30 YEARS
PREVIOUS SOLVED PAPERS

GATE 2020

COMPUTER SCIENCE & INFORMATION TECHNOLOGY

(Fully Solved with Explanations)

By Team of Engineers Academy

ENGINEERS ACADEMY®
Your GATEway to Professional Excellence
IES • GATE • PSUs • JTO • IAS • NET

E info@engineersacademy.org W www.engineersacademy.org
National Board Helpline Number : + 91 809 444 1777

CONTENTS

CACHE AND MAIN MEMORY

OBJECTIVE QUESTIONS

NOTES

1. A block-set associative cache memory consists of 128 blocks divided into four block sets. The main memory consists of 16,384 blocks and each block contains 256 eight bit words.

(i) How many bits are required for addressing the main memory?

(ii) How many bits are needed to represent the TAG, SET and WORD fields?

[GATE-1990 : 2 Marks]

2. In a two-level virtual memory, the memory access time for main memory, $t_{A1} = 10^{-8}$ 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.

[GATE-1990 : 2 Marks]

3. State whether the following statements are TRUE or FALSE with reason. Transferring data in blocks from the main memory to the cache memory enables an interleaved main memory unit to operate unit at its maximum speed.

[GATE-1990 : 2 Marks]

4. The total size of address space in a virtural memory system is limited by

(a) the length of MAR

(b) the available secondary storage

(c) the available main memory

(d) all of the above

[GATE-1991 : 2 Marks]

5. The principle of locality justifies the use of

(a) Interrupts (b) DMA

(c) Polling (d) Cache Memory

[GATE-1995 : 1Marks]

6. The main memory of a computer has 2 cm blocks while the cache has 2 c blocks. If the cache uses the set associative mapping scheme with 2 blocks per set, then block k of the main memory maps to the set

(a) (k mod m) of the cache

(b) (k mod c) of the cache

(c) (k mod 2 c) of the cache

(d) (k mod 2 cm) of the cache

[GATE-1999 : 1 Mark]

7. Which of the following is/are advantage of virtual memory?

(a) Faster access to memory on an average.

(b) Processes can be given protected address spaces.

(c) Linker can assign addresses independent of where the program will be loaded in physical memory.

(d) Programs larger then the physical memory size can be run.

[GATE-1999 : 2 Marks]

8. A graphics card has on board memory of 1 MB. Which of the following modes can the card not support?

(a) 1600 × 400 resolution with 256 colours on a 17 inch monitor

(b) 1600 × 400 resolution with 16 million colours on a 14 inch monitor.

(c) 800 × 400 resolution with 16 million colours on a 17 inch monitor.

(d) 800 × 800 resolution with 256 colours on a 14 inch monitor.

[GATE-2000 : 2 Marks]

9. Which of the following requires a device driver?

(a) Register (b) Cache

(c) Main memory (d) Disk

[GATE-2001 : 1 Mark]

10. More than one word are put in one cache block to

(a) exploit the temporal locality of reference in a program

(b) exploit the spatial locality of reference in a program

(c) reduce the miss penalty

(d) None of the above

[GATE-2001 : 1 Mark]

NOTES

NOTES

11. Which of the following statements is false?

(a) Virtual memory implements the translation of a program's address space into physical memory address space

(b) Virtual memory allows each program to exceed the size of the primary memory

(c) Virtual memory increases the degree of multiprogramming

(d) Virtual memory reduces the context switching overhead

[GATE-2001 : 1 Mark]

12. Consider a system with 2 level cache. Access times of Level 1 cache, Level 2 cache and main memory are 1 ns, 10 ns, and 500 ns, respectively. The hit rates of Level 1 and Level 2 caches are 0.8 and 0.9 respectively. What is the average access time of the system ignoring the search time within the cache?

(a) 13.0 ns (b) 12.8 ns (c) 12.6 ns (d) 12.4 ns

[GATE-2004 : 1 Mark]

13. Consider a small two-way set-associative cache memory, consisting of 4 blocks. For choosing the block to be replaced, use the least recently used (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

[GATE-2004 : 2 Marks]

14. A dynamic RAM has a memory cycle time of 64 nsec. It has to be refreshed 100 times per msec and each refresh takes 100 nsec. What percentage of the memory cycle time is used for refreshing?

(a) 10 (b) 6.4 (c) 1 (d) 0.64

[GATE-2005 : 1 Mark]

15. Increasing the RAM of a computer typically improves performance because

(a) Virtual memory increases

(b) Larger RAMs are faster

(c) Fewer page faults occur

(d) Fewer segmentation faults occur

[GATE-2005 : 1 Mark]

16. Consider a direct mapped cache of size 32 KB with block size 32 bytes. The CPU generates 32 bit addresses. The number of bits needed for cache indexing and the number of tag bits are respectively

(a) 10, 17 (b) 10, 22 (c) 15, 17 (d) 5, 17

[GATE-2005 : 2 Marks]

17. A cache line is 64 bytes. The main memory has latency 32ns and bandwidth 1 GB/s. The time required to fetch the entire cache line from the main memory is

(a) 32 ns (b) 64 ns (c) 96 ns (d) 128 ns

[GATE-2006 : 2 Marks]

18. A computer system has a level-1 instruction cache (1-cache), a level-1 data cache (D-cache) and a level-2 cache (L2-cache) with the following specifications:

	Capacity	Mapping Method	Block Size
1-cache	4K words	Direct mapping	4 Words
D-cache	4K words	2-way set-associative mapping	4 Words
L2-cache	64K words	4-way set-associative mapping	16 Words

Capacity Mapping method Block size 1-cache 4K words Direct mapping 4 Words D-cache 4 K words 2-way set-associative mapping 4 Words L2-cache 64K words 4-way set-associative mapping 16 Words The length of the physical address of a word in the main memory is 30 bits. The capacity of the tag memory in the I-cache, D-cache and L2-cache is, respectively,

(a) 1K × 18-bit, 1 K × 19-bit, 4K × 16-bit

(b) 1K × 16-bit, 1 K × 19-bit, 4K × 18-bit

(c) 1K × 16-bit, 512 × 18-bit, 1K × 16-bit

(d) 1K × 16-bit, 512 × 18-bit, 1K × 18-bit

[GATE-2006 : 2 Marks]

