Karlsruhe Institute of Technology

HEPiX Benchmarking Working Group Report

Matthias J. Schnepf on behalf of the HEPiX Benchmarking WG | 4. November 2024

The Benchmark Working Group

- active members
  - M. Michelotto, D. Giordano (co-chairs)
  - L. Atzori, C. Hollowell, M. Schnepf, A. Sciaba, E. Simili, R. Sobie, D. Southwick, T. Sullivan, N. Szczepanek, A. Valassi, E. Vamvakopoulos
  - Contributors needed. Feel free to join
- meetings
  - presentations on various topics
    - once per month
    - announced to to the hepix-cpu-benchmark list
  - Sprint meetings
    - 2nd and 4th week each month
    - for developers

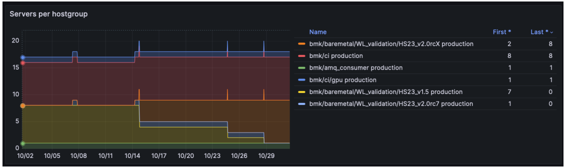
October 2024

- 30 Oct HEP-Workloads Sprint meeting
- 09 Oct HEPiX Benchmarking Working Group
- 02 Oct HEP-Workloads Sprint meeting

https://indico.cern.ch/category/1806/

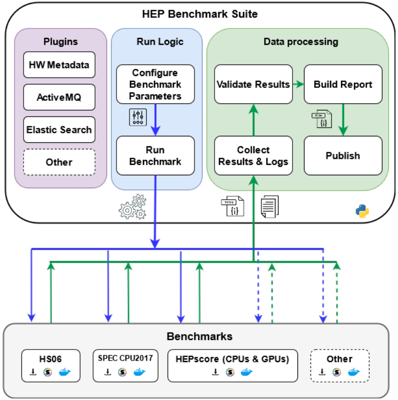
Tasks of the Working Group

- software development
  - HEP Workloads
  - HEPScore
  - HEP Benchmark suite
- operation
  - maintain the infrastructure used to
    - build workload images using VMs for gitlab runners
    - validate workloads on bare metal nodes
    - collect benchmark data into OpenSearch, HDFS DB and HEPiX benchmark table
    - analysis (more later)
- organize exchanges on CPU/GPU benchmarking topics



HEP Benchmarks project

- HEP Workloads
  - individual reference physics applications
  - also GPU and gravitational-wave applications
- HEPScore
  - uses the workloads of experiments
  - combines them in a single benchmark score
- HEP Benchmark Suite
  - orchestrator of multiple benchmark (HEPScore, HS06, SPEC CPU2017, DB12)
  - report benchmark results to HEPiX BWG



HEPScore23

- current benchmark for pledged resources (since April 2023)
- 7 workloads from 5 experiments
  - 3 single process (SP) workloads + 4 four thread/process workloads
  - container images based on Linux CC7
- support for x86 and aarch64
- 1:1 normalization with HS06 for the reference CPU model (Intel Xeon Gold 6326 CPU 2.90 GHz (HT=on)
- will provide new versions of workloads based on ALMA 9 images for comparison

VO	Workload	SW version
ALICE	DIGI Reco	O2/nightly-20221215-1
ATLAS	Gen SHERPA (SP)	Athena 23.0.3
	Reco	Athena 23.0.3
Belle II	Gen Sim Reso (SP)	release-06-00-08
CMS	Gen Sim	CMSSW_12_5_0
	Reco	CMSSW_12_5_0
LHCb	Sim (SP)	v3r412

Reported Benchmark Results

- HEP Benchmark Suite can report results to BWG
- about 25 sites contributed Thanks, and please contribute further
- 150 different configuration (CPU models, RAM, SMT conf, ...)
- x86 (AMD, Intel) and ARM (Neoverse-N1, Neoverse-V2)
- data available in a public table

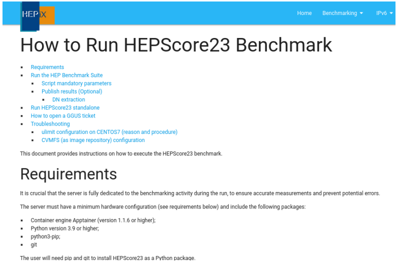
Show 10 entries

Search:

