HEPScore – a new CPU benchmark for the WLCG

Randall Sobie (Victoria) and Domenico Giordano (CERN)
Co-chairs of the WLCG Task Force on CPU Benchmarks

On behalf of the
HEPiX CPU Benchmarking Working Group
WLCG HEPScore Deployment Task Force

Motivation

CPU benchmarks are an important part of the WLCG infrastructure

Experiment requests and site pledges
Accounting of CPU usage
Many sites use them for procurements

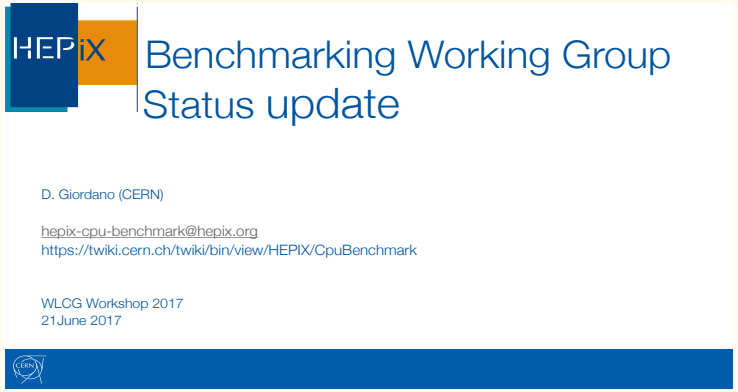
The current WLCG benchmark, HEPSpec06 (2009), has several drawbacks

Not representative of HEP workloads (HEP workloads are more performant on newer hardware)
HEPSpec06 is the 32bit version
SPEC stopped supporting the underlying SPEC-CPU 2006 benchmark (2018)

WLCG needs a benchmark for other processors (ARM and GPUs)

We have HEP workloads for ARM from a number of experiments
Workloads with GPUs are just emerging

History I


HEPiX Benchmarking Working Group
Status update

D. Giordano (CERN)

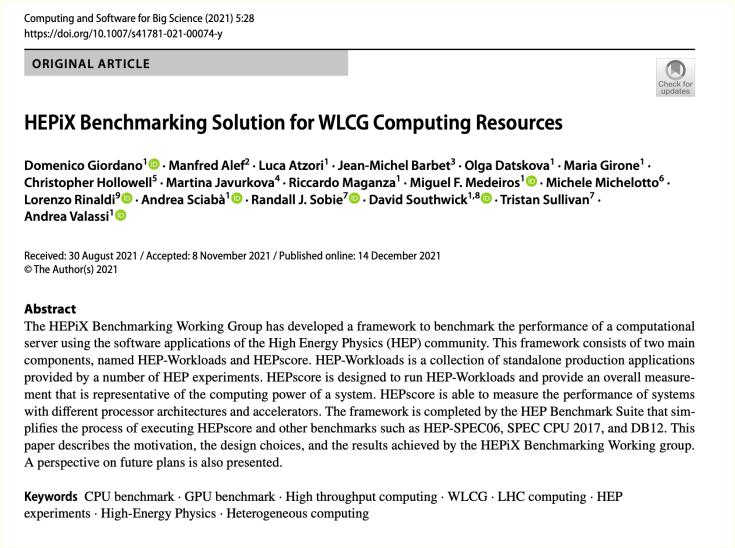
hepix-cpu-benchmark@hepix.org
https://twiki.cern.ch/twiki/bin/view/HEPIX/CpuBenchmark

WLCG Workshop 2017
21June 2017

WLCG Workshop Manchester 2017

HEPiX Benchmarking WG: First proposal of HEP Benchmark with containerized HEP applications

https://indico.cern.ch/event/609911/contributions/2620190/attachments/1480455/2295576/WLCG_Workshop_2017_benchmarking_giordano.pdf


Computing and Software for Big Science (2021) 5:28
https://doi.org/10.1007/s41781-021-00074-y

ORIGINAL ARTICLE

HEPiX Benchmarking Solution for WLCG Computing Resources

Domenico Giordano[1] · Manfred Alef[2] · Luca Atzori[1] · Jean-Michel Barbet[3] · Olga Datskova[1] · Maria Girone[1] · Christopher Hollowell[5] · Martina Javurkova[4] · Riccardo Maganza[1] · Miguel F. Medeiros[1] · Michele Michelotto[6] · Lorenzo Rinaldi[9] · Andrea Sciabà[1] · Randall J. Sobie[7] · David Southwick[1,8] · Tristan Sullivan[7] · Andrea Valassi[1]

Received: 30 August 2021 / Accepted: 8 November 2021 / Published online: 14 December 2021

Abstract
The HEPiX Benchmarking Working Group has developed a framework to benchmark the performance of a computational server using the software applications of the High Energy Physics (HEP) community. This framework consists of two main components, named HEP-Workloads and HEPscore. HEP-Workloads is a collection of standalone production applications provided by a number of HEP experiments. HEPscore is designed to run HEP-Workloads and provide an overall measurement that is representative of the computing power of a system. HEPscore is able to measure the performance of systems with different processor architectures and accelerators. The framework is completed by the HEP Benchmark Suite that simplifies the process of executing HEPscore and other benchmarks such as HEP-SPEC06, SPEC CPU 2017, and DB12. This paper describes the motivation, the design choices, and the results achieved by the HEPiX Benchmarking Working group. A perspective on future plans is also presented.