19. A CPU has a cache with block size 64 bytes. The main memory has k banks, each bank being c bytes wide. Consecutive c-byte chunks are mapped on consecutive banks with warp-around. All the k banks can be accessed in parallel, but two accesses to the same bank must be serialized. A cache block access may involve multiple iterations of parallel bank accesses depending on the amount of data obtained by accessing all the k banks in parallel. Each iteration requires decoding the bank numbers to be accessed in parallel and this takes k/2ns. The latency of one bank access is 80ns. If c = 2 and k = 24, then latency of retrieving a cache block starting at address zero from main memory is

(a) 92 ns (b) 104 ns (c) 172 ns (d) 184 ns

[GATE-2006 : 2 Marks]

Common Data for Q.20 & 21

Consider two cache organization : The first one is 32 KB 2-way set associative with 32-byte block size. The second one is of the same size but direct mapped. The size of an address is 32 bits in both cases. A 2-to-1 multiplexer has latency of 0.6 ns while a k-bit comparator has a latency of k/10 ns. The hit latency of the set associative organization is h_1 while that of the direct mapped one is h_2.

20. The value of h_1 is

(a) 2.4 ns (b) 2.3 ns (c) 1.8 ns (d) 1.7 ns

[GATE-2006 : 2 Marks]

NOTES

21. The value of h_2 is

(a) 2.4 ns (b) 2.3 ns (c) 1.8 ns (d) 1.7 ns

[GATE-2006 : 2 Marks]

Common Data for Q.22 & 23

A CPU has a 32 KB direct mapped cache with 128-byte block size. Suppose A is a two dimensional array of size 512 × 512 with elements that occupy 8-bytes each. Consider the following two C code segments. P_1 and P_3.

```
P1 :    for (i = 0; i < 512; i ++) {
            for (j = 0; j < 512; j ++) {
                x + = A [i] [j];
            }
        }
P2 :    for (i = 0; i < 512; i++) {
            for (j = 0 ; j < 512; j++) {
                x + = A [j] [i];
            }
        }
```

P1 and P2 are executed independently with the same initial state, namely, the array A is not in the cache and i, j, x are in registers. Let the number of cache misses experienced by P1 be M_1 and that for P2 be M_2.

22. The value of M_1 is

(a) 0 (b) 2048 (c) 16384 (d) 262144

[GATE-2006 : 2 Marks]

23. The value of the ratio M_1/M_2 is

(a) 0 (b) 1/16 (c) 1/8 (d) 16

[GATE-2006 : 2 Marks]

24. Consider a 4-way set associative cache consisting of 128 lines with a line size of 64 words. The CPU generates a 20-bit address of a word in main memory. The number of bits in the TAG, LINE and WORD fields are respectively

(a) 9, 6, 5 (b) 7, 7, 6 (c) 7, 5, 8 (d) 9, 5, 6

[GATE-2007 : 1 Mark]

Common Data for Q.25 & 26

Consider a machine with a byte addressable main memory of 2^{16} bytes. Assume that a direct mapped data cache consisting of 32 lines of 64 bytes each is used in the system. A 50 × 50 two-dimensional array of bytes is stored in the main memory starting from memory location 1100 H. Assume that the data cache is initially empty. The complete array is accessed twice. Assume that the contents of the data cache do not change in between the two accesses.

NOTES

25. How many data cache misses will occur in total?

(a) 48 (b) 50 (c) 56 (d) 59

[GATE-2007 : 2 Marks]

26. Which of the following lines of the data cache will be replaced by new block in accessing the array

(a) line 4 to line 11 (b) line 4 to line 12

(c) line 0 to line 7 (d) line 0 to line 8

[GATE-2007 : 2 Marks]

Common Data for Q.27 & 28

Consider a computer with a 4-way set-associative mapped cache of the following characteristics; a total of 1 MB of main memory, a word size of 1 byte, a block size of 128 words and a cache size of 8 KB.

27. The number of bits in the TAG, SET and WORD fields, respectively are:

(a) 7, 6, 7 (b) 8, 5, 7 (c) 8, 6, 6 (d) 9, 4, 7

[GATE-2008 : 2 Marks]

28. While accessing the memory location 0C795H by the CPU, the contents of the TAG field of the corresponding cache line is

(a) 000011000 (b) 110001111

(c) 00011000 (d) 110010101

[GATE-2008 : 2 Marks]

29. For inclusion to hold between two cache level L1 and L2 in a multilevel cache hierarchy, which of the following are necessary?

1. L1 must be a write-through cache
2. L2 must be write-through cache
3. The associativity of L2 must be greater that of L1
4. The L2 cache must be at least as large as the L1 cache

(a) 4 only (b) 1 and 2 only

(c) 1, 2 and 4 only (d) 1, 2, 3 and 4

[GATE-2008 : 2 Marks]

30. In an instruction execution pipeline, the earliest that the data TLB (Translation Look a side Buffer) can be accessed is

(a) before effective address calculation has started

(b) during effective address calculation

(c) after effective address calculation has completed

(d) after data cache lookup has completed

[GATE-2008 : 2 Marks]

NOTES

NOTES

Common Data for Q. 31 & 32

Consider a machine a 2-way set associative data cache of size 64 Kbytes and block size 16 bytes. The cache is managed using 32 bit virtual addressed and the page size is 4 Kbytes. A program to be run on this machine begins as follows:

```
double APR [1024] [1024]
int i, j ;
/ * Initalize array APR to 0.0 */
for (i = 0; i < 1024; i ++)
    for (j = 0; j < 1024; j ++)
            APR [i] [j] = 0.0;
```

The size of double 8 bytes. Array APR is in memory starting at the beginning of virtual page 0×FF000 and stored in row major order. The cache is initially empty and no pre-fetching is done. The only data memory references made by the program are those to array APR.

