A comparison of HEPSPEC benchmark performance on ATLAS Grid-Sites versus ideal conditions

Michael Boehler[1*], David Cameron[2], David South[3]

[1]Albert Ludwigs Universitaet Freiburg (DE), [2]University of Oslo (NO),[3] Deutsches Elektronen-Synchrotron (DE)

*michael.boehler@cern.ch

Introduction

- HEP-SPEC06 is the HEP-wide benchmark for CPU performance since 2009.
- The metric hs06 (unit HEP-SPEC06 seconds per event) is calculated as follows:

$$\text{hs06} = \frac{t_{walltime} \times n_{cores} \times corepower}{n_{events}} \quad (1)$$

- The corepower has to be extracted via the HEP-SPEC06 benchmark suite.

Objective

- The ATLAS Software Performance Group (SPOT) makes frequent measurements of the hs06 value of $t\bar{t}$ event simulation on dedicated hardware.
- Comparable simulation processes take much longer on the world-wide ATLAS computing grid (WLCG).
  - SPOT measures (2020): **3 kHS06 sec/ event**
  - Average of ATLAS jobs on WLCG (2020): **4.7 kHS06 sec / event**
- Where does this **discrepancy** come from**?**

Methods

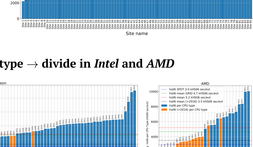
- Run **hundreds of identical** mini simulation **jobs** with HammerCloud (HC) on all computing sites
  → get reproducible data, directly comparable with SPOT results.

	Cut	Jobs	Sites	diff. CPU types
0	total	102066	96	154
1	exclude TEST queues	98196	89	154
2	(nJobs per CPU & site) >= 25	96805	86	125
3	(total nJobs per site) >= 50	96757	85	125

- In total 96 sites have been probed operating 154 different CPU types. **Selection:** Test queues are rejected. Only CPU types with more than 25 jobs per site considered. **Final selection:** Sites with at least 50 jobs

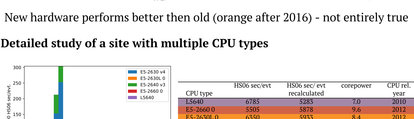
Results

The Global Picture



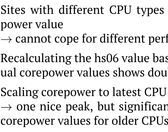
- Each job reports hs06 value.
- Each bar is the average hs06 value per site.
- Dashed lines show SPOT result, mean Grid Prod 2020, and HC mean.

Selecting the results by CPU type → divide in *Intel* and *AMD*



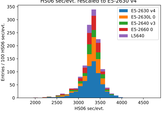
- New hardware performs better then old (orange after 2016) - not entirely true

Detailed study of a site with multiple CPU types

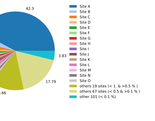


CPU type	HS06 sec/evt	HS06 sec/ evt recalculated	corepower	CPU rel. year
L5640	6785	5283	7.0	2010
E5-2660 0	5505	5878	9.6	2012
E5-2630L 0	6350	5933	8.4	2012
E5-2640 v3	3177	4028	11.4	2014
E5-2630 v4	3295	3665	10.0	2016
site overall	4470	4595	9.0	

- Sites with different CPU types but one corepower value
  → cannot cope for different performances
- Recalculating the hs06 value based on individual corepower values shows double peak
- Scaling corepower to latest CPU type
  → one nice peak, but significant reduction of corepower values for older CPUs

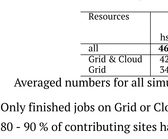


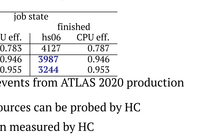
CPU type	corepower	corepower re-scaled	decrease [%]	CPU rel. year
E5-2630 v4	10.0	10.00	0	2016
E5-2640 v3	11.4	11.8	-3	2014
E5-2630L 0	8.4	4.3	48	2012
E5-2660 0	9.6	5.0	48	2012
L5640	7.0	3.4	51	2010



- hs06 does not scale linearly over different CPU generations

Closure - SPOT ⇔ HC ⇔ Grid Production





- HC jobs on identical hardware used by SPOT team show nice agreement.
- Contribution of sites to ATLAS event simulation production in 2020.

Resources	job state			
	all		finished	
	hs06	CPU eff.	hs06	CPU eff.
all	**4664**	0.785	4127	0.787
Grid & Cloud	4215	0.946	3987	0.946
Grid	3462	0.955	3244	0.953

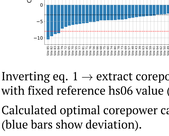
Averaged numbers for all simulated events from ATLAS 2020 production

- Only finished jobs on Grid or Cloud resources can be probed by HC
- 80 - 90 % of contributing sites have been measured by HC
- Weight each site according to its contribution to 2020 production (pie plot)
  → values from HC tests comparable to results from 2020 production

Resources	ATLAS grid prod. 2020	HC benchmarks tot. frac	hs06	hs06 w	rel dev [%]
all resources	4127	80.3	3585	3906	5.4
Grid & Cloud	3987	90.5	3618	3921	1.6
Grid	3244	85.3	3595	3265	-0.6

- Closure for Grid (& Cloud) resources agree **within 2%**

Detecting site performance deviations



- Inverting eq. 1 → extract corepower value from the benchmark jobs with fixed reference hs06 value (here: 3kHS06)
- Calculated optimal corepower can be compared with published value (blue bars show deviation).
  - positive values: site underestimates performance
  - negative values: site overestimates performance

Conclusion

- HammerCloud is an ideal tool to make world wide benchmarks
- Good closure: HC benchmarks reproduce both SPOT & Grid production results
- hs06 does not scale linearly over different CPU generations
- Discrepancy mainly induced by failed jobs and non Grid and Cloud resources
- HC benchmark results can be used to verify published corepower values

References

[1] Henning, J. SPEC CPU2006 benchmark descriptions. ACM SIGARCH Computer Architecture News. 34, 1-17 (2006,9)
[2] HEPiX, HEPiX Benchmarking WG web site, [Online; accessed 24-June-2021], 2017, url: https://w3.hepix.org/benchmarking.html
[3] J. Elmsheuser et al., Grid site testing for ATLAS with HammerCloud, Journal Of Physics: Conference Series. 513, 032030 (2014,6)

21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research
24.–28. Oct 2022 Villa Romanazzi Carducci, Bari, Italy

Bundesministerium für Bildung und Forschung