Memory Hierarchy Performance Characterization

Grupo de Investigación en Arquitectura de Computadores (gaZ) Universidad Zaragoza



Agustín Navarro-Torres, Jesús Alastruey-Benedé, Pablo Ibáñez-Marín, Víctor Viñals-Yúfera {agusnt, jalastru, imarin, victor}@unizar.es



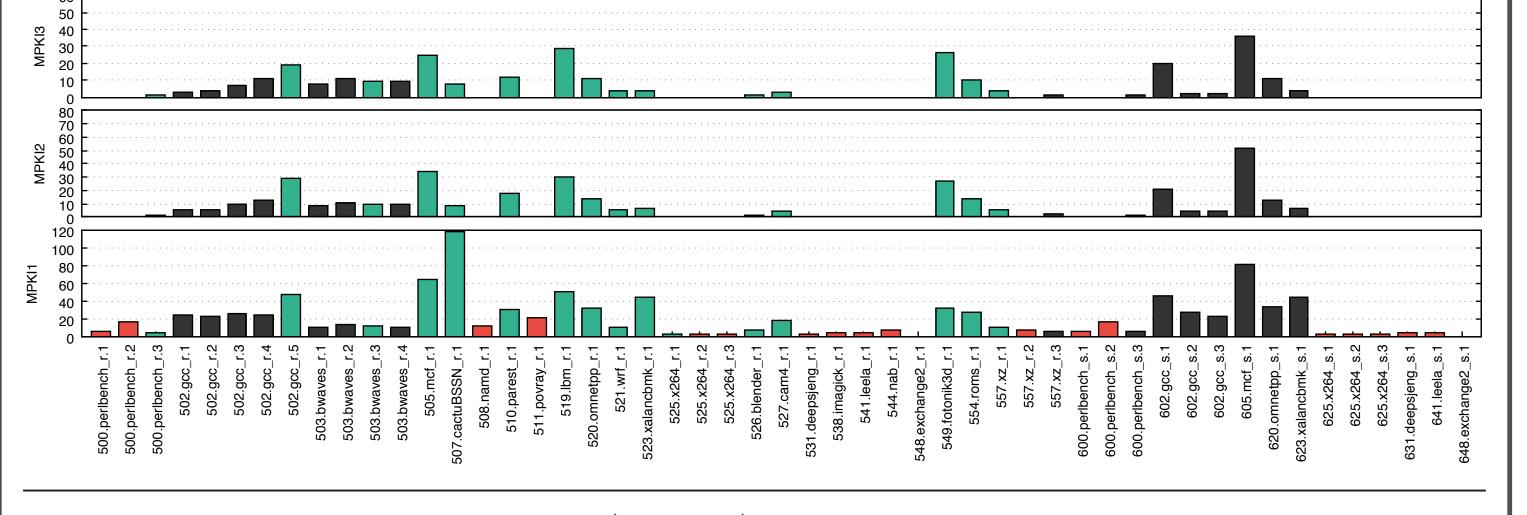
1. Introduction

SPEC CPU is one of the most widely used benchmark suites for high performance computing research on academia and industry. The last version, CPU2017 [1], was released on June 2017. In this work we analyze the memory hierarchy the study all the single-thread benchmarks, identify the memory intensive ones, and analyze their sensitivity to the last-level cache (LLC) size and to the different hardware prefetchers.

The characterization was performed on an Intel Xeon Skylake-SP Gold 5120. Several hardware performance counters were collected with the Perf profiler. We used the Intel Cache Allocation Technology (CAT) [2] to modify the available LLC capacity.

