ELEC 5200/6200 Computer Architecture and Design
Fall 2013
Chapter 5 Problems
Due: Wednesday, 10/30/13

Problem 1:

(a) A one-level SRAM cache has n times shorter cycle time compared to the main memory. Show that it will provide an average memory access speed up of n/(1 + mn), where m is the miss rate. Show that irrespective of how fast the cache is the speed up cannot be greater than 1/m.

(b) Show that the average memory access speed up has an upper bound $1/(m_1 \times m_2)$, where m_1 and m_2 are the miss rates of L1 and L2 caches, respectively.

Problem 2: To meet the data access time requirement of a computer its cache system is to have a 95% hit rate. However, to reduce the cost and to match the speeds of the processor and memory, we are forced to use a smaller size SRAM, which when placed in a 1-level cache can only provide a 50% hit rate.

(a) If the main memory is 50 times slower than the SRAM cache hardware, find the average data access time for the 1-level cache with 50% hit rate, expressed in terms of T1, the cycle time of the SRAM cache. Show that this data access time is at least seven times longer than that if the hit rate was 95%, i.e., cache was larger but had the same cycle time.

(b) We add a level-2 DRAM cache to bring the data access time to the required value. The cycle time of this DRAM is two times slower than the L1 SRAM. Determine the minimum hit rate for the L2 cache.

Problem 3: For a two-level cache, the cycle times for L1 and L2 caches and main memory are 1, 10 and 100 clock cycles, respectively. Both L1 and L2 caches have the same hit rate h. What should h be so that the average data access time of this cache system is 3 cycles?

Problem 4: Consider a processor that uses a 32-bit virtual memory address. The memory consists of 32-bit words and is byte addressable. Determine,

a. How many bytes the virtual memory can have?

b. If the page size is 16KB how many records should the page table hold?

c. How much data space physical memory should have to hold 32K pages.

Problem 5: A cache must accommodate extra bits besides the data bits. Each block contains tag bits and one valid bit. Show that for 32 bit data words, the ratio of the hardware storage bits to data bits in cache can be expressed as,

$$\frac{\text{Cache hardware bits}}{\text{Data bits in cache}} = 1 + \frac{B}{bw}\left(1 + \log_2 \frac{W}{w}\right)$$

Where W = number of words in the main memory
w = number of data words in cache
B = number of blocks in cache
b = word size in bits