Computer Architecture

Fall, 2023

Week 14

2023.12.11

組別：_____ 簽名：_____

[group8]

1. Which of the following statements are true?
   (a) SRAM is slower but cheaper than DRAM.
   (b) Loop is an example of spatial locality.
   (c) Most of the cost of the memory hierarchy is at the highest level.
   (d) The unit of swapping data between cache and memory is blocks, and is managed by the compiler.

Ans:

(a) DRAM is slower but cheaper than SRAM.
(b) Loop is an example of Temporal locality.
(c) True.
(d) It's managed by the cache controller.

[group5]

2. Assume there is a cache with 8 blocks and each block represents one word. All valid bits are initialized to 0. Please try to put the data into the cache and complete the following table.

Ans:

Word address	Binary address	Hit/miss	Cache block
18	10 010	miss	10
13	01 101	miss	101
18	10 010	hit	10
21	10 101	miss	101

Index	Valid bit	Tag	Data
000	N		
001	N		
010	Y	10	Mem[10 010]
011	N		
100	N		
101	Y	10	Mem[10 101]
110	N		
111	N		

[group12]

3. Which of the following statements are true about cache write?
   (a) On write hit, write back policy will only update the data in cache, while write through policy will only update the data in memory.
   (b) On write hit, write through policy typically use a write buffer which is a LIFO with 4 entries.
   (c) On write hit, CPU might get stalled even if there is a write buffer in write through policy.
   (d) On write hit, in write back policy, the data in cache will write back to memory only if the dirty block is replaced.
   (e) On write miss, write through policy will always fetch the block, while write back policy will always not fetch the block.

Ans:

(a) False
Write back will also update the data in memory if the block is dirty.
Write through will update the data in both cache and memory.

(b) False, LIFO -> FIFO

(c) True
Store frequency > 1 / DRAM write cycle may result to write buffer saturation,
which makes CPU stalled.

(d) True
If the block is replaced and it is dirty, write it back to the memory.
If the block is replaced and it is not dirty, just update the block in cache.

(e) False
Write through policy:
Allocate on miss => fetch the block
Write around => don't fetch the block.
Write back policy : Usually fetch the block.

[group14]

4. Subdivide the memory address for a 32-bit word into tag, index, and offset for the following cache configurations:
   (a) i. 32 bytes per block, 256 blocks in a cache
   ii.16 bytes per block, 64 blocks in a cache
   (b) What is block address and block number with address 1400? (for the above two configurations)

Ans:

(a) i. tag = 19bits, index = 8 bits, offset = 5 bits
ii. tag = 22bits, index = 6 bits, offset = 4 bits

(b) i. block address = 43, block number = 43

ii. block address = 87, block number = 23

[group1]

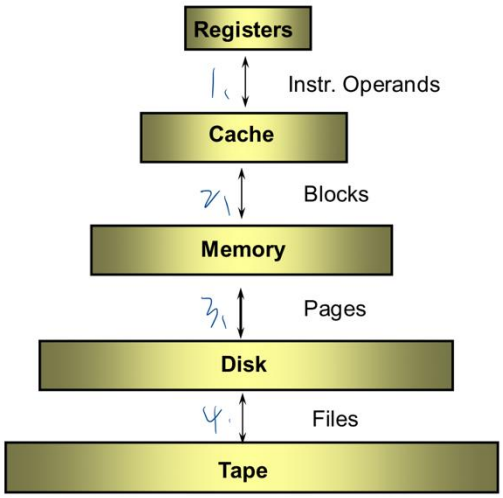
5. Select the correct statements
   (a) For a fixed-size cache, miss rate tends to decrease as the block size increases.
   (b) Write-through caches have better consistency between main memory and cache.
   (c) Miss penalty is the time to replace a block in the upper level plus the time to deliver the block to the processor.
   (d) Both DRAM and SRAM need to be refreshed regularly.
   (e) Array data structure is one of temporal locality.

Ans:

   (a) false. For a fixed-size cache, miss rate will increase as the block size increase.
   (b) true.
   (c) true.
   (d) false. Only DRAM need.
   (e) false. Array data structure is one of spatial locality.

[group4]

6. Please match the following options to their corresponding numbers and show the fastest part and the lowest level.
   (a) OS
   (b) User
   (c) Compiler
   (d) Cache controller



Ans:

1. C
2. D
3. A
4. B

Fastest part: Register

Lowest lever: Tape

[group13]

7. Define the following terms:
   (a) Memory hierarchy
   (b) hit rate
   (c) hit time
   (d) miss rate
   (e) miss penalty

Ans:

   (a) An expanded view of memory system
   (b) Fraction of memory access found in the upper level
   (c) Time to access the upper level
   (d) Miss rate = 1 – hit rate
   (e) Time to access a block in the lower level + time to deliver the block to the processor.

[group6]

8. Please justify the following statements (T/F)
   (a) Small memories fast, and large memories slow.
   (b) 90/10rule means that 10% of code execute 90% of time.
   (c) DRAM is slow but cheap and dense, which is good for presenting users with a BIG memory system.

Ans:

   (a) True.
   (b) True.
   (c) True.