2. Identification of Memory Intensive Benchmarks

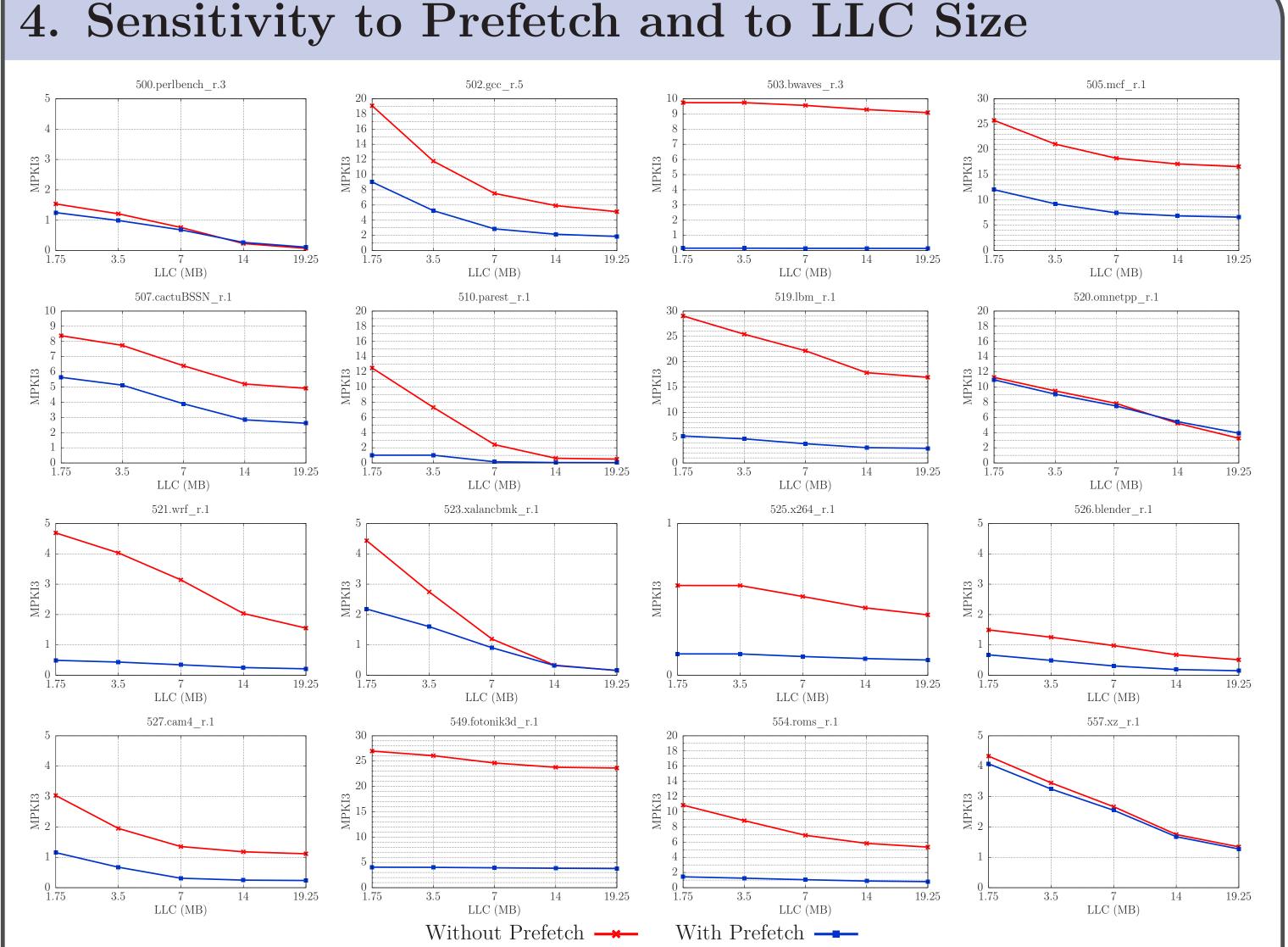
3. Performance Impact of Prefetch/LLC Size