Keywords CPU benchmark · GPU benchmark · High throughput computing · WLCG · LHC computing · HEP experiments · High-Energy Physics · Heterogeneous computing

Benchmark Suite and feasibility studies 2020

"**Suite**" is the infrastructure for running containerized HEP workloads

"**Workloads**" are experiment developed applications (gen, sim, reco)

"**HEPScore**" a single number based on a combination of workloads

See article in CSBS and conference proceedings

History II

WLCG Management Board launched a benchmark Task Force (Nov 2020)

H. Meinhard CERN/IT (Chair)

D. Giordano CERN/IT and R.Sobie Victoria (Co-Chairs since July 2022)

Experts in benchmarks, accounting, sites and representatives from the HEP community

Four LHC experiments plus Belle-II, LIGO/VIRGO/KAGRA, JUNO/BES-III, DUNE

Bi-weekly meetings

CERN in-person Workshop in September 2022

Reports to the WLCG MB and GDB, and contributions to the HEPiX Workshops

Initial focus on finding a CPU benchmark for x86-based systems

And a transition plan from HEP-SPEC06 to HEPScore (Next talk)

Timescales are driven by the WLCG cycle for pledges (scrutiny group)

Pledges for FY2025 are made in Oct 2023

HEPScore needs to be finalized by early summer 2023 to be used in FY2025 pledge cycle

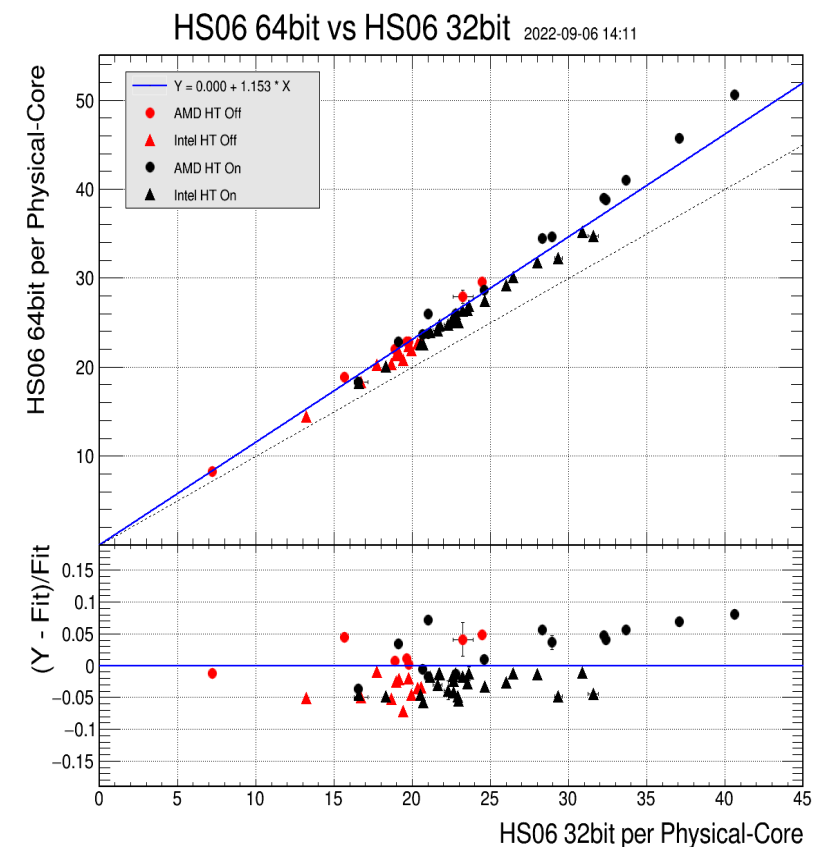
Workloads need to be finalized by April 2023

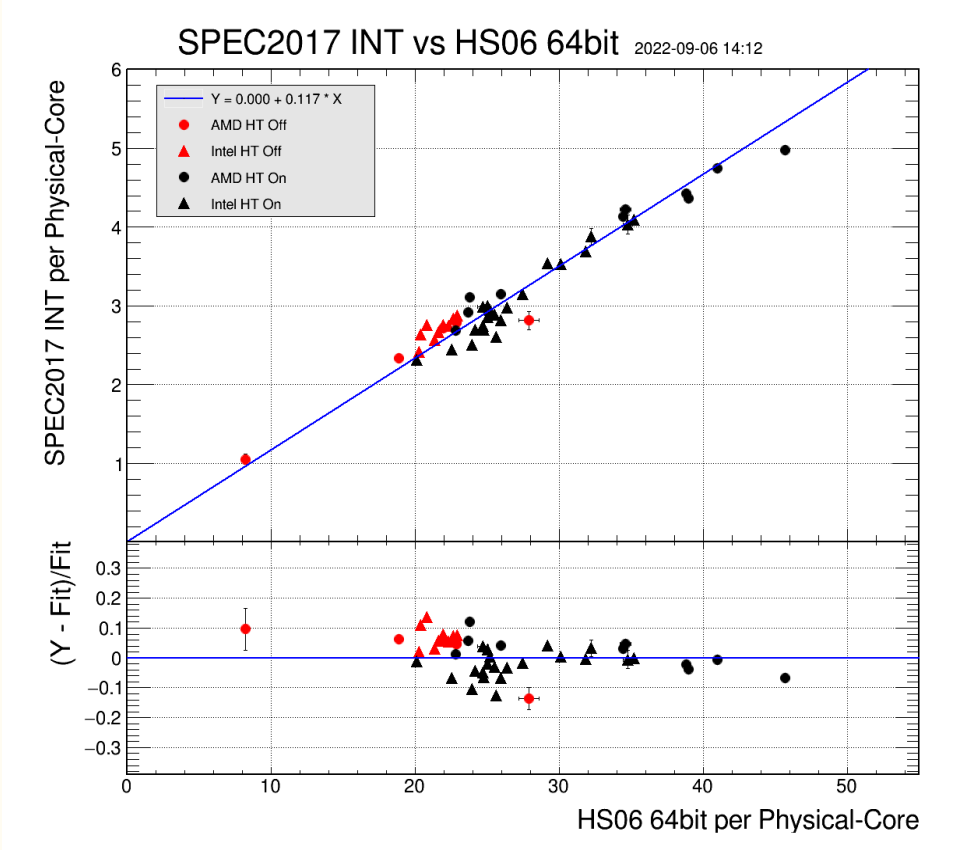
Review of HEP-SPEC06 and validation of HEP Workloads

Extensive campaign 2021-2022 to accumulate a large set of measurements

HEP-SPEC06, SPEC2017 and HEP Workloads

Approximately 70 different “systems” (CPU, cores, site, hyper-threading) around the world





SPEC2017
newer version of SPEC-CPU 2006
(integer/FP versions)

HEP-SPEC06
is based on a subset of SPEC-CPU2006
(32 and 64 bit versions)
(32 bit version used by WLCG)

