

HEPscore benchmark

D. Giordano (CERN)

on behalf of

HEPiX Benchmarking WG & WLCG HEPscore Deployment TF

HEPiX Autumn 2022 Workshop

01 Nov 2022

Two teams collaborate for this project

HEPiX Benchmarking WG

Roles

- Evaluation of benchmark alternatives
- Design and development of the **HEP Benchmarks project**
- Validation of the HEP workloads
- Analysis of benchmark measurements

Team of ~13 people

Active (again) since 2016

WLCG HEPscore deployment TF

Roles

- Recommend the HEPscore composition
- Strategy for HS06->HEPscore migration
- Coordinate the collection of new workloads
- Onboard WLCG sites for validation

Team of ~20 people

Started on Nov 4. 2020

HEP Benchmarks project

HEPscore has been proposed by the HEPiX Benchmarking WG as replacement of HS06

- Uses the workloads of the HEP experiments
- Combine them in a single benchmark score

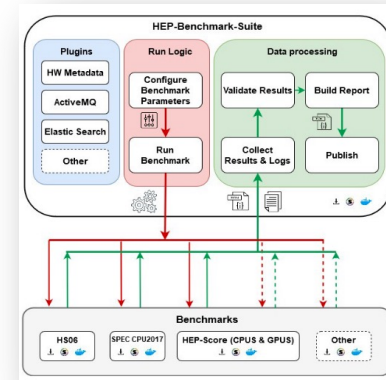
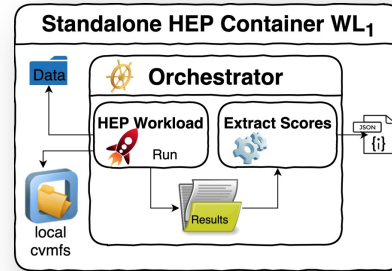
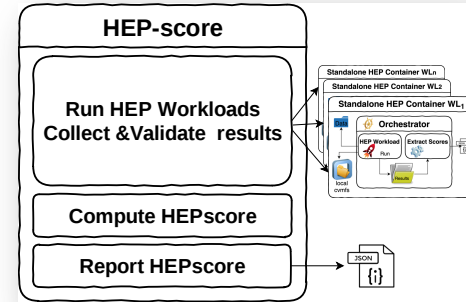
HEPscore relies on *HEP Workloads*

- Individual **reference** HEP applications

In addition, *HEP Benchmark Suite*

- Orchestrator of multiple benchmark (HEPscore, HS06, SPEC CPU2017)
- Central collection of benchmark results on OpenSearch

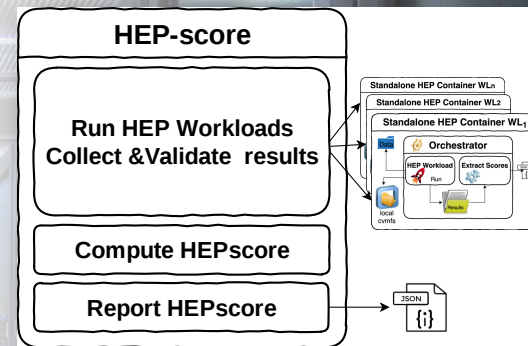
All released under GPLv3 license



HEPscore definition

Ingredients:







- a set of reference workloads (WLS)
- a measure of performance per WL (m_i): work done in unit of time
- a reference server



The score S of a server (srv) is defined as the **geometric mean** of the **speed factors** $x_i(srv,ref) = m_i(srv)/m_i(ref)$ respect to the reference server (ref)

$$\bar{x} = \left(\prod_{i=1}^n x_i^{w_i} \right)^{1/\sum_{i=1}^n w_i}$$

https://en.wikipedia.org/wiki/Weighted_geometric_mean

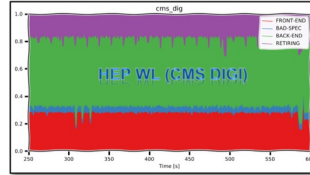
	WL ₁ 	WL ₂ 	WL _n 	Score $\left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$	S(A,B)			
Ref. Srv 	$m_1(ref)$	1 (by def)	$m_2(ref)$	1 (by def)	$m_n(ref)$	1 (by def)	1 (by def)	
Srv A 	$m_1(A)$	$x_1(A,ref)$	$m_2(A)$	$x_2(A,ref)$	$m_n(A)$	$x_n(A,ref)$	$S(A,ref)$	$\frac{S(A,ref)}{S(B,ref)}$
Srv B 	$m_1(B)$	$x_1(B,ref)$	$m_2(B)$	$x_2(B,ref)$	$m_n(B)$	$x_n(B,ref)$	$S(B,ref)$	

Brief history

WG activities - short overview (3)

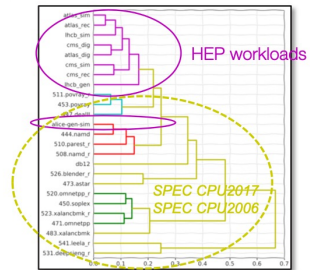
Summer 2018

- Proposal of building a set of **HEP reference workloads** (WLCG MB)
 - Enable feature studies of the experiments' workloads
 - Build a HEP benchmark suite



Fall 2018

- Collect instructions from LHC experiments to run reference WLS
- **Prototype** the build of HEP reference benchmarks in containers
- Studies on hardware **performance counters** (using Trident)
 - HEP WLS have same characteristics and differ more respect to HS06 and SPEC CPU 2017 workloads



DENDROGRAM OF WLS SIMILARITY

2019

- Start the **HEP Benchmarks project**



From a challenge to a routine

Collect, maintain, extend **workloads** from several HEP experiments

- Not affordable with ad-hoc recipes for each workload
 - More than **30 workloads from 7 experiments** prepared in the last years
- Experts from the Experiments focus on providing the workloads: software, data, result parser
- Experts on benchmarking focus on implementing a **unified** approach

