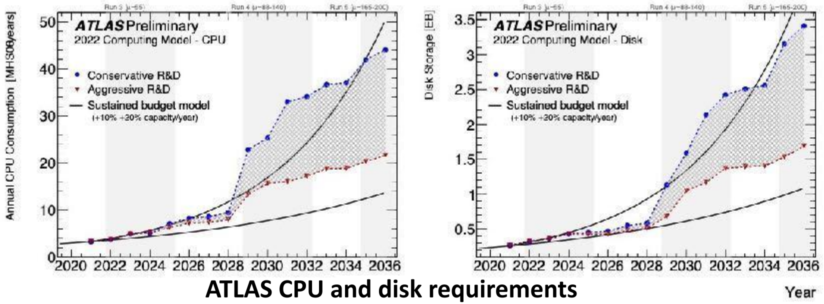
- HL-LHC example: 10X data, 10X complexity
- In 2030, LHC experiments will need:
  - > O(100) PFlops in sustained compute performance
  - O(10) Exabyte/year data throughput

HEP Long-Term Involvement in HPC

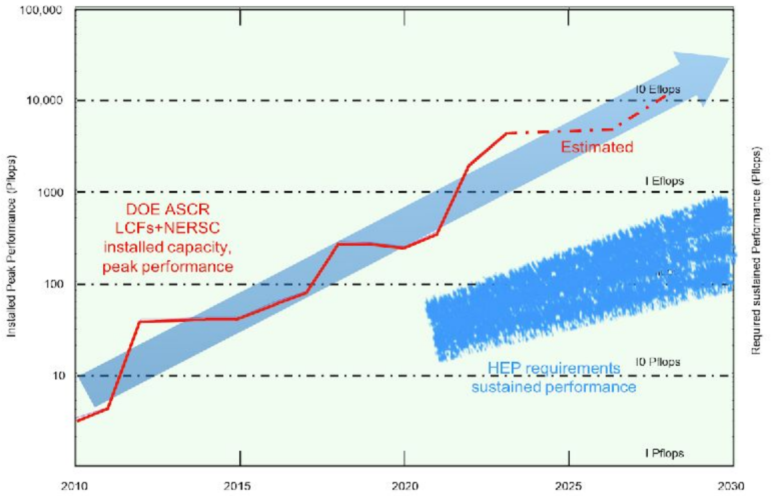
- Chosen platform for Cosmic Frontier (CMB-S4, DESI, LSST DESC, LZ, —)
- LHC experiments among top users of NERSC's Cori system (~10% of available cycles in 2021)
- DUNE has identified HPC resource utilization as a major component of its computing strategy

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

Slide by Salman Habib

What do Experiments Need from HEP-CCE?

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 pose a portability challenge — complex HEP workflows need to run on grid/cloud/HPC systems in full-scale 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 ML methods are increasingly incorporated – HPC facilities offer unique large-scale AI/ML training opportunities

ASCR HPC systems (2022/2023)

Slide by Salman Habib

HEP-CCE Structure Evolution (Technical)

HEP-CCE

Past: 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

Current: 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
  - Frontier AI applications
- **Accelerating HEP Simulation**
  - Event generators for accelerated systems
  - Optical photon transport on GPUs
  - GPU-enabled tracking geometries

Slide by Salman Habib

HEP-CCE Phase 1 Organization

HEP-CCE

Leadership

- **PI/Deputy PI:** Salman Habib, Paolo Calafiura
- **Institutional Leads:** Salman Habib (Argonne), Adam Lyon (Fermilab), Kerstin Kleese Van Dam (BNL), Peter Nugent (LBNL)
- **Experiment Reps:** Julian Borrill (CMB-S4), Jim Chiang (LSST DESC), Oli Gutsche (CMS), Andrew Norman (DUNE), Torre Wenaus (ATLAS)