Workloads
Run3 workloads for LHC experiments

alice_gen_sim_reco

atlas_gen_sherpa
atlas_sim_mt
atlas_reco_mt

belle2_gen_sim_reco

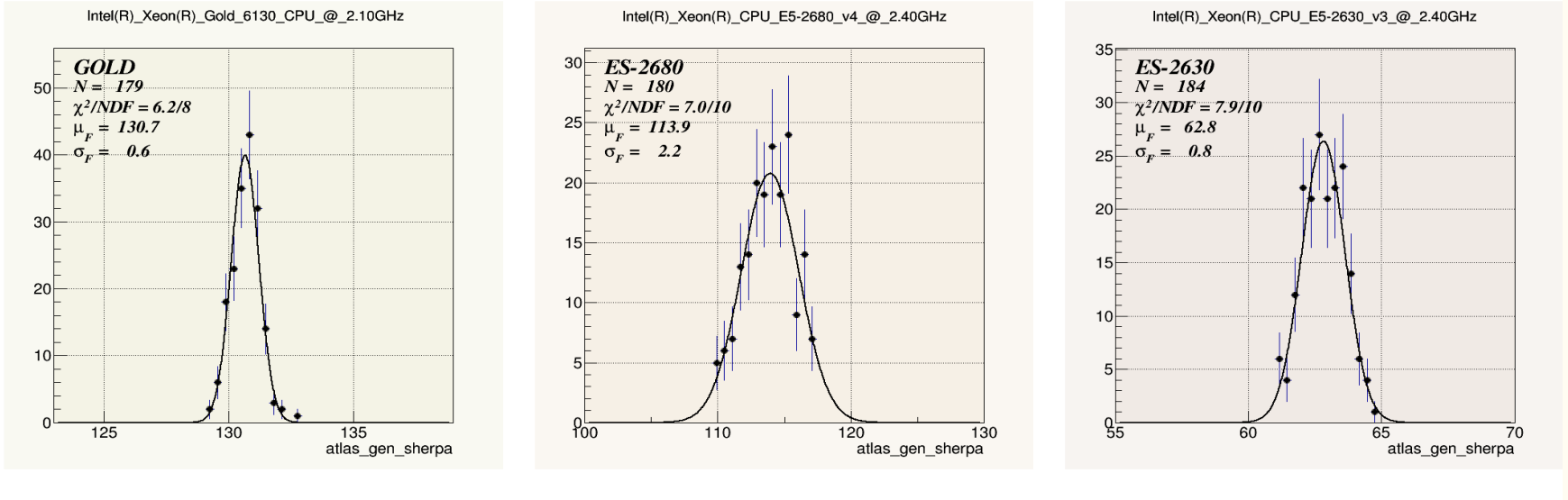
cms_reco
cms_digi
cms_gen_sim

juno_gen_sim_reco

igwn_pe (Gravity Wave)

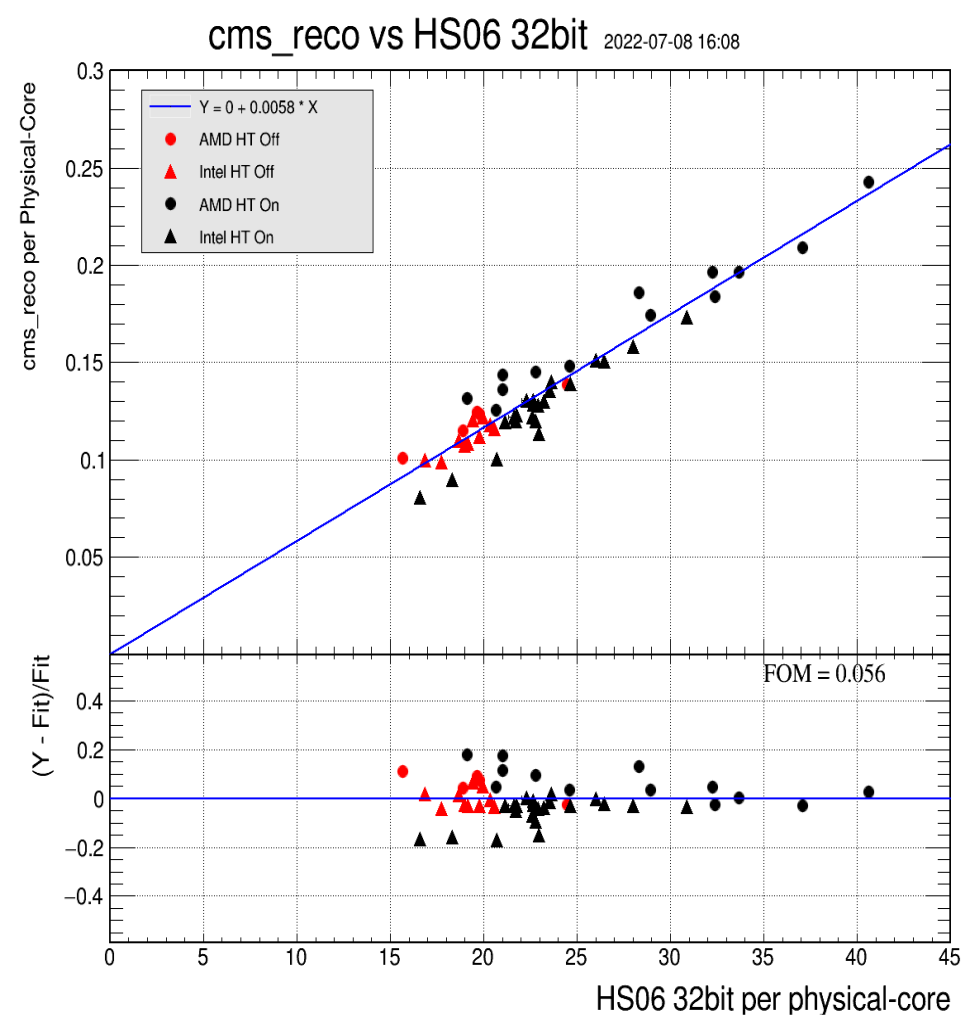
lhcb_gen_sim

Each workload has been run and validated on a set of CERN servers
Reliable/reproducible to < 1%



Some of the workloads were found to have technical issues (some are still being resolved)

Workloads vs HEPSpec06 and SPEC2017



Plots of Workload results (events/s) vs HEPSpec06
Normalized to the number of physical cores