31. The total size of the tags in the cache directory is

(a) 32 kbits (b) 34 kbits (c) 64 kbits (d) 68 kbits

[GATE-2008 : 2 Marks]

32. Which of the following array elements has the same cache index as APR[0] [0]?

(a) APR[0][4] (b) APR[4][0] (c) APR[0][5] (d) APR[5][0]

[GATE-2008 : 2 Marks]

33. The cache hit ratio for this initialization loop is

(a) 0% (b) 25%

(c) 50% (d) 75%

[GATE-2008 : 2 Marks]

34. How many 32K × 1 RAM chips are needed to provide a memory capacity of 256 K-bytes?

(a) 8 (b) 32 (c) 64 (d) 128

[GATE-2009 : 1 Mark]

35. Consider a 4-way set associative cache (initially empty) with total 16 cache blocks. The main memory consists of 256 blocks and the request for memory blocks is in the following order:

0, 255, 1, 4, 3, 8, 133, 159, 216, 129, 63, 8, 48, 32, 73, 92, 155

Which one of the following memory block will NOT be in cache if LRU replacement policy is used?

(a) 3 (b) 8 (c) 129 (d) 216

[GATE-2009 : 2 Marks]

NOTES

36. A main memory unit with a capacity of 4 megabytes is built using 1M × 1-bit DRAM chips. Each DRAM chip has 1K rows of cells with 1K cells in each row. The time taken for a single refresh operation is 100 nanoseconds. The time required to perform one refresh operation on all the cells in the memory unit is

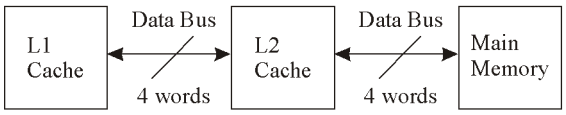
(a) 100 nanoseconds (b) $100*2^{10}$ nanoseconds

(c) $100*2^{20}$ nanoseconds (d) $3200*2^{20}$ nanoseconds

[GATE-2010 : 1 Mark]

Common Data for Q.37 & 38

A computer system has an L1 and L2 cache, an L2 cache, and a main memory unit connected as shown below. The block size in L1 cache is 4 words. The block size in L2 cache is 16 words. The memory access times are 2 nanoseconds, 20 nanosecond and 200 nanoseconds for L1 cache, L2 cache and main memory unit respectively.



37. When there is a miss in L1 cache and a hit in L2 cache, a block is transferred from L2 cache to L1 cache. What is the time taken for this transfer?

(a) 2 nanoseconds (b) 20 nanoseconds

(c) 22 nanoseconds (d) 88 nanoseconds

[GATE-2010 : 2 Marks]

38. When there is a miss in both L1 cache and L2 cache, first a block is transferred from main memory to L2 cache, and then a block is transferred form L2 cache to L1 cache. What is the total time taken for these transfers?

(a) 222 nanoseconds (b) 888 nanoseconds

(c) 902 nanoseconds (d) 968 nanoseconds

[GATE-2010 : 2 Marks]

39. An 8KB direct-mapped write back cache is organized as multiple blocks, each of size 32 bytes. The processor generates 32 bit addresses. The cache controller maintains the tag information for each cache block comprising of the following.

1 Valid bit

1 Modified bit

As many bits as the minimum needed to identify the memory block mapped in the cache.

What is the total size of memory needed at the cache controller to store meta-data (tags) for the cache?

(a) 4864 bits (b) 6144 bits (c) 6656 bits (d) 5376 bits

[GATE-2011 : 2 Marks]

NOTES

Linked Answer Q.40 & 41

A computer has a 256 Kbyte, 4-way set associative, write back data cache with block size of 32 Bytes. The processor sends 32 bit address to the cache controller. Each cache tag directory entry contains, in addition to address tag, 2 valid bits, 1 modified bit and 2 replacement bit.

40. The number of bits in the tag field of an address is

(a) 11 (b) 14 (c) 16 (d) 27

[GATE-2012 : 2 Marks]

41. The size of the cache tag directory is

(a) 160 Kbits (b) 136 Kbits (c) 40 Kbits (d) 32 Kbits

[GATE-2012 : 2 Marks]

42. In a k-way set associative cache, the cache is divided into v sets, each of which consists of k lines. The line of a set are placed in sequence on after another. The lines in set s are sequenced before the lines in set (s + 1). The main memory blocks numbered 0 onwards. The main memory block numbered 'j' must be mapped to any one of the cache lines from

(a) (j mod v) * k to (j mod v) * k + (k – 1)

(b) (j mod v) * (j mod v) + (k – 1)

(c) (j mod k) to (j mod k) + (v – 1)

(d) (j mod k) * v (j mod k) * v + (v – 1)

[GATE-2013 : 1 Mark]

43. A RAM chip has capacity of 1024 words of 8 bits each (1K × 8). The number of 2 × 4 decoders with enable line needed to construct a 16 K × 16 RAM from 1K × 8RAM is

(a) 4 (b) 5 (c) 6 (d) 7

[GATE-2013 : 2 Marks]

44. An access sequence of cache block address of length N and contains *n* unique block addresses. The number of unique block addresses between two consecutive accesses to the same block address is bounded above by *k*. What is the miss ratio if the access sequence is passed through a cache of associativity $A \geq k$ exercising least recently used replacement policy?

(a) n/N (b) $1/N$ (c) $1/A$ (d) k/n

[GATE-2014 : 2 Marks]

45. A 4-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 size of the physical address space is 4 GB. The number of bits for the TAG field is _____ .

[GATE-2014 : 1 Mark]

46. In designing a computer's cache system, the cache block (or cache line) size is an important parameter. Which one of the following statements is correct in this context?

(a) A smaller block size implies better spatial locality

(b) A smaller block size implies a smaller cache tag and hence lower cache tag overhead

(c) A smaller block size implies a larger cache tag and hence lower cache hit time

(d) A smaller block size incurs a lower cache miss penalty

[GATE-2014 : 2 Marks]

47. If the associativity of a processor cache is double while keeping the capacity and block size unchanged, which one of the following guaranteed to be NOT affected?

(a) Width of tag comparator

(b) Width of set index decoder

(c) Width of way selection multiplexor

(d) Width of processor to main memory data bus

[GATE-2014 : 2 Marks]

48. Consider a main memory system that consists of 8 memory modules attached to the system bus, which is one word wide. When a write request is made, the bus is occupied for 100 nanoseconds (ns) by the data, address, and control signals. During the same 100 ns, and for 500 ns thereafter, the addressed memory module executes one cycle accepting and storing the data. The (internal) operation of different memory modules may overlap in time, but only one request can be on the bus at any time. The maximum number of stores (of one word each) that can be initiated in 1 millisecond is _____.