CPU	SMT enabled	Online CPUs	# Sockets	Cores/ Socket	Threads/ core	Ncores	L2 cache	L3 cache	# Meas	Score	Score/ Ncores	sem	Spread	RAM	SWAP	Site	hash	min_score	max_score	first_date	last_date
filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter	filter
AMD EPYC 9754 128-Core Processor	1	0-511	2	128	2	512	256 MiB (256 instances)	512 MiB (32 instances)	5	7450.248	14.6	21.642	1.378	1 TiB	4 GiB	UKI-SCOTGRID-GLASGOW	71892	7389.8486	7497.4875	2023-09-21 13:10:00	2023-09-27 11:57:38
AMD EPYC 9654 96-Core	1	0-383	2	96	2	384	192 MiB (192 instances)	768 MiB (24 instances)	3	7268.25	18.9	7.37	0.312	820 GiB	24 GiB	JP-KEK-CRC-02 (KEKCC-2024)	71892	7254.4787	7279.6869	2024-08-24 02:05:29	2024-08-24 10:33:44

Documentation

- https://w3.hepix.org/benchmarking/how_to_run_HS23.html
- installation description
- dependencies
- GGUS user support (32 of 34 closed)
- instructions for accounting reports



HEPiX Home Benchmarking IPv6

How to Run HEPScore23 Benchmark

- Requirements
- Run the HEP Benchmark Suite
  - Script mandatory parameters
  - Publish results (Optional)
    - DN extraction
- Run HEPScore23 standalone
- How to open a GGUS ticket
- Troubleshooting
  - ulimit configuration on CENTOS7 (reason and procedure)
  - CVMFS (as image repository) configuration

This document provides instructions on how to execute the HEPScore23 benchmark.

Requirements

It is crucial that the server is fully dedicated to the benchmarking activity during the run, to ensure accurate measurements and prevent potential errors.

The server must have a minimum hardware configuration (see requirements below) and include the following packages:

- Container engine Apptainer (version 1.1.6 or higher);
- Python version 3.9 or higher;
- python3-pip;
- git

The user will need pip and git to install HEPScore23 as a Python package.

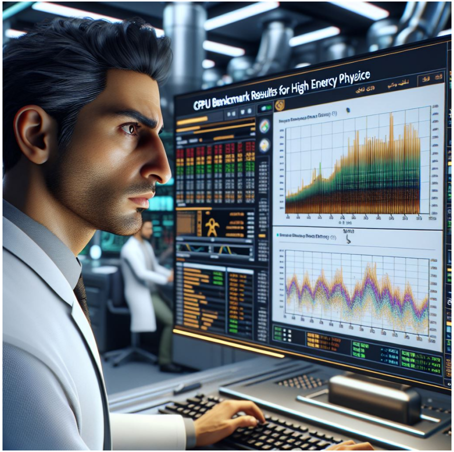
Improvements

- release candidate of new HEPScore version
  - fix discovered issues
    - ALICE digi-reco, to reduce memory footprint and improve the event configuration
    - ATLAS gen, to better accout of all the processing steps on the MC application
  - new features
    - configurable number of cores to be used by workloads
    - support multiple container registries in the same configuration
    - provision of tarball with container images
- new version of HEP Benchmark Suite
  - release candidate for version 3.0
  - support of plugins to collect system information
    - load
    - power consumption (ipmitool)
    - CPU frequency

generated with Microsoft copilot

Analysis by the Benchmarking WG

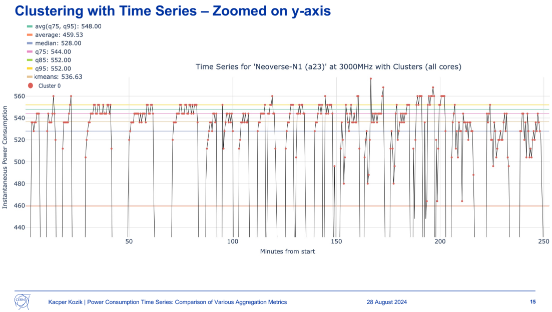
- power metric for HEPScore23
- performance measurements on the Grid
- power efficiency
- architecture specific workload performance



generated with Microsoft copilot

Analysis: Power Metric at CERN

- benchmark suite records time series of power measruement
- single number needed for comparison (power efficiency)
- study of several metrics
- 85% quantil will be used in benchmark group (simple, stable, close to normal opperation)



