國立中山大學資訊工程學系

94 學年度第 1 學期博士班資格考試　計算機結構

1. When designing memory systems it becomes useful to known the frequency of memory reads versus writes and also accesses for instructions versus those for data. Using the average instruction mix information for MIPS in the following table, find
   - The percentage of all memory accesses for data
   - The percentage of data accesses that are reads
   - The percentage of all memory accesses that are reads

   Ignore the size of a datum when counting accesses.　(10%)

Instruction class	Frequency
Arithmetic	30%
Logical	15%
Load	25%
Store	10%
Conditional branch	18%
Jump	2%

2. (1) Please briefly explain the following branch prediction schemes: 2-bit predictor, correlating branch predictor, and Tournament predictor.　(7%)

   (2) Consider the following simplified code fragment:

   ```
   If (d==0)   d=1;
   If (d==1)   …
   ```

   The branches corresponding to the two if statements are labeled b1 and b2. The possible sequences for an execution of this fragment, assuming d has values 0, 1, and 2, are shown in Table 1. Assume the sequence above is executed repeatedly and ignore other branches in the program and a (1,1) predictor (1-bit predictor with 1 bit of correlation) is used. The pair of prediction bits are written together, with the first bit being the prediction if the last branch in the program is not taken (NT) and the second bit being the prediction if the last branch in the program is taken (T). If the predictor is initialized to NT/NT, please fill Table 2.　(8%)

Table 1

Initial value of d	d==0?	b1	value of d before b2	d==1?	b2
0	yes	not taken	1	yes	not taken
1	no	taken	1	yes	not taken
2	no	taken	2	no	taken

Table 2

d=?	b1 prediction	b1 action	New b1 prediction	b2 prediction	b2 action	New b2 prediction
0	**NT**/NT	NT	NT/NT	**NT**/NT	NT	NT/NT
2						
0						
2						

3. (1) Please explain as detailed as possible the difference between dynamic scheduling and hardware-based speculation. (8%)

   (2) Assume that the reorder buffer (ROB) in a hardware-based speculative processor has four buffer entries, named 0, 1, 2, and 3. For the following code fragment, assume that ADD.D, SUB.D, and ADDI instructions execute for 1 cycle, and MUL.D executes for 10 cycles. Assume that the processor has sufficient function units to avoid stalling instruction issue.

```
ADD.D   F0, F8, F8
SUB.D   F1, F9, F0
MUL.D   F2, F8, F9
ADD.D   F3, F0, F1
SUB.D   F4, F0, F2
ADDI    R10, R12, R12
```

Fill in the following table to show ROB contents and history as it would exist on the cycle that the MUL.D instruction is ready to write its result. Assume that F8, F9 and R12 are initialized and that the ROB is initially empty. Because a ROB is implemented as a circular queue, the entry number labels repeat modulo 4 reading down the table. If a ROB entry would be reallocated during the simulated execution time, write the details of the new allocation in the next available correspondingly numbered table row. Use the rightmost column to indicate if the instruction has been committed. (12%)

	ROB fields			Committed?
Entry	Instruction	Destination	Value	Yes/no
0	ADD.D			
1				
2				
3				
0				
1				
2				

4. Which has the lower miss rate: a 32 KB instruction cache with a 32 KB data cache or a 64 KB unified cache? Use the miss rates in Table 3 to help calculate the correct answer, assuming 50% of the instructions are data transfer instructions. Assume a hit takes 1 clock cycle and the miss penalty is 100 clock cycles. A load or store hit takes 1 extra clock cycle on a unified cache if there is only one cache port to satisfy two simultaneous requests (a structural hazard). What is the average memory access time in each case? Assume write-through caches with a write buffer and ignore stalls due to the write buffer. (15%)

Table 3: Miss per 1000 instructions

Size	Instruction cache	Data cache	Unified cache
16 K	4	42	51
32 K	2	40	43
64 K	1	36	40

5. As caches increase in size, blocks often increase in size as well.

(1) If a large instruction caches has larger data blocks, is there still a need for prefetching? Explain the interaction between prefetching and increased block size in instruction caches. (5%)

(2) Is there a need for data prefectch instructions when data blocks get larger? Explain? (5%)

(3) Cache misses can be sorted into three simple categories: Compulsory, Capacity, and Conflict. Please explain why they occur and how to reduce them. (5%)

6. Answer the following questions.

(1) The major complication in implementing predicated instructions is deciding when to annul an instruction. Please explain the advantages and disadvantages when predicated instructions are annulled early or later in the pipeline. (5%)

(2) Discuss the advantages and disadvantages of loop unrolling and software pipelining. (8%)

(3) The following three techniques have been developed to reduce cache miss penalty or miss rate via parallelism: Nonblocking Caches to Reduce Stalls on Cache Misses, Hardware Prefetching of Instructions and Data, and Compiler-Controlled Prefetching. Please briefly explain these techniques and how they work. (12%)