[GATE-2014 (Set-2): 2 Marks]

49. The memory access time is 1 nanosecond for read operation with a hit in cache, 5 nanosecond for a read operation with a miss in cache, 2 nanosecond for a write operation with a hit in cache and 10 nanosecond for a write operation with a miss in cache. Execution of a sequence of instructions involves 100 instruction fetch operations, 60 memory operand read operations and 40 memory operand write operations. The cache hit-ratio is 0.9. The average memory access time (in nanosecond) in executing the sequence of instruction is _____ .

[GATE-2014 : 2 Marks]

50. Assume that for a certain processor, a read request takes 50 nanoseconds on a cache miss and 5 nanoseconds on a cache hit. Suppose while running a program, it was observed that 80% of the processor's read requests result in a cache hit. The average read access time in nanoseconds is _____ .

[GATE-2015 : 1 Mark]

NOTES

NOTES

51. Consider a machine with a byte addressable main memory of 2^{20} bytes, block size of 16 bytes and a direct mapped cache having 2^{12} cache lines. Let the addresses of two consecutive bytes in main memory be $(E201F)_{16}$ and $(E2020)_{16}$. What are the tag and cache line address (in hex) for main memory address $(E201F)_{16}$?

(a) E, 201 (b) F, 201 (c) E, E20 (d) 2, 01F

[GATE-2015 : 1 Mark]

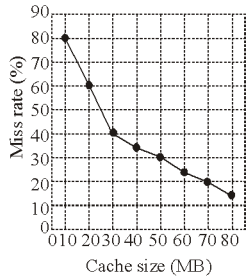
52. A processor can support a maximum memory of 4 GB, where the memory is word-addressable (a word consists of two bytes). The size of the address bus of the processor is at least ____ bits.

[GATE-2016 : 1 Mark]

53. The width of the physical address on a machine is 40 bits. The width of the tag field in a 512 KB 8-way set associative cache is _____ bits.

[GATE-2016 : 2 Marks]

54. A file system uses an in-memory cache to cache disk blocks. The miss rate of the cache is shown in the figure. The latency to read a block from the cache is 1 ms and to read a block from the disk is 10 ms. Assume that the cost of checking whether a block exists in the cache is negligible. Available cache sizes are in multiples of 10 MB.



The smallest cache size required to ensure an average read latency of less than 6 ms is ____ MB. **[GATE-2016 : 2 Marks]**

55. Consider a 2-way set associative cache with 256 blocks and uses LRU eplacement. Initially the cache in empty. conflict misses are those misses which occur due to contention of multiple blocks for the same cache set. Compulsory misses occur due to first time access to the block. The following sequence of accesses to memory blocks (0,128,256,128, 0, 128, 256, 128, 1, 129, 257, 129, 1, 129, 257, 129) is repeated 10 times. The number of conflict misses experienced by the cache is_____.

[GATE-2017 : 2 Marks]

56. A cache memory unit with capacity of N words and block size of B words is to be designed. If it is designed as direct mapped cache, the length of the TAG field is 10 bits. If the cache unit is now designed as a 16-way set-associative cache, the length of the TAG field is _____ bits.

[2 Marks : GATE-2017]

NOTES

57. Consider a two-level cache hierarchy with L_1 and L_2 caches. An application incurs 1.4 memory accesses per instruction on average. For this application, the miss rate of L_1 cache is 0.1; the L_2 cache experiences, on average, 7 misses per 1000 instructions. The miss rate of L_2 expressed correct to two decimal places is_____. **[1 Mark : GATE-2017]**

58. Consider a machine with a byte addressable main memory of 2^{32} bytes divided into blocks of size 32 bytes. Assume that a direct mapped cache having 512 cache lines is used with this machine. The size of the tag field in bits is _____. **[2 Marks : GATE-2017]**

59. The read access times and the hit ratios for different caches in a memory hierarchy are as given below.

Cache	Read access time (in nanoseconds)	Hit ratio
I – cache	2	0.8
D – cache	2	0.9
L2 – cache	8	0.9

The read access time of main memory is 90 nanoseconds. Assume that the caches use the referredword- first read policy and the write back policy. Assume that all the caches are direct mapped caches. Assume that the dirty bit is always 0 for all the blocks in the caches. In execution of a program, 60% of memory reads are for instruction fetch and 40% are for memory operand fetch. The average read access time in nanoseconds (up to 2 decimal places) is _____. **[2 Marks : GATE-2017]**

60. In a two-level cache system, the access times of L_1 and L_2 caches are 1 and 8 clock cycles, respectively. The miss penalty from the L_2 cache to main memory is 18 clock cycles. The miss rate of L_1 cache is twice that of L_2. The average memory access time (AMAT) of this cache system is 2 cycles. The miss rates of L_1 and L_2 respectively are:

(a) 0.111 and 0.056 (b) 0.056 and 0.111

(c) 0.0892 and 0.1784 (d) 0.1784 and 0.0892

[2 Marks : GATE-2017]

61. A 32 -bit wide main memory unit with a capacity of 1GB is built using 256M×4-bit DRAM chips. The number of rows of memory cells in the DRAM chip is 2^{14}. The time taken to perfome one refresh operation is 50 nanoseconds. The refresh period is 2 milliseconds. The percentage (rounded to the closest integer) of the time available for performing the memory read/ write perations in the main memory unit is _____.

[1 Mark: GATE-2018]

NOTES

62. The size of the physical address space of a processor is 2^P bytes. The word length is 2^w bytes. The capacity of cache memory is 2^N bytes, the size of each cache block is 2^M words. For a k- way set-associative cache memory, the length (in number of bits) of the tag field is

(a) $P-N-\log_2 k$

(b) $P-N+\log_2 k$

(c) $P-N-M-W-\log_2 k$

(d) $P-N-M-W+\log_2 k$

[2 Marks: GATE-2018]

63. A certain processor uses a fully associative cache of size 16 kB. The cache block size is 16 bytes. Assume that the main memory is byte addressable and uses a 32-bit address. How many bits are required for the Tag and the Index fields respectively in the addresses generated by the processor?