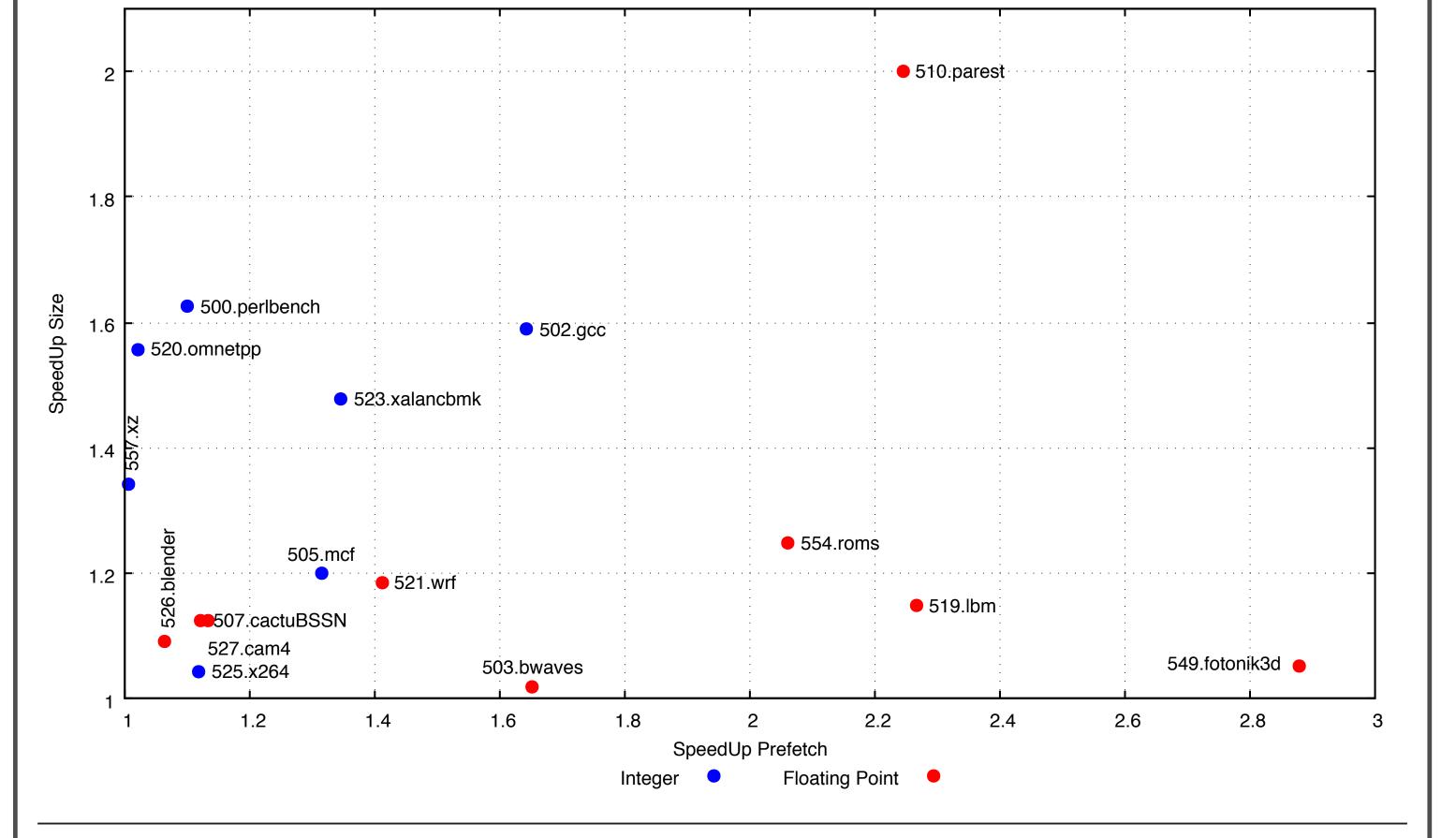


Misses per kilo-instruction (MPKI) for the three cache levels. All CPU2017 single-thread benchmark-input pairs were executed without hard-ware prefetching and 1.75MB of LLC. Benchmarks that have very low MPKI2 and MPKI3 ratios are plotted in red.

