CSC 362 Homework Assignment #4
Due Date: Tuesday, June 27

Note: an answer key will be provided on the due date so no late assignments will be accepted. Word Processor all answers, show work for partial credit. Answer 10 of the 11 problems or all 11 for extra credit.

1. Suppose a computer has 16 Gigabytes of memory and a cache of 1024 blocks where each block can store 8 words. The word size is 64 bits and memory is word addressable.
   a. What is the format of memory addresses if the cache is direct-mapped?
   b. What is the format of memory addresses if the cache is 2-way set associative?
   c. What is the format of memory addresses if the cache is 8-way set associative?
   d. What is the format of memory addresses if the cache is fully associative?

2. Read problem #11 on page 393. Use the same memory and cache layout and determine for the 20 memory references below which ones are hits and which ones are misses, and compute the hit rate. Assume the cache is already loaded with the following:
   - block 0: tag 5
   - block 1: tag A
   - block 2: tag 5
   - block 3: tag 0
   The memory references are: 5A, 5B, 5C, 5D, 5E, 32, 33, 34, A0, A1, 50, 51, 52, 26, 27, 2A, 2B, 5B, 5C, 0C. Apply the references in order. A miss causes the cache to change. Your answer need only list which of the 20 accesses are hits and which are misses (you don’t have to draw the cache as it updates). Also include the cache’s overall hit rate.

3. A computer uses an on-chip cache backed up by an off-chip cache backed up by main memory. The on-chip cache has a hit time of .33 ns, the off-chip cache has a hit time of 2 ns and main memory has an access time of 20 ns. The on-chip cache has a hit rate of 90.4%, the off-chip cache has a hit rate of 96.2% and main memory has a hit rate of 100%.
   a. What is the average memory access time for this machine?
   b. We add a second cache to the chip itself (between the current on-chip and off-chip caches) with a hit time of .6 ns and a hit rate of 94.1%. What is the new average memory access time for this machine?
   c. We add to part b virtual memory with a TLB with a hit time of .33 ns hit time and hit rate of 98.7%, a page table in memory with the same hit time as main memory (20 ns) and a 100% hit rate, and main memory now has a hit rate of 99.9973% where swapping takes 1,250,000 ns. What is the average memory access time for this machine now?

4. Aside from faster access, the cache improves on the EAT because of locality of reference.
   a. What is locality of reference?
   b. How does grouping multiple words in a line/block improve cache hit rate?

5. Answer the following two questions about cache types.
   a. Why does a direct-mapped cache have a lower hit rate than any form of cache with associativity?
   b. Why is a direct-mapped cache faster and cheaper than any form of cache with associativity?

6. Assume a computer is word addressable with 2G of addresses and is running a process whose page table is given below. Answer the following questions.
   a. How big is a page/frame in words? NOTE: look at b-d to help answer this
   b. What is the physical address for the logical address 1000000000000000?
   c. What is the physical address for the logical address 0110011001100111?
   d. What is the physical address for the logical address 0101111000011111?
   e. How large is this program (approximately) in words?
   f. How much of the program is in memory?
7. Architects are considering two changes to a processor. First, they are thinking of making a superpipeline which means that memory access stages will now be divided into 2, which allows them to double the speed of the system clock and thus increase the speed by a factor of 2. The superpipeline will incur 25% more stalls so that the process will be idle 25% of the time. Or, the architects are going to add a larger on-chip cache which will mean that memory accesses decrease. The cache is 25 times faster than memory and the expanded cache will improve on the previous cache’s hit rate by 10%. Use Amdahl’s law to compute each change’s speedup. Which should the architects adopt?

8. Examine figure 7.10 on page 418. List the order that the signals are sent over these control lines to a disk controller: error, ready, request, reset, write/read

9. Explain under what circumstance(s) the CPU is interrupted in each of the four I/O forms: programmed, interrupt-driven, DMA, channel.

10. In a RAID 5 system, a disk block is broken into 5 stripes plus parity. For a given block, the first byte of each stripe is as follows, compute the parity byte for these five (assuming even parity): 11110000, 00000011, 10101011, 01000101, 11000011.

11. Answer the following regarding the RAID levels:
   a. We do not use RAID 4, why not?
   b. We want to purchase RAID for our webserver. We want redundancy but do not want to use RAID 1. Which RAID level would you recommend?
   c. Rank the RAID levels that we actually use in terms of expense from cheapest to most expensive.