Computer Architecture I Mid-Term II

Chinese Name: ______________________________

Pinyin Name: ______________________________

Student ID: ______________________________

E-Mail ... @shanghaitech.edu.cn: ______________________________

Question	Points	Score
1	1	
2	18	
3	10	
4	25	
5	13	
6	10	
7	14	
8	9	
Total:	100	

- This test contains 12 numbered pages, including the cover page, printed on both sides of the sheet.
- We will use Gradescope for grading, so only answers filled in at the obvious places will be used.
- Use the provided blank paper for calculations and then copy your answer here.
- Please turn **off** all cell phones, smartwatches, and other mobile devices. Remove all hats and headphones. Put everything in your backpack. Place your backpacks, laptops, and jackets out of reach.
- The total estimated time is 105 minutes.

- You have 105 minutes to complete this exam. The exam is closed book; no computers, phones, or calculators are allowed. You may use two A4 pages (front and back) of handwritten notes in addition to the provided RISC-V green sheet.
- There may be partial credit for incomplete answers; write as much of the solution as you can. We will deduct points if your solution is far more complicated than necessary. When we provide a blank, please fit your answer within the space provided.
- Do **NOT** start reading the questions/ open the exam until we tell you so!
- Unless otherwise stated, always assume a 32-bit machine for this exam.

$\boxed{1}$ 1. First Task (worth one point): Fill in you name
Fill in your name and email on the front page and your ShanghaiTech email on top of every page (without @shanghaitech.edu.cn) (so write your email in total 12 times).

2. **$ Cache!**

Notice: We assume a 32-bit machine by default.

6 (a) This section involves T / F questions. Incorrect answers on T / F questions are penalized with negative credit (in total no less than 0 point). Circle the correct answer. **Notice: NO selection will be treated as a wrong choice.**

T / F: Cache benefits from temporal and spatial locality.
T / F: Cache replacement policy is used for choosing which SET should be evicted.
T / F: Using multi-level cache will increase miss penalty.
T / F: Larger cache will decrease the miss rate.
T / F: Larger cache will decrease hit time and achieve higher performance.
T / F: Write-back cache has no write allocate.

6 (b) This section involves cache calculations. You should show the process of your calculation. Only giving a solution will receive no point.

1. An 8-way set-associative cache's total size is 4096 Bytes, the block size is 32 Bytes. Calculate the *index* and *tag* fields length.

2. A direct-mapped cache has 8-bit index field, the block size is 16 Bytes. Calculate the cache size and *tag* field length.

3. There is a computer with 3-level caches. L1 cache: local miss rate is 25%, hit time is 2 cycles. L2 cache: local hit rate is 90%, hit time is 15 cycles. L3 cache: local miss rate is 5%, hit time is 100 cycles. It takes 400 cycles to directly access memory. Calculate the AMAT and global miss rate of the given computer.

6 (c) This section involves several single choice questions. **Choose the best-fit choice** for every underlined space.

1. For a cache with fixed size and associativity, enlarging the block size will _____ miss and _____ miss.

   (a) increase compulsory (b) decrease compulsory
   (c) increase capacity (d) decrease capacity
   (e) increase conflict (f) decrease conflict

2. For a cache with fixed block size and number of sets, increasing associativity will _____ miss and _____ miss.

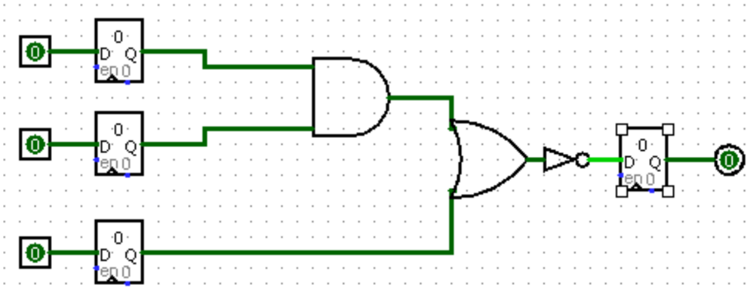
   (a) increase compulsory (b) decrease compulsory
   (c) increase capacity (d) decrease capacity
   (e) increase conflict (f) decrease conflict