Requirements

- Provide consistent CLI, report structure, error logging
- Reproducible results
- Zero burden from accessing remote data, databases, etc
- Not too large package distribution
- Portable
- Long term support

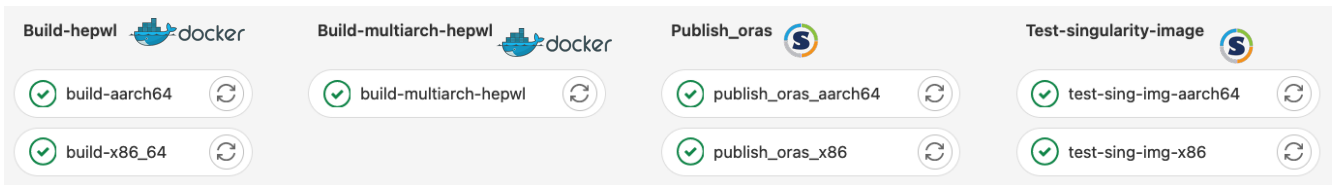
Experiment	WL repo	Description	Latest Container	Latest Built Version	Latest Pipeline status
alice	alice-digi-reco-core-run3	link	docker	v1.0	pipeline passed
alice	alice-gen-sim	link	docker	v2.1	pipeline passed
alice	alice-gen-sim-reco-run3	link	docker	ci-v0.6-aod	pipeline passed
atlas	atlas-digi-reco	link	docker	v2.1	pipeline passed
atlas	atlas-gen	link	docker	v2.1	pipeline passed
atlas	atlas-gen_sherpa	link	docker	ci-v2	pipeline passed
atlas	atlas-gen_sherpa-ma	link	docker	v2.0	pipeline passed
atlas	atlas-kv	link	docker	v2.1	pipeline passed
atlas	atlas-reco_mt	link	docker	v2.1	pipeline passed
atlas	atlas-sim	link	docker	v2.1	pipeline passed
atlas	atlas-sim_mt	link	docker	v2.1	pipeline passed
atlas	atlas-sim_mt	link	docker	v2.1	pipeline passed
belie2	belie2-digi-reco	link	docker	ci-v0.4	pipeline passed
belie2	belie2-digi-reco	link	docker	v2.1	pipeline passed
cris	cris-digi-reco	link	docker	ci-v0.7	pipeline passed
cris	cris-digi-reco	link	docker	v2.1	pipeline passed
cris	cris-digi-reco	link	docker	v2.1	pipeline passed
cris	cris-digi-reco	link	docker	v2.1	pipeline passed
cms	cms-digi-reco-run3	link	docker	v1.0	pipeline passed
cms	cms-digi-reco-run3	link	docker	v0.1	pipeline passed
cms	cms-digi-reco-run3	link	docker	v0.1	pipeline passed
cms	cms-reco	link	docker	v2.1	pipeline passed
cms	cms-reco-run3	link	docker	ci-v0.6	pipeline passed
cms	cms-reco-run3-aarch64	link	docker	v0.1	pipeline passed
cms	cms-reco-run3-ma	link	docker	v1.1	pipeline passed
dune	dune-reco-fd	link	docker	v2.1	pipeline passed
hell	hell-reco	link	docker	v2.1	pipeline passed
ligo	ligo-reco	link	docker	v2.1	pipeline passed
jun	jun-reco	link	docker	v2.1	pipeline passed
lhcb	lhcb-reco	link	docker	v2.1	pipeline passed
lhcb	lhcb-gen-sim-2021	link	docker	v2.1	pipeline passed
mg5amc	mg5amc-madgraph4gpu	link	docker	v2.1	pipeline passed
mg5amc	mg5amc-madgraph4gpu-2022	link	docker	v2.1	pipeline passed



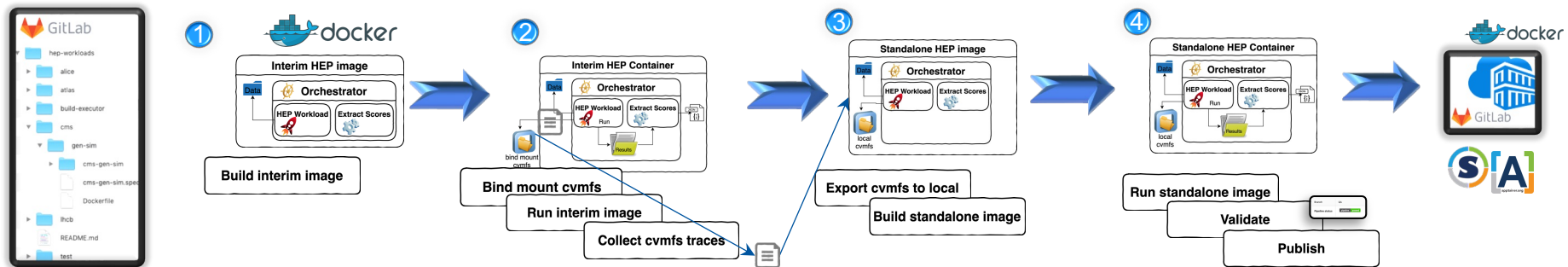
Workloads' repository

Realized an effective infrastructure to **build and distribute** the HEP workloads

Containers are built for multiple architectures: x86, **aarch64**, (potentially) Power, **GPUs**

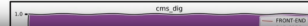


- GitLab CI/CD for fully **automated** continuous integration
- GitLab **Registry** for container distribution (Docker & Singularity/Apptainer)



Brief history

WG activities - short overview (3)



