ELEC 5200/6200 (Fall 2021)
Homework 6
Assigned 11/12/21, due 11/19/21
(35 points possible)

Problem 1: Consider a 16-bit address space with the following list of memory addresses, given as word addresses.

3, 180, 43, 2, 191, 88, 190, 14, 181, 44, 186, 253

(a) For each address given, provide the full 16-bit binary memory **byte** address along with the binary tag(s) and index for a direct-mapped cache with 16 one word blocks. Assuming these addresses are accessed sequentially as listed above, determine if each memory access is a hit or miss assuming the cache is initially empty. (4 points)

>>4 last 4 bits

Word	16-Bit Byte Address	Tag	Index	Access
3	0000 0011	0000	0011	Miss
180	1011 0100	1011	0100	miss
43	0010 1011	0010	1011	miss
2	0000 0010	0000	0010	miss
191	1011 1111	1011	1111	miss
88	0101 1000	0101	1000	miss
190	1011 1110	1011	1110	miss
14	0000 1110	0000	1110	miss
181	1011 0101	1011	0101	miss
44	0010 1100	0010	1100	miss
186	1011 1010	1011	1010	miss
253	1111 1101	1111	1101	miss

(b) For each of the above memory references, provide the full 16-bit binary memory **byte** address along with the binary tag(s) and index for a direct-mapped cache with 8 two-word blocks. Assuming these addresses are again accessed sequentially, determine if each memory access is a hit or miss assuming the cache is initially empty. (4 points)

>>4 last 3 bits

Word	16-Bit Byte Address	Tag	Index	Access
3	0000 0011	0000	001	Miss
180	1011 0100	1011	010	miss
43	0010 1011	0010	101	miss
2	0000 0010	0000	001	Hit
191	1011 1111	1011	111	miss
88	0101 1000	0101	100	miss
190	1011 1110	1011	111	Hit
14	0000 1110	0000	111	miss
181	1011 0101	1011	010	Hit
44	0010 1100	0010	110	miss
186	1011 1010	1011	101	miss
253	1111 1101	1111	110	miss

Problem 2:

(a) A one-level cache system contains a SRAM cache that is n times faster than the main memory. The cache has an access time T and its miss rate is m. Show that this system will provide an average memory access speed up of n/(1 + mn). Irrespective of how fast the cache hardware is, show that the speed up has an upper bound 1/m. (4 points)

$t = T + (1-h)nT = T + mnT$

- Speed up $= \frac{nT}{t} = \frac{nT}{T+mnT} = \frac{n}{1+mn}$

Cache infinately faster: $\lim_{n\to\infty} \frac{n}{1+mn} = \lim_{n\to\infty} \frac{1}{\frac{1}{n}+m} = \frac{1}{m}$

(b) For a two-level cache, show that the average memory access speed up has an upper bound $1/(m_1 \times m_2)$, given that m_1 and m_2 are the miss rates of L1 and L2 caches, respectively. (4 points)

$t = T + m_1(KT + m_2 nT)$

Speedup: $\frac{nT}{t} = \frac{nT}{T + m_1(KT m_2 nT)} = \frac{1}{\frac{1}{n} + m_1\left(\frac{K}{n}\right) + m_1 m_2}$

- $n \to \infty$, speed up $= \frac{1}{m_1 m_2}$

Problem 3: For a direct mapped cache design with a 32-bit byte address, the following bits are used to access the cache:

Tag	Index	Offset
31- 10	9 -5	4 - 0

(a) What is the cache block size (in words)? (2 points)

2^5 bytes = $\frac{2^5}{2^2} = 2^3$ words = 8 words

(b) How many entries does the cache have? (2 points)

5 bits (9-5) = 2^5 = 32 entrys

(c) Assuming a 32 bit data word, what is the ratio of total bits required to implement the cache to the number of data storage bits? (2 points)

Assume 1 valid bit

data bits = 8X32 = 256 bits

tag bits = 22 bits

Valid bits = 1

$\Rightarrow$ 256 + 22 + 1 = 279 bits

ratio = $\frac{\text{total bits}}{\text{data bits}} = \frac{279}{256} = 1.089$

Problem 4: To satisfy the architectural requirement of a computer system with a one-level cache the hit rate must be 95%. An early prototype for cost reasons uses a SRAM of limited capacity as a one-level cache that only provides a 70% hit rate. The main memory is 70 times slower than the SRAM cache.