3. For a cache with fixed block size and length of the *tag* field, increasing associativity will _____ and _____.

   (a) increase hit time (b) decrease hit time
   (c) increase miss rate (d) decrease miss rate
   (e) increase miss penalty (f) decrease miss penalty

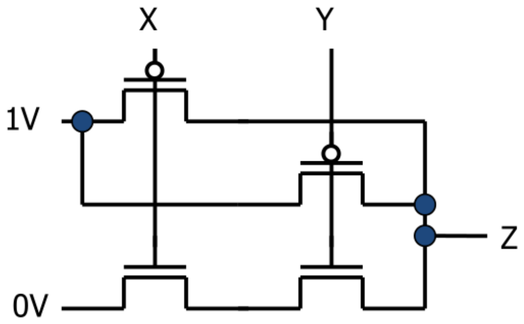
3. **FSM and SDS**

2 (a) Consider the following circuit. Assume registers have a CLK to Q time of 50ps, a setup time of 30ps, and a hold time of 30ps. Assuming that all gates have the same propagation delay, what is the maximum propagation delay each individual gate could have to achieve a clock rate of 2GHz.



2

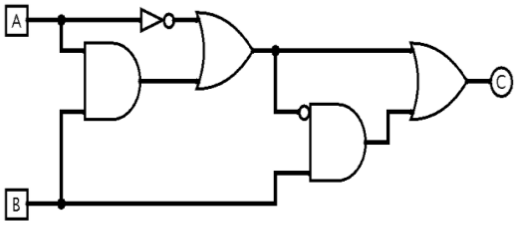
(b) Consider the following CMOS Transistors circuit. Fill the chart.



X	Y	Z
0 Volt	0 Volt	
0 Volt	1 Volt	
1 Volt	0 Volt	
1 Volt	1 Volt	

2

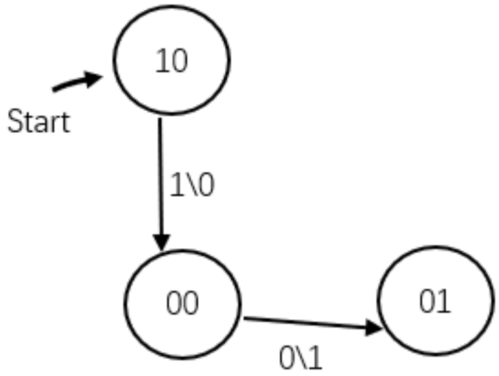
(c) The circuit shown below can be simplified. Write a Boolean expression that represents the function of the simplified circuit using the minimum number of AND, OR and NOT gate.



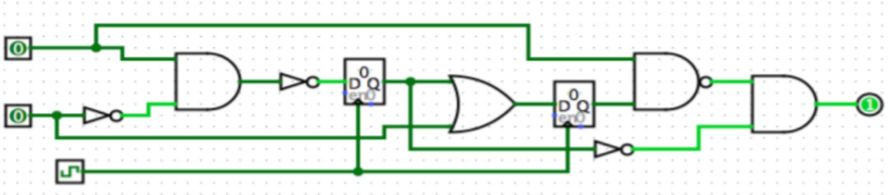
C = ______________________

(You must show your work above to earn points.)

2 (d) Suppose we feed in digits of a number to a FSM most significant bit first(e.g. it would see 1 first if we input 1000). Complete the below diagram so that it outputs 1 exactly when the number is divisible by 2, but not by 4. Assume that the machine's starting state is such that it has seen more than four 0's.

