

Multicore Computer Architecture - Storage and Interconnects

Tutorial 3

Cache Memory Optimizations



NPTEL Online Course : Multi-Core Computer Architecture and Interconnects Dr. John Jose, IIT Guwahati

Dr. John Jose

Assistant Professor

**Department of Computer Science & Engineering
Indian Institute of Technology Guwahati, Assam.**

Tutorial Problem-1

❖ The address of a word in a byte addressable 16MB physical memory is 0xAA0C2A. This word upon bringing to the cache is mapped to set 48. What is the block size of the cache memory ?

❖ A A 0 C 2 A

❖ 1010 1010 0000 1100 0010 1010

❖ 1010 1010 0000 **1100 0010 1010** offset → 64bytes

Tutorial Problem-2

- ❖ A cache has access time (hit latency)=10 ns and miss rate is 5%. An optimization was made to reduce the miss rate to 3 % but the hit latency was increased to 15 ns. Under what condition this change will result in better performance (Lower avg. memory access time)?
- ❖ $AMAT\ 1 = HT1 + MR1 \times MP$ $HT1 = 10ns; MR1=0.05$
- ❖ $AMAT\ 2 = HT2 + MR2 \times MP$ $HT2 = 15ns; MR1=0.03$
- ❖ $AMAT2 < AMAT1$

Tutorial Problem-3

- ❖ A cache has hit rate of 90%, 64 byte block, cache hit latency of 5ns. Main memory takes 150 ns to return first word (32 bits) of a block and 10 ns for each subsequent word.
 - (a) What is the miss latency of the cache?
 - (b) If doubling the cache block size reduces the miss rate to 3%, does it reduce average memory access time?

Tutorial Problem-3

- ❖ A cache has hit rate of 90%, 64 byte block, cache hit latency of 5ns. Main memory takes 150 ns to return first word (32 bits) of a block and 10 ns for each subsequent word.
 - (a) What is the miss latency of the cache?
 - (b) If doubling the cache block size reduces the miss rate to 3%, does it reduce average memory access time?

Tutorial Problem-4

- ❖ For a cache, that has a miss rate of 3% and miss penalty of 500 cycles. In a program 50% of the instructions are memory accesses (load-store)
- ❖ (a) Find the misses per 1000 instruction (MPKI)
- ❖ (b) Find memory stall cycles per miss
- ❖ Miss rate: $\text{miss/mem access} = (\text{miss / instruction}) / (\text{mem acc / instruction})$

$$\text{MR} = \text{MPI} / \text{MAPI}$$

$$\text{MPI} = \text{MR} \times \text{MAPI}$$

$$\text{MAPI} = 1.5$$

Tutorial Problem-5

- ❖ Consider a cache system with miss rate of an I-cache is 2% and that of D-cache is 4%. The processor CPI=2 without memory stalls and miss penalty =100 cycles for all misses. Determine how much faster the processor would run with a perfect cache that never missed. Assume frequency of all loads and store is 36 %.
- ❖ $\text{Actual CPI}_{\text{real}} = \text{Base CPI} + \text{stall CPI}$ $\text{CPI}_{\text{ideal}} = \text{Base CPI} = 2$
- ❖ $\text{Stall CPI} = (\% \text{ use of IC} \times \text{stall of IC}) + (\% \text{ use of DC} \times \text{stall of DC})$

Tutorial Problem-5

- ❖ miss penalty =100 cycles for all misses. Assume frequency of all loads and store is 36 %.
- ❖ Actual CPI $\text{real} = \text{Base CPI} + \text{stall CPI}$ CPI $\text{ideal} = \text{Base CPI} = 2$
- ❖ Stall CPI = (% use of IC x stall of IC) + (% use of DC x stall of DC)

Tutorial Problem-6

- ❖ Consider a 32 bit processor with 16KB direct mapped L1-cache that uses a block size of 4 words. It has an L2-cache of 256 KB with 4-way associativity and block size of 8 words. The system uses a byte addressable 256 MB DRAM system. Upon running a program, 16 consecutive fixed length instructions (each instruction is one word) starting at main memory address 0x 8226620 are executed. These instructions operate on an array A of 8 words, with starting address 0x 42AF5F8 Assuming caches are initially empty; indicate the non empty sets on L1 cache and L2 cache after the execution of the program.

Tutorial Problem-6

- ❖ 32 bit processor: 1 word → 4 bytes: 256 MB DRAM → 28 bit address
- ❖ L1 Cache: 16KB, direct mapped, block size= 4 words (16B)

NPTEL Online Course : Multi-Core Computer Architecture - Storage and Interconnects Dr. John Jose, IIT Guwahati

- ❖ L2 Cache : 256 KB, 4-way, block size= 8 words (32B).
- ❖ Instruction 0x 8226620, 16 consecutive fixed length instructions (each instruction is one word) Data 0x 42AF5F8 , array of 8 words.

Tutorial Problem-6

- ❖ L1 Cache: 16KB, direct mapped, block size= 4 words (16B)
- ❖ Instruction 0x 8226620, 16 consecutive fixed length instructions (each instruction is one word) Data 0x 42AF5F8 , array of 8 words.

Tutorial Problem-6

- ❖ L2 Cache : 256 KB, 4-way, block size= 8 words (32B).
- ❖ Instruction 0x 8226620, 16 consecutive fixed length instructions (each instruction is one word) Data 0x 42AF5F8 , array of 8 words.

Tutorial Problem-6

- ❖ Non-Empty Blocks
- ❖ L1: Sets 610, 611, 612, 613 (4 words x 4 = 16 instructions)
Sets 863, 864, 865 (2 + 4 + 2 words of data array A)
- ❖ L2: Sets 817, 818 (8 words x 2 = 16 instructions)
Sets 1967, 1968 (2 + 6 words of data array A)

**THANK
YOU!**



**FOR
YOUR
PRESENCE**

NPTEL Online Course : Multi-Core Computer Architecture - Storage and Interconnects Dr. John Jose, IIT Guwahati

johnjose@iitg.ac.in
<http://www.iitg.ac.in/johnjose/>