**2016 Computer Architecture – Midterm Exam**

1. (50%) True or false questions
2. Dependability, scalability, and though put are characteristics of supercomputers.
3. Price-performance and power are critical to warehouse-scale computers.
4. Code density is a key requirement of printers.
5. Computer organization includes the high-level aspects like the memory system, the bus architecture, and the design of internal CPU.
6. The growth rate of capacity of flashes is faster than that of DRAM.
7. A key advantage of benchmarks is to lessen the weakness of any one benchmark.
8. Combining independent loops that have same looping and some variables overlap can reduce miss rates.
9. Victim caches are used to reduce miss rates.
10. In write allocate policy, blocks stays out of the cache until the program tries to read the blocks.
11. We can predict cache performance of one program form another.
12. (18%) Assume that we make an enhancement to a computer that improves some mode of execution by a factor of 10. Enhanced mode is used 50% of the time, measured as a percentage of the execution time when the enhanced mode is in use. Recall that Amdahl’s law depends on the fraction of the original, unenhanced execution time that could make use of enhanced mode.

Thus, we cannot directly use this 50% measurement to compute speedup with Amdahl’s law.

1. (5%) What is the Amdahl’s law?
2. (6%) What is the speedup we have obtained from fast mode?
3. (6%) In the textbook, Amdahl’s law is for uniprocessor. Justify if it can be applied to multicore and multiprocessor architectures.
4. (18%) We can break the uniprocessor miss rate into the three C’s classification (capacity, compulsory, and conflict) and provides insight into both application behavior and potential improvements to the cache design.
5. (6%) Explain these miss categories.
6. (6%) Give optimization techniques for capacity miss.
7. (6%) Give optimization techniques for conflict miss.
8. (12%) Computer designers predicted the programmers would want unlimited amounts of fast memory. An economical solution to that desire is a memory hierarchy, which takes advantages of locality and cost/performance of memory technologies.
9. (6%) Explain what the locality is.
10. (6%) For registers, cache, and main memory, whom are they managed by?
11. (20%) Instruction prefetch is frequently done in hardware outside of the cache. To have value, we need a processor that can allow instructions to execute out-of-order. One approach is to fetch items before they are requested by the processor. Another approach can be also applied by software.
12. (6%) Explain how hardware prefetching works.
13. (6%) Explain how software prefetching works.
14. (8%) Compare the advantages and disadvantages of hardware prefetching and software prefetching.
15. (15%) The pressure of both a fast clock cycle and power limitations encourages limited size for first level caches. One of key issues focuses on the reduction of the critical timing path in a cache hit.
16. (6%) Give three techniques to reduce hit time of L1 cache.
17. (9%) Explain the three techniques above how to reduce hit time of L1 cache.