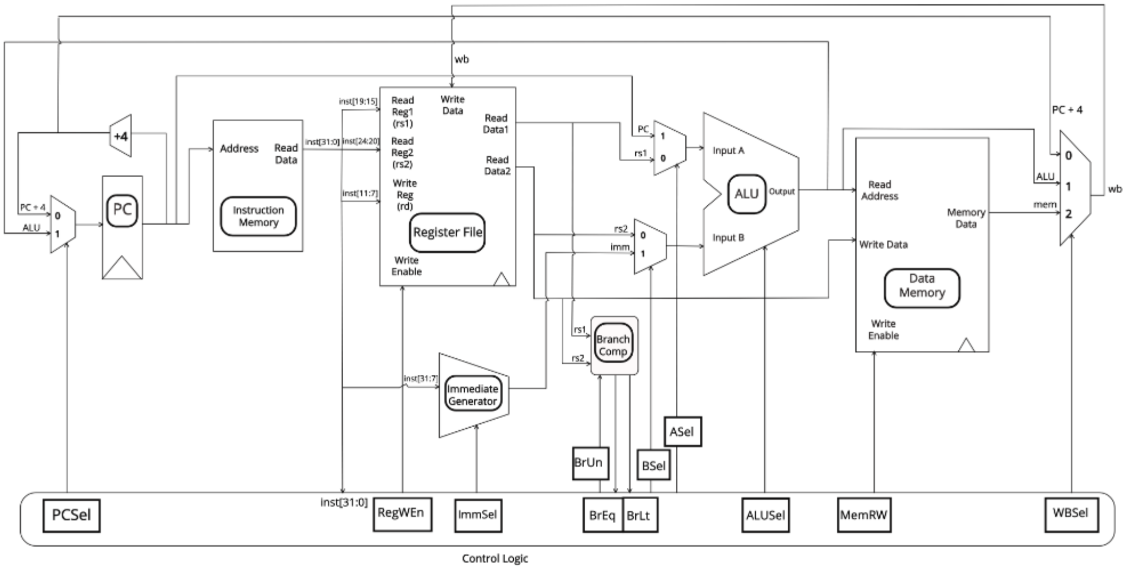


2 (e) Consider the following circuit:



You are given the following timing parameters: Register Clk-To-Q: 2ps, Register Setup: 2ps, NOT Gate: 1ps, AND Gate: 4ps, OR Gate: 3ps, NAND Gate: 4ps. **Assume the 2 inputs comes from registers and the output is connected to a register as well.**
What is the minimum clock period at which this circuit can be run?

4. **Control Single Cycle and Pipeline**
   In this part, we will be working with the single-cycle CPU datapath on the figure below.



2 (a) Explain what happens in each datapath stage.
IF Instruction Fetch

ID Instruction Decode

EX Execute

MEM Memory

WB Writeback

Clocking Methodology

- A state elementis an element connected to the clock (denoted by a triangleat the bottom). The input signal to each state element must stabilize before each rising edge.
- The critical path is the longest delay path between state elements in the circuit. If we place registers in the critical path, we can shorten the period by reducing the amount of logic between registers.

For this exercise, assume the delay for each stage in the datapath is as follows:
IF: 150 ps ID: 100 ps EX: 200 ps MEM: 250 ps WB: 100 ps

2 (a) Which instruction(s) exercise the critical path?

2 (b) What is the fastest you could clock this single cycle datapath?

2 (c) How can you improve its performance? What is the purpose of pipelining?

Pipeline
In order to pipeline, we add registers between the five datapath stages. Then we do Performance Analysis for the pipeline.
Datapath components delays are listed below.

Register clk-to-q	30ps	Branch comp.	75ps	Memory write	200ps
Register setup	20ps	ALU	200ps	RegFile read	150ps
Mux	25ps	Memory read	275ps	RegFile setup	20ps

3 (a) With the delays provided above for each of the datapath components, what would be the fastest possible clock time for a single cycle datapath?

3 (b) What is the fastest possible clock time for a pipelined datapath?

3 (c) What is the speedup from the single cycle datapath to the pipelined datapath? Whyis the speedup less than 5?

Hazards

4 (d) Given the RISC-V code below and a pipelined CPU with **no forwarding**, how many hazards would there be? What types are each hazard? Consider all possible hazards from all pairs of instructions.

Instruction	C1	C2	C3	C4	C5	C6	C7	C8	C9
1. sub t1, s0, s1	IF	ID	EX	MEM	WB				
2. or s0, t0, t1		IF	ID	EX	MEM	WB			
3. sw s1, 100(s0)			IF	ID	EX	MEM	WB		
4. bgeu s0, s2, l				IF	ID	EX	MEM	WB	
5. add t2, x0, x0					IF	ID	EX	MEM	WB

(you should answer as this: between instructions 1 and 5 (data hazard from reg t4))

4 (e) How would you fix each hazard? How many stalls would need to be added?

