

CPU Benchmarking: Present and Future

Helge Meinhard, CERN-IT
HEPiX Benchmark WG chair
GDB, 09-Nov-2011



- In 2006/7, on IHEPCCC's request, HEPiX created a working group on benchmarking
 - Procurements: performance rather than boxes
 - Experiment requests and site pledges
- In 2007/8, detailed analysis of standard benchmarks and experiment applications
- WG result: HEP-SPEC06
 - 3 applications of SPECcpu 2006 integer, 4 applications of SPECcpu 2006 FP
 - Defined conditions: OS (SL 64 bit), compiler (gcc), compilation flags (32-bit app, conservative)
- WLCG adopted benchmark for its purposes

- HEP-SPEC06 (HS06) well established
 - Used for procurements at most major sites
 - Standard measure for requests, capacities and pledges
 - Worries about potential difficulties (price of SPEC licence, mechanics of running) proved unjustified
- Adopted far beyond HEP
 - Other communities in need for a standardised CPU benchmark jumped on the bandwaggon
 - ... even though their application's behaviour may be different
- HEPiX October 2011: Vendor presentations (AMD, Dell) mostly based on HS06 results and comparisons
 - ... even though there are still obvious misconceptions



- HS06 obviously a success story
 - Recognised as such by HEPiX board in October 2011
- Why worry?
 - (cf. presentations by Michele Michelotto and Manfred Alef at HEPiX October 2011)
 - Defined conditions no longer adequate
 - Experiments moved to 64-bit applications
 - OSes and compilers evolved
 - Need to recheck compilation flags
 - “Whole-node” scheduling suggests SPEC rate rather than concurrent multiple speed runs
 - SPEC released SPECcpu 2006 v1.2

- 32-bit vs. 64-bit applications:

CPU	HEP-SPEC06			Site
	32-bit apps	64-bit apps	difference	
intel-5160	38.47	41.47	+ 7.8 %	GridKa
intel-e5345	58.98	60.57	+ 2.7 %	GridKa
intel-l5420	70.66	73.70	+ 4.3 %	GridKa
intel-l5430	72.99	74.50	+ 2.1 %	GridKa
intel-e5520 (HT on)	121.38	142.11	+ 17.1 %	GridKa
intel-l5520 (HT off)	93.27	108.22	+ 16.0 %	GridKa
intel-l5520 (HT on)	115.69	134.05	+ 15.9 %	GridKa
amd-6168	183.08	210.20	+ 14.8 %	GridKa
amd-4x6174	379.50	430.67	+ 13.5 %	GridKa
amd-2x6128	128.48	151.28	+ 17.7 %	INFN PD
amd-2x6136	148.87	173.33	+ 16.4 %	INFN PD
amd-2x6174	197.42	225.48	+ 14.2 %	INFN PD
intel-2x5650 HT on	210.99	238.27	+ 12.9 %	INFN PD
intel-2x5650 HT off	168.23	195.23	+ 16.0 %	INFN PD

(<https://hep.caspr.it/processors/dokuwiki/doku.php?id=benchmarks:64bit>)

- 32-bit vs. 64-bit applications
 - Difference initially small, now reaching 15...20%
 - Not all applications are 64 bit already
 - Variation of new systems much smaller than in comparison with 3 years old ones
 - Age profile of sites' machine parks probably similar
 - Even if not reflecting reality at 100%, still good enough for requests, pledges, accounting, procurements
- OS and compiler, compilation flags
 - A few percent influence of OS and compiler, similar arguments as above hold true



- “Whole-node” (multi-core) scheduling: still at a marginal level
 - Multiple speed runs still reflect current reality
- SPECcpu 2006 v1.2
 - Improvements are in scripts and platform support
 - Results 1:1 compatible
 - No need to adapt HS06
- Conclusion: HS06 no longer very accurate, but probably “good enough”
 - Better carry on with it as defined rather than introducing a new benchmark now or (even worse) re-define existing one



- SPEC working on SPECcpu v6
 - Expected to be released next year as SPECcpu 2012
- In about a year's time, will need to re-assess benchmarking for HEP anyway
- No need for urgent action now
- Sites and experiments need to be prepared to make a similar effort as back in 2008

