

**Question 1 (20 points):** Consider a processor with a 2 GHz clock frequency, and separate instruction and data L1 caches. The base CPI with no memory stalls is 1. The instruction cache has a miss rate of 1% and the data cache has a miss rate of 4%. In the program running on this machine, approximately 25% of the instructions are loads and 15% are stores. On a cache miss, the processor must stall for 100 ns to complete a memory access.

1. (6 marks) What percentage of the total time spent by the program is memory stalls?

Assume that the program executes 1000 instructions.

For the ideal machine (without any misses), the accesses to the data cache happen in parallel with the accesses to the instruction cache. With a 2 GHz clock frequency, the clock cycle is 0.5 ns Therefore:

$$\begin{aligned}
 \text{Optimal time} &= 1000 \text{ cycles} \times 0.5 \text{ ns} = 500 \text{ ns} \\
 \text{Actual time} &= \text{hit time} + \text{miss time} \\
 \text{Actual time} &= 500 \text{ ns} + (1000 \times 0.01 + 400 \times 0.04) \times 100 \text{ ns} \\
 \text{Actual time} &= 500 \text{ ns} + 2600 \text{ ns} = 3100 \text{ ns} \\
 \text{Percentage time} &= \frac{2600}{3100} \times 100 = 84\%
 \end{aligned}$$

2. (6 marks) What is the average memory access time (in nanoseconds) for this machine?

Even though there are only 1000 cycles to access both instruction and data L1 cache, there are actually 1400 access, thus this number has to be used when computing AMAT

$$\begin{aligned}
 \text{AMAT} &= \frac{\text{hit time} + \text{miss rate} \times \text{miss penalty}}{\# \text{ of accesses}} \\
 &= \frac{500 \text{ ns} + 2600 \text{ ns}}{1400} = 2.21 \frac{\text{ns}}{\text{access}} \quad (1)
 \end{aligned}$$

3. (8 marks) To improve performance, your machine is being redesigned to include an L2 unified cache. The L2 cache has an access time of 10 ns, but will reduce the number of accesses to memory to 1% of all L1 accesses. How many times faster is this machine than the one in Part (b)?

See diagram on next page.

$$\begin{aligned}
 \text{AMAT} &= \frac{1000 \times 0.5 \text{ ns} + 26 \times 10 \text{ ns} + 14 \times 100 \text{ ns}}{1400 \text{ accesses}} = 2.21 \frac{\text{ns}}{\text{access}} \\
 &= \frac{500 \text{ ns} + 260 \text{ ns} + 1400 \text{ ns}}{1400 \text{ accesses}} = \frac{2160 \text{ ns}}{1400 \text{ accesses}} = 1.54 \frac{\text{ns}}{\text{accesses}} \\
 \text{Speedup} &= \frac{3100 \text{ ns}}{2160 \text{ ns}} = 1.44
 \end{aligned}$$

This machine is 1.44 times faster than the one in Part (b).

