Question: A system with 10 MHz clock uses a two level cache. The first level cache (L1) is direct–mapped, with 4KB capacity and 8 Byte blocks. The second level cache (L2) is a four way set associative, with 4MB of total capacity and 16-Byte blocks. L1 cache has a miss rate of 5%, whereas the L2 cache has a miss rate of 10%. L1 hits cause no stalls with access time of 1 clock cycle. The L2 hit time is 10 cycles. Main memory access time is 100 cycles. There is a 64-bit bus between memory and L2 cache, and between L1 and L2.

a. How many bits are used for line/set address for each cache?

b. What is the average memory access time for any instruction?

Solutions

a.

L1 cache capacity = 4KB
Line size of L1 = 8 B
No. of Lines = 4 KB/ 8B= 2^9
No of bits for Line Address = 9
L2 cache capacity = 4 MB
Line Size of L2 4-way set associative cache= 16 B
No. of Sets in L2 cache = 4 MB/ (4*16 B) = 2^{16}
No. of bits for set address = 16

b.

Hit Ratio L1 = 0.95
Hit Ratio L2 = 0.90
L1 Access Time = 1 cc
L2 Access Time 10 cc
Memory Access Time = 100 cc
Average Memory Access Time = $HitRatio_{L1}$* $AccessTime_{L1}$ + (1- $HitRatio_{L1}$)* ($MissPenalty_{L1}$)
Miss Penalty $_{L1}$ = $HitRatio_{L2}$($AccessTime_{L1}$ +$AccessTime_{L2}$) + (1- $HitRatio_{L2}$) *(Miss $Penalty_{L2}$)
Miss Penalty $_{L2}$= 1cc+ 10 cc+ 2*100 cc; two transfer from Memory to L2
= 211 cc
Miss Penalty $_{L1}$ = 0.9 *11 + 0.1 * 211
= 31 cc
Average Memory Access Time= 0.95 * 1+ 0.05* 31
= 2.5 cc
= 2.5 *0.1 us
= 0.25 us