

Status and Prospects of the ~~WLCG HEP-SCORE~~ Deployment ~~Task Force~~ CPU Benchmarking

Helge Meinhard / CERN-IT
HEPiX Online Spring 2022
26 April 2022

Why a CPU Benchmark?

- From WLCG perspective, most importantly
 - Experiment requests and site pledges
 - Accounting of CPU usage
- Many sites also use them for procurements



Why a New CPU Benchmark?

- Current benchmark, HEP-SPEC06 introduced in January 2009, has several drawbacks
 - No longer modelling performance scaling of HEP applications sufficiently well (order of 10%) in all cases
 - Running conditions good in 2009, now totally obsolete
 - SPEC stopped supporting underlying SPEC-CPU 2006 in 2018



How to Benchmark?

- Benchmarking for HEP is **not** about what can be obtained from a given CPU model optimising all possible running conditions
- It is about how a complete machine configuration performs under realistic batch service and experiment conditions at HEP sites
 - Depends not only on CPU model, but also on memory size and configuration, disk configuration, ...
 - Depends on experiment environments, compilers, flags, optimisations, ...



Future HEP Benchmarking (1)

- Starting in 2018, benchmark experts got together and worked on a new benchmark: HEPiX Benchmarking Working Group co-chaired by Manfred Aef (KIT), Domenico Giordano (CERN) and Michele Michelotto (INFN Padua)
 - Several reports to GDB, HEPiX, WLCG MB
 - Domenico's report to HEPiX 26-Oct-2021:
<https://indico.cern.ch/event/1078853/contributions/4576275/>
 - Large parts of Michele's report to HEPiX 26-Apr-2022 (in a few minutes!)
 - Without fixing the details of how to use it for a given purpose



Future HEP Benchmarking (2)

- Result of the HEPiX benchmarking working group:
 - “HEP Benchmark Suite”: Framework for running containerised benchmarking workloads
 - Automatises benchmark runs
 - Ensures structured delivery and storage of results
 - “HEP Workloads”: Collection of (mostly) HEP workloads for which it is desirable to obtain performance information
 - Rather dynamic – add improved workloads, new compilers/flags/OS, ...
 - Sometimes called “the matrix” or “the basket”
 - “HEP Score”: Single number based on a defined, stable combination of defined, stable reference workloads; sample implementation (“HEP-SCORE_beta”) using various LHC experiment workloads (from Run 2)
 - Good consistency with HEP-SPEC06



HEP-SCORE Deployment Task Force

- WLCG Management board discussed and decided to launch a task force
- Started in November 2020, bi-weekly meetings since then
- Membership:
 - Experts on pledge etc. process, procurements, accounting
 - Experiment experts
 - Four LHC experiments
 - Belle 2, DUNE, LIGO/Advanced VIRGO(/KAGRA), JUNO/BES III etc.
 - Site experts
 - Some MB members



Topics to Cover

- Compute facilities at WLCG sites
 - Still very much x86 dominated (still mostly Intel, increasingly AMD)
- Compute facilities used (quasi-)opportunistically
 - Various processors (x86, POWER, ARM), e.g. at HPC sites
 - Various GPUs in various relations with CPUs, e.g. at HPC sites
 - May see some of this soon at WLCG sites, too – even as part of the pledges
 - More may be coming, e.g. FPGAs



Strategy

- Start with CPU benchmarking on x86-based systems; look at other CPUs and/or GPUs later
 - Aim: single benchmark with a stable definition for at least (a typical CPU server lifetime cycle | a complete LHC machine cycle)
- Framework by benchmarking WG is very attractive
 - Use it to record behaviour of (wide) selection of workloads (not limited to HEP) across machines
- Study behaviour of large range of workloads over large range of benchmarking platforms
 - Aim: Select a sample of workloads (possibly with weights)
 - Sufficiently representative of real workloads on installations we (WLCG or HEP) use
 - Sufficiently small to be practical as a benchmark
 - Sufficiently precise and reproducible
- Propose a transition scenario from HEP-SPEC 06 to HEP-SCORE 2x for WLCG



