# Computer Architecture
# Fall, 2017
# Week 15
# 2017.12.18

[group2]

1. 請說明如何改善 cache performance，並舉例說明之。

**Ans:**

- Reduce the time to hit in the cache
- Decreasing the miss ratio(可利用 associative cashes 來取代 direct mapped 的 block placement)
- Decreasing the miss penalty(適度的 block size 因為過大容易造成 pollution)

[group1]

2. Count the actual CPI

   I-cache miss rate = 3 %

   D-cache miss rate = 4 %

   Miss penalty = 200 cycles

   Base CPI(ideal cache) = 3

   Load and stores are 48 % of instructions

**Ans:**

0.03 * 200 + 0.04 * 0.48 * 200 + 3 = 12.84

[group3]

3.

A. 請簡述 virtual address 和 physical address 的特點與差異。

B. 請問三種 Memory Organization 分別為何？簡單介紹他們個別的特色。

**Ans:**

A.

Physical addresses refer to hardware addresses of physical memory.

Virtual addresses refer to the virtual store viewed by the process.

- virtual addresses might be the same as physical addresses
- might be different, in which case virtual addresses must be mapped into physical addresses.
- virtual space is limited by size of virtual addresses (not physical addresses)
- virtual space and physical memory space are independent

B.

One-word-wide memory organization: 每次到 memory 可以 access 一個 word，bus 一次也是傳輸一個 word。

Wide memory organization：一次可以到 memory access 多個 word，bus 一次也可以傳送多個 word。

interleaved memory organization ：將 memory 分成多個 bank，可以同時 access 多筆 data，但是 bus 還是一次只能傳送一筆 data

[group6]

4. The data cache has 80% hit rate and miss penalty is 120 cycles. 30% of instructions are load and store . The instruction cache has a hit rate of 90% with a miss penalty of 50 cycles. Assume the base CPI in perfect memory system is 1.0. Calculate the overall CPI for the machine?

**Ans:**

The overall CPI = 1+1*(1-0.8)*50+0.3*(1-0.9)*120 = 14.6

[group10]

5.

A. If a direct mapped cache has a hit rate of 80%, a hit time of 4ns, and a miss penalty of 80ns, what is the average memory access time(AMAT)?

B. If increases the hit rate to 95%, but increases the hit time to 7ns, what's the new AMAT?

**Ans:**

A. AMAT = Hit time + Miss rate * Miss penalty
= 4 + (1-0.8) * 80
= 20 (ns)

B. AMAT = Hit time + Miss rate * Miss penalty
= 7 + (1 - 0.95) * 80
= 11 (ns)
The new AMAT is 11ns.

[group4]

6. How many one-word blocks can be cached in a 4-way set-associative cache whose size is 4MB? How many sets are there? (32-bit processor)

**Ans:**

4 Bytes = 1 Word

1 Block = 1 Word = 4 Bytes

4 MB = 1 M Blocks

Set Size = 4 Blocks

Sets=(1 M Blocks)/(4 Blocks/Set)=256 K Sets

[group7]

7. 請問在 Multilevel Cache 的設計中，Primary cache 及 L-2 cache 分別有什麼特色並造成何種結果？

**Ans:**

**Primary cache**

Focus on minimal hit time

**L-2 cache**

Focus on low miss rate to avoid main memory access

Hit time has less overall impact

**Results**

L-1 cache usually smaller than a single-level cache

L-1 block size smaller than L-2 block size

[group8]

8.

A. What are the approaches to reducing the average memory access time in a computer with memory hierarchy?
   (a) reducing the miss rate of cache
   (b) using direct mapped cache
   (c) increasing the associativity of cache
   (d) increasing the cache size

B. Consider a 1 KB, 4-way set associative cache (initially empty) with block size of 64 bytes. The main memory consists of 256 blocks and the request for memory blocks is in the following order: 0, 255, 1, 4, 3, 8, 142, 133, 159, 216, 113, 129, 63, 8, 17, 48, 32, 73, 92, 155. Which one(s) of the following memory blocks will NOT be in the cache if LRU replacement policy is used?
   (a) 3 (b) 8 (c) 133 (d) 216

**Ans:**

A. (a)

B. (c) (d)

[group14]

9. Fill in the blank

_____ cache : each memory block can be placed in any cache location.

Direct mapped cache : cache block is available _____ (Before / After) HIT/MISS.

