Computer Architecture Homework 5

Spring 2023, April

1 True or False

Please fill your answer (T or F) in the table below:(10 pts)

1	2	3	4
F	T	F	F

1. Any cache miss that occurs when the cache is full is a capacity miss.
2. Cache replacement policy is used for choosing which cache line should be evicted.
3. For a 1-level, LRU cache, cache blocking will greatly speed up matrix transposition when the cache size is much larger than the matrix size.
4. In a 32-bit machine (word size is 32 bit), the clock frequency is 2GHz. It takes 3 cycles to access L1, L1 cache has a hit rate of 30%.It takes 40 cycles to access L2, L2 cache has a hit rate of 80%. It takes 400 cycles to access physical memory. The average memory access time is 87ns.

T=0.5ns 1.5ns 20ns 200ns

miss penalty : 200ns
L1 hit : 1.5ns
L2 hit : 20ns

AMAT = 1.5 + 0.7 × (20 + 0.2 × 200) = 43.5ns

2 AMAT Analysis

A program is running on a system with a single-level, write-back, 512B direct-mapped cache that has the block size of 4B. It traverses an array with step size = 1. Choose how the modification changes each component of AMAT. If there are multiple choices, select them all.(15 pts)

N=1

100%

Modification	Hit Time	Miss Rate	Miss Penalty
Change to 2-way Associativity (cache size and block size keeps the same)	A.Increase B.Decrease C.No effect	A.Increase B.Decrease C.No effect	A.Increase B.Decrease C.No effect
Your Choice	A	C	C

tag index offset
? 9 2
↓
? 8 2

3 N-way Set Associative Cache

A program is running on a byte-addressed system with a single-level cache, where memory addresses are 10 bits long. The entire cache is shown below. The block size of this cache is 16 B.

Index	Tag1	Tag2
0b00	1011	0011
0b01	0011	
0b10	0110	1110
0b11	0010	1111

1. Is this cache direct-mapped, set-associative, or fully-associative? If it is associative, also write down its associativity.(5 pts)

set - associative 2-ways

2. Write down the width (in bits) of Tag, Set index and Block offset.(5 pts)

Tag	Set index	Block offset
4	2	4

3. How many bytes of data can this cache contain? Please show your process.(5 pts)

Byte = $\frac{\text{blocks}}{\text{sets}}$ * sets * $\frac{\text{Bytes}}{\text{Block}}$

Total cache capacity = associativity * # of sets * block size = 2 x 4 x 16 = 128 Bytes

4. We will access the data of addresses as follows, one by one. Fill in the blanks, and determine if each access would be a hit or miss based on the cache state shown above (suppose the cache is empty at first). This cache uses LRU replacement policy. If it's a miss, classify the possible miss type(s), select all possible miss types.(20 pts)

Address				Your Choice
0b0011010000	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b0011011011	A.Hit	B.Compulsory miss	C.Conflict miss	A
0b1101000100	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b0011001100	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b0110100010	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b0010111100	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b1110100010	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b1011000000	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b1111111111	A.Hit	B.Compulsory miss	C.Conflict miss	B
0b0011001101	A.Hit	B.Compulsory miss	C.Conflict miss	A

4 Cache Friendly Programming

Assume a 32-bit machine, suppose the cache has the following settings:

Cache levels	1
Block size	16 bytes
Number of sets	4
Cache size	128 bytes
Block replacement policy	LRU

Suppose the following code is running on a system with the above cache. Answer the questions below and show your progress/explanation.

```
#define array_size    64    //line 1
#define repeat_times  1             2
#define step_size     2             3
int main(){                         4
    int array[array_size] = {};     5
    for (int r = 0; r < repeat_times; r++){          6
        for (int i = 0; i < array_size; i += step_size){   7
            array[i] = array[i] + 2333;                8
        }
    }
    return 0;
}
```

tag index offset
26 2 4

1. What is the total number of accesses to the cache?(5 pts)

line 8 : read + write ⇒ 2 access

total = ~~reape~~ repeat_times * $\lfloor \frac{array\_size}{step\_size} \rfloor * 2 = 1 \times \lfloor \frac{64}{2} \rfloor \times 2 = 64$

∴ there are total 64 accesses.

2. What is the hit rate?(10 pts)

		read	write
... 000 00 0000	i = 0	miss	hit
... 000 00 1000	i = 2	hit	hit
... 000 01 0000	i = 4	miss	hit
... 000 01 1000	i = 6	hit	hit

So for every two i, there are 1 miss, 3 hits

hit rate = $\frac{3}{4} = 75\%$

3. Which type(s) of miss occur(s)(5 pts)? i

all compulsory miss

4. Suppose repeat_times **goes to infinity** (only for this question), what number will the hit rate converge to?(5 pts)

The repeat-times is in the outer loop,
it does not effect the hit rate

So the hate rate remains (converges) to 75%

5. If repeat_times **is changed to 2** (only for this question), try to swap two lines of the above code to maximize the hit rate without disturbing the results. Which two lines will you choose and what is the maximized hit rate?(10 pts)

swap line ~~5~~ 6 and line 7

the repeat-time is now in the inner loop

So for every two i, the effect becomes

	read	write	read	write
i=0	miss	hit	hit	hit
i=2	hit	hit	hit	hit

So the maximized hit rate is $\frac{7}{8} = 87.5\%$

6. **For the modified code in the previous question,** suppose again repeat_times **goes to infinity** (only for this question), what number will the hit rate converge to?(5 pts)

since the repeat-time is in the inner loop,
so as repeat-times goes to infinity,
for every two i, there is only the first access is miss,
the others are all cache hit

So the hit rate is $\lim_{n\to\infty} \frac{2n-1}{2n} = 1$

So the hit rate converge to 100%