Computer Architecture

Fall, 2022

Week 15

2022.12.19

組別：_____ 簽名：_____

_____

Group10

In this exercise, we will look at the different ways capacity affects overall performance. Assume that main memory accesses take 70 ns and ~~that 36% of all instructions access data memory~~. The following table shows data for L1 caches attached to each of two processors, P1 and P2.

	L1 Size	L1 Miss Rate	L1 Hit Time
P1	2 KiB	8.0%	0.66 ns
P2	4 KiB	6.0%	0.90 ns

= 1 cycle

What is the Average Memory Access Time for P1 and P2 (in cycles)?

Ans: AMAT = hit time + L1 miss rate × memory access time

$P_1 = 0.66 + 70 \times 0.08 \times 0.36 = 2.676$

$P_2 = 0.9 + 70 \times 0.06 \times 0.36 = 2.412$

$P_1 = 1 + \left[\frac{70}{0.66}\right] \times 0.08 = 9.56$

$P_2 = 1 + \left[\frac{70}{0.9}\right] \times 0.06 = 5.68$

Group7

Which option(s) is/are true? Explain if it's false.

(a) We can access and transfer data in every memory bank simultaneously using interleaved memory organization.

√ (b) Swap space is a program's virtual memory space in disk.

√ (c) Using dynamically scheduled CPU, we can pend the load/store instruction and proceed to execute every other instruction when encountering a cache miss. 有些指令會 dependent 若 independent 就可以

independent

√ (d) We can improve the performance of caches by increasing its associativity or using multilevel caches.

T (e) We can implement speculation about data from multi-way associative caches.

√ (f) For a processor with clock rate of 2GHz and base CPI of 1, its D-cache miss rate is 8%, I-cache miss rate is 3%, and hit time of both is 1 cycle. If main memory access time is 50ns, and there are 100 load/save instructions in a 400-instruction program without jumps/branches, its effective CPI is 6 for this program.

√ (g) With the statement in (f), its AMAT is 3ns for this program.

(h) If we add a secondary D-cache into (f) with a miss rate of 2% to main memory and hit time = 10ns, its effective CPI becomes 3.9.

Ans: b c d f g

a. in a clock, sequencial

c. 不行，需要有一个 MUX Delay

f. $\frac{1}{2\times10^9} = 0.5$ ns , $\frac{50}{0.5} = 100$ cycles

I catch $= 0.03 \times 100 = 3$

D catch $= 0.08 \times 100 \times \frac{1}{4} = 2$

CPI $= 3+2+1 = 6$

g. 1 cycle $= 0.5$ ns

AMAT: $0.5 + 0.08 \times 50 \times \frac{1}{4} + 0.03 \times 50 = 3$ ns

AMAT: $0.5\text{ns} + \frac{400}{400+400\times0.25}(0.03\times50) + \frac{400\times0.25}{400+400\times0.25}(0.08\times50) = 2.5$ ns

h. $\frac{10}{0.5} = 20$ cycles

$0.08 \times 20 \times \frac{1}{4} = 0.4$, $0.02 \times 100 \times \frac{1}{4} = 0.5$

CPI $= 3+0.4+0.5+1 = 4.9$

Group1

Please select the correct options and explain the incorrect ones.

√ a. When CPU performance increased, miss penalty becomes more significant.

✓ b. The fully associative cache searches all entries in a given set at once.

c. The number of virtual pages is always equal to the number of physical pages.

d. Increasing associative will reduce miss rate and hit time. maybe increase

✓ e. Virtual memory "block" is called a page.

f. The design of an interleaved memory organization allows data inside different memory banks' can be sent to the cache at the same time.

g. When calculating average memory access time, hit time is not important.

Ans: a b e

c. 每个 program 的 translator 都不一樣 smaller or equal

d. hit time 不會變

f. 只能寄一个到 cache

g. hit time 都要考慮

Group3

Choose the correct answers and explain wrong ones

(A) In a multi-level cache system, the lower level should have lower miss rate.