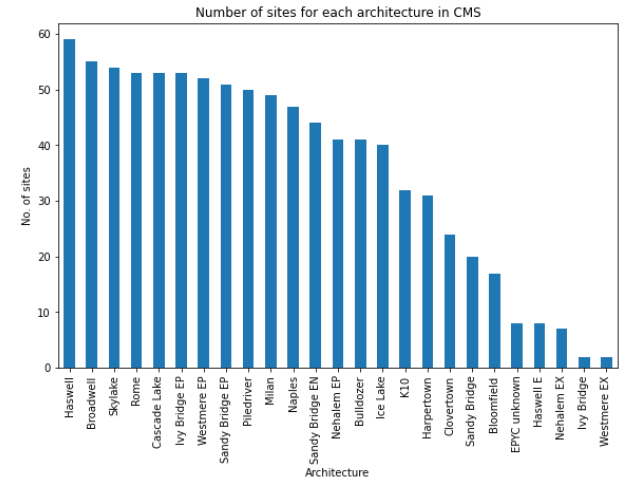
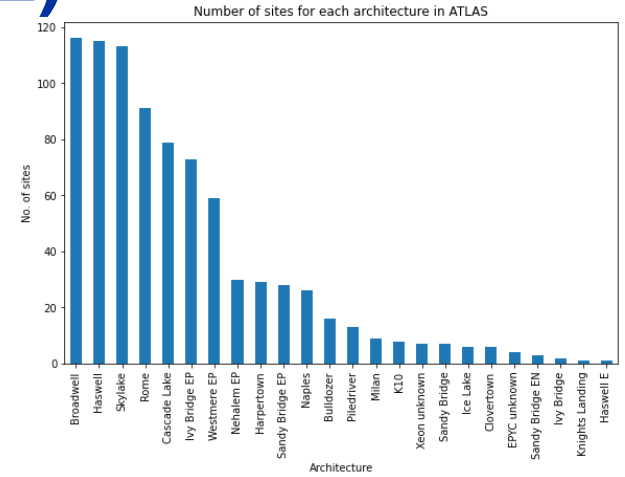
Progress since Autumn 2021

- Workloads:
 - LHC experiments: event generation, detector simulation and digitisation, reconstruction; (analysis)
 - Other (mostly HEP) experiments: Belle II, JUNO, GW experiments, (DUNE)
 - Non-HEP workloads: HS06, SPEC-CPU 2017, DB12
 - All (but one late-coming) workloads packaged and tested, almost all verified
 - Packages available of verified workloads for benchmarking runs
- Infrastructure: Upgraded wherever needed, fully tested, functional, used for collecting numerous results
- Benchmarking platforms: plethora from BNL, CERN, FNAL, IHEP, IJCLab, KIT, LIGO, NDGF-T1, Nikhef, PIC, RAL, ...
 - Long list of server configurations
 - Lots of results on individual workloads available
- Completed cross-check of available benchmark server configurations with server configurations used by ATLAS and CMS workloads
- Analysis making good progress (cf. next talk)



Most common CPU architectures (2/2)

- **Five architectures are responsible for about 85% of the jobs**
 - Rome, Broadwell, Haswell, Cascade Lake, Skylake
 - Proportions remarkably similar between ATLAS and CMS
 - All already tested by the HEPiX benchmarking WG
- **Many architectures (most for CMS) are present at a large number of sites**



Courtesy: Andrea Sciabà / CERN

(My Personal) Perspectives

- Realistic scenario:
 - End June: Matrix of workloads and server configurations completely filled
 - Soon after: Analysis of individual results completed
 - End September: Workloads (and, if necessary, weights) for HEP-SCORE 2x selected
 - Present it at HEPiX in autumn
 - End November: Migration scenario proposed to and discussed with WLCG MB
 - January 2023: Start accounting both using HS06 and HEP-SCORE 2x
 - Other migration details to be sorted out until end November



(My Personal) Conclusions

- Defining a HEP-SPEC 06 successor is a complex, ambitious task
 - Has taken and will take many months rather than weeks
 - ... and then we will have to tackle non-x86 CPUs and GPUs
- Discussions and work in a constructive and collaborative spirit
- Steady progress leading to ... HEP-SCORE 22 (HS22)

- Thanks to all working group and task force members!



Questions? Comments?



(after Michele's presentation)

