1. (10%) Assume that the variables f, g, h, i, j, and k are assigned to register $s0, $s1, $s2, $s3, $s4, and $s5, respectively. And, assume further that the base address of the array A and B are in registers $s6 and $s7, respectively, and the elements of the array A and B are 4-byte words. Given the following two MIPS code sequences:

(I)

```
Loop:  sll    $t1, $s3, 2
       add    $t1, $t1, $s6
       lw     $t0, 0($t1)
       bne    $t0, $s5, Exit
       addi   $s3, $s3, 1
       j      Loop
Exit:
```

(II)

```
       sll    $t1, $s3, 2
       add    $t1, $t1, $s6
       lw     $t0, 0($t1)
       bne    $t0, $s5, Exit
Loop:  addi   $s3, $s3, 1
       sll    $t1, $s3, 2
       add    $t1, $t1, $s6
       lw     $t0, 0($t1)
       beq    $t0, $s5, Loop
Exit:
```

(a) (5%) Please verify that the corresponding C statement of the given two MIPS code sequences are the same.

(b) (5%) Followed by (a), which code sequence is faster? Why? (You should give a detailed analysis or explanation to get the full credit.)

2. (10%) Assume for a given processor, the CPI of arithmetic instruction is 1, the CPI of load/store instruction is 10, and the CPI of branch instruction is 4. Assume a program has the following instruction breakdown: 500 million arithmetic instructions, 300 million load/store instructions, and 100 million branch instructions.

(a) (3%) what is the execution time for the processor if the operation frequency is 5 GHz?

(b) (7%) Suppose that new, more powerful arithmetic instructions are added to the ISA. On average, through the use of these more powerful arithmetic instructions, we can reduce the number of arithmetic instructions needed to execute the program by 25%, and the cost of increasing the clock cycle time by only 10%. Is this a good design choice? Why? (You should give a detailed analysis or explanation to get the full credit.)

3. (10%) For the classical 5-stage pipelined MIPS datapath as shown in the textbook, some forwarding paths are necessary for solving data hazard problem. The basic two forwarding paths are controlled by the following two forwarding conditions:

   (a) *EX hazard*:

```
if  (EX/MEM.RegWrite
and (EX/MEM.RegisterRd ≠ 0)
and (EX/MEM.RegisterRd=ID/EX.RegisterRs))
    ForwardA = 10

if  (EX/MEM.RegWrite
and (EX/MEM.RegisterRd ≠ 0)
and (EX/MEM.RegisterRd = ID/EX.RegisterRt))     ForwardB = 10
```

   (b) *MEM hazard*:

```
if  (MEM/WB.RegWrite
and (MEM/WB.RegisterRd ≠ 0)
and (MEM/WB.RegisterRd = ID/EX.RegisterRs))
    ForwardA = 01

if  (MEM/WB.RegWrite
and (MEM/WB.RegisterRd ≠ 0)
and (MEM/WB.RegisterRd = ID/EX.RegisterRt))
    ForwardB = 01
```

   Let's consider the following MIPS code when summing a vector of numbers in a single register, a sequence of instructions will all read and write to the same register, for example

```
add    $s1, $s1, $s2
add    $s1, $s1, $s3
add    $s1, $s1, $s4
...
```

   In this case, the above forwarding paths will be conflict and need to be solved. Suppose that we decide to modify the forwarding conditions for the *MEM hazard*, please add some additional controls for the *MEM hazard* such that the result in the MEM stage will be the more recent result.

4. (12%) Consider the classic 5-stage pipelined MIPS processor. Assume that the average breakdown of different instructions for a test benchmark is: 50% for R-type instructions, 15% for conditional beq instructions, 10% for unconditional jump instructions, 15% for load instructions, and 15% for store instructions. Assume further that the designed branch predictors have the following accuracy performance: 40% for always-taken static predictor, 60% for always not-taken static predictor, and 80% for 2-bit dynamic predictor.

注意:背面有試題

(a) (3%) Stall cycles due to mispredicted branches increase CPI. Assume the branch outputs are determined in the Exe stage, that there are no data hazards, and that no delay slots are used. What is the extra CPI due to mispredicted branches with the always-taken predictor?

(b) (3%) Repeat (a) for the 2-bit dynamic predictor.

(c) (3%) With the 2-bit predictor, what speedup would be achieved if we could convert half of the branch instruction in a way that replaced each branch instruction with two ALU instructions? Assume that correctly and incorrectly predicted instructions have the same chance of being replaced.