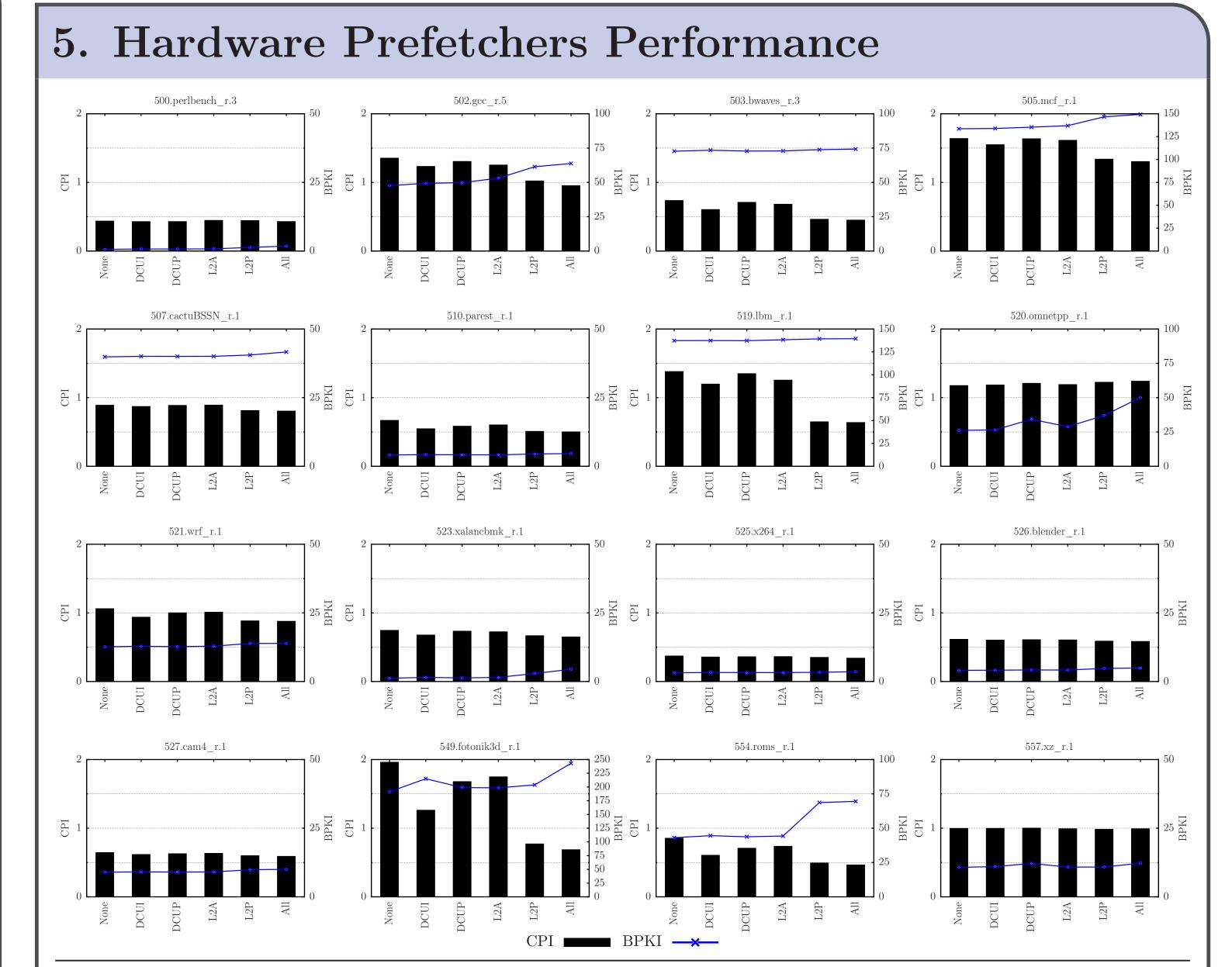
Only 16 out of 23 benchmarks are memory intensive

For these benchmarks only one input will be considered in the following experiments (green bars).





Performance impact of hardware prefetching (x axis) and LLC size (y axis). Speed-ups with respect to the smallest LLC without hardware prefetching.



Sensitivity of the memory intensive benchmarks to the hardware prefetching and to the LLC size.

Hardware prefetching is very effective in reducing MPKI3 even with the smallest LLC size Cycles per instruction (CPI) and bytes read from main memory per kiloinstruction (BPKI) for different prefetching configurations.

With hardware prefetching, increasing LLC size translates to significant MPKI3 reductions for 9 out of 16 benchmarks

6. Conclusions

- Several benchmark-input pairs show very low MPKI2/3 ratios even with 1.75MB of LLC and no hardware prefetching.
- Hardware Prefetching is very effective in 13 out of 16 benchmarks.
- Increasing LLC size with hardware prefetching reduces MPKI3 in 9 out of 16 benchmarks.
- Hardware prefetching is bandwidth-efficient.
- L2P prefetcher achieves most of the CPI reduction.

Hardware prefetching is very efficient because it achieves noticeable CPI reductions with low bandwidth overhead

7. References

- 1. SPEC CPU2017, Standard Performance Evaluation Corporation, https://www.spec.org/cpu2017/
- 2. Introduction to Cache Allocation Technology in the Intel® Xeon® Processor E5 v4 Family, Intel, *https://goo.gl/oZaVQ7*

This work was supported in part by grants TIN2016-76635-C2-1-R (AEI/ERDF, EU) and gaZ: T58_17R research group (Aragón Gov. and European ESF).