5. **Floating Point Numbers**

There is another representation for floating point numbers: hexadecimal floating-point format (HFP). IBM System/360 computers and subsequent machines based on that architecture (mainframes) support it. Details of single-precision 32-bit HFP are shown below:

S	Exponent (7 bits)	Significand (24 bits)

The number is represented as the following formula: $(-1)^S \times 0.\text{significand} \times 16^{\text{exponent}-64}$. In this question, we only consider normalized numbers (i.e. the first 4 bit of the significand should not be all zero).

3 (a) Convert -118.625 into single-precision 32-bit HFP. Write the answer in hexadecimal format. Show key steps.

2 (b) What is the largest representable number? What is the smallest positive normalized number? Write the answer in hexadecimal format.

Largest representable number: 0x________________________________

Smallest positive normalized number: 0x____________________________

4 (c) Compared with IEEE 754 single-precision floating-point format, what is the advantage and disadvantage of single-precision 32-bit HFP **regarding** the range of representable number and precision? Please briefly explain why.

____

____

____

____

2 (d) In x86 architecture, there is a `fabs` instruction that returns the absolute value of a floating-point number. This instruction is faster than using branches. Finish the following C code that emulates this instruction without using `if-else` statements or ternary operators. Suppose `int` and `float` have the same size. (x86 architecture uses IEEE 754 format)

```
float fabs (float a) {
   int tmp = *(int *) &a;

   ________________________________________

   ________________________________________
}
```

2 (e) Can we use the same method to get the absolute value of a single-precision 32-bit HFP number? Why?

____

____

6. **Performance**

1 (a) What is the relationship among CPU time per program, Instruction count, CPI and Clock rate?

____

1 (b) Provide the formula for Amdahl's Law:

____

6 (c) Name all elements of the Flynn Taxonomy (full names instead of abbreviations) and provide an example for each if there exists one.

____

____

____

____

2 (d) Explain the difference between `_mm_load_pd` and `_mm_loadu_pd`. (Hint: both of them operate on an 128-bit vector)

____

____

7. **Datapathology**

Consider the single cycle datapath as it relates to a new RISC-V instruction, memory add:

madd rd, rs, rt

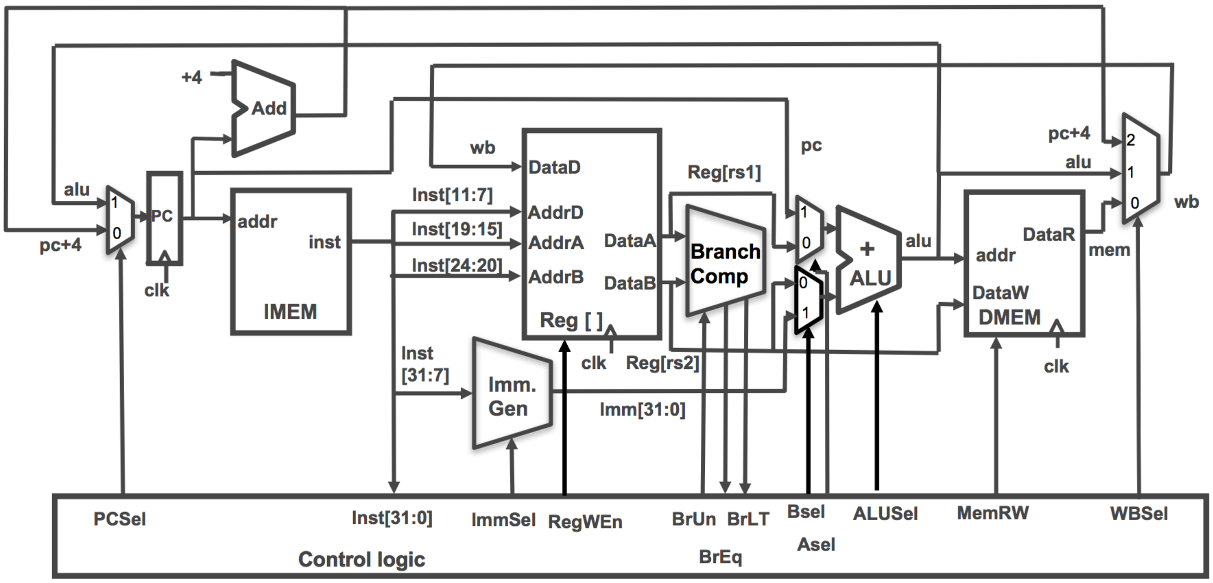
The instruction does the following:
1) Reads the value of memory at the address stored in rs.
2) Adds the value in the register specified by rt to the memory value and stores the resulting value in rd.