(d) (3%) Some branch instructions might be much more predictable than others. If we know that 80% of all executed branch instructions are easy-to-predict loop-back branches, what is the accuracy of the 2-bit predictor on the remaining 20% of the branch instructions?

5. (8%) Suppose we want to perform two sums: one is sum of 10 scalars and the other is a matrix sum of a pair of 2-D arrays, with dimension 10×10. Let assume that only the matrix sum is parallelizable.

   (a) (3%) What speedup do you get with 10 versus 40 processors for the load was perfectly balanced?

   (b) (5%) Consider a larger problem with the matrix dimension grows to 20×20. And for the case of 40 processors. Suppose that if one processor has 12.5% of the parallel load and the other 30 processors share the remaining load. What is the speedup do you get with 40 processors?

6. [10%] Consider two different implementations of the same instruction set architecture. The instructions can be divided into four classes according to their CPI (class A, B, C, and D). P1 with a clock rate of 2.5GHz and CPIs of 1, 3, 2, and 2, and P2 with a clock rate of 3 GHz and CPIs of 2, 2, 2, and 2.

   (a) [3%] The result of the benchmark running on the machine has an instruction count of 1.6E12, and execution time of 700 s, and a reference time of 9000 s. Find the CPI if the clock cycle time is 0.25 ns.

   (b) [3%] Find the increase percentage in CPU time if the number of instructions of the benchmark is increased by 30% and the CPI is increased by 20%.

   (c) [4%] Given a program with a dynamic instruction count of 1.0E6 instructions divided into classes as follows: 20% class A, 10% class B, 50% class C, and 20% class D. What is the ratio of CPI for P1/P2? 1.6

7. [13%] For the following calculation:

   (a) [3%] Assume 151 and 212 are signed 8-bit decimal integers store in two's complement format. Calculate 151+214. The result should be written in decimal.

   (b) [3%] Assume 151 and 212 are unsigned 8-bit integers. Calculate 151+214 using saturating arithmetic. The result should be written in decimal.

(c) [3%] Given by a floating-point number 427D0000(hex) that is represented by IEEE 754 standard. What decimal number does it?

(d) [4%] Given by a binary representation of the decimal number 158.75 What is the biggest normalized value and smallest normalized value of the IEEE 754 single precision?

8. [10%] For a direct-mapped cache design with 32-bit address, the following bits of the address are used to access the cache:

	Tag field	Index field	Offset field
Case I	[31—10]	[9—4]	[3—0]
Case II	[31—12]	[11—6]	[5—0]

(a) [3%] What is the block size (in words) of each case?

(b) [3%] For the total of two cases, how many entries does the cache have??

(c) [4%] If the processor has a base CPI of 1, and clock rate of 4 GHz. Assume the memory access time is 100 ns, including al the miss handling. Suppose the miss rate per instruction at the primary cache is 2%. Now we add a secondary data cache that has a 5 ns access time for either a hit or a miss, and is large enough to reduce the miss rate to main memory to 0.4%. What is the total CPI of this two-level cache?

9. [10%] Caches are important to providing a high-performance memory hierarchy to processors. Below is a list of 32-bit memory address references, given as word addresses.

3, 180, 43, 2, 191, 88, 190, 14, 181, 44, 186, 253, 0, 23

(a) [3%] Assuming the cache is initially empty. For each of these references, given a direct-mapped cache with two-word blocks and a total size of 8 blocks. What is the hit rate?

(b) [3%] If the miss stall time is 25 cycles and 4 cycles for access time, what is the total cycles for this cache?

(c) [4%] For a three-way set associative cache with two-word blocks and a total size of 24 words. Use LRU replacement. What is the hit ratio? (5-7-1)

10. [7%] Listed below are key page table parameters.

Virtual address size	Physical DRAM installed	Page size	PTE size
44 bits	8GiB	4 KiB	4 bytes

(a) [3%] For a single-level page table, how much physical memory by byte is needed for storing the page table?

(b) [4%] In a memory hierarchy system, we can integrate virtual memory, TLB and page table. A memory reference can encounter some different types of misses: a TLB miss and a page fault. Considering all the combinations of these four events where the related techniques are used respectively. Please rank the event by the highest to lowest performance.

注意:背面有試題

Event number	TLB	Page table
1	Direct-mapped	Write back
2	Direct-mapped	Write through
3	Fully associated	Write through
4	Fully associated	Write back