POSTAL Book Package 2023

Computer Science & IT

Objective Practice Sets

Computer Organization & Architecture

Contents

MADE EASY Publications

4 CHAPTER

Memory Hierarchy Design

Multiple Choice Questions & NAT Questions

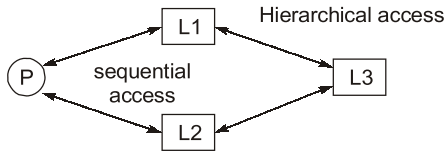
Q.1 Why more than one words are put in one block in order to exploit?
(a) Spatial locality
(b) Temporal locality
(c) More misses
(d) None of the above

Q.2 A certain computer system design has a single CPU, a two-level cache, and supports memory mapped I/O for output-only controllers. Which of the following is true?
(a) The design *i* impossible, since memory mapped I/O will prevent cache coherence.
(b) In two-level caches, the L_1 cache is generally built from SRAM.
(c) In two-level caches the L_1 cache is generally larger then L_2 cache.
(d) In two-level caches, the L_2 caches the L_2 cache generally has a lower latency than the L_1 cache.

Q.3 A 2 level memory has average access time 30 ns with cache and memory access time as 20 ns and 150 ns respectively. What is the hit ratio?
(a) 80% (b) 93%
(c) 70% (d) 99%

Q.4 The access time of cache is 100 μs, the access time of main memory is 90 μs, and hit ratio is 95%, then access efficiency related to cache is
(a) 0.9569 (b) 0.9869
(c) 0.9469 (d) 0.9969

Q.5 Consider a processor with two caches which it can access directly in parallel L1(80% hit rate) and L2 (90% hit rate) with access times as 100 ns and 200 ns respectively. In case of miss in any of those if fetches the data from the main memory L3 (100% hit rate) which has a latency of 500 ns. What is average access time (7 average) of the organization?



Q.6 A two way set associative cache memory unit with a capacity of 16 kB is built using a block size of 8 words. The word length is 32 bits. The physical address space is 4 GB. The number of bits in the TAG, SET fields are
(a) 20, 7 (b) 19, 8
(c) 20, 8 (d) 21, 9

Q.7 You have an L1 data cache; L2 cache, and main memory. The hit rates for each are:

50% hit rate, 2 cycle hit time to L1
70% hit rate, 15 cycle hit time to L2
100% hit rate, 200 cycle hit time to main memory.

What fraction of access are required from L2 from main memory respectively?
(a) 35% to 15% (b) 15% and 35%
(c) 21% and 15% (d) None of these

Q.8 Let access times of cache is 10^{-8}. Time require for block access is 10^{-6} and hit ratio is 0.9000. The average time for the CPU to access a word in two level memory is given by
(a) 0.11×10^{-6} (b) 0.91×10^{-6}
(c) 1.009×10^{-6} (d) 1.001×10^{-6}

Q.9 The percentage of times that a page number is found in the associative register is called the hit ratio.
Assume it takes 50 nanoseconds to search the associative registers and 100 nanoseconds to access main memory. What is the percentage of slowdown in memory-access time if the hit ratio is 90%?
(a) 24 (b) 40
(c) 45 (d) 60

Q.10 In a two level memory hierarchy, the access time of the cache memory is 12 nsec and theaccess time of the main memory is 1.5 msec. The hit ratio is 0.98. What is the average access time of the two level memory system?

(a) 13.5 nsec (b) 42 nsec
(c) 7.56 nsec (d) 41.76 nsec

Q.11 Given below are some statements associated with cache memory. Identify the correct statement.

(a) The Level 1 cache is always faster than the Level 2 cache.
(b) The Level 2 cache is used to mitigate the dynamic slowdown every time a Level 1 cache miss occurs.
(c) Level 2 cache comes as on board only.
(d) In modern day computers, the Level 2 cache is considered an internal cache.

Q.12 Consider a small way set associative mapping with a total of 4 blocks. LRU replacement policy is used for choosing the block to be replaced. The number of cache misses for the following sequence of block addresses 8, 12, 0, 12, 8 is _____.

Q.13 Suppose that a direct mapping cache has 2^9 lines with 2^4 bytes per cache line. If cache items of a byte addressable memory space of 2^{29} bytes. How many bits of space will be required for storing tags (do not include bits for validity or other tags; only consider the cost of the tags themselves) [in bits].

(a) 2^8 bits (b) 2^{11} bits
(c) 2^{12} bits (d) 2^{13} bits

Q.14 Two control signals in microprocessor which are related to Direct Memory Access (DMA) are

(a) HOLD & HLDA (b) S0 and S1
(c) RD and WR (d) INTR & INTA

Q.15 Let a cache has 128 lines each with 16 words and is 2-way set associative. Suppose the main memory has 16 bit address. Both cache and memory is word addressable. How many bits are in the "tag" field in the cache?