Analysis: Power Metric at CERN

Karlsruhe Institute of Technology

- benchmark suite records time series of power measruement
- single number needed for comparison (power efficiency)
- study of several metrics
- 85% quantil will be used in benchmark group (simple, stable, close to normal opperation)

Averaged Metric Ratio of Power Consumption for Different Machines over Available Frequencies

	average	median	kmeans	q75	q85	q95
AMD EPYC 7452 32-Core Processor (d20-10)	-11.7%	-6.4%	-6.0%	-1.7%	-0.2%	1.7%
AMD EPYC 7513 32-Core Processor (d22)	-10.0%	-6.7%	-6.1%	-3.3%	0.5%	3.3%
AMD EPYC 7513 32-Core Processor (d22-10)	-10.1%	-4.0%	-1.9%	-1.3%	-0.4%	1.3%
AMD EPYC 7643 48-Core Processor (amd-e22)	-6.0%	-4.2%	-4.3%	-1.8%	0.2%	1.8%
AMD EPYC 8534P 64-Core Processor (Siena)	-7.9%	-4.3%	-4.3%	-1.0%	0.4%	1.0%
AMD EPYC 9754 128-Core Processor (Bergamo)	-10.5%	-4.1%	-3.0%	-1.0%	0.3%	1.0%
Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz (h17)	-8.2%	-2.7%	-2.5%	-1.0%	0.4%	1.0%
Neoverse-N1 (Max128-28)	-15.0%	-5.0%	-1.9%	-1.4%	-0.0%	1.4%
Neoverse-N1 (Max128-30)	-12.2%	-4.8%	-2.8%	-1.3%	-0.2%	1.3%
Neoverse-N1 (a23)	-14.5%	-4.0%	-2.8%	-1.8%	0.3%	1.8%
Neoverse-V2 (Grace)	-15.0%	-4.8%	-2.6%	-1.6%	-0.2%	1.6%
arm-e22	-10.0%	-3.9%	-1.8%	-1.6%	-0.8%	1.6%

$4\% < |x|$

$2\% < |x| \leq 4\%$

$0\% \leq |x| \leq 2\%$

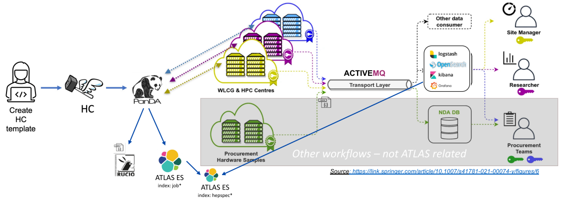
$$\frac{\sum_{freq}^{n_{freq}} \left[\frac{metric - avg(Q_{75}, Q_{95})}{avg(Q_{75}, Q_{95})} \cdot 100\%\right]_{freq}}{n_{freq}}$$

Kacper Kozik | Power Consumption Time Series: Comparison of Various Aggregation Metrics 28 August 2024 19

Analysis: Performance Measurements on the Grid at CERN

Karlsruhe Institute of Technology

- automated benchmarks on the Grid
- compare declaired and runtime corepower



Source: https://link.springer.com/article/10.1007/s41781-021-00074-y/figures/6

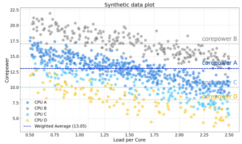
Runtime Corepower

- Runtime corepower per site:
  - For each CPU model on each site calculate the weight as:

$$w_x = \frac{\sum_{i\ on\ jobs} \text{walltime\_x\_core}_i^x}{\sum_{x\ on\ cpu} \sum_{i\ on\ jobs} \text{walltime\_x\_core}_i^x}$$