(a) 28 bits and 4 bits (b) 24 bits and 4 bits

(c) 24 bits and 0 bits (d) 28 bits and 0 bits

[1 Mark : GATE-2019]

64. A certain processor deploys a single - level cache. The cache block size is 8 words and the word size is 4 bytes. The memory system uses a 60-MHz clock. To service a cache miss, the memory controller first takes 1 cycle to accept the starting address of the block, it then takes 3 cycles to fetch all the eight words of the block, and finally transmits the words of the requested block at the rate of 1 word per cycle. The maximum bandwidth for the memory system when the program running on the processor issues a series of read operations is _____ $\times 10^6$ bytes/sec. **[2 Marks : GATE-2019]**

65. Assume that in a certain computer, the virtual addresses are 64 bits long and the physical addresses are 48 bits long. The memory is word addressible. The page size is 8 kB and the word size is 4 bytes. The Translation Look-aside Buffer (TLB) in the address translation path has 128 valid entries. At most how many distinct virtual addresses can be translated without any TLB miss?

(a) 4×2^{20} (b) 6×2^{10}

(c) 256×2^{10} (d) 8×2^{20}

[2 Marks : GATE-2019]

❍❍❍

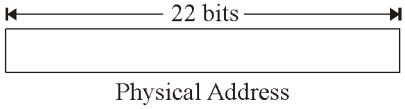
ANSWERS AND EXPLANATIONS

1. *Sol.*

(i) Main memory size

$$= 16384 \text{ blocks}$$
$$= 16384 \times 256 \text{ words}$$
$$= 2^{14} \times 2^{8} \text{ words}$$
$$= 2^{22} \text{ words}$$

Number of bits required for addressing the main memory is 22 bits.



(ii) Set associative cache

Block size = 256 word = 2^8 words

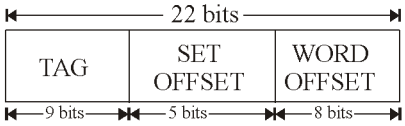
$\therefore$ No. of bits in WORD OFFSET is 8 bits

No. of blocks in set-associative cache = 128

Number of blocks in one set = 4

$$\therefore \text{ No. of sets in cache } = \frac{128}{4} = 32 = 2^5$$

Number of bits in SET OFFSET is 5.



Number of bits in TAG = 22 – (8 + 5) = 9 bits

2. Sol.

Main memory access time,

$$t_{al} = 10^{-8} \text{ sec}$$

Secondary memory access time,

$$t_{a2} = 10^{-3} \text{ sec}$$

Access efficiency,

$$\eta = 80\% = 0.8$$

Average access time,

$$t_{avg} = \eta \times t_{a2} = 0.8 \times 10^{-3} \text{ sec}$$

Hit ratio, H

$$t_{avg} = H \times t_{a1} + (1 - H)t_{a2}$$
$$0.8 \times 10^{-3} = H \times 10^{-8} + (1 - H) \times 10^{-3}$$
$$0.8 = H \times 10^{-5} + (1 - H)$$
$$H(1 - 10^{-5}) = 0.2$$
$$H = \frac{0.2}{(1-10^{-5})}$$
$$= 0.200002 \approx 20\%$$

The hit ratio H must be 20%.

3. *Sol.*

Statement is true, main memory transfers the data in the form of block to cache & cache transfer the data in the from of words, so it enable the interleaved main memory to operate unit at maximum speed.

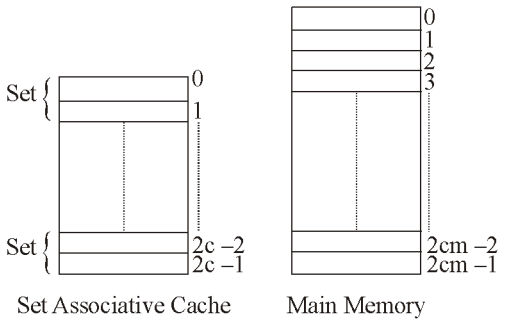
4. *Ans. (b)*

Since virtual memory io implemented or secondary storage.

5. *Ans. (d)*

The principle of locality justifies the use of cache memory.

6. *Ans. (b)*



# of set in set associative cache

$$= \frac{\text{\# of blocks in cache}}{\text{\# of blocks in one set}}$$

$$= \frac{2c}{2} = c$$

# of sets in cache = c

Therefore, the block k of the main memory maps to the set (k mod c) of the cache.

7. *Ans. (d)*

Virtual memory is a concept which allows programs larger than the physical memory size to run.

8. *Ans. (b)*

1 MB of memory supported by graphics card

(a) [1600 * 400 * $\log_2$ 256 bits] < 1 MB

(b) [1600 * 400 * $\log_2 10^6$ bits] > 1 MB

(c) [800 * 400 * $\log_2 10^6$ bits] < 1 MB

(d) [800 * 800 * $\log_2$ 256 bits] < 1 MB

$\therefore$ Option (b) can not supported by card.

9. *Ans. (d)*

Disk is the IO device attached externally to the processor. Therefore, disk requires a device driver.

10. *Ans. (b)*

There are two types of locality of reference:

1. **Temporal :** A recently executed instruction is likely to be executed again very soon.
2. **Spatial:** Instructions in close proximity to a recently executed instruction are also likely to be executed soon.

The spatial aspect suggest that instead of fetching just one item from the main memory to the cache, it is useful to fetch several items that reside at adjacent address as well.

Therefore more than one work are put in one cache block to exploit the spatial locality of reference in the program.

11. *Ans. (d)*

Techniques that automatically move program and data blocks into the physical main memory when they are required for execution are called virtual memory techniques. Programs, and hence the processor, reference an instruction and data space that is independent of the available physical main memory space. Virtual memory increases the degree of multi-programming.

However, virtual memory increases the context switching overhead.