(a) 6 (b) 8
(c) 5 (d) 7

Q.16 Consider three processors with three cache configurations:

Processor 1 : Direct mapped *i*-cached and *d*-cache with one-word blocks
Instruction miss-rate = 4%, data miss-rate = 6%

Processor 2: Direct mapped *i*-cache and *d*-cache with four-word blocks
Instruction miss rate = 2%, data miss-rate = 4%

Processor 3: Two-way set associative *i*-cache and d-cache with four word blocks.
Instruction miss rate = 2%, data miss-rate = 3%
for these processors, 50% of the instructions contain a data reference. Assume that the cache penalty is 6 + blocks size in words. Determine which processor spends the most cycles on cache misses (*i*-cache refers to instruction cache and *d*-cache refers to data cache).

(a) Processor 1 (b) Processor 2
(c) Processor 3 (d) None of these

Q.17 Assume that we are having a computer with the following characteristics: 1 MB of main memory; word size of 1 byte; block size of 16 bytes; and cache size of 64 kB. For the main memory addresses of FF010, the corresponding set number for a four-way set-associative cache is _____.

Q.18 In a two-level virtual memory, the memory access time for main memory, $t_{A_1} = 10^{-8}$ sec, and the memory access time for the secondary memory, tag = 10^{-3} sec. What must be the hit ratio, H such that the access efficiency is within 80 percent of its maximum value.

Q.19 Consider a small 2-way set-associative memory, consisting of four blocks. For choosing the block to be replaced, use the least recently (LRU) scheme. The number of cache misses for the following sequence of block addresses is 8, 12, 0, 12, 8

(a) 2 (b) 3
(c) 4 (d) 5

Q.20 Consider the following fragment of code:
for ($i = 0$; $i <= 100$; i ++)
{$A[i] = B[i] + C$}
Assume A and B are arrays of 64-bit integers and C and i are 64-bit integers. Assume that all data values and there addresses are kept in memory (at addresses 0, 5000, 1500 and 2000

Answers Memory Hierarchy Design

1.	(a)	2.	(b)	3.	(b)	4.	(a)	5.	(126)	6.	(b)	7.	(a)	8.	(a)	9.	(d)
10.	(d)	11.	(b)	12.	(4)	13.	(d)	14.	(a)	15.	(a)	16.	(a)	17.	(769)	18.	(20)
19.	(c)	20.	(a)	21.	(b)	22.	(768)	23.	(d)	24.	(c)	25.	(a)	26.	(a)	27.	(a)
28.	(b)	29.	(200)	30.	(b)	31.	(248)	32.	(a)	33.	(a)	34.	(a)	35.	(c)	36.	(b)
37.	(b)	38.	(d)	39.	(b)	40.	(b)	41.	(a)	42.	(b)	43.	(d)	44.	(a)	45.	(d)
46.	(a)	47.	(a)	48.	(a)	49.	(7)	50.	(8)	51.	(d)	52.	(b)	53.	(a)	54.	(c)
55.	(b)	56.	(b)	57.	(c)	58.	(a)	59.	(a)	60.	(d)	61.	(c)	62.	(b)	63.	(c)
64.	(76)	65.	(b)	66.	(a)	67.	(a)	68.	(c)	69.	(c)	70.	(b)	71.	(c)	72.	(1.411)
73.	(d)	74.	(a)	75.	(c)	76.	(d)	77.	(c)	78.	(b)	79.	(a)	80.	(c)	81.	(a)
82.	(d)	83.	(a)	84.	(d)	85.	(480026)	86.	(d)	87.	(b)	88.	(b)	89.	(1.2)	90.	(d)
91.	(b, c)	92.	(a, b, c, d)	93.	(a, b, d)												

Explanations Memory Hierarchy Design

1. (a)
Spatial locality is the one which is to exploit.

3. (b)

$T_c = 20$ ns H : Hit ratio

$T_m = 150$ ns

$T_{avg} = T_C \times H + (1 - H) \times (T_C + T_M)$

or, $30 = 20 \times H + (1 - H) \times 170$

$\therefore \quad H \equiv 93\%$

4. (a)

$T_1 = 100\ \mu s$, $T_2 = 90\ \mu s$, $H = 95\%$

$\therefore$ Access efficiency $= \dfrac{T_1}{T_S}$

$$= \frac{1}{1 + (1 - H)\dfrac{T_2}{T_1}} = \frac{1}{1 + (1 - 0.95)\dfrac{90\ \mu s}{100\ \mu s}}$$

$$= \frac{1}{1 + 0.05 \times 0.9} = 0.9569$$

5. (126)