Summer 2018

- Proposal of building
 - Enable feature set
 - Build a HEP benchmark

Fall 2018

- Collect instructions from
- **Prototype** the build
- Studies on hardware
- HEP WLS have similar characteristics and differ more

2019

- Start the **HEP Benchmark**

Extensive validation process of WL

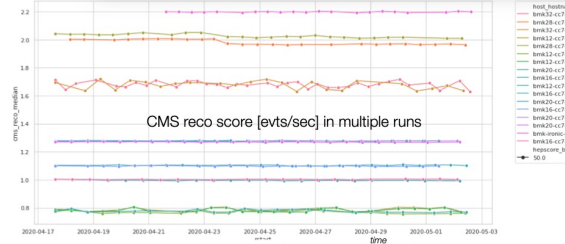
- Validating reproducibility, robustness, run duration, disk space
- Continuously running in a number of virtual & physical machines
- Evaluated a different number of events per WL to shorten the runtime

Workload	ATLAS gen	ATLAS sim	ATLAS digi-reco	CMS gen-sim	CMS digi	CMS reco	LHCb gen-sim
Robustness	✓	✓	✓	✓	✓	✓	✓
Reproducibility	0.8%	2%	0.6%	1.5%	1%	1%	1%
Memory	✓	✓	✓	✓	✓	✓	✓
Image size (unpacked)	1.65 GB	6.0 GB	6GB	5.4 GB	11 GB	8.4 GB	2.6 GB
Readiness	✓	✓	✓	✓	✓	✓	✓

✓ okay
✗ blocker

CPU Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz (32 cores, SMT ON)

WL	# threads or proces. (default)	# Evt/thread (default)	Duration of a single WL run on ref machine [h:mm]	Wdir size (per running copy)
Atlas gen	1 (SP)	200	~12	50MB
Atlas sim	4 (MP)	10	~1:32	100 MB
CMS gen-sim	4 (MT)	20	~0:15	70 MB
CMS digi	4 (MT)	50	~0:09	400 MB
CMS reco	4 (MT)	50	~0:15	100 MB
LHCb gen-sim	1 (SP)	5	~0:40	15 MB
Total			~3:30	



D. Giordano (CERN)

HEPiX Autumn 2020 Online workshop

13/10/2020

6



Brief history

WG activities - short overview (3)

Summer 2018

- Proposal of building
- Enable feature s
- Build a HEP ben

Fall 2018

- Collect instructions fr
- **Prototype** the build
- Studies on hardware
- HEP WLS have s
- and differ more r

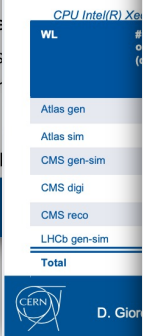
2019

- Start the **HEP Bench**

Extensive validation process of WL

- ❑ Validating r
- ❑ Continuous
- ❑ Evaluated a
- runtime

Demo: results in the central DB @ CERN



The screenshot shows a database search interface with the following details:

- Search Results:** A table with columns for 'Time', 'message host logs url', 'message host logs description', 'message software version', 'message profile happens access', 'message software version', and 'message host hostname'. A bar chart above the table shows a peak in activity around 18:00.
- Expanded Document:** A JSON document showing metadata for a 'Digi' workload, including 'name', 'description', 'version', and 'software' fields.
- Code Snippets:** Several code blocks are visible, likely representing the configuration or data for the workload, including fields like 'name', 'description', 'version', and 'software'.

D. Giordano (CERN)

Brief history

WG activities - short overview (3)

Summer 2018

- Proposal of building
- Enable feature st
- Build a HEP ben

Fall 2018

- Collect instructions fr
- **Prototype** the build
- Studies on hardware
- HEP WLS have s
- and differ more r

2019

- Start the **HEP Bench**

D. Giordano (CERN)

Extensive validation process of WL

- Validating r
- Continuous
- Evaluated a
- runtime

CPU Intel(R) Xe

WL

Atlas gen

Atlas sim

CMS gen-sim

CMS digi

CMS reco

LHCb gen-sim

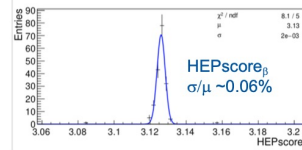
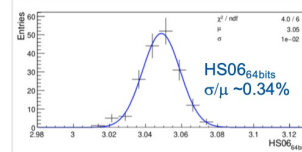
Total

D. Giordano

Demo: results in the central DB @ CERN

HEPscore_β vs HS06_{64bits} measurements (II)

- In a few occasions hundreds of servers with same HW have been benchmarked and compared
- E.g.: measurements done on 180 identical servers from a CERN delivery
 - One measurement per server
 - HS06_{64bits} also scaled to the reference server
 - HEPscore_β resolution is narrower than HS06_{64bits}



D. Giordano (CERN)

HEPiX Autumn 2021 Online workshop

26/10/2021

12



D. Giordano (CERN)

HEPiX Workshop 01/11/2022

10

Brief history

WG activities - short overview (3)

Summer 2018

- Proposal of building
 - Enable feature st
 - Build a HEP ben

Fall 2018

- Collect instructions fr
- **Prototype** the build
- Studies on hardware
- HEP WLS have s
- and differ more r

2019

- Start the **HEP Bench**

Extensive validation process of WL

- Validating r
- Continuous
- Evaluated a runtime

WL
Atlas gen
Atlas sim
CMS gen-sim
CMS digi
CMS reco
LHCb gen-sim
Total

Demo: results in the central DB @ CERN

Discover interface showing search filters and expanded document content.