12. *Ans. (c)*

Access time

$$= t_1h_1 + (1-h_1)h_2t_2 + (1- h_1)(1-h_2)t_m$$

$$= 1*0.8 + 0.2*0.9*10 + 0.2*0.1*500$$

$$= 0.8 + 1.8 + 10$$

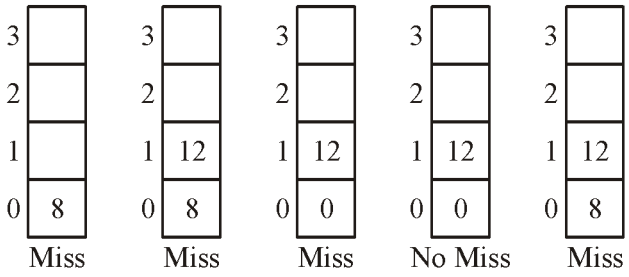
$$= 12.6 \text{ ns.}$$

13. *Ans. (c)*

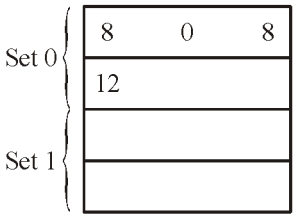
Sequence 8, 12, 0, 12, 8

Apply the LRU as follows

Page Frame



OR



Total number of miss = 4.

14. *Ans. (c)*

Time for Memory cycle time = 64nsec

Time for refresh = 100nsec

Number of refreshes in a memory cycle

So, in 64nsec = (64nsec×100 Times)/1msec

$$= 64\times10^{-4}$$

So, time for refresh in a memory cycle

$$= 100 \text{ ns}*64*10^{-4}$$

$$= 64 * 10^{-2} \text{ ns} = .64\text{nsec}$$

So, percentage of time spent for refresh

$$= (64 * 10^{-2}/64) * 100 = 1\%$$

15. *Ans. (c)*

Increasing the RAM means increase the primary memory which reduce the swapping so fewer page faults occur.

16. *Ans. (a)*

Size of cache = 32 KB

$$= 32 \times 2^{10} \text{ byte}$$

$$= 2^5 \times 2^{10} \text{ byte}$$

$$= 2^{15} \text{ byte} = 15 \text{ bits}$$

Size of tag = 32 – 15

= 17 bits

Cache indexing size = 10 bits

17. *Ans. (c)*

For 1 sec it is 10^9 bytes, for 64 bytes ?

$= 64*1/10^9$

$= 64$ ns

MM latency is 32 ns.

Total time required to place cache line

$= 64 + 32 = 96$ ns.

18. *Ans. (a)*

1. I-cache : Number of blocks in cache

$= 4\ K/4 = 2^{10}$ blocks

Bits to represent blocks = 10

No. of words in a block = $4 = 2^2$ words

Bits to represent words = 2

Tag bits = 30 – (10 + 2) = 18

Each block will have it's own tag bits. So, total tag bits = 1 K × 18 bits.

2. D-cache : Number of blocks in cache

$= 4K/4 = 2^{10}$ blocks

No. of sets in cache = $2^{10}/2 = 2^9$ sets

Bits to represent sets = 9

No. of words in a block = $4 = 2^2$ words

Bits of represents words = 2

Tag bits = 30 – (+2) = 19

Each block will have it's own tag bits. So, total tag bits = 1 K × 19 bits.

3. L2 cache : Number of blocks in cache

$= 64\ K / 16 = 2^{12}$ blocks

No. of sets in cache = $2^{12}/4 = 2^8$ sets

Bits to represent sets = 8

No. of words in cache $16 = 2^4$ words

Bits to represent words = 4

Tag bits = 30 – (8 + 4) = 18

Each block will have it's own tag bits. So, total tag bits = 2^{12} × 18 bits = 4 K × 18 bits

19. *Ans. (d)*

Cache block size = 64 bytes

Main memory has K banks or k = 24

Each bank is 2 byte long because c = 2

Total time for one parallel access

T = K/2 + Latency

= 24/2 + 80

= 12 + 80 = 92 ns

Total latency time = CT

= 2 × 92 = 184 ns

20. *Ans. (a)*

Tag	Set	Block
18	9	9

So, $h_1 = 18/10 + 0.6$ ns

$= 1.8 + 0.6$ ns $= 2.4$ ns

21. *Ans. (d)*

Tag	Set	Block
17	10	5

So, $h_1 = 17/10$ ns $= 1.7$ ns

22. *Ans. (c)*

$$\begin{bmatrix} a[0][0], a[0][1], \ldots\ldots\ldots\ldots, a[0][511] \\ a[1][0], a[1][1], \ldots\ldots\ldots\ldots, a[0][511] \\ . \\ . \\ . \\ \ldots\ldots\ldots\ldots \\ a[511][0], a[511][1], \ldots\ldots\ldots\ldots, a[511][511] \end{bmatrix}$$

One element size = 8 B

Block size = 128 B

$\therefore$ One block holds $= \frac{128}{8} = 16$ element

1 row contains = 512 elements

$\therefore$ 32 block are required to carry row

Cache memory size = 32 kB

Block size = 128 B

$$\#\ \text{lines} = \frac{32\ \text{kB}}{128\ \text{B}} = \frac{2^{15}}{2^7} = 2^8 = 256$$

P1 : In this array is accessed in row-major order like

a[0][0], a[0][1], a[0][2], ...

Accessing a[0][0] = Miss

Accessing a[0][1] = hit

Accessing a[0][2] = hit

⋮

Accessing a[0][15] = hit

Accessing a[0][16] = Miss

Accessing a[0][17] = hit

⋮

Accessing a[0][31] = hit

Accessing a[0][32] = Miss

⋮

∴ One row accessing = 32 miss operations

Total miss (M_1) = 32 × 512 = 16384

23. ***Ans. (b)***

P_2 : In this array is accessed in column-major order like.

a[0][0], a[1][0], a[2][0], ...

