

國立交通大學試題紙

科目：計算機結構 (B)

日期：96 年 7 月 26 日 第 1 頁 共 2 頁

請 “✓” 明 ✓不可看書 可看書

* 請將答案依題號順序寫入答冊

1. (11%) The following loop is a code sequence for multiplying a scalar s to a vector of floating-point numbers. Assume that $R1$ is initially the address of the array element with the highest address, $F2$ contains the scalar value s , and $8(R2)$ is the last element to operate on.

```

loop:  L.D      F0, 0(R1)      ;F0 = vector element
        MUL.D   F4, F0, F2     ;multiply scalar from F2
        S.D     F4, 0(R1)     ;store result
        DADDUI  R1, R1, #-8    ;decrement pointer 8B (double word)
        BNE     R1, R2, loop   ;branch R1!= R2
    
```

Consider the following machine:

- It is a single-issue processor.
- The integer pipeline is the standard 5-stage pipeline with forwarding and bypassing.
- The function units are fully pipelined or replicated.
- Branches are resolved in the ID stage.
- There are as many registers, both FP and integer, as needed.
- The latencies of operations are shown as follows:

Instruction producing result	Instruction using result	Latency in clock cycles
FP ALU op	Another FP ALU op	2
FP ALU op	Store double	2
Load double	FP ALU op	1
Load double	Store double	0
Branch delay slots	Any op	1

- (a) Without rescheduling the code, how many stalls are there in the code and how many cycles are required per iteration? Please show your work by putting a stall (or stalls) where it would occur in the code. (2%)
- (b) Repeat (a) with rescheduling the code. (2%)
- (c) Unroll the loop twice and reschedule it to eliminate the stalls in the loop body as many as possible. Are all stalls eliminated in your answer and how many cycles are required per (original) iteration in average? (4%)
- (d) Describe three limits to the gains of loop unrolling. (3%)

國立交通大學試題紙

科目：計算機結構 (B)

日期：96 年 7 月 26 日 第 2 頁 共 2 頁

2. (6%) Evaluate the impact of two different cache organizations on the performance of a CPU. Assume that the CPI with a perfect cache is 2.0, the clock cycle time is 1.0 ns, and there are 1.5 memory references per instruction. One cache is direct mapped and the other is two-way set associative. The hit time is 1 clock cycle and cache miss penalty is 60 ns for either cache organization. The miss rate of the direct-mapped cache is 1.5%, and that of the two-way set-associative cache is 1.0%. For the two-way set-associative cache, the CPU cycle time must be stretched 1.2 times.
 - (a) Calculate the average memory access time of each cache organization. (3%)
 - (b) Calculate the CPU performance of each cache organization. Which one of the caches has better performance? (3%)
3. (8%) The three main metrics to evaluate the performance of a cache memory are miss rate, miss penalty, and hit time.
 - (a) Describe the characteristic of a trace cache and the metrics could be improved by it. (3%)
 - (b) Repeat (a) for a victim cache. (3%)
 - (c) For a cache with larger size, which of the metrics could be improved and which of the metrics could be worsened, if any, by it? (2%)
4. (8%) For a virtual memory system,
 - (a) Where can a block be placed in main memory? Why? (2%)
 - (b) Draw the block diagram for the address translation hardware of a paged virtual memory with fully-associative translation lookaside buffer (TLB), and describe the way to find a block in the main memory. (4%)
 - (c) What happens on a write? Why? (2%)
5. (8%) Describe and compare the following two cache coherence protocols: *directory-based* and *snooping*. Which one of the protocols is more suitable to be applied to symmetric shared-memory architectures and distributed shared-memory architectures, respectively? Why?
6.
 - (a) (4%) Draw an 8-node Omega network and give an example to show that the network is blocking.
 - (b) (5%) Draw a 2D torus of 16 nodes and calculate its bisection bandwidth, ports per switch, and total number of lines.