(B) We can increase the miss penalty to improve the cache performance.

(C) The minimum unit of the cache is called “page” and the minimum unit of the virtual memory is called “block”.

(D) If we increase associativity, miss rate will increase too.

(E) The performance of fully associative cache is always better than direct mapped cache, so we don't need direct mapped cache at all.

(F) The n-way set associative cache is better than the fully associative cache because the n-way set associative cache has the best hit rate. Assume send the address : **1 memory bus clock ;** each DRAM access initiated : **15 memory bus clocks ;**

send a word of data : **1 memory bus clock ;**

A cache block : **4 words .**

Miss penalty of a four-word-wide bank of DRAMs = 1 + 15 (+1)= 17

(G)

1. Transfer time overlap with cycle time
2. 1 memory bus clock to send the address
3. 10 memory bus clocks for each DRAM access initiated
4. 1 memory bus clock to send a word of data
5. 4 words/per cache block

For a two-word-wide bank of DRAM, the miss penalty is 1+10+1=12

Ans:

Group4

Please calculate the effective CPI with the given conditions.

CPU base CPI = 3, clock rate = 5GHz

L1 miss rate to main memory or L2 = 1.5%

L2 miss rate to main memory = 0.01%

access time to main memory = 80ns

access time to L2 = 10ns

a. Use only one cache(L1)

b. Use two caches(L1、L2)

$\frac{1}{5\times10^9} = 0.2ns$

Ans: a. penalty to access memory $= \frac{80}{0.2} = 400$ cycles

$CPI = 3 + 0.015 \times 400 = 9$

b. penalty to access L2 $= \frac{10}{0.2} = 50$ cycles

$CPI = 3 + 0.015 \times 50 + 0.0001 \times 400 = 3.79$

3.75 0.04

Group14

Assume we are using a fully associative 4-block cache with an LRU replacement policy. Given the following ordered reference addresses, fill out the cache table and note if each step is a hit or miss:

	0	1	7	7	6	5	8	7	4	0	1	4	5	7	4	8	4	3	2	0
Blocks	**Cache**																			
1st	0	0	0	0	0	5	5	5	5	0	0	0	0	7	7	7	7	7	2	2
2nd		1	1	1	1	1	8	8	8	8	1	1	1	1	1	8	8	8	8	0
3rd			7	7	7	7	7	7	7	7	7	7	5	5	5	5	5	3	3	3
4th					6	6	6	6	4	4	4	4	4	4	4	4	4	4	4	4
H/M	M	M	M	H	M	M	M	H	M	M	M	H	M	M	H	M	H	M	M	M

Ans:

Group8

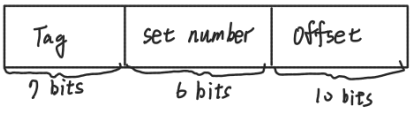
$2^9 = 512$

Consider a 8-way set associative mapped cache of size 512 KB with block size 1 KB. There are 7 bits in the tag. Find Size of main memory and Tag directory size.

- Set size = 8
- Cache memory size = 512 KB $= 2^{19}$B
- Block size = Frame size = Line size = 1 KB $= 2^{10}$B
- Number of bits in tag = 7 bits

Ans:

Block size $= 2^{10}$ B → Number of bits for block size = 10



$$\text{number of block} = \frac{\text{Cache size}}{\text{block size}} = \frac{2^{19}\text{B}}{2^{10}\text{B}} = 2^9 \text{ lines}$$

$$\text{number of sets in cache} = \frac{\text{number of block}}{\text{set associative}} = \frac{2^9}{2^3} = 2^6$$ number of bits in set = 6

Number of bits in physical address = Number of bits in tag + Number of bits in set number + offset

$= 7 + 6 + 10 = 23$ bits

Thus Main memory $= 2^{23}$ bytes $= 8$ MB

Tag directory size = Num. tags x tag size = Num. of lines in cache x Num. of bits in tag

$= 512 \times 7$ bits $= 3584$ bits $= 448$ bytes