1. After graduating, you are asked to become the lead computer designer at Speedo Computers, Inc. Your study of usage of high-level language constructs suggests that procedure calls are one of the most expensive operations. You have invented a scheme that reduces the loads and stores normally associated with procedure calls and returns. The first thing you do is run some experiments with and without this optimization. Your experiments use the same state-of-the-art optimizing compiler that will be used with either version of the computer. These experiments reveal the following information:

- The clock rate of the unoptimized version is 5% higher.
- 30% of the instructions in the unoptimized version are loads or stores.
- The optimized version executes 2/3 as many loads and stores as the unoptimized version. For all other instructions the dynamic counts are unchanged.
- All instructions (including load and store) take one clock cycle.

Which is faster? Justify your decision quantitatively.

2. You are building a computer with a hierarchical memory system that consists of separate instruction and data caches followed by main memory. You are using a RISC-V multicycle processor running at 5 GHz.

(i) The instruction cache is perfect (i.e., always hits) but the data cache has a 15% miss rate. On a cache miss, the processor stalls for 200ns to access main memory, then resumes normal operation. Taking cache misses into account, what is the average memory access time?

(ii) How many clock cycles per instruction (CPI) on average are required for loa5d and store word instructions considering the non-ideal memory system?

(iii) Consider a benchmark application that has 25% loads, 10% stores, 11% branches, 2% jumps, and 52% R-type instructions. Taking the nonideal memory system into account, what is the average CPI for this benchmark?

(iv) Suppose that the instruction cache is also nonideal and has a 10% miss rate. What is the average CPI for the benchmark in part (c)? Take into account both instruction and data cache misses

3. Assume we have a computer where the CPI is 1.0 when all memory accesses (including data and instruction accesses) hit in the cache. The cache is a unified (data + instruction) cache of size 256 KB, 4-way set associative, with a block size of 64 bytes. The data accesses (loads and stores) constitute 50% of the instructions. The unified cache has a miss penalty of 25 clock cycles and a miss rate of 2%. Assume 32-bit instruction and data addresses.

3.1 What are the Number of bits used for block offset?

3.2 What are the Number of sets in the cache?

3.3 What are the Number of bits for the cache index?

3.4 What are the Number of bits for the tag?

3.5 Calculate the number of Stall Cycles per instruction

3.6 How much faster would the computer be if all memory accesses were cache hits?

4. One difference between a write-through cache and a write-back cache can be in the time it takes to write. During the first cycle, we detect whether a hit will occur, and during the second (assuming a hit) we actually write the data.

Let's assume that 50% of the blocks are dirty for a write-back cache. For this question, assume that the write buffer for the write through will never stall the CPU (no penalty). Assume a cache read hit takes 1 clock cycle, the cache miss penalty is 50 clock cycles, and a block write from the cache to main memory takes 50 clock cycles. Finally, assume the instruction cache miss rate is 0.5% and the data cache miss rate is 1%. Assuming that on average 26% and 9% of instructions in the workload are loads and stores, respectively, **estimate the performance of a write-through cache with a two-cycle write versus a write-back cache with a two-cycle write.**

5. How many cycles are required to run the following program on the multicycle RISC-V processor? What is the CPI of this program? `[15 points]`

```
      addi x8, x0, 0        # i = 0
      addi x9, x0, 0        # sum = 0
      addi x10, x0, 10           # x10 = 10
Loop:
      beq x8, x10, L2            # if i == 10, go to L2
      add x9, x9, x8             # sum = sum + i
      addi x8, x8, 1             # i = i + 1
      j Loop
L2:
```

6. A processor from Fasto Technologies (FT) has 32 registers, uses 16-bit immediates and has 142 instructions in its ISA. In a given program,

- 20 % of the instructions take 1 input register and have 1 output register.,

- 30 % have 2 input registers and 1 output register,

- 25 % have 1 input register, 1 output register and take an immediate input as well,

- The remaining 25 % have one immediate input and 1 output register.

(1) For each of the 4 types of instructions , how many bits are required? *Assume that the ISA requires that all instructions be a multiple of 8 bits in length.*

(2) How much less memory does the program take up if variable-length instruction set encoding is used as opposed to fixed-length encoding?