AMAT = $0.8 \times 100 + 0.2(0.9 \times 200) + (0.2)(0.1 \times 500)$
$= 80 + 36 + 10 = 126$

6. (b)

Offset field (block size) = 8 words × Size of each word
$= 8 \times 4$ bytes $= 32$ bytes

$$\text{Number of blocks} = \frac{\text{Size of cache}}{\text{Block size}}$$

$$= \frac{16\ \text{kB}}{32\text{B}} = 512$$

$$\text{Number of sets} = \frac{512}{2} = 256$$

Bits required for set field $= \log_2(256) = 8$ bits

Tag bits $= 32 - 5 - 8 = 19$ bits

7. (a)

Fraction serviced from 12 = (Miss in L1) × (Hit in L2) $= 0.5 \times 0.7 = 35\%$

Fraction serviced from memory = (L1 miss) × (2 Miss) × Hit in memory
$= 0.5 \times 0.3 = 15\%$

So, option (a) is correct.

8. (a)

$t_A = t_{A_1} + (1 - H)t_B$
$= 10^{-8} + (1 - 0.9000)10^{-6}$
$t_A = 0.11 \times 10^{-6}$

t_{A_1} = access time of $M1$ or cache is 10^{-8}

t_B = Time required for the block transfer is called the block access or block transfer time

$t_B = 10^{-6}$

Hit ratio $= H = 0.9000$

By using above formula

$t_A = 0.11 \times 10^{-6}$

9. (d)

Effective access time = (hit ratio)*(memory access time+ search time in associative registers) + (fail ratio) * (search time in associative registers = 2* memory access time)

= (0.90 * 150) + (0.10 * 250) =160 nanoseconds.

Now percentage slowdown = 160 – 100 = 60 ns

10. (d)

Average access time = Hit Ratio × Cache access time + Miss Ratio × Memory access time

Here, Memory access time

= 1.5 μs i.e. 1500 ns

So, average access time

= 0.98 × 12 ns + 0.02 × 1500 ns

= 11.76 ns + 30 ns = 41.76 ns

11. (b)

The level-1 cache may be of same speed as level-2 the level-2 cache used to mitigate the dynamic slow down every time a level-1 cache miss occur level-1 cache comes on processor chip while level-2 cache comes on external chip.

12. (4)

Given block addresses are 8, 12, 0, 12, 8 and initially cache is empty.

8 mod 2 = 0

12 mod 2 = 0 miss

0 mod 2 = 0 miss

12 is inside the cache hit.

8 mod 2 = 0 cache miss

Line 0	8, 0, 8	Set 0
Line 1	12	
Line 2		Set 1
Line 3		

Total, 4 misses and 1 hit.

So, answer is 4.

13. (d)

Number of memory blocks which can be mapped to a given cache line

$$= \frac{\text{Total number of blocks}}{\text{Number of cachelines}}$$

(As direct mapping is used)

$$= \frac{2^{25}}{2^{9}} = 2^{16}$$

So, we need 16 tag bits for each cache line.

So, total tag storage required = $2^9 \times 16$

= 2^{13} bits

So, option (d) correct.

14. (a)

The DMA controller sends a hold request line to the CPU and waits for the CPU to assert the HLDA.

15. (a)

Tag(6 bits)	Set(6 bits)	Word(4 bits)

16. (a)

For processor 1:

Miss penalty = 6 + 1 = 7 cycles

Stall cycles per instruction

= 4% × 7 + 50% × 6% × 7

= 0.28 + 0.21 = 0.49

For processor 2:

Miss penalty = 6 + 4 = 7 cycles

Stall cycles per instruction

= 4% × 7 + 50% × 4% × 10

= 0.2 + 0.2 = 0.4

For processor 3:

Miss penalty = 6 + 4 = 10 cycles

Stalls cycles per instruction

= 2% × 10 + 50% × 3% × 10

= 0.2 + 0.15 = 0.35

Therefore, processor 1 spends the most cycles on cache misses.

So, option (a) is correct.

17. (769)

Given that the cache size = 64 kB

One word = 1 Byte and each block contains

= 16 bytes ⇒ 16 words

Number of cache lines = 64 kB/16 B

= 4 K = 4096

$$\text{Number of cache sets} = \frac{4096}{4} = 1024$$

Size of main memory = 1 MB

Total number of address bits are = 20(2^{20} = 1 M)

TAG	Set index	Offset
6(14-19)	10(4-13)	4(0-3)

Given memory address = FF010

⇒ 1111 1111 0000 0001 0000

Offset = 0000

Set index = 11 0000 0001

TAG Bits are = 111111

Set index is = 11 0000 0001

⇒ 769 (mod 1024) ⇒ 769th set