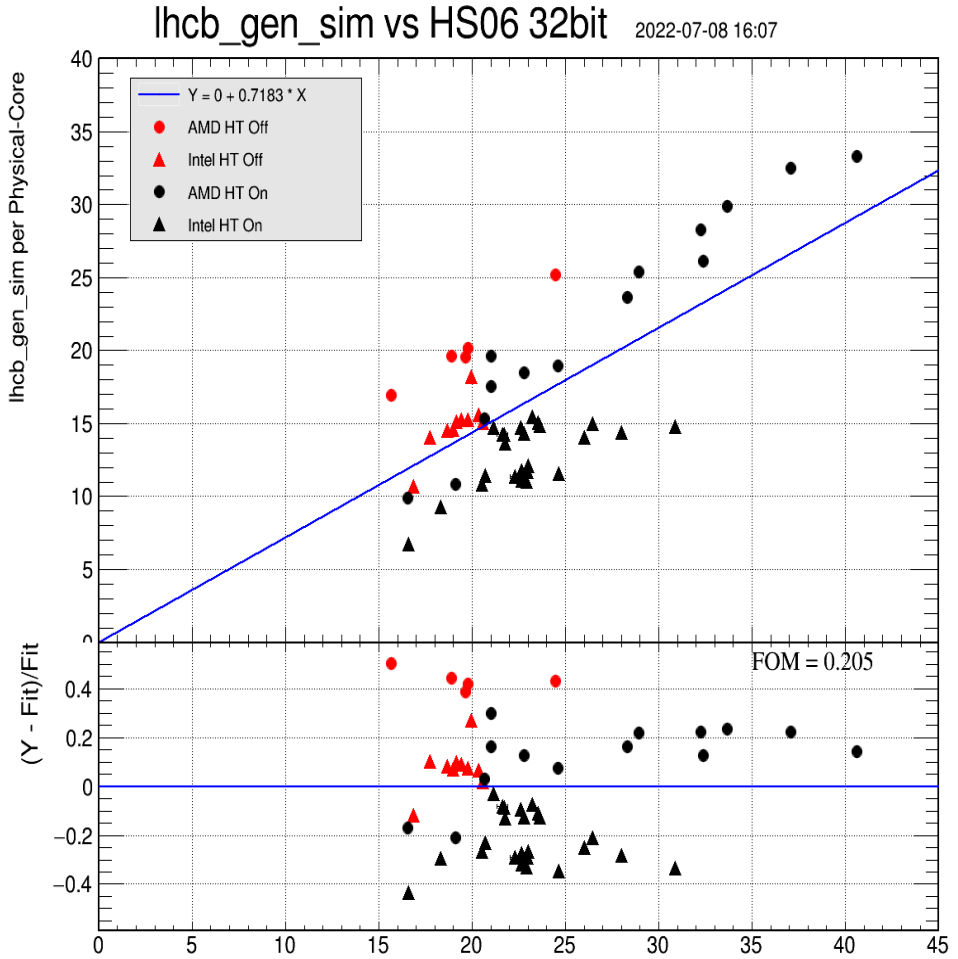
Red	HT Off
Black	HT On
Circle	AMD
Triangle	Intel

Measured deviation in y-axis is used as a metric (FOM)

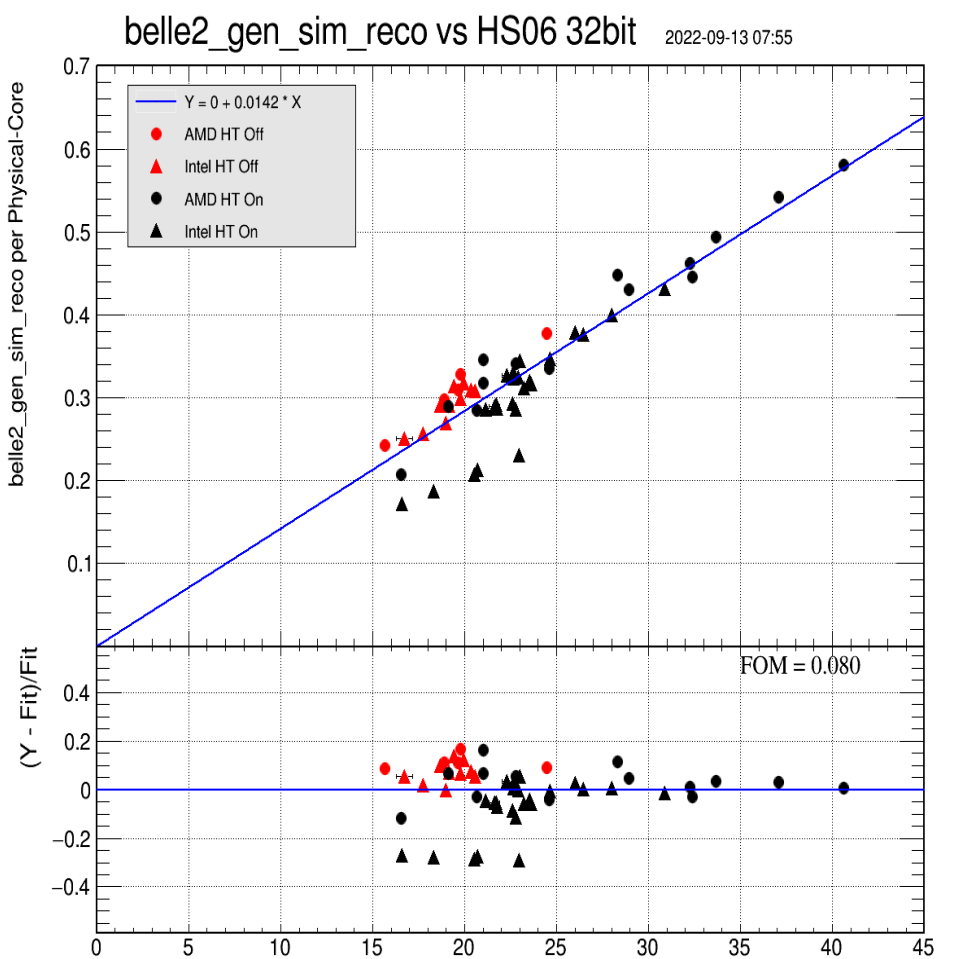
Plots and tables of each of the 11 workloads vs 4 benchmarks

Workloads vs HEP-SPEC06 (32bit)

