CS110 sp24 HW5

Due: 14^{th} May

You should finish this homework either by writing it **neatly** by hand or using LaTeX (**highly recommended!!!**). You can find the `.tex` file on Piazza.

1 T/I/O Breakdown

1. Given that we have a direct-mapped byte-addressed cache with capacity 32B and block size of 8B. Of the 32 bits in each address, which bits are offset bits? Which bits are index bits? What about tag? *Note: Please provide your answer in the format [n:m] to denote the range from the m^{th} bit to the n^{th} bit (e.g., [1:0] represents the two lowest bits).*

Tag	Index	Offset
[31:5]	[4:3]	[2:0]

2. Given the cache in question 1.1, assuming that we will access memory addresses in the following order, classify each of the accesses as a cache hit (H), cache miss (M) or cache miss with replacement (R). Ignore miss types for now. *Note: The distinction of M and R here is just for your understanding, and that the cache doesn't behave differently for these cases.*

Address	Hit, Miss, Replace	Miss Type
0x00000004	Miss	Compulsory
0x00000005	Hit	
0x00000068	Miss	Compulsory
0x000000C8	Replace	Conflict
0x00000068	Replace	Conflict
0x000000DD	Miss	Compulsory
0x00000045	Replace	Conflict
0x000000CF	Replace	Confilct
0x000000F3	Miss	Compulsory

2 Set-Associative Caches

Given that we have a 2-way set associative cache. This time we have an 8-bit address space, 8B blocks, and a cache size of 32B. Classify each of the accesses as a cache hit (H), cache miss (M) or cache miss with replacement (R). Assume that we have an LRU replacement policy. Ignore miss types for now.

Address	Hit, Miss, Replace	Miss Type
0b0000 0100	Miss	Compulsory
0b0000 0101	Hit	
0b0110 1000	Miss	Compulsory
0b1100 1000	Miss	Compulsory
0b0110 1000	Hit	
0b1101 1101	Replace	Conflict
0b0100 0101	Miss	Compulsory
0b0000 0100	Hit	
0b0011 0000	Replace	Capacity
0b1100 1011	Replace	Capacity
0b0100 0010	Replace	Capacity

3 The 3C's Cache Misses

Go back to question 1 and 2 and classify each miss as one of the three types of misses.

4 Code Analysis

Consider the following function that takes in two integer arrays, a (of length a_len) and b (of length b_len), and returns the 1D convolution of a and b. Assume results is properly allocated. Let a=0x1000, b=0x2000, results=0x3030, a_len=4, and b_len=2. *Note: The register keyword in C provides a hint to the compiler to consider storing a variable in a processor register.*

```
void convolve_1d(int* a, int a_len, int* b, int b_len,
    int* results) {
    for (int i = 0; i < a_len - b_len + 1; i++) {
        register int sum = 0;
        for (int j = 0; j < b_len; j++) {
            sum += b[j] * a[i + j];
        }
        results[i] = sum;
    }
}
```

1. Given that we have a single-level, direct-mapped 64B cache with 16B blocks and 16-bit addresses. What is the overall hit rate for a call to convolve_1d?
Answer:
The overall hit rate is $\frac{2}{15}$

2. Given that we have a 2-way set associative cache of the same size with a LRU replacement policy. What is the overall hit rate for a call to convolve_1d?
Answer:
The overall hit rate is $\frac{4}{5}$

3. Given that we have a fully associative cache of the same size with a LRU replacement policy. What is the overall hit rate for a call to convolve_1d?
Answer:
The overall hit rate is $\frac{4}{5}$

5 AMAT

1. In a 2-level cache system, if L1 has a local miss rate of 50% and the global miss rate of L2 is 20%, what is the local miss rate of L2?
Answer:
$50\% \cdot x = 20\%$, so the local miss rate of L2 is 40%

Suppose your system consists of:

1. An L1 that has a hit time of 2 cycles and has a local miss rate of 20%.
2. An L2 that has a hit time of 15 cycles and has a global miss rate of 5%.
3. Main memory where accesses take 100 cycles.

2.What is the AMAT of the system?
Answer:
$\text{AMAT} = 2 + 20\% \times (15 + 25\% \times 100)$
So the AMAT of the system is 10 cycles.

3. Suppose we want to reduce the AMAT of the system to 8 cycles or lower by adding in a L3. If the L3 has a local miss rate of 25%, what is the largest hit time that the L3 can have?
Answer:
$2 + 20\% \times (15 + 25\% \times (x + 25\% \times 100)) \leq 8$
$x \leq 35$
So the largest hit time that the L3 can have is 35 cycles

The following TA(s) are responsible for this homework:
Xinxin Yu: yuxx@shanghaitech.edu.cn
Lei Jia: jialei2022@shanghaitech.edu.cn