HEPscore_β vs HS06_{64bits} measurements (II)

- In a few occasio
- HW have been b
- E.g.: measureme
- servers from a C
- One measureme
- HS06_{64bits} also s
- the reference ser
- HEPscore_β resul
- than HS06_{64bits}

Action 1: Extend list of HEP workloads

- The WG to build the standalone containers
 - In close contact with the experiments' experts
- 10 workloads for x86 to enter in the matrix
 - 8 ready; 4 are Single Process or Single Thread
- In addition, 2 prototype workloads for GPU (Madgraph generator and CMS HLT-like)
 - Demonstrate the HEPscore usability on other arch. (longer term goal!)

	A	B	C	D	E	F	G	H	I	J
	WL	Responsible	OS	Platform	WL developed in a git repo (if relevant)	Merged in HEP Workloads repo	Built	Validated	Reference score	Ready for the "matrix"
1	WL	Responsible	OS	Platform	WL developed in a git repo (if relevant)	Merged in HEP Workloads repo	Built	Validated	Reference score	Ready for the "matrix"
2	Atlas Gen-Sim-Reco	S. Plano	o17	o85						
3	Atlas gen-therpa	W. Lamp	o17	o85						
4	Atlas simMT	W. Lamp	o17	o85						
5	LHCb gen-sim 2023	A. Valsani	o17	o85						
6	CMS gen-sim Run3	A. Scibba	o17	o85/arm			x86/arm			
7	CMS Digi Run3	A. Scibba	o17	o85/arm			x86/arm			
8	CMS Reco Run3	A. Scibba	o17	o85/arm			x86/arm			
9	CMS HLT-like	A. Scibba	o17	o85 & GPU			x86			
10	Belu2	R. Sobiech	o17	o85						
11	Dune	A. Wu Nish	o17	o85	https://github.com					On hold for lack of time from Dune experts
12	Juno	X. Yan	o17	o85	https://github.com					
13	IQMNN	J. Wells	o17	o85	https://git.ligo.org					
14	Madgraph	A. Valsani	o17	o85 / GPU						



D. Giordano (CERN)



D. Giordano



D. Giordano (CERN)



D. Giordano (CERN)



M. Michelotto (INFN)

HEPIX Spring 2022

26/04/2022

15



D. Giordano (CERN)

HEPIX Workshop 01/11/2022

Progress in the last 6 months

- ❑ Finalized the analysis of the 2022 large measurement campaign
- ❑ Agreed on the composition of HEPscore
- ❑ Converging on a procedure to migrate from HS06 to HEPscore
- ❑ Progress on x86/aarch64 HEP workloads
- ❑ **HEPscore workshop** <https://indico.cern.ch/event/1170924/>

HEPscore workshop

2 days of workshop (19–20 Sept 2022)

- Good representation of the different bodies involved:
Experiments, Sites, WLCG Collaboration
- Registered participants: 51

Valuable feedback from our WLCG community on

- Proposed composition of HEPscore
- Usability of the HEP Benchmark Suite
- Strategy for the adoption of HEPscore as WLCG benchmark

HEPscore Workshop

19–20 Sept 2022
CERN
Europe/Zurich timezone

- Overview
- Timetable
- Contribution List
- My Conference
 - My Contributions
- Registration
- Videoconference

The goal of the workshop is to define the strategy for the adoption of HEPscore as replacement of HepSpec06.

The HEPspec06 benchmark has been a reliable estimate of CPU performance for many years, and is currently used by the WLCG for accounting and pledges. However, HEPspec06 is based on the SPEC2006 benchmark that is no longer supported. Further, it uses applications that do not reflect those used by the HEP community and will not provide benchmark for the new CPU+GPU systems.

The HEPscore Workshop will consist of several sessions:

Monday September 19

The **first session** will summarize the work of the WLCG HEPscore Task Force and the HEPiX Benchmark Working Group. During the session, potential candidates for the new HEPscore22 benchmark will be introduced.

The **second session** is devoted to the presentation of the current HEP-Workloads provided in the past year by the LHC experiments, Belle2, Juno and IWGN. Besides a description of the applications and of their performance, representatives of each experiment will highlight their expectations about the HEPscore22 composition and its lifetime.

The **last session** will introduce the framework developed to run HEP benchmarks and collect benchmark measurements ensuring traceability and monitoring. The HEP Benchmark Suite will be described and feedback about its usability will be provided by WLCG sites having used the suite.

Tuesday September 20

The **morning session** will focus on the policy and strategy foreseen to evolve from HepSpec06 to HEPscore, and the implications to accounting, pledging and procurement. The session will start with presentations from members of the Accounting Task Force, regarding the status of the development and the proposed deployment strategies. A round table with all the stakeholders will follow the presentations.

The **last session** is named "beyond x86". It will cover R&D work done by the Benchmarking Working Group in the area of heterogeneous computing, in order to extend HEPscore also to the benchmarking of servers with GPUs. In addition, there will be presentations on CPU power consumption and its relation to HEPscore.

Please register to the workshop to communicate if you will attend via zoom or in person.

Notes of the workshop are available for the registered participants as codiMD document in CERNbox at [this url](#)

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

Agenda

- 5 sessions to address interconnected topics:
 - HEPscore TF results
 - HEP Workloads
 - HEP Suite
 - HEPscore Deployment
 - Beyond x86
- 30 contributions
- Allocated time for discussions