Ignore pipelining for parts(a)-(c).

2 (a) Write the Register Transfer Level (RTL) corresponding to madd rd, rs, rt

____

4 (b) Change as little as possible in the datapath below to enable madd. **Draw your changes right in the figure and list all your changes below**. Your modification may use muxes, wires, constants, and new control signals, but nothing else. (You may not need all the provided boxes.)



(i)	
(ii)	
(iii)	

7 (c) We now want to set all the control lines appropriately. List what each signal should be, either by an intuitive name or {0, 1, *("don't care")}. Include any new control signals you added.

PCSel	ImmSel	RegWEn	BrUn	BrLT	BrEq	Asel
Bsel	ALUSel	MemRW	WBSel			

1 (d) Briefly (one sentence) explain why madd **CANNOT** be run on the standard 5-stage RISC-V pipeline.

8. Superscalar Processors

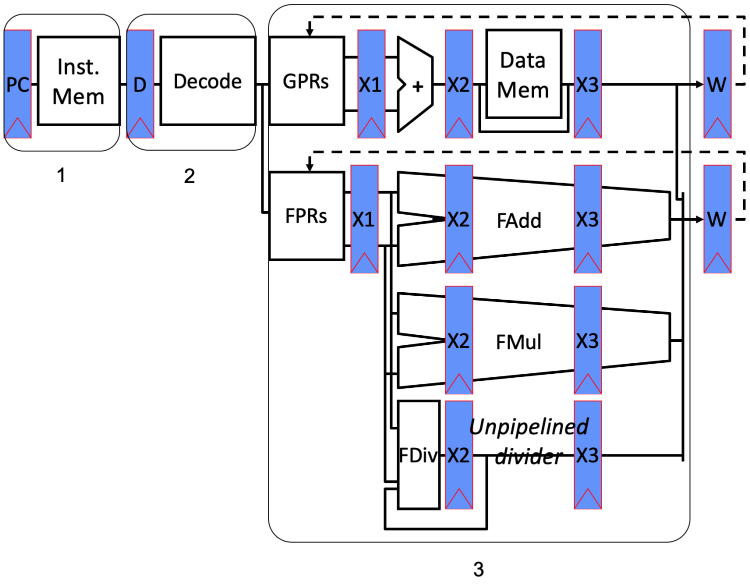
2 (a) Choose (underline) the correct descriptive phrases corresponding to superscalar processors in the brackets.

A superscalar processor can execute [at most one / more than one] instructions per clock cycle. It allows performance gain in [throughput / latency] at a given clock rate. A single-core superscalar process without support for vector operations is classified as an [SISD / SIMD / MISD / MIMD] processors according to Flynn's Law.

2 (b) Calculate the CPI (cycle per instruction) of a program with following parameters.

Operation	$Freq_i$	CPI_i
ALU	40%	2
Load	25%	6
Store	20 %	4
Branch	15 %	3

3 (c) Here is a simplified datapath schematic diagram of a superscalar processor. Fill in the following blanks with the stage number given in the diagram.



1. Issue buffer sits between stage _____ and stage _____.

2. In stage _____ instructions are executed in parallel.

2 (d) What can be done to deal with write hazards (marked by dotted lines) without equalizing all pipeline depths and without bypassing?

Scrap paper - do not submit!

Scrap paper - do not submit!