Next Generation of HEP CPU Benchmarks

Domenico Giordano (CERN)
on behalf of the HEPiX CPU Benchmarking WG

CHEP 2018 Conference, Sofia, Bulgaria
Performance measurement

From “Art of Computer Systems Performance Analysis Techniques For Experimental Design Measurements Simulation And Modeling”

- by Raj Jain, Wiley Computer Publishing, John Wiley & Sons, Inc
- 1992 Computer Press Award Winner

“Performance is a key criterion in the design, procurement, and use of computer systems […] to get the highest performance for a given cost.”

“The types of applications of computers are so numerous that it is not possible to have a standard measure of performance […] for all cases.”

“The first step in performance evaluation is to select the right measures of performance, the right measurement environments, and the right techniques.”
Select the right measures

Typical HEP application consists of
- A cluster of several hundred algorithms
- Complex framework interconnecting these algorithms
- Linear instruction spread (no hotspots)
- Non existence of classical numerical loops
- Average runtime of several hours

From 50 to 80% of WLCG CPU time spent in simulation
- CPU requirements will change and increase at HL-LHC
  - more data to process (higher luminosity)
  - more complex events (higher pileup)

Requirement:
the HEP benchmark must scale with
the average performance of the job mix running in WLCG
A reminder about HS06

- Subset of SPEC CPU® 2006 benchmark
  - SPEC's industry-standardized, CPU-intensive benchmark suite, stressing a system’s processor, memory subsystem and compiler.

- HS06 is suite of 7 C++ benchmarks
  - In 2009, proven high correlation with experiment workloads
    "<<CPP showed a good match with average lxbatch e.g. for FP+SIMD, Loads and Stores and Mispredicted Branches>> [1]
  - Execution time of the full HS06 suite: O(4h)

- Since 2009 HS06 has been used for
  - Performance studies
  - Procurement procedures
  - Pledges & Accounting
HS06 & recent CPU models

- Reported by Alice and LHCb that their workloads did not scale anymore with HS06
  - *J. Phys. : Conf. Ser.* 898 (2017) 082011

- Independent studies still show agreement within 10% for Atlas and CMS workloads
  - CPU benchmarking with production jobs

- Some “understood” differences:
  - HS06 compiled at 32 bits, whereas experiment applications at 64 bits (10%-20% effect) [ref]
New emerging scenarios

Replace HS06 by an equally accurate benchmark suite
- SPEC CPU® 2006 is being retired by SPEC, replaced by CPU® 2017
- Prepare for the future changes
  • Evolution of experiment software and better usage of the full CPU potential

Use fast benchmarks in some specific cases
- Commercial Cloud and HPC opportunistic resources
  • Assessment of the delivered performance & forecasting the job slot duration
  • Contexts where changing conditions require prompt feedback but not necessary high accuracy
SPEC CPU2017

The SPEC CPU2017 benchmark suite features updated and improved workloads, use of OpenMP to accommodate more cores and threads, and an optional metric for measuring power consumption.

Gainesville, Va., June 20, 2017 -- The Standard Performance Evaluation Corp. (SPEC) today released the SPEC CPU2017 benchmark suite, an all-new version of the nonprofit group's software for evaluating compute-intensive performance across a wide range of hardware systems.

The SPEC CPU2017 benchmark suite is the first major update of the worldwide standard CPU performance evaluation software in more than 10 years. The new suite includes updated and improved workloads with increased size and complexity, the use of OpenMP to allow performance measurement for parallelized systems with multiple cores and threads, and an optional metric for measuring power consumption.

Current SPEC CPU subcommittee members include AMD, ARM, Dell, Fujitsu, HPE, IBM, Inspur, Intel, Nvidia, and Oracle.

Larger suite, more complex code, shaped for multi-core and multi-threads

same application area as in HS06
Defining the working point

Different configurations to be investigated

- Running mode: as in HS06 or as multiple rate runs
- List of benchmarks
  - All (23)
  - Only C++ benchmarks (8)
  - Selecting the sub-set best matching the HEP workloads
- Compiler flags
  - So far: -g -O3 -fPIC -pthread

Running SPEC CPU2017

- Benchmark metrics:
  - CPU2017 suite comes with 2 metrics (like CPU2006):
    - SPEED (single benchmark run, OpenMP supported)
    - RATE (multiple benchmark copies running in parallel)
  - HS06: running multiple SPEED benchmarks in parallel (better simulation of batch system than pure RATE)
Are SPEC CPU®2017 and HS06 correlated?