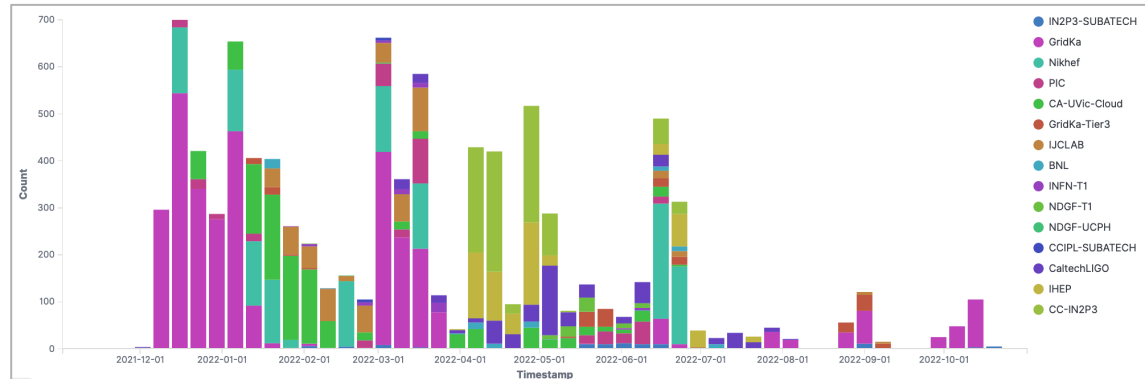
HEP Workloads 2021/22

SW from LHC experiments (Run 3) + Belle2, Juno, IGWN

- Alice (gen-sim), Atlas (gen, sim_mt, reco_mt), CMS (gen-sim, digi, reco), LHCb (sim)
- Belle2 (gen-sim-reco), Juno (gen-sim-reco), IGWN

Repeated measurements

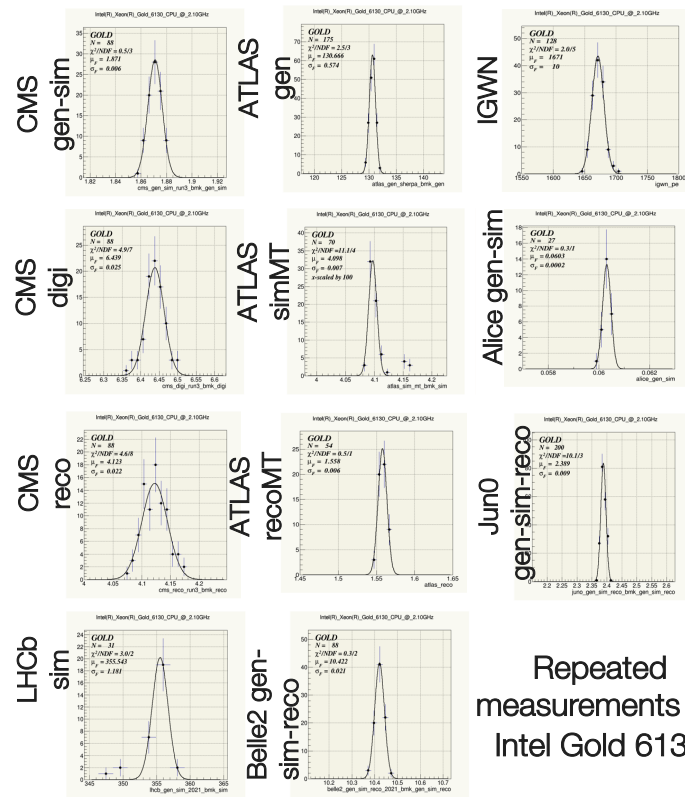
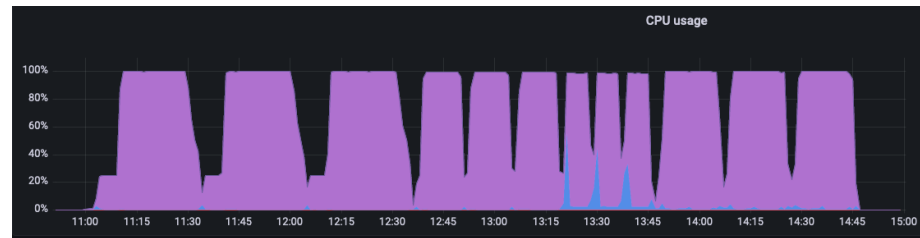
- On ~40 different CPU models from ~15 WLCG sites
- Thanks to all site admins that made this possible!



2022 measurement campaign

Measured

- Robustness against failures
- Resolution (σ/μ typically $< 1\%$)
- Performance



Repeated measurements on Intel Gold 6130

Analysis results at a glance

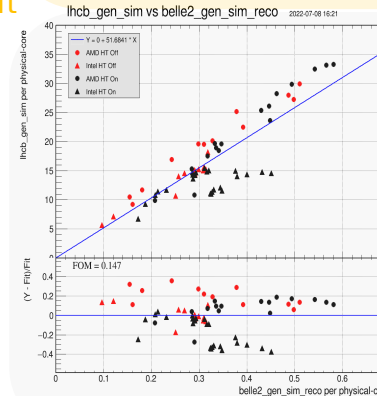
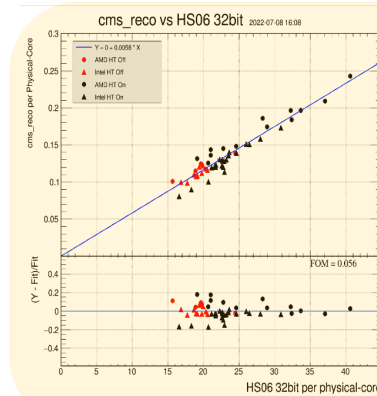
– (Re)confirmed HEPSpec06 (HS) and SPEC2017 (SP) findings, i.e. high correlation

– HEP Workloads vs HS/SP

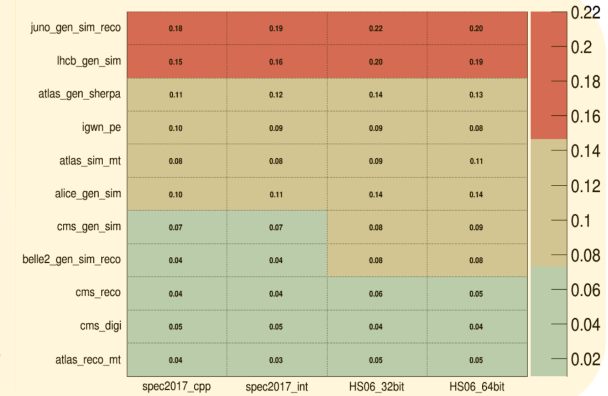
- Measure scaling and mean deviation from fit