Technical Teams

- **PPS: Portable Parallelization Strategies** (Leads: Oli Gutsche, Charles Leggett, Meifeng Lin)
  - Exploit massive concurrency
  - Portability requirements
- **IOS: HEP I/O and HPC Storage** (Leads: Rob Ross, Peter van Gemmeren)
  - Accelerator-friendly data models
  - Fine-grained I/O, event batching
- **EG: Event Generators** (Leads: Taylor Childers, Stefan Hoeche)
  - Optimizing event generators for accelerated systems
- **CW: Complex workflows** (Leads: Kyle Chard, Shantenu Jha)
  - Porting and running HEP workflows on HPC systems

Slide by Salman Habib

Current HEP-CCE Organization

HEP-CCE

Leadership

- **PI/Deputy PIs:** Salman Habib, Paolo Calafiura (Technical)
- **Institutional Leads:** Salman Habib (Argonne), Meifeng Lin (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

Slide by Salman Habib

HEP-CCE Structure Evolution (Management)

HEP-CCE

Management Team

- PI, Deputy PIs (technical coordination, outreach/EDI), admin lead
- Technical Leads
- Institutional Council
  - PI, Deputy PIs, admin lead
  - Institutional leads

Steering Group

- PI, Deputy PIs, admin lead
- 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), admin lead
- Technical Leads (Data storage, Portable Applications/Workflows, Scaling up HEP ML, Accelerating HEP Simulation)
  - Task leads

Slide by Salman Habib

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

BNL Activities in HEP-CCE Phase 2

HEP-CCE

- HPC systems are getting more diverse and complex, in terms of both the hardware and software environments
    - **Hardware:** Processing units: CPU, GPU, APU, etc, diverse memory hierarchies
    - **Software:** Different programming models, software modules, job manager, authentication/authorization
    - There is a need to enhance the current software ecosystem to make HPC more accessible to NPP users

- **Goal**: Develop cross-cutting solutions to meet HEP computing needs with advanced HPC capabilities, leveraging ASCR and HEP technologies and incorporating latest AI/ML advancements.

- Part of the **HEP-CCE** project, BNL is investigating portable workflow management solutions that can enable seamless workflow executions across Grid and LCF sites (and cloud computing in the future)
    - ATLAS and DUNE workflows as initial test cases
    - For ATLAS workflows, leverage the **PanDA** workflow management system and **Globus Compute** (FuncX)
    - For DUNE workflows, investigate integrating **Superfacility API** with existing workflow management systems
    - Also closely monitor and coordinate with the IRI development

- We also need application software that can take advantage of the massive parallelism on HPC systems
    - Portable programming models are preferred, but also open to vendor APIs if proven advantageous.
    - **Manual porting is time consuming => Leveraging LLMs to automate part of the process**
    - Ongoing work evaluates using open-source/commercial LLMs with RAG to perform various coding tasks: **code documentation, summarization, explanation, porting and optimization.**

Atif, Mohammad, et al. "Evaluating portable parallelization strategies for heterogeneous architectures in high energy physics." *arXiv preprint arXiv:2306.15869* (2023).

Brookhaven National Laboratory

Argonne National Laboratory · Brookhaven National Laboratory · Fermilab · Berkeley Lab

BNL Personnel

Name	FTE (FY25)	Task 1	Task 2	Task 3
Meifeng Lin	0.4	Institutional PI	PAW Lead	
Mohammad Atif	0.65	PAW-FastCaloSim miniapp	PAW-Cookbook/AI Code Assistant	SML-scaling up training
Dennis Bollweg	0.2	SML-scaling up training		
Kriti Chopra	0.32	PAW-Cookbook/AI Code Assistant		
Ozgur Kilic	0.31	PAW-DUNE Workflow	PAW-Cookbook/AI Code Assistant	
Mikhail Titov	0.16	PAW-ATLAS Workflow		
Matteo Turilli	0.03	PAW (currently on leave)		
Tianle Wang	0.24	PAW-ATLAS Workflow		
Doug Benjamin	0.1	PAW-ATLAS Workflow		
Wen Guan	0.15	PAW-ATLAS Workflow		
Total	*2.56*			

Brookhaven National Laboratory