  - For each site calculate the weighted average (using available benchmarking CPU Models):

$$\text{corepower\_runtime}^{site} = \frac{\sum_{x on\ cpu} w_x \cdot \text{corepower\_runtime}_x^{site}}{\sum_{x on\ cpu} w_x}$$



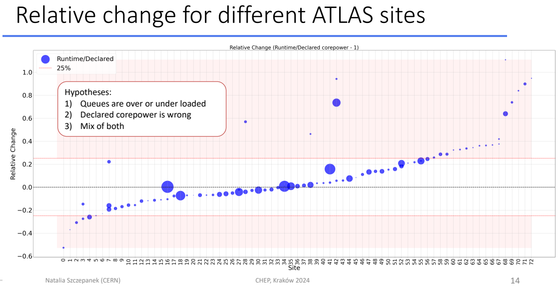
- Relative change:

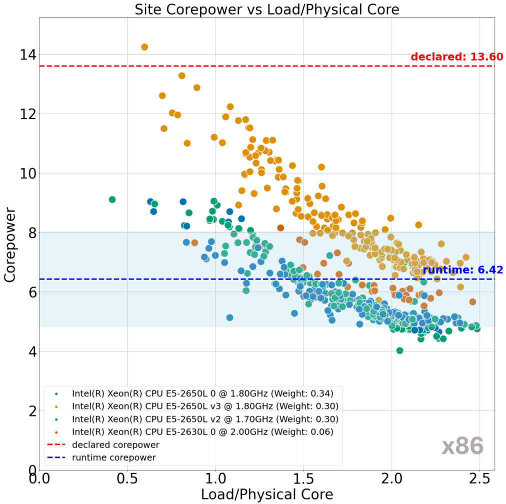
$$Relative\ change = \frac{\text{corepower\_runtime}_s}{\text{corepower\_declared}_s} - 1$$

Natalia Szczepanek (CERN) CHEP, Kraków 2024 6

Analysis: Performance Measurements on the Grid at CERN

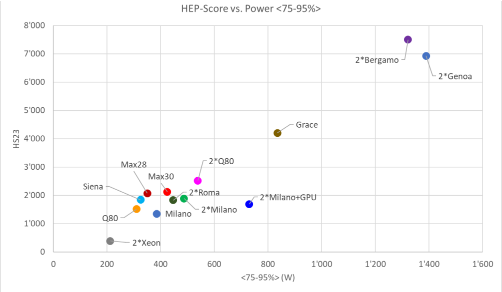
- automated benchmarks on the Grid
- compare declaired and runtime corepower
- system load dependent performance
- identify under-/over-performing sites due to wrong declaired corepoer
- CHEP presentation

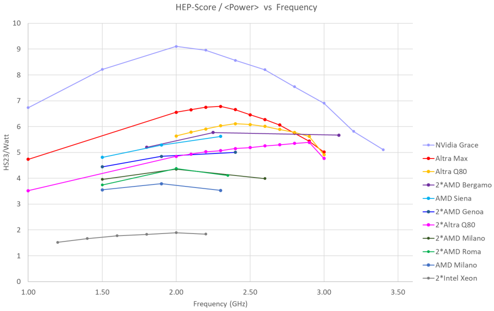




Analysis: Power Efficiency at Glasgow

- several systems benchmarked (ARM and x86)
- power efficiency at different CPU frequencies

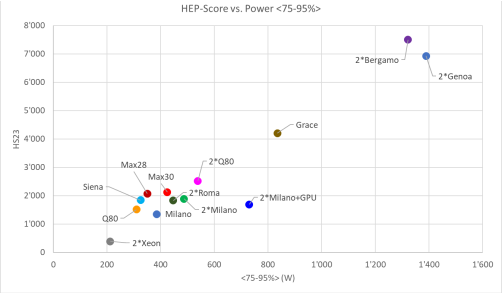




Karlsruhe Institute of Technology

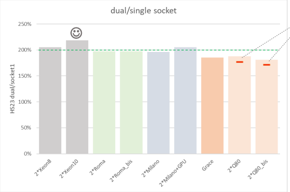
Analysis: Power Efficiency at Glasgow

- several systems benchmarked (ARM and x86)
- power efficiency at different CPU frequencies
- performance of single vs dual socket configuration
- CHEP presentation



Single vs Dual Socket

We have compared the performance of dual socket configuration vs. single socket (on the available dual socket machines: **Ampere Altra Q80**, **AMD Epyc Roma** & **Milano**, and **Intel Xeon**).



