

HPC Demonstrators: Status Update

HEP CPU/GPU benchmarks as first demonstrators

David Southwick (CERN / IT-DI-OPL)

on behalf of

HPC integration in WLCG / openlab (access of resources)

HEPIX CPU Benchmarking WG (benchmarking activity)

CERN WLCG GDB

8 July 2020

Introduction

Common challenges for HPC integration are driving the technical program of demonstrators:

❑ HPC Benchmarking

❑ Data Access demonstrator

- Demonstrate I/O can effectively use local storage and cached data delivery at the scale of HPC
- Working with WLCG DOMA and the Datalake prototypes

❑ AAI demonstrator

- Demonstrate authentication and authorization needs can be met within the stricter security requirements of HPC sites
- Working with WLCG Authorization working group

Common Challenges: <https://zenodo.org/record/3647548#.XwMUYS2Q3aY>

HPC benchmark demonstrator

First task for my contract; supervised by M. Girone.

HPC Benchmark Demonstrator – result of a collaboration of efforts:

- ❑ HEPiX benchmarking working group
- ❑ HPC integration in WLCG working group
- ❑ Support from SDSC, DEEP, CINECA, and Subatech HPC centers

Exploiting synergies between these groups advances development on HPC!

Concepts

HPC Benchmarking Demonstrator adopts the HEP Benchmark Suite - developed by the HEPiX Benchmarking W.G.

HEP-Benchmarks: <https://gitlab.cern.ch/hep-benchmarks>

❑ HEP-workloads:

- Production WLCG workloads from the experiments
- Measures events/sec (throughput) as the relevant metric
- Standalone (Self Contained), container-ised (Docker/Singularity)

❑ HEP-score:

- Benchmarks resources using HEP-Workloads
- Orchestrates these workloads

❑ HEP-Benchmark-Suite:

- Meta-orchestrator of multiple benchmark suites (HEP-score, HS06, SPEC2017)

HEP Benchmarks: <https://indico.cern.ch/event/908146/timetable/#7-benchmarking-and-accounting>

Development process

Standalone containers with GPU Workloads are developed, tested, packaged, and distributed using CERN Infrastructure (as all the other HEP-workloads)

- ❑ **Experiment workloads** packaged as OCI-standard Docker & Singularity images using gitlab-CI of hep-workloads-GPU project
- ❑ **Distributed** via gitlab registry and CVMFS

In addition, testbeds run on:

- ❑ heterogeneous resources @ CERN (IT-CM, IT-CF, openlab, ...)
- ❑ **Scale deployments across HPC resources** at SDSC, MARCONI, Subatech

Access to HPC resources

Support from HPC sites has been crucial for testing and enhancement of HEP-benchmarks on HPC centers

- ❑ Access to large pool of nodes and various hardware configurations
- ❑ Ability to scale across clusters and partitions, integrating with job scheduling tools (SLURM)
- ❑ +Functional tests on Subatech & Cineca



SDSC SAN DIEGO SUPERCOMPUTER CENTER

CPU Nodes	Skylake 8168 (144x2), 6148 (72x2) Cascade 8268 (216x2)
Cores	~28K + 128 GPUs
DRAM/node	768 GB
GPU	4x NVIDIA V100/node

DEEP
Extreme Scale Technologies

CPU Nodes	Skylake 6146 (50x2) Cascade 4125 (75), 8260M (16x2)
Cores	2568 + 91 GPUs + 16 FPGA
DRAM/node	192GB / 48GB / 348GB
GPU	1x NVIDIA V100/node

https://fz-juelich.de/ias/jsc/EN/Expertise/Supercomputers/DEEP-EST/_node.html

https://www.sdsc.edu/support/user_guides/popeye-simons.html

HEP-Workloads (CPU/GPU) run on HPC

Successful deployment using variety of methods (pip, wheel, tarball etc.)

- CPU HEP-workload benchmarks (Atlas, CMS, LHCb; gen, sim, digi, reco) using HEPscore
- GPU Workloads (Work in progress)
 - Simple track (O. Datskova, R. De Maria), Patatrack (CMS + G. Moreno, D. S. D. Giordano. A.Sciaba) running on GPUs
 - not yet as tuned as benchmark
 - First focus was on interoperability
- ROOT Analysis Workloads (Work in progress)
- Summer student (D. Kankowska) activity (X. Valls)
- The HEPiX Benchmarking WG is working to deliver HEP-Workloads-GPUs based on the experiments workloads developed for GPU

CMS GPU workloads seminar <https://indico.cern.ch/event/927838/>

Challenges faced on HPC

Assumptions valid for WLCG sites (and vendor servers) are not valid in HPC centers -
HPC sites generally more restrictive.