– HEP Workload Vs Workload matrix

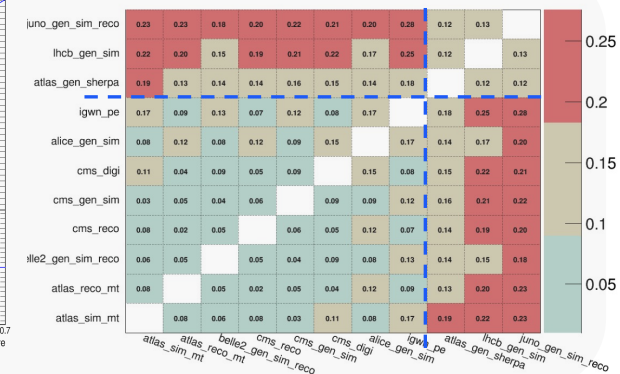
- Indication of the correlation between the workloads



Mean deviation from fit
2022-07-08 16:18



Mean deviation from fit
All CPU-arch 2022-07-08 16:43



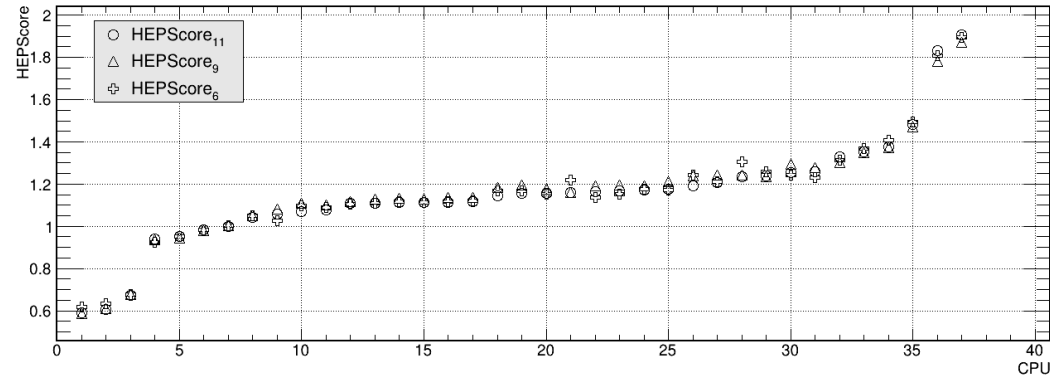
HEPscore candidates

Several combinations investigated:

- All workloads/Exclude the long-running ones/Use a subset/No weights/Weights from Grid fraction of jobs
- Little difference (**few %**) between the candidates

Preference for

- A small workload set for shorter runtime
- The simplest approach of unweighted WLS
- HEPscore composition of **7 workloads**:
 - Alice (digi_reco),
Atlas (gen_sherpa, reco),
CMS (gen_sim, reco),
LHCb (sim), Belle2 (gen_sim_reco)
 - Important to include the Alice Reco workload: reconstruction of Pb-Pb events. Workload being prepared



Desiderata: **at least some workloads should run on ARM**

Benchmark DB

OpenSearch instance @ CERN

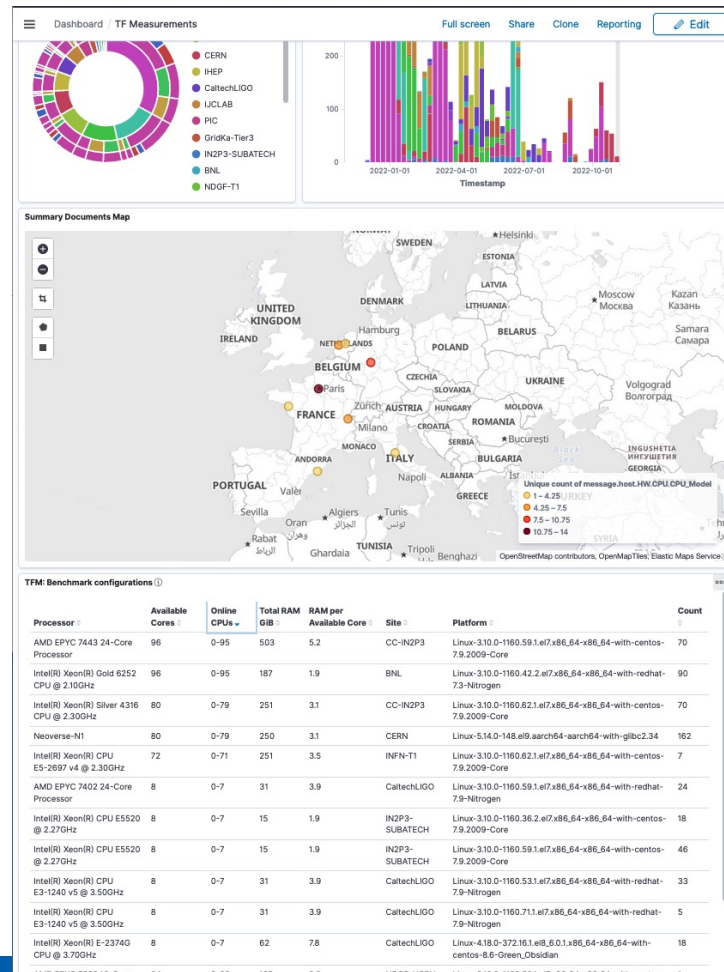
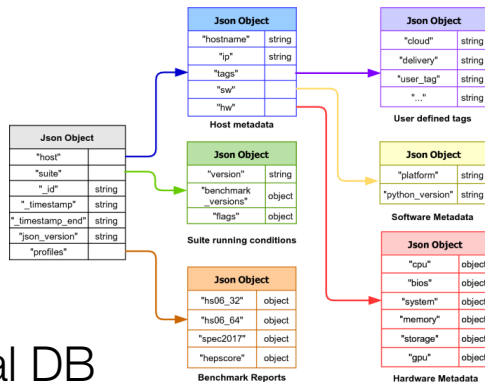
– <https://es-hep-bmk.cern.ch/>

