Discussion on Next Steps

Timeframe for HEP-CCE 2

Circa 2024-2028

HL-LHC scheduled to start in 2029, DUNE in 2030(?)

Next DOE HPC iteration is 2028+

From cross-cutting Research to experiment-specific Development

Portable Parallelization Strategy

Short-term deliverables

      Publish recommendations → Define workflow to provide experiment-dependent recommendations

      Recommendations will be time-dependent: track evolution of software tools, hardware, and -likely- paradigms

Future deliverables:

      Add use cases, identify patterns, build tutorial examples,

      monitor parallel computing evolution → develop parallelization benchmark (and/or contribute to WLCG HEPScore)

      Continue to provide feedback to portable parallelization tool developers

      Actively participate in use case collection for next gen HPCs

New ideas:

      Green computing

      Integrate ML inference, distributed training

IOS

Bring current work (emulator, Darshan, HDF-5 EDM) to experiments

HPC-friendly data model

1) document summarizing talks from experts,
2) performance studies as new approaches become available in production.
3) extract design (anti-)patterns
4) Potentially foundation data model classes usable by future experiments.

Lossy compression

Demonstrated 20% savings for ATLAS/CMS, resistance from physics community → understandable given potential biases

Investigating "intelligent" lossy compression schemes from other domains.

Domain-specific compression algorithms may be an easier sell,

Potential gains order of magnitude larger

Event Generators

Current effort part of a complex sociology

Need to make a strategic choice:

    Focus on portability, leaving the "physics" to other

    Develop new functionality e.g. GPU version for cpu-intensive high-priority NLO processes

        For that need resource-loaded plan with a HL-LHC/DUNE timescale

            Probably too large for CCE but could play role

            Similar to Celeritas/ADEPt

Complex Workflows

Looking forward to SRS aimed at capture characteristics of HEP workflows (scale, structure, portability, etc)

For example

Run hundreds of thousands of jobs 24/7/365 in 20 countries

Synchronize data movement (PB/day) with execution

Recently: Complex ML/optimization workflows (simulation-driven inference active learning, etc)

How well do they map to ExaWorks products?

HEP-CCE PIER Plan

We mainly interact with experiments, facilities, and with ASCR researches

Not much yet in terms of interactions with Universities.

    No formal fellowship or training effort

    Nascent role in CompHEP traineeship program

Need to setup a strong team to work in detail our future EDI program