Compared to a single socket Q80

Consistent findings show that **ARM** machines in dual-socket configurations exhibit over **10% performance degradation** compared to two single-socket machines, or even compared to the same machine with only one socket enabled.

This effect is a almost absent on **x86** CPUs, where the dual socket configuration is better optimized …
… so much that for our Intel machines 1+1 > 2 ☺

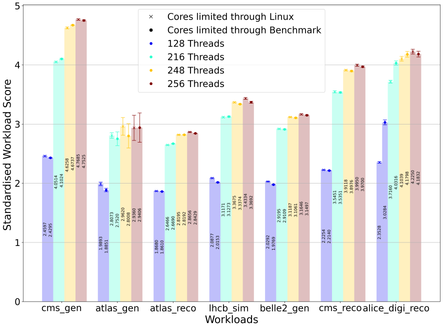
This is a known issue for both **Ampere Altra** and **Altra Max**:

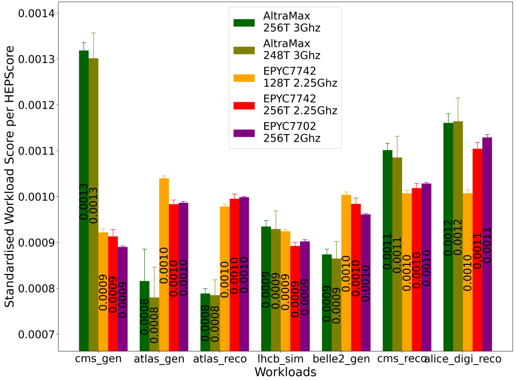
https://www.anandtech.com/show/16315/the-ampere-altra-review/3

https://www.anandtech.com/Show/Index/16979?cPage=2&all=False&sort=0&page=3&slug=the-ampere-altra-max-review-pushing-it-to-128-cores-per-socket

Analysis: Power Efficiency by KIT

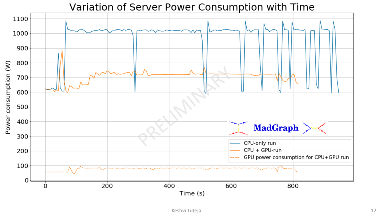
- some systems benchmarked (ARM and x86)
- power efficiency at different CPU frequencies
- x86 vs aarch64 performance and thread scan for different workloads
- CHEP poster

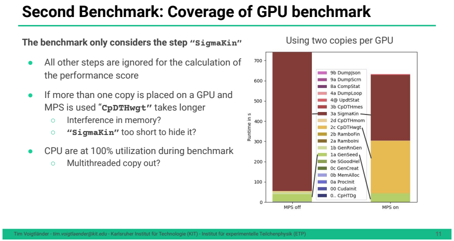




GPU Benchmarking

- more and more workloads are available for GPUs
- need to study
  - % of utilization of the GPU vs CPU
  - using of multi GPUs
  - energy consumption of the system and its components
- Workloads available (still evolving)
  - CMS HLT
  - CMS MLPF
  - MadGraph4gpu@NLO
  - further workloads can be added
- no further progress since last HEPiX
- contributions highly welcome





Summary

- HEPScore23 successfully in production for more than one year
- improvements and new features will be released soon in HEP Benchmark Suite v3.0 and HEPScore v2.0
- several ongoing analysis
  - performance on Grid
  - sustainability/power efficiency
- GPU benchmarks on the To-Do list
- looking forward to seeing more HEPScore23 benchmarks in the HEPiX database
- looking for contributers

Backup