(a) Find the average data access time for the one-level cache with 70% hit rate, expressed in terms of the cycle time T1 for the cache. Show that the data access will become almost five times faster if the hit rate could be raised to 95%. (4 points)

$T_m = 70 T_1$

Av Access time $T(h) = T_1 + (1-h)T_m$

for h = 0.7 $\quad T(0.7) = T_1 + 0.3T_m = T_1(1 + 0.3 \times 70) = 22\, T_1$

for h = 0.95 $\quad T(0.95) = T_1 + 0.05T_m = T_1(1 + 0.05 \times 70) = 4.5 T_1$

Access Time Ratio $\frac{T(0.7)}{T(0.95)} = \frac{22}{4.5} = 4.89 \approx 5$

(b) Suppose a level-2 cache brings the data access time to the required value when the cycle time of the L2 cache is 1.5T1. Determine the minimum hit rate for the L2 cache. (4 points)

Two-level cache access time $= T1 + (1-h1)[T_2 + (1-h_2)T_m]$

$T1 + 0.05Tm \geq T1 + 0.3[T_2 + (1-h_2)T_m]$

$\Rightarrow 3.5T1 \geq 0.45T1 + 21(1-h_2)T1$

$\Rightarrow \frac{3.05}{21} \geq (1-h_2)$

$\Rightarrow h_2 \geq 1 - 0.145 = 0.855$

Hit rate of L2 cache needs to be at least 85.5%

Problem 5: In a two-level cache system, cycle times of L1 and L2 caches and main memory are 1, 8 and 64 clock cycles, respectively. The miss rate of L2 cache is twice that of L1. What should the two miss rates be so that the average data access time of this cache system is 2 cycles? (3 points)

L1 hit rate $= 1 - \text{miss rate} = 1-m$

L1 miss Rate $= m$

L2 hit rate $= 1-2m$

L2 miss rate $= 2m$

AMAT $= 2$

$2 = [(1-m) \times 1] + (m \times [1 + (1-2m) \times 8 + 2m \times (8+64)]$

$2 = 1 + 8m + 128m^2$

$(16m-1)(8m+1) = 0$

$16m - 1 = 0$ or $8m+1=0$

$m = 0.0625$ $\quad m = -0.125$ ⇒ discard neg value

miss rate for L1 $= 0.0625$ & L2 $= 2m \Rightarrow 2(0.0625) = 0.125$

Problem 6: A cache must accommodate extra bits besides the data bits. Each block contains tag bits and one valid bit. Show that the overhead for a one-level direct mapped cache is;

$$\text{Cache overhead} = \frac{\text{Total bits in cache} - \text{Data bits in cache}}{\text{Data bits in cache}} = \frac{1}{Bb}\left(1 + \log_2 \frac{W}{w}\right)$$

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

(2 points)

Data bits in Cache = wb

Total bits in Cache = data bits + (# block in cache) x (1 valid bit + # tag bits)

$$= wb + \left(\frac{W}{B}\right) \times \left(1 + \log_2\left(\frac{W}{w}\right)\right)$$

$$\text{Cache overhead} = \frac{\left[wb + \left(\frac{W}{B}\right) \times \left(1 + \log_2\left(\frac{W}{w}\right)\right) - wb\right]}{wb}$$

$$= \frac{\left(1 + \log_2\left(\frac{W}{w}\right)\right)}{Bb}$$