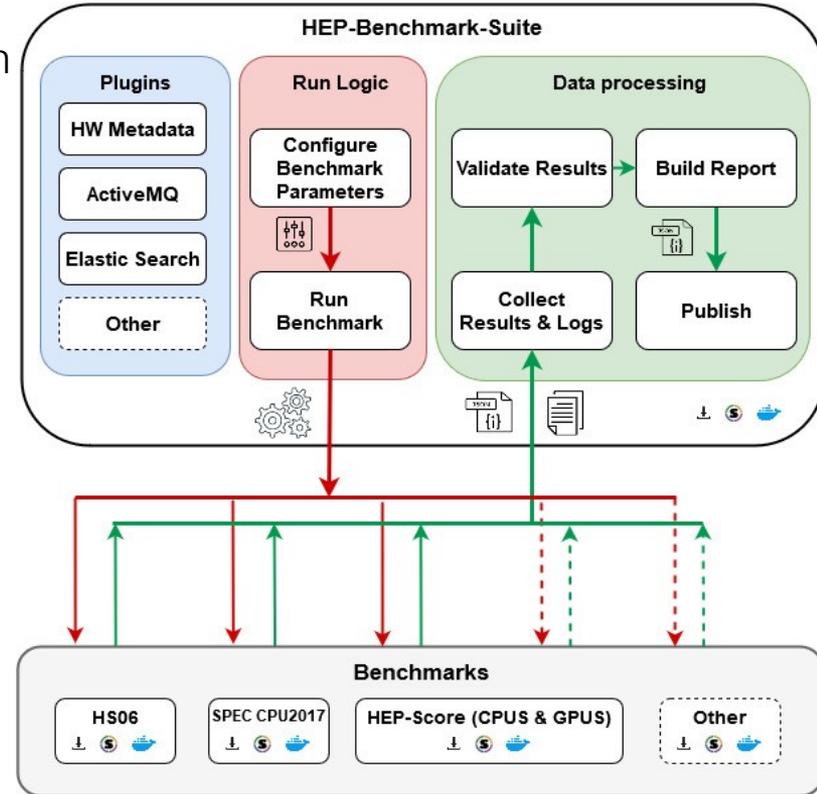
Typical constraints include:

- no Docker
- no unprivileged user namespace executions
- no nested containers
- (potentially) no external connectivity on worker nodes
 - Or limited to/with CERN

Overcoming HPC constraints

The HEPiX Benchmarking WG aims for simple installation and running of all HEP benchmarks (HS06, HEPscore, etc.), therefore is working to overcome the HPC constraints:

- ❑ HEP Benchmark Suite will be installed with a minimum set of python modules
- ❑ is modular (in plugin packages) and declarative (in configuration)
- ❑ **Unprivileged** and root-less



M. Medeiros, D. Giordano, D. Southwick

Initial benchmark results

Successful collection of benchmarking reports across partitions at SDSC:

- ❑ Starting small with individual benchmarks
- ❑ Extending to large-scale CPU+GPU jobs via slurm

```
{
  "status": "success",
  "environment": {
    "date": "Wed May 13 06:03:23 2020",
    "system": "Linux workergpu06 3.10.0-1062.9.1.el7.x86_64 #1 SMP Fri Dec 6 15:49:00 EST 2019",
    "container_exec": "singularity"
  },
  "wl-scores": {
    "cms-gen-sim-bmk": {
      "gen-sim": 1.319
    },
    "cms-reco-bmk": {
      "reco": 3.5363
    },
    "atlas-sim-bmk": {
      "sim": 0.1239
    },
    "lhcb-gen-sim-bmk": {
      "gen-sim": 152.5704
    },
    "cms-digi-bmk": {
      "digi": 5.4014
    },
    "atlas-gen-bmk": {
      "gen": 529.9112
    }
  },
  "hash": "b8810078c312c43caab5ba859c994b151eb0c743c368482a1c650dc7553705a6",
  "name": "HEPscore19",
  "repetitions": 3,
  "scaling": 355,
  "version": "0.x",
  "reference_machine": "CPU Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz",
  "registry": "gitlab-registry.cern.ch/hep-benchmarks/hep-workloads",
  "benchmarks": {
    "lhcb-gen-sim-bmk": {
      "run1": {
        "duration": 1050,
        "report": {
          "wl-scores": {
            "gen-sim": 152.328,
            "gen": 821.7859,
            "sim": 196.0516
          }
        }
      }
    }
  },
  "score": 584.3714982864489,
  "method": "geometric_mean",
  "container_exec": "singularity"
}
```