HEP-SPEC06 (32-bit)
is used for pledges and accounting

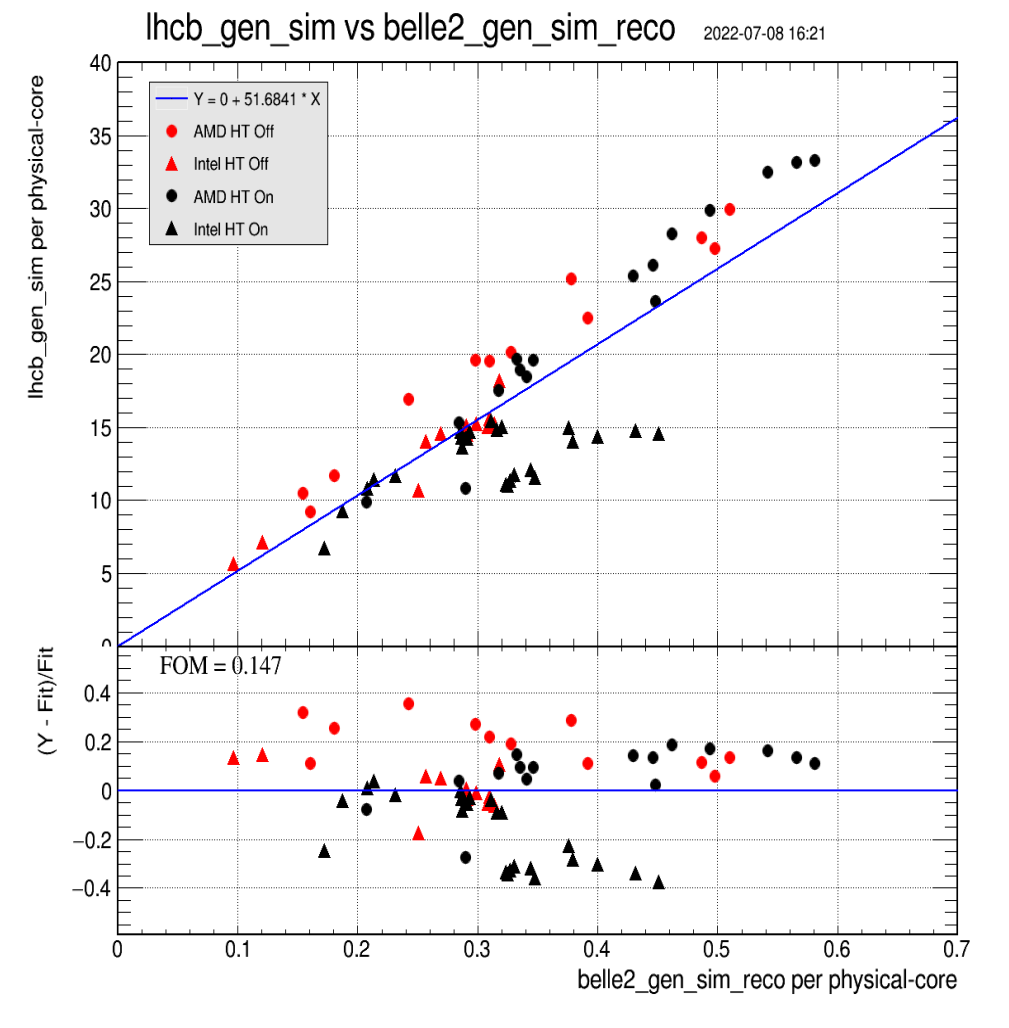


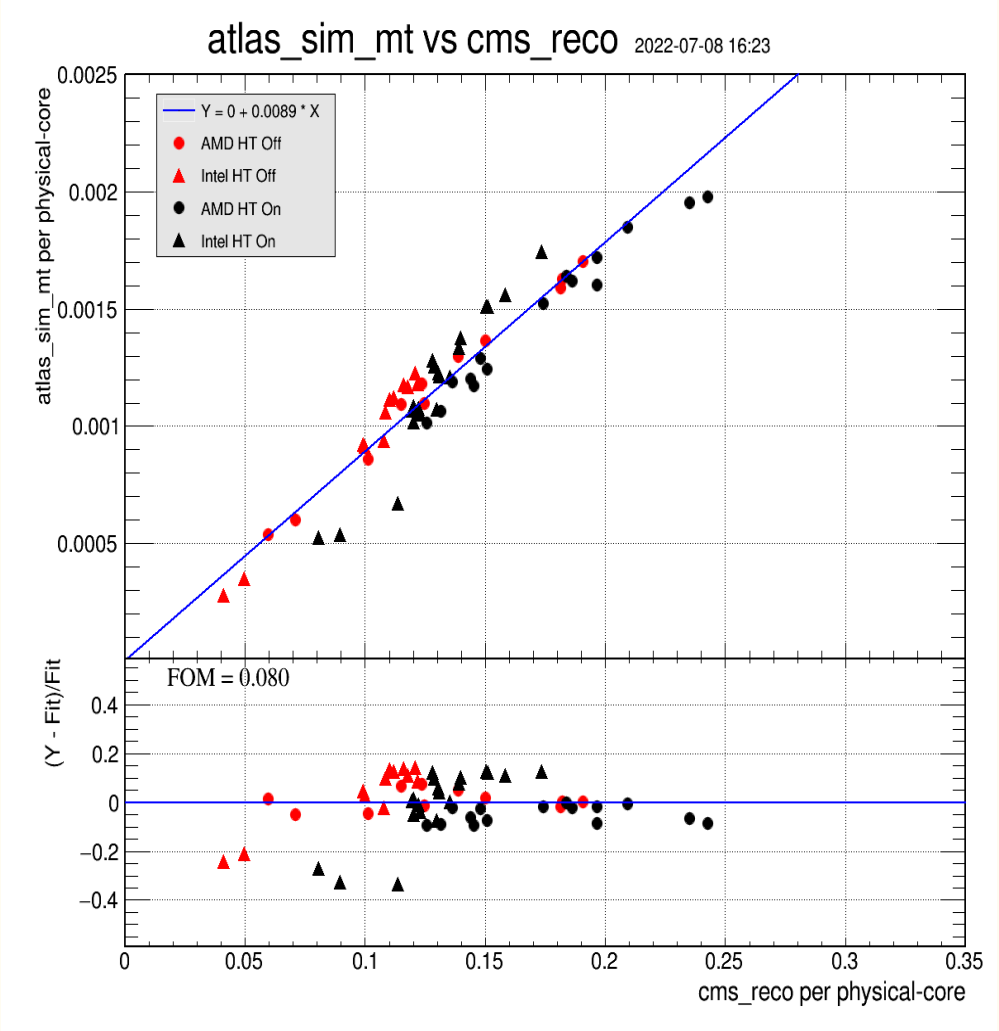
44 plots



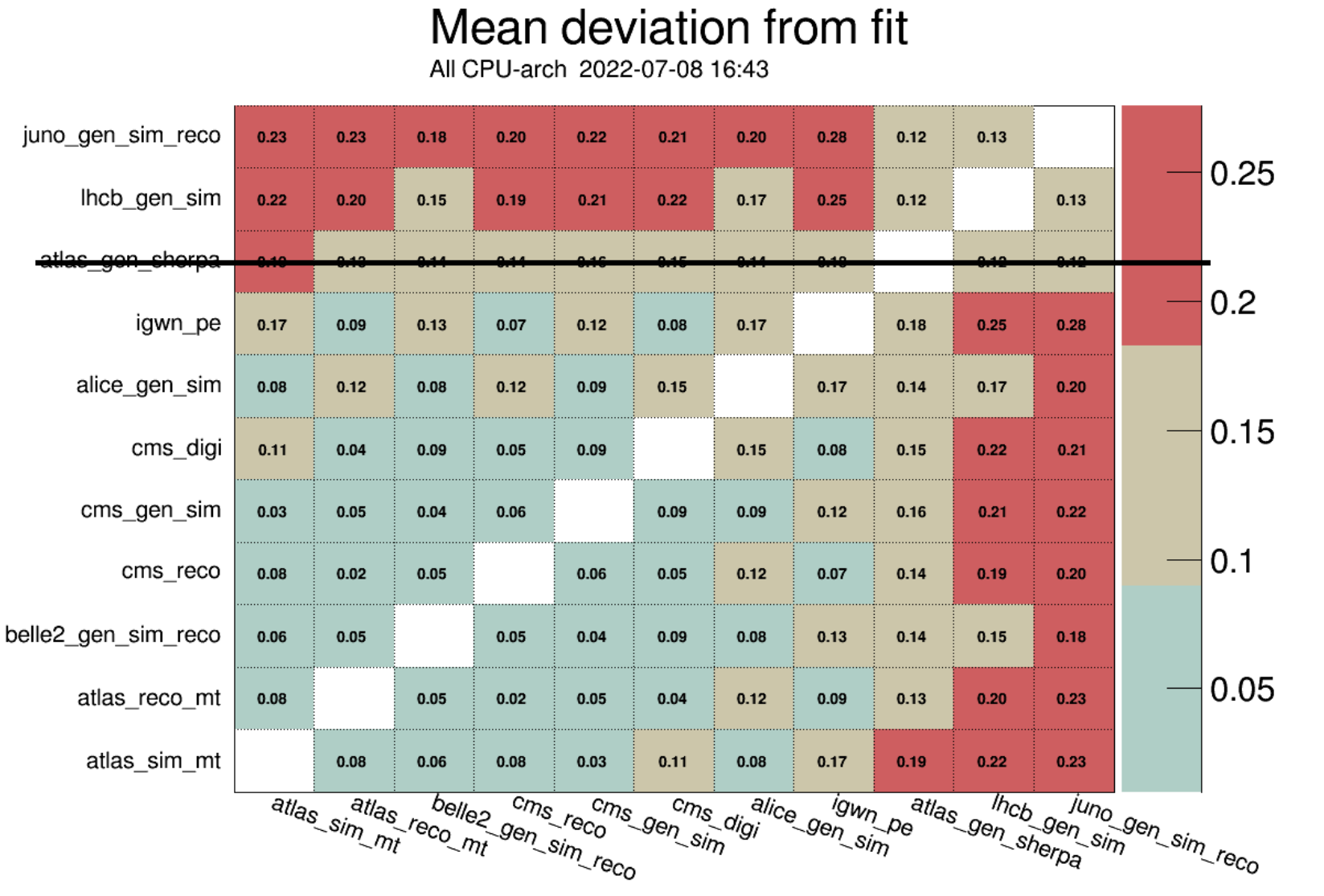
Workloads vs Workloads

50 combinations





Workload vs Workload - mean deviation from fits



Matrix gives an indication of the correlation between the workloads

Larger deviations are considered as a "feature" of the workload

Note that atlas_gen_sherpa had an issues and a new version was validated this week

Task Force Survey

Surveyed the TF members for thoughts on how to select HEPScore

1. Support for a HEPScore benchmark based on LHC and other experimental workloads
2. **HEPScore should reflect the relative CPU usage of the experiments and application**
3. **HEPScore should run in a timely manner 3-6 hours**
4. **HEPScore should be valid for one or more LHC beam period**
5. Interest in a “fast HEPscore” and a “CPU+GPU HEPScore” in the long term

Task Force Workshop held at CERN on September 19, 20

https://indico.cern.ch/event/1170924/

Summary:

https://cernbox.cern.ch/s/DRhWUzgDQowRZwT

Potential x86 HEPScore candidate sample

The HEPiX WG and WLCG TF converged on an initial set of 7 workloads

ALICE	(reco)
ATLAS	(gen_sherpa and reco_mt)
Belle II	(gen_sim_reco)
CMS	(digi and gen_sim)
LHCb	(gen_sim)

Key criteria: Reliable workload, short-runtime, complementarity, lesser correlation
Time to run this set is 3-6 hours depending on the server performance (see backup)

Weighting: We considered different ways of combining the workloads
We found that equally weighting the workloads was as good as other option
"Equal weighting" is close to the observed CPU usage on the WLCG Grid