– Access restricted: contact me for read access

Benchmark reports

– Modular JSON structure

– Suitable for non-relational DB



HEP BMK Suite in Procurement and Enrollment

- During the tendering and evaluation phase:
- Bidders usage to provide HS06 score
 - Sample evaluation

Lifecycle of HW at CERN - IroniC enrollment process:

- Benchmarking as IroniC clean step
- Where data goes: ElasticSearch, inventory

Conclusion

- Future work
- Wishlist
- Feedback

Luca Atzori

Conclusions

- We are running smoothly the benchmark components during the measurement campaign.
- We are enthusiastic! with the data population and aggregation mechanism.
- If we have in the future Jupyter notebook direct access to data will be very helpful for custom graphs and better inspection of the data.
- Sanity check of the wallclock, job efficiency and system I/O are important to detect corner cases in hardware configuration, runtime environment condition or benchmark nature (e.g. cpu bound or I/O bound).

Emmanouil Vamvakopoulos



Conclusion / Suggestions

- HEPscore has changed for the better what we decide to purchase.
 - Compared to HS06 it's easier to run, more flexible and more insightful.
 - The 32bit requirement of HS06 was skewing results.
- HEPscore will allow us to compare diverse types of hardware.
 - E.g. x86 vs ARM vs GPU
- HEPscore is still "just" a benchmark and doesn't account for operational reality.
 - E.g. RAL has purchased more memory per core than VOs specify for 5+ years.
- We don't want to over optimize for one fixed benchmark.
Suggestion:
 - Release a new HEPscore annually (in April) that sites can use to guide procurement if they so choose.
 - Pick a Golden HEPscore for official use per X years / LHC Run.

Alastair Dewhurst

Alastair Dewhurst, 19th September 2022



Science and Technology Facilities Council



Task Force Measurements

Much easier

- A little trial and error to get started
- Supplied scripts were simple to customise and run
- Easily repeatable for the different scripts

Work flow:

- run all repetitions for a script with publish false
- then upload afterwards

Ran the HS06 benchmarks – straight forward

- Create the HS06 tarball
- Copy to benchmark machine
- Run benchmark script

Andrew Pickford

Envisage the WLCG sites running benchmarks via the Suite with the data (and its metadata) written to OpenSearch at CERN

Metrics gathering and analysis

Mattieu Puel

Publishing to the message bus

- Extensive data structure reported (host configuration, results...).
- An API key might be handier to manage compared to SSL certs (generation proc...

ELK

- Published data is useful to sites: processor powers, memory population... → that's a net improvement compared to former Hepix website publication.
- Some new smart graphics recently (TF Measurements, processors distribution...) ! Thank you for that.
- Ability to save searches, vizualizations and dashboards would be a plus.



Concluding remarks

Carles Acosta Silva

- We run the benchmarks in PIC WNs: Intel Xeon E5v3 and v4 architectures and more modern AMD EPYC 7452 machines
- We consider a great improvement to run the benchmarks using containers and collecting the data in a DB to be visualized in the kibana dashboard

Power consumption measurements

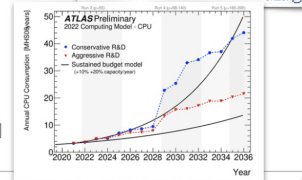
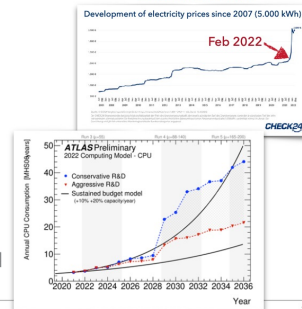
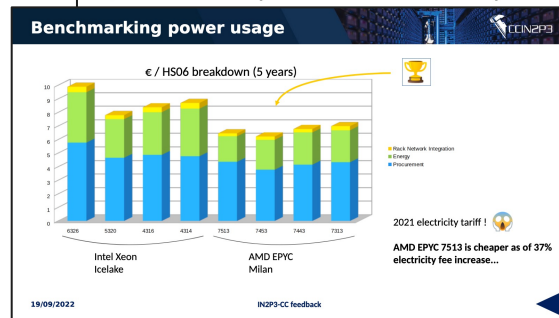
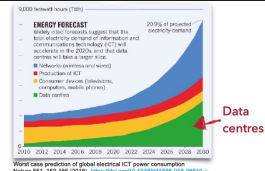
- ❑ Several similar initiatives ongoing at multiple sites
 - Same motivation: measure “events/watt”

- ❑ Opportunities for collaboration

- Which tools to use, pros/cons, caveats
- Use the HEPiX Benchmarking WG as exchange forum

WHY

- ▶ Data centres are significant consumers of electricity
- ▶ Power consumption for data centres is expected to rise further



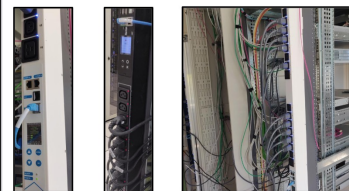
Outline (transitional talk)

- Power comparison **x86_64 & arm64**
 - Available hardware (Glasgow + Leichster)
 - Power reading & exporter tools
 - Benchmarks & results (ATLAS)
 - Limitations
- Outlook
 - New hardware at Glasgow
 - Looking forward to HEP-Score

Original slides:
<https://indico.cern.ch/event/1128343/contributions/4787174/attachments/2412950/41296>



Methods III: Smart PDUs



- Pros:**
- good level of accuracy
 - integration with monitoring infrastructure
- Cons:**
- Need the physical infrastructure!