Accessing a[0][0] = Miss

Accessing a[1][0] = Miss

Accessing a[2][0] = Miss

⋮

Accessing a[511][0] = Miss

Accessing a[0][1] = Miss

⋮

Accessing a[511][1] = Miss

⋮

One column accessing = 512 miss operation

Total miss (M_2) = 512 * 512

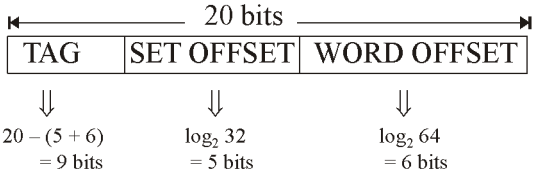
$$\therefore \quad \frac{M_1}{M_2} = \frac{16384}{512 \times 512} = \frac{1}{16}$$

24. ***Ans. (d)***

4-way set associative

lines = 128

Block size = 64 words



$$\text{Number of set} = \frac{\#\text{lines}}{4} = \frac{128}{4} = 32$$

25. ***Ans. (c)***

$$\begin{bmatrix} a[0][0], a[0][1], \ldots\ldots\ldots\ldots, a[0][49] \\ a[1][0], a[1][1], \ldots\ldots\ldots\ldots, a[1][49] \\ . \\ . \\ . \\ a[49][0], a[49][1], \ldots\ldots\ldots\ldots, a[49][49] \end{bmatrix}$$

One element size = 1 B

Total array elements = 50 × 50 = 2500

So, total array size = 2500 × 1 B = 2500 B

Block size = 64 B

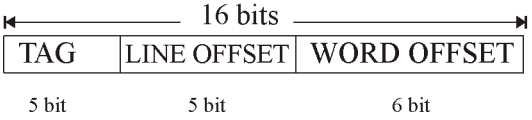
∴ # block required to store the array

$$= \frac{2500\ \text{B}}{64\ \text{B}} = 39.04 \approx 40$$

lines in cache = 32

Block size = 64 B

Physical address = 1100 H



$(1100)_H = (0001000100000000)_B$

TAG	LO	WO
00010	00100	000000

∴ $LO = (00100)_B = 4$

∴ Array is stored in the main memory starting from memory location of line 4.

26. ***Ans. (a)***

The lines of the data cache that is replaced by new blocks in accessing the array = line 4 to line 11.

27. ***Ans. (d)***

Cache memory size = 8 KB

Block size = 128 word × 1 byte = 128 byte

$$\text{No. of block in cache} = \frac{2^{13}\ \text{B}}{2^7\ \text{B}} = 2^6$$

$$\text{No. of size} = \frac{\text{No. of block}}{\text{No. of ways}}$$

$$= \frac{2^6}{2^2} = 2^4$$

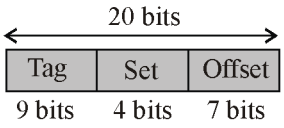
Tag bits = 20 –(4 + 7)

= 20 – 11 = 9

So, Tag set and word fields are 9, 4, 7

28. ***Ans. (a)***

We have 16 sets in cache and correspondingly 16 regions in physical memory to which each set is mapped.



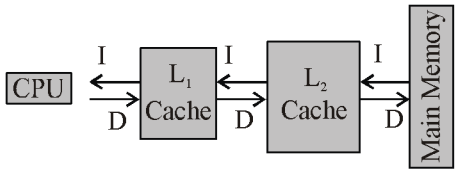
WORD bit size is 7 as we need 7 bits to address 128 possible words in a cache block.

So, the lowest 7 bits of 0C795H will be used for this giving us the remaining bits as 0000 1100 0111 1 of these bits, the lower 4 are used for addressing the 16 possible sets so tab

bits: 000011000.

29. ***Ans. (b)***

Consider the following diagram :



L_1 and L_2 both are write through caches. This two conditions are necessary.

30. ***Ans. (b)***

In instruction execution pipeline, the earliest that the data TLB can be accessed during effective address calculation has started.

31. ***Ans. (d)***

2 way set associative.

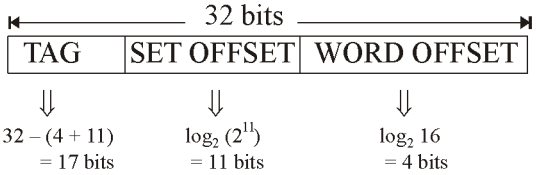
Cache memory size = 64 kB

Block size = 16 B

$$\#\ \text{lines} = \frac{64\,\text{kB}}{16\,\text{B}} = 4\,\text{K} = 2^2 \times 2^{10} = 2^{12}$$

$$\#\ \text{sets} = \frac{2^{12}}{2} = 2^{11}$$

32-bit virtual address



Tag memory size

= Number of set * No. of lines in a set * Number of tag bits

$= 2^{11} \times 2 \times 17$ bits

$= 68 \times 2^{10}$ bits

= 68 K bits

32. ***Ans. (b)***

Array is APR [1024] [1024]

$$\begin{bmatrix} a[0][0], a[0][1], \ldots\ldots\ldots\ldots, a[0][1023] \\ a[1][0], a[1][1], \ldots\ldots\ldots\ldots, a[1][1023] \\ . \\ . \\ . \\ a[1023][0], a[1023][1], \ldots\ldots\ldots\ldots, a[1023][1023] \end{bmatrix}$$

1 element size = 8 B

Block size = 16 B

$\therefore$ One block holds 2 elements

One row size = 1024 element

$\therefore$ No. of blocks required to store one row

= 512 blocks

APR [0][0] has the same cache index as APR [4][0]

33. ***Ans. (c)***

Row major access

a[0][0] = Miss

a[0][1] = hit

a[0][2] = Miss

a[0][3] = hit

.

.

.

Every miss is followed by a hit

$$\therefore \text{Hit ratio} = \frac{1024}{1024+1024} = \frac{1024}{2048} = \frac{1}{2} = 50\%$$

34. ***Ans. (c)***

Basic RAM = 32 K × 1

Design RAM = 256 K × 8

$$\text{No. of RAM} = \frac{256\,\text{K} \times 8}{32\,\text{K} \times 1}$$

$$= \frac{2^{18} \times 2^3}{2^{15} \times 2^0} = 64$$

It require 8 parallel line and in each parallel line 8 serial RAM chip are required.

35. ***Ans. (d)***

There are total 4 sets in the cache and each set contains 4 blocks.

