INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR
Class Test 1 Solutions

Date: 11-04-23 **Timing:** 3 pm - 4:30 pm **Place:** NR223 **Total Marks =** 30
Subject No : CS60003 **HIGH PERFORMANCE COMPUTER ARCHITECTURE**
Department/Centre/School : Computer Science and Engineering
Answer all questions. In case of reasonable doubt, make assumptions and state them upfront. Marks will be deducted for claims without proper reasoning. Write your answers (with all analysis, justification, calculation etc) within the allotted time.

1. **[10]**

Instruction	Issue	FU / EX	MEM	CDB	Commit	ROB/RS
L.D F1, 0(R1)	1	A1 / 2-2	3	4	5	#1 / LD1
L.D F2, 0(R2)	1	A2 / 2-2	4	5	6	#2 / LD2
ADD.D F2, F1, F2	2	FA1 / 6-8	-	9	10	#3 / FAD1
MUL.D F4, F1, F2	2	FM1/10-14	-	15	16	#4 / FMUL1
ADD.D F5, F2, F3	3	FA2/10-12	-	13	16	#5 / FAD2
S.D F5, 0(R3)	3	A1 / 4-4	14	-	17	#6 / ST1

[10]

2. Let us first extract all the available information from the problem description. Here is a summary:

- **CPU:** 1.1 GHz (0.909 ns equivalent), CPI of 1.35 (excludes memory accesses)
- **Instruction mix:** 75% non-memory access instructions, 20% loads, 10% stores
- **Caches:** Split L1 with no hit penalty (i.e. the access time is the time it takes to execute the load/ store instruction)
  - *L1 I-cache:* 2% miss rate, 32-byte blocks (requires 2 bus cycles to fill), miss penalty is 15 ns + 2 cycles
  - *L1 D-cache:* 5% miss rate, write-through (no-write allocate), 95% of all writes do not stall because of a write buffer, 16-bytes blocks (requires 1 bus cycle to fill), miss penalty is 15 ns + 1 cycle
- **L1/ L2 bus:** 128-bit, 266 MHz bus between the L1 and L2 caches
- L2 (unified) cache, 512 KB, write-back (write-allocate), 80% hit rate, 50% of replaced blocks are dirty (must go to main memory), 64-byte blocks (requires 4 bus cycles to fill), miss penalty is 60 ns + 7.52 ns = 67.52 ns
- Memory, 128 bits (16 byes) wide, first access takes 60 ns, subsequent accesses take 1 cycle on 133 MHz, 128-bit bus.

(a) The average memory access time for instruction accesses:

- L1 (inst) miss time in L2: 15 ns access time plus two L2 cycles (two = 32 bytes in inst. cache line/ 16 bytes width of L2 bus) = $15 + 2 \times 3.75 = 22.5$ ns (3.75 is equivalent to one 266 MHz L2 cache cycle)
- L2 miss time in memory: 60 ns + plus four memory cycles (four = 64 bytes in L2 cache/ 16 bytes width of memory bus) = $60 + 4 \times 7.5 = 90$ ns (7.5 is equivalent to one 133 MHz memory bus cycle)
- Avg. memory access time for inst. = avg. access time in L2 cache + avg. access time in memory + avg. access time for L2 write-back = $0.02 \times 22.5 + 0.02 \times (1 - 0.8) \times 90 + 0.02 \times (1 - 0.8) \times 0.5 \times 90 = 0.99$ ns (1.09 CPU cycles)

(b) The average memory access time for data reads: Similar to the above formula with one difference: the data cache width is 16 bytes which takes one L2 bus cycles transfer (versus two for the inst. cache), so

- L1 (read) miss time in L2: $15 + 3.75 = 18.75$ ns
- L2 miss time in memory: 90 ns
- Avg. memory access time for read $= 0.02 \times 18.75 + 0.02 \times (1 - 0.8) \times 90 + 0.02 \times (1 - 0.8) \times 0.5 \times 90 =$ 0.92 ns (1.01 CPU cycles)

(c) The average memory access time for data writes: Assume that writes misses are not allocated in L1, hence all writes use the write buffer. Also, assume that the write buffer is as wide as the L1 data cache.

- L1 (write) time to L2: $15 + 3.75 = 18.75$ ns
- L2 miss time in memory: 90 ns
- Avg. memory access time for data writes $= 0.05 \times 18.75 + 0.05 \times (1 - 0.8) \times 90 + 0.05 \times (1 - 0.8) \times 0.5 \times 90 = 2.29$ ns (2.52 CPU cycles)

(d) The overall CPI, including memory accesses:

- Components: base CPI, instruction fetch CPI, read CPI or write CPI, instruction fetch time is added to data read or write time (for load/ store instructions).

$$CPI = 1.35 + 1.09 + 0.2 \times 1.01 + 0.10 \times 2.52 = 2.84 \; cycles/inst.$$

[10]

3. **Assuming the contents of R1 and R31 at the beginning of both Load and Store instructions to be 0**:-

(a) Physical page to be written to disk is **6** or **0x06**.

(b) 32-bit physical memory address of LD instruction is **0x0FFF**.

(c) 32-bit physical memory address of data read by LD is **0x64C8**.

(d) 32-bit physical memory address of ST instruction is **0x4000**.

(e) 32-bit physical memory address of data written by ST is **0x7004**.
