國立中山大學資訊工程學系

93 學年度第 1 學期博士班資格考試　計算機結構

1. (10%) The design of MIPS provides for 32 general-purpose registers and 32 floating-point registers. If registers are good, are more registers better? List and discuss as many trade-offs as you can that should be considered by instruction set architecture designers examining whether to, and how much to, increase the numbers of MIPS registers.

2. Consider adding a new index addressing mode to MIPS. The addressing mode adds two registers and an 11-bit signed offset to get the effective address. Our compiler will be changed so that code sequences of the form

```
ADD    R1, R1, R2
LW     Rd, 100(R1)    (or store)
```

will be replaced with a load (or store) using the new addressing mode. If loads and stores comprise 25% and 15% of the average instruction mix respectively, use the average instruction frequencies in evaluating this addition.

(1) (7%) Assume that the addressing mode can be used for 10% of the displacement loads and stores (accounting for both the frequency of this type of address calculation and the shorter offset). What is the ratio of instruction count on the enhanced MIPS compared to the original MIPS?

(2) (8%) If the new addressing mode lengthens the clock cycle by 5%, which machine will be faster and by how much?

3. (1) (5%) What is instruction-level parallelism? Please explain the limitations of instruction-level parallelism.

(2) (10%) Please describe the Tomasulo's algorithm and hardware-based speculation, and then explain the key difference between them.

(3) (10%) Please explain and compare the following branch prediction schemes: 2-bit predictor, correlating branch predictor, and Tournament predictor.

4. (1) (5%) Show why the following loop is not parallel.

```
for (i=2; i<100; i=i+1)    {
        a[i] = b[i] + a[i];          /* S1 */
        c[i-1] = a[i] + d[i];        /* S2 */
        a[i-1] = 2 * b[i];           /* S3 */
        b[i+1] =2 * b[i];            /* S4 */
}
```

(2) (10%) Please describe and compare loop unrolling and software pipelining.

5. (1) (10%) Fig. 1 is the data cache in the Alpha 21264 microprocessor. Please explain the three portions of the address in Fig. 1 and describe the cache of Fig. 1 (e.g. associativity, cache size, write strategy, etc.)

(2) (15%) Assume that the CPI with a perfect cache is 2.0, the clock cycle time is 1.0 ns, there are 1.5 memory references per instruction, the size of both caches is 64 KB, and both have a block size of 64 bytes. One cache is direct mapped and the other is two-way set associative. Since the speed of the CPU is tied directly to the speed of a cache hit, assume the CPU clock cycle time must be stretched 1.25 times to accommodate the selection multiplexor of the set-associative cache. To the first approximation, the cache miss penalty is 80 ns for either cache organization. Assume the hit time is 1 clock cycle, the miss rate of a direct-mapped 64 KB cache is 1.5%, and the miss rate for a two-way set-associative cache of the same size is 1.0%. Calculate the average memory access time and then CPU performance.

$$\text{Average memory access time} = \text{Hit time} + \text{Miss rate} \times \text{Miss penalty}$$

$$\text{CPU time} = \text{IC} \times \left(\text{CPI}_{\text{execution}} + \frac{\text{Misses}}{\text{instruction}} \times \text{Miss penalty} \right) \times \text{Clock cycle time}$$

6. (1) (5%) Existing MIMD multiprocessors fall into two classes by their memory organization: centralized shared-memory architecture and distributed-memory architecture. Please explain and discuss their advantages and disadvantages.

(2) (5%) Please describe two alternative architectural approaches and associated communication mechanisms for communicating data among processors in the large-scale multiprocessor.

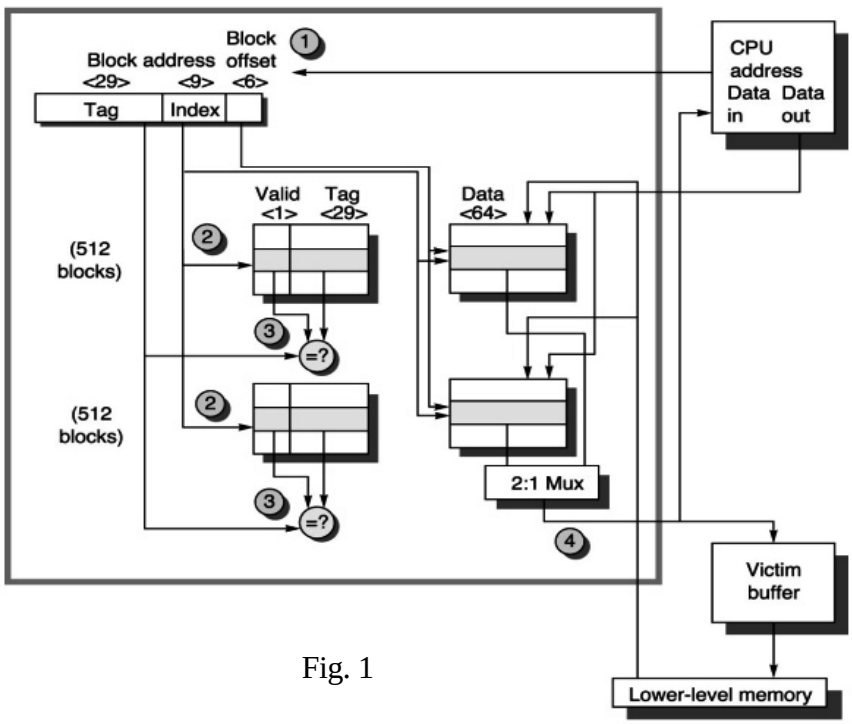

Fig. 1