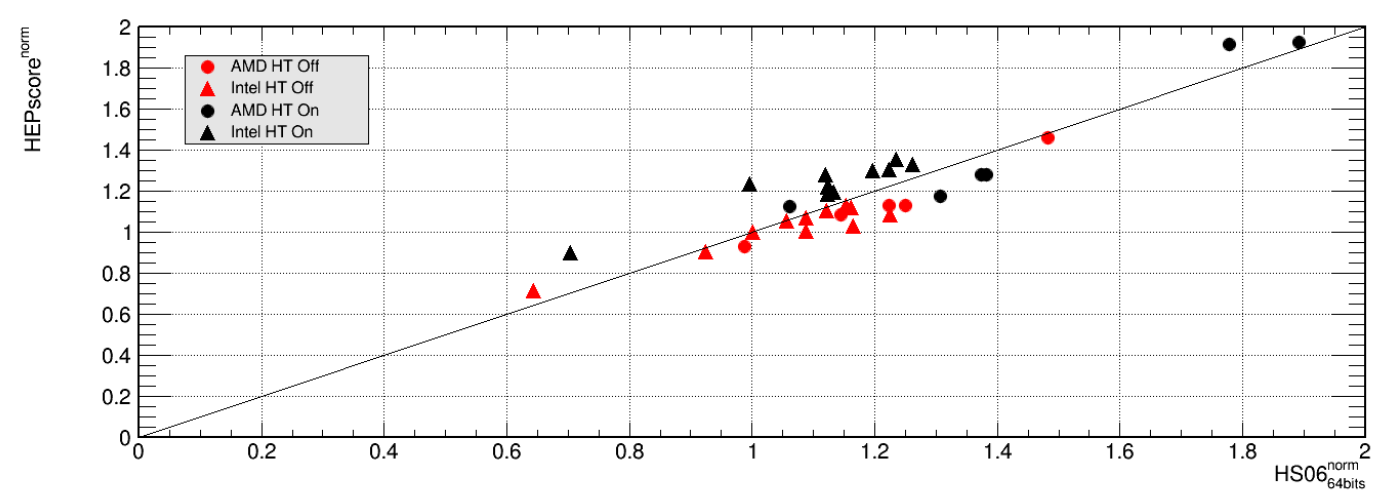
Cross checks: Removing one workload did not make a significant impact
We looked at the results using "newer" CPUs and found little difference

Calculation of HEPScore

The workloads are measured in a similar manner as HEP-SPEC06
- Each workload is run 3 times and we take the geometric mean

We renormalize the measurement (events/second) to the results of a "Reference Machine"
- Currently the Reference Machine is older Intel that is being replaced with a current Ice Lake server

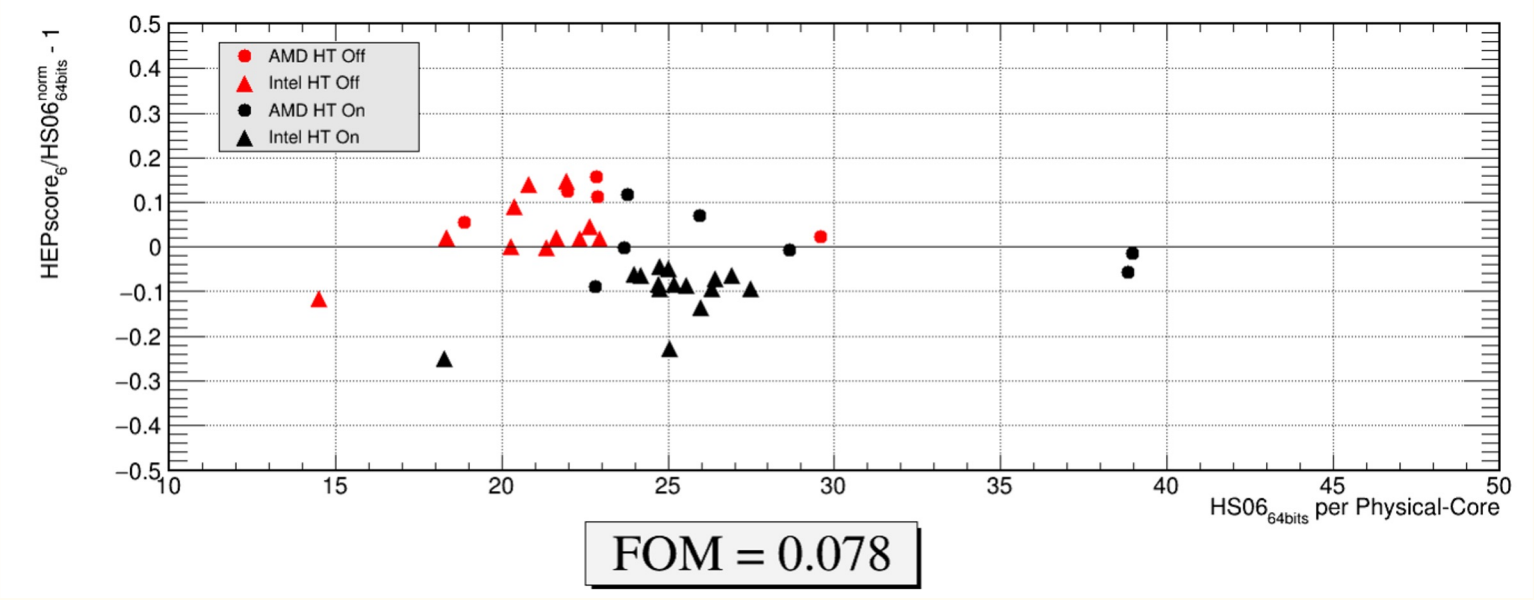
The candidate workloads are combined by taking the geometric mean