0 mod 4 = 0set 0

255 mod 4 = 3set 3

1 mod 4 = 1set 1

4 mod 4 = 0set 0

3 mod 4 = 3set 3

8 mod 4 = 0set 0

133 mod 4 = 1set 1

159 mod 4 = 3set 3

216 mod 4 = 0set 0

129 mod 4 = 1set 1

63 mod 4 = 3set 3

8 mod 4 = 0set 0 (already in cache)

48 mod 4 = 0set 0 48 will replace block 0 using LRU

32 mod 4 = 0set 0 32 will replace block 4

73 mod 4 = 0set 1

92 mod 4 = 0set 0 92 will replace block 216

155 mod 4 = 3set 3 155 will replace 255

36. ***Ans. (d)***

Main memory unit has capacity = 4 MB

$$\text{No. of DRAM chips} = \frac{4\,\text{MB}}{1\text{M} \times 1\,\text{bit}}$$

$$= \frac{4\,\text{MB} \times 8\,\text{bit}}{1\text{M} \times 1\,\text{bit}} = 32$$

1 DRAM chip has 1K (2^{10}) Rows and 1 K (2^{10}) cells.

In each row

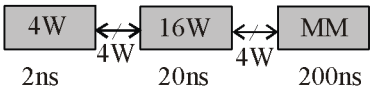
$$= 32 \times 2^{10} \times 2^{10} = 32 \times 2^{20}$$

Time take in one refresh operation = 100 ns

Time required to perform one refresh operation on all the cells in MM unit

$= 32 \times 2^{20} \times 100 = 3200 \times 2^{20}$ ns

37. ***Ans. (d)***



Memory access time to access L1=2 nanoseconds

Memory access time to access L2 = 20 nanoseconds

When there is a miss in L1 cache and hit in L2 cache so we have to access first L1 then after it L2 and the data bus size is 4 W. Now the access time is calculated as

2 Access time of L1 cache
20 Access time of L2 cache
2 Access time of L1 cache
20 Access time of L2 cache
2 Access time of L1 cache
20 Access time of L2 cache
2 Access time of L1 cache
20 Access time of L2 cache

88ns

38. ***Ans. (d)***

When there is a miss in L1 and L2 cache then block transfer for mm to L2 cache, and then block is transferred from L2 cache to L1 cache.

So, total time for these transfer

Total Access time = Block transfer time from main memory to L2 cache + access time of L2 + Access time of L1.

2 Access time of L1 cache
20 Access time of L2 cache
200 Access time of MM
20 Access time of L2 cache
200 Access time of MM
20 Access time of L2 cache
200 Access time of MM
20 Access time of L2 cache
200 Access time of MM
20 Access time of L2 cache
2 Access time of L1 cache
20 Access time of L2 cache
2 Access time of L1 cache
20 Access time of L2 cache
2 Access time of L1 cache
20 Access time of L2 cache

968 ns

39. *Ans. (d)*

Cache memory size = 8 KB

Block size = 32 bytes

Mapping technique is Direct.

Physical address size is 32 bits

cache lines = 2^{13} / 2^5 = 2^8 = 256 lines

Memory address interpretation in direct mapping is :

Modified bit	Valid bit	Tag	Line offset	Word offset

Word offset requires = 5 bits

Line offset requires = 8 bits

Tag requires = 32bits–(8+5) bits=19 bits

the metadata present in cache memory is = data memory + tag size

Tag size = No of cache lines* No of bits in tag field for any maping.

Hence the total size of memory needed at the cache controller to store metadata tags for the write back cache is :

= Number of cache lines*(tag bits+valid bit + modified bit)

256*(19+1+1)=256*21

= 5376 bits.

40. *Ans. (c)*

$$\text{No. of cache lines} = \frac{2^8 \times 2^{10}}{2^5} = 2^{13}$$

$$\text{No. of sets} = \frac{2^{18}}{2^2} = 2^{11}$$

16	11	5
Tags	So	Word offset

No. of tag bits are 16 bits

∴ option 'C' is correct.

41. *Ans. (a)*

Tag directory size = Number of sets × no. of tag s in each sets × (no. of tag bits in each tag + valid bits + modified bit + replacement bit)

= 2^{11} × 4 × (16 + 2 + 1 + 1) = 160 KBits

∴ Option (a) is correct.

42. *Ans. (a)*

Main memory block 'j' mapped to any of cache line. Cache line range is:

(j mod v)*k → (j mod v)*k + (k – 1)

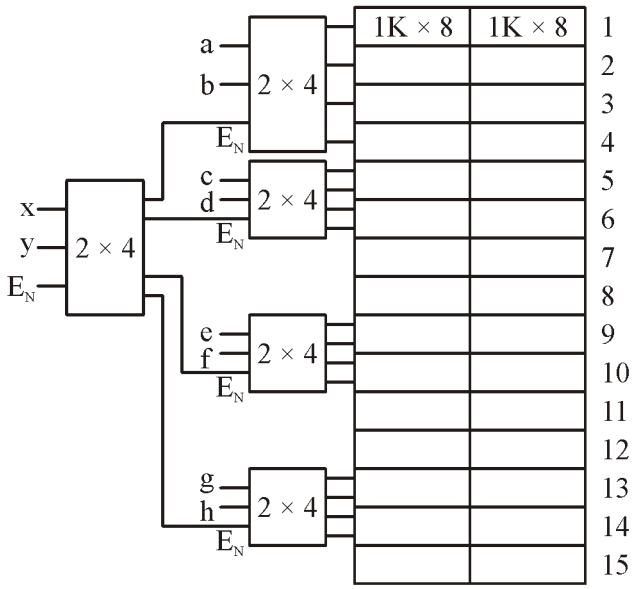
43. *Ans. (b)*

We need 16 K × 16 RAM from 1 K × 8 RAM

So, number of chips required

$$= \frac{16K \times 16}{1K \times 8} = 16 \times 2$$

So we need 16 output lines



∴ We need 5 decoders.

44. *Ans. (a)*

Miss ratio = n/N, if the access sequence is passed through a cache of associativity A ≥ k exercising least-recently used replacement policy.

45. **Sol.**

TAG	SET	WORD
20	7	5

← 32 bit →

WORD offset = 5 bits

[∵ word length = 32 length]

SET offset = 7 bits

$$[\because \#\text{blocks} = \frac{16\text{ kB}}{8\text{ Words}} = 512 \text{ blocks}$$

#sets = 512/4 = 128]

∴ #TAG bits = 32 – (7 + 5) = 20 bits