Very high correlation (0.975) between SPEC CPU®2017 score and HS06 (64 bits)

- Measured on 7 different Intel CPU models, 1 AMD Opteron and 1 (Desktop) AMD Ryzen
- SMT on and off
- NB in the plot
  - error bars are [5%,95%] values
  - Marker size \( \Rightarrow \) amount of data collected

Residuals ratio of the linear fit
- \(|\text{extr./meas.}-1| < 5\%\)

Scaling factor respect to the number of running slots/core is less representative of the HEP workloads

- Caveat: study (M. Alef KIT) done on few CPU models so far
Are the individual benchmarks independent?

Are all SPEC CPU® 2017 benchmarks needed?
- Less benchmarks => Shorter runtime
  - Currently a run of the 8 C++ benchmarks takes <2.5 hours/iteration> in the 7 tested CPU models
- Better control of benchmark score Vs HEP job mix

Found subsets of the rate C++ still well representative of the full performance score
- Example:
  - 508.namd_r, 520.omnetpp_r, 526.blender_r
  - Discrepancy
    - max= 0.06
    - mean= -0.003 ± 0.004

Limitation of the study: focusing on x86 arch, mainly Intel CPUs
**Toward a common tool to run SPEC CPU2017**

**Script** to trigger SC17
- Very similar to the HS06 script (runspec.sh)
- Produces json output with results of each running benchmark
  - Includes configuration information
- Enables the sharing/comparison of the measurements

---

The profile report

```
"message": {
  "host": {
    "machine": "machine",
    "model": "model",
    "class": "class",
    "features": "features",
    "version": "version",
    "hwp": "hwp",
    "os": "os",
    "compiler": "compiler"
  },
  "tstamp_end": "2018-06-11T03:03:16Z",
  "profiles": [
    {
      "benchmark": "benchmark",
      "category": "category",
      "name": "name",
      "version": "version",
      "options": "options"
    }
  ],
  "tstamp_start": "2018-06-10T16:30:03Z"
}
```

**Host metadata**

**SC17 report, including config options**

**SC17 individual benchmark scores**

At CERN this json doc is transferred to ElasticSearch via a message broker
Compare applications by IPC and memory usage

Unveil the (dis)-similarities between HEP workloads and proposed benchmarks (HS06, SPEC CPU2017,...)
- Percentage of time spent in Front-End Vs Back-End Vs Bad speculation (Instructions per Cycle)
- Memory transactions and bandwidth usage

More details in [CHEP18 Poster 72: Trident]
Fast benchmarks investigations

Among several applications proposed and studied, two are the preferred by the WLCG collaborations

- ATLAS KV (KitValidation)
  - Mainly GEANT4. Default workload: 100 single muon event simulation
  - Well in agreement with Atlas and CMS simulation [ref]

- DIRAC Benchmark 2012 (DB12)
  - In agreement with Alice and LHCb job performance when DB12 runs at job time (single slot)

Not robust enough to replace long-running benchmarks

- DB12 dominated by the front-end call and branch prediction unit
- SMT is not beneficial at all for DB12 when loading all CPU threads
- KV shortness and event simplicity affected by systematics
  - Found in the performance studies of Meltdown/Spectre patches
  - Can be improved extending the test duration and event complexity
Conclusions

HS06 is a decade old suite used for benchmarking
- Well know by vendors, site managers, funding agencies
- Stable, reproducible, accurate
- But is reaching end of life

Looking for future alternatives in HEP
- SPEC CPU 2017
  - Preliminary studies do not show much advantage respect to the HS06
  - Suite of C++ benchmarks score highly correlated with HS06 score
- Fast benchmarks can play a role in cloud contexts, where re-benchmarking is needed
  - But the current fast benchmark cannot replace HS06 in procurement and accounting tasks

A suite of HEP workloads could be an alternative to industrial standard benchmarks
- Provided that distribution, maintenance and license issues are properly sorted out
- Work in progress