HEP Workloads containers for x86 and aarch64

- ❑ Increasing interest to build/run experiment's software on aarch64
- ❑ Having the HEP-workloads capable to run on x86/aarch64 enables studies like the one presented at ACAT by E. Simili

HEP-Score Containers

HEP-Score containers can run on **Singularity** (or **Docker**, which we do not use):

(x86) *Singularity-CE 3.9.9-1.el7* (previous version 3.8.7-1.el7 disappeared from EPEL and replaced with *AppTainer 1.1.0-1.el7*)

(arm) *singularity version 3.8.4-1.el7* (still available in EPEL)

Example execution of a containerised HEP-Score job:

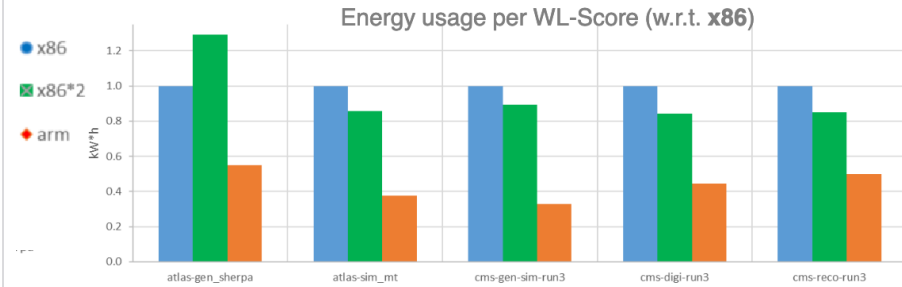
```
$ mkdir -p /tmp/test/results
$ chmod a+rw /tmp/test/
$ singularity run -B /tmp/test:/results oras://registry.cern.ch/hep-workloads/cms-gen-sim-run3-bmk:latest
```

We used 5 HEP-Score containers from the [container_registry](#) (prev. slide):

```
gitlab-registry.cern.ch/hep-benchmarks/hep-workloads-sif/atlas-sim_mt-ma-bmk:v2.0
gitlab-registry.cern.ch/hep-benchmarks/hep-workloads-sif/atlas-gen_sherpa-ma-bmk:ci-v1.0
gitlab-registry.cern.ch/hep-benchmarks/hep-workloads-sif/cms-reco-run3-ma-bmk:v1.1
gitlab-registry.cern.ch/hep-benchmarks/hep-workloads-sif/cms-digi-run3-ma-bmk:v1.0
gitlab-registry.cern.ch/hep-benchmarks/hep-workloads-sif/cms-gen-sim-run3-ma-bmk:v1.0
```

} ATLAS (2x)
} CMS (3x)

Note: the HEP workloads are designed to scale with the number of available threads, therefore power consumption cannot be directly compared among machines with a different number of cores (/threads), as the machine with more threads would have done more work ...



3 scenarios for HEPscore

1. **Conservative:** use the workloads agreed during the workshop
 - HEPscore just runs on x86. None of the composing benchmark run on aarch64
 - BUT: new workloads are already available (see case-2)
2. **Reasonable:** HEPscore just runs on x86, some composing workloads run also on aarch64
 - It's already the case, **with the updated workload versions for CMS and Atlas**
3. **Desirable:** HEPscore runs on x86 and aarch64
 - The composing workloads are all built for x86 and aarch64
 - Achievable if Alice, LHCb and Belle2 workloads are also built for aarch64

Discussion ongoing to evaluate the feasibility and timescale of case-3

Accounting transition HS06->HEPscore

Involves the Accounting Task Force

Strategy being designed around some principles/constraints

- Accounting/pledges transparency/continuity
- Lower impact on sites
- Progressive adoption
- Time constraint: draft resource requests made by Sept 2023 for year 2025

More details at the WLCG workshop (on Wed Nov 09, 9:00-10:00)

- Your feedback is welcome!

Extend HEPscore to heterogeneous resources

In the future WLCG resources will include GPUs

- This is already true for the online farms
- HEP experiments have/are re-writing their offline applications to use also **GPUs**

HEP Benchmark project:


growing support for heterogeneous workloads

- Madgraph4gpu
- CMS HLT-like
- ML/AI train AI model (e.g. MLPF)

Prototypes of analysis workloads are also available

All this is still too premature to be included in a production HEPscore

GPU workload performance



Preliminary testing on HPC enables direct comparison of same codebase and same hardware:

- Xeon Gold 6148 @ 2.4Ghz, Nvidia V100

Workload	CPU only	GPU only	Speedup	Time(CPU)	Time(GPU)
MadGraph5	0.026(float)	0.744	28x	29m 8s	11m 8s
CMS-HLT	525	9,450	18x	23m 9s	17m 15s
ML particle flow (epoch time)	659s	138s *1 GPU	4.8x	33m 36s	8m 29s

Non-production development values
Results likely to improve*

D.Southwick - HEPscore workshop 2022

Conclusions

The replacement of HS06 with HEPscore for CPUs will very likely happen in 2023

- The HEPscore TF shall submit to the WLCG MB, by the end of 2022, a recommendation about the strategy for the adoption of HEPscore

There are still few open questions, in the agenda of the WLCG workshop

- Can we adopt rapidly workloads that run on x86 and aarch64?
- Converge on the procedure for the accounting transition

Input from this workshop is greatly appreciated!