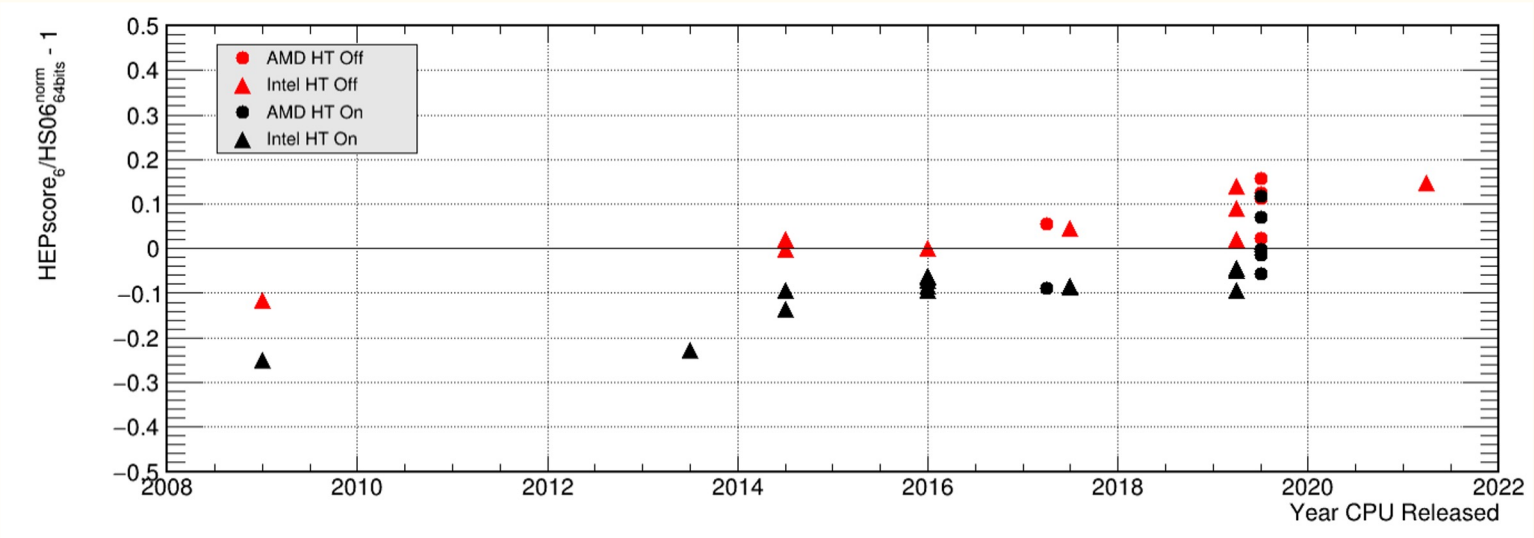
We normally look at the residuals plot but comparing the normalized HEPScore vs HEP-SPEC06 is valuable

HEPScore results

HEPScore used in plots does not include ALICE workload



HEPScore
Residuals plot vs
HS06



HEPScore
Residuals plot vs
Year CPU Released

HEPScore gives a higher score relative to HEPSpec06 with new CPUs

Work to do

x86 HEPScore
Fix some issues with **ATLAS_gen_sherpa** (done)
Validate and analyze the new **ALICE_reco** (Pb-Pb collision) workload

Commission a new reference machine at CERN: Intel IceLake 6326
It is the newest intel model at CERN, expect long lifetime in the data centre

x86 and ARM HEPScore
We have 6 workloads that can run on ARM processors (3 CMS and 3 ATLAS)
Validated the new workloads including data on two ARM servers (results are consistent with x86)
More detail on next slide

Should we release an x86 version of HEPScore?

An x86 version of HEPScore could be finalized very soon

We need to correct results from **atlas_gen_sherpa workload** and validate a new **alice_reco** workload

The timescale for x86/ARM HEPScore is dependent on getting ARM-compatible workloads

ARM Workload status:

CMS	3 Workloads
ATLAS	3 Workloads
ALICE	Workload ready in a few weeks
LHCb	Workload ready in a few months
Belle II	No estimate (discussions ongoing and tests being planned)

Plan to review the status of the workloads in March-April 2023

Recent talk at ACAT Conference showed that power consumption of an ARM processor was 45% lower than x86 and processing time was shorter for the atlas_sim workload

https://indico.cern.ch/event/1106990/contributions/4991256/attachments/2534801/4362468/PoW_ACAT2022.pdf

How to utilize HEPScore?

Emerging consensus from the Task Force (and the experiments/sites) is that site should not re-benchmark existing hardware with HEPScore

Initially sites would run both HEPScore and HEP-SPEC06 on their new hardware

Consequences:

Sites will need to calculate an **"average HEPScore"** if they have heterogenous systems
(They do that today with HEPSpec06)

Sites would be provided with a conversion factor relating HEPScore and HEPSpec06
(Dependent on the relative normalization)

Site will likely be asked to initially quote their site capacity in HEPSpec06 units, and optionally, HEPScore units
("Transition period")

Once agreed, then sites will only publish their site capacity in HEPScore units

Normalization of HEPScore

Should HEPScore be normalized to HEP-SPEC06?

Normalize HEPScore to HEP-SPEC06 on the reference machine at CERN
The normalization of HEPScore is trivial to set from a technical perspective

Preference for HEPScore == HEP-SPEC06 with some reservations and concerns from TF

Benefits:

It would simplify the estimate of the capacity of a site trivial
Accounting tables and plots would not require correction factors

Concerns:

Some worry about the risk of confusing the HEPScore and HEPSpec06 benchmarks
Easier to miss an error

Transition plan

Recall that computing pledges are made in October for the fiscal-year that starts in 18 months
(Pledges in Oct 2023 for the FY2025 (April 2025-March 2026)

Goal is release HEPScore (x86, and hopefully ARM) for production use around April 2023

Sites asked to benchmark new hardware with both HEP-SPEC06 and HEPScore

Site capacity provided to the Accounting Team and Scrutiny Group with either or both benchmarks
(If HEPScore is normalized to HEP-SPEC06, then the numbers should be identical)

The Task Force and WLCG MB will confirm (April 2024) whether HEPScore will be benchmark used by the Scrutiny Group in October 2024 for the FY2026 Pledges

Summary

HEPiX Benchmark Working Group has developed a Suite for measuring HEP workloads and a containerized method for deploying the workloads to remote sites

Extensive measurement campaign using workloads from LHC and non-LHC experiments
Significant ongoing effort investigating benchmarks for systems with GPUs

WLCG HEPScore Deployment Task Force has been engaged in the development of HEPScore

<u>Facing key questions on</u>
Inclusion of ARM processor systems
The deployment schedule for HEPScore
The implications of a new benchmark to the sites and Accounting Team

Acknowledgements

Collaborative effort
Individuals
Sites providing resources for benchmark studies
Experiments providing workloads
HEPiX CPU Benchmark Working Group
WLCG CPU Benchmark Task Force

Many publications and conference presentations:
HEPiX benchmarking solution for WLCG computing resources
Computing and Software for Big Science (2021) 5, 28

October 2022 HEPiX and ACAT