More increased associativity always being more decreases miss rate is _____ (right / false).

_____ (Decreasing/Increasing) miss rate may decrease AMAT (Average Memory Access Time)?

For a fixed size cache, increasing associativity _____ index, _____ tag.

**Ans:**

Fully associative cache

Before

false
decreasing
shrink, expands

[group5]

10. Assume A cache block = 4 words
    1 memory bus clock to send the address
    18 memory bus clocks for each DRAM access initiated
    2 memory bus clock to send a word of data
    Compute the miss penalty of three memory organizations : A one-word-wide bank of DRAMs, A four-word-wide bank of DRAMs and A four-bank, one-word-wide bus of DRAMs respectively.

**Ans:**

A one-word-wide bank of DRAMs : 1 + 4 x 18 + 4 x 2 = 81
A four-word-wide bank of DRAMs : 1 + 18 + 2 = 21
A four-bank, one-word-wide bus of DRAMs : 1 + 1 x 18 + 4 x 2 = 27

[group9]

11. Assume that the miss rate of an instruction cache is 3% and the miss rate of the data cache is 6%. If a processor has a CPI of 4 without any memory stalls and the miss penalty is 100 cycles for all misses. Assume the frequency of all loads and stores is 30%. How much faster a processor will run with a perfect cache that never missed.

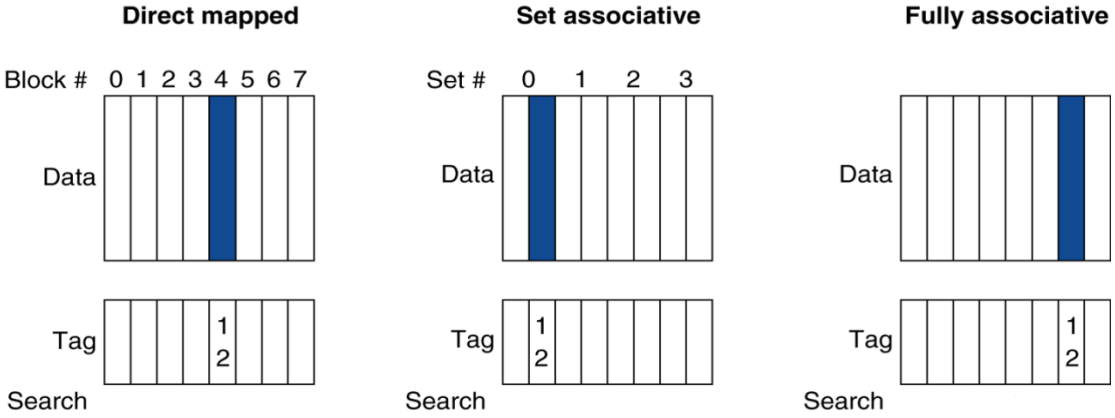
**Ans:**

CPI(new) = 4 + 1*0.03*100 + 0.3*0.06*100 = 8.8
8.8/4 =2.2
2.2 times faster

[group12]

12. 有一天，老王在家裡研讀計算機結構的時候，才去上個廁所回來，課本就被旁邊的屁孩弟弟亂塗鴉，傷腦筋的老王只好把這一頁剪下來，並詢問各位同學。老王的問題如下：
    (1) 請簡單說明 Direct mapped、Set associative 和 Fully associative 三者的特性及差異。
    (2) 請問三者需要檢索的範圍為何？在圖(b)畫上箭頭作答。


圖(a) 老王的課本




圖(b)

**Ans:**

(1) **Direct mapped :** 直接指派 data 到指定的位置，較無彈性，且容易 miss，但檢索較快。

**Set associative :** 每個 Block 都可以找到屬於自己的 Set，而且 Comparator 的數量取決於一個 set 有幾個 entry，相較於 Fully associative 就沒那麼昂貴，也不用檢索所有 entry。

**Fully associative :** data 可以放在任意的 cache entry 內，但是每次都需要檢索所有的 entry，而且有多少 entry 就需要多少 comparator，較昂貴。

**(2)**