Producing meaningful results

Scaling and Memory usage study using

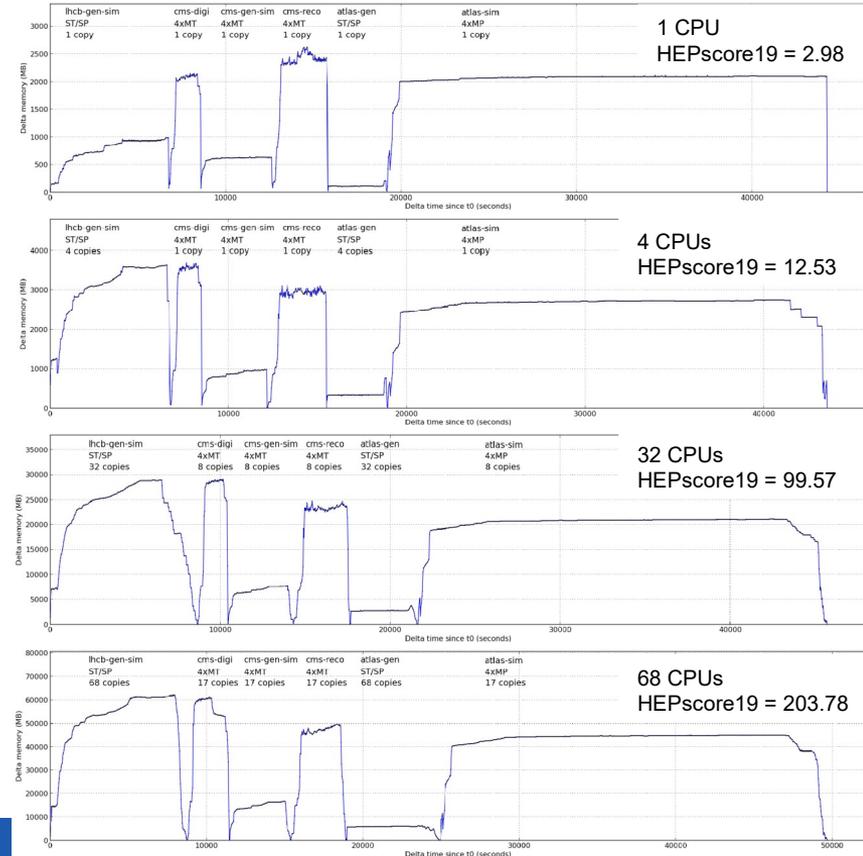
HEPscore on Marconi-A2 HPC at

CINECA:

- ❑ non-traditional x86 CPUs (KNL)
 - Spinoff of LHCb activities on KNL
- ❑ 6 WLCG workloads in order
- ❑ Memory usage is sustained

Memory profiles

HEPscore on Marconi-A2 HPC at CINECA (1x 68-core 1.4GHz KNL with 4x HT)



Study by A. Valassi: <https://indico.cern.ch/event/872154/>

Conclusions & next steps

- ❑ First demonstrator for HPC are HEP benchmark workloads
 - very good progress thanks to the work already done for CPU benchmarking
 - finalize the development of the GPU CMS Patatrack WL as standalone benchmark container
 - prepare the build infrastructure for other HEP GPU workloads or for other architectures (IBM Power)
 - HEP workloads running on GPUs and usable for offline benchmarking are still a few. Need to increase the list
- ❑ Future Work
 - Run at large scale on HPC centers, different CPU models and data centers (SDSC, MARCONI100 ...)
 - Next challenge will be work on a Data Access demonstrator, in close collaboration with WLCG DOMA DataLake prototypes
 - Would like to run in more HPC centers – Please contact us if you want to help in this activity!
- ❑ **More reports in the next few months!**



Questions?

Further Information

- ❑ Challenges for HPC integration: <https://zenodo.org/record/3647548#.XwMUYS2Q3aY>
- ❑ HEP Benchmarks: <https://indico.cern.ch/event/908146/timetable/#7-benchmarking-and-accounting>
- ❑ CMS GPU seminar: <https://indico.cern.ch/event/927838/>
- ❑ Scaling and memory study A. Valassi: <https://indico.cern.ch/event/872154/>
- ❑ DEEP-EST: https://fz-juelich.de/ias/jsc/EN/Expertise/Supercomputers/DEEP-EST/_node.html
- ❑ SDSC: https://www.sdsc.edu/support/user_guides/popeye-simons.html
- ❑ MARCONI: <http://www.hpc.cineca.it/hardware/marconi>
- ❑ Subatech: <http://www-subatech.in2p3.fr/en>