國立中央大學資訊工程學系 114 學年度第二學期博士班資格考試題紙

科目： **計算機結構** （Computer Architecture） 第一頁 共三頁(page 1 of 3 ）

1. (10%) A modern processor features a two-level cache hierarchy. The L1 cache has a hit time of 1 clock cycle and a miss rate of 4%. The L2 cache has a hit time of 12 clock cycles and a local miss rate of 25%. The main memory access time is 120 clock cycles.
    a. Calculate the Average Memory Access Time (AMAT) for this memory hierarchy.
    b. If the designers want to reduce the AMAT to exactly 2.0 cycles by optimizing the L2 cache hit time, what must the new L2 hit time be?

2. (10%) Describe the "3 Cs" (Compulsory, Capacity, Conflict) model of cache misses. Suppose you are analyzing a memory trace and decide to double the cache associativity (e.g., from 4-way to 8-way) while keeping the total cache capacity and block size strictly constant. How will each of the 3 Cs be affected, and what are the potential trade-offs of this architectural change?

3. (10%) When a CPU generates a virtual memory address that results in a TLB miss followed by a Page Fault (assuming a hardware-managed TLB architecture), a specific sequence of hardware and software events occurs. Please arrange the following events (A through H) in the correct chronological order. Write down the correct sequence of letters.

    A. The OS allocates a free physical page frame in main memory (possibly evicting another page to disk if RAM is full).
    B. The OS restores the CPU state and restarts the instruction that caused the fault.
    C. The hardware page table walker reads the Page Table Entry (PTE) from memory and discovers that the "Valid" bit is set to 0.
    D. Once the disk I/O completes, the OS updates the PTE with the new physical page number and sets the Valid bit to 1.
    E. The CPU generates a virtual address, checks the hardware TLB, and experiences a TLB miss.
    F. The hardware triggers a Page Fault Exception, saving the CPU state and transferring control to the Operating System's page fault handler.
    G. The OS initiates a disk I/O operation to bring the missing page from secondary storage into the newly allocated physical frame.
    H. The restarted instruction executes again; the hardware fetches the now-valid PTE, loads it into the TLB, and the memory access succeeds.

背面還有 Please Turn Over

4. (10%) Modern server architectures and operating systems often support "Huge Pages" (e.g., 2MB or 1GB) in addition to the standard 4KB pages. For memory-intensive workloads (such as large-scale database servers or training deep learning models), compare the use of Huge Pages against standard 4KB pages. Specifically, explain the primary performance advantage of using Huge Pages (focusing on the role of the TLB) and discuss at least one significant drawback of this approach.

5. (10%) Explain why true Least Recently Used (LRU) replacement policies are rarely implemented in highly associative caches (e.g., 16-way L2/L3 caches). Given this hardware limitation, modern processors often use either a Random replacement policy or a Pseudo-LRU (approximated LRU) policy. Briefly compare the trade-offs between True LRU, Random, and Pseudo-LRU in terms of hardware complexity and cache miss rate.

6. (10%) What are the advantages of PC relative address for conditional branch?

7. (10%) In the context of computer architecture and instruction-level parallelism, which of the following statements are true regarding Dynamic Scheduling?
   (A) It allows the hardware to rearrange the execution order of instructions to reduce stalls caused by data dependencies.
   (B) It eliminates the need for a compiler to perform any instruction scheduling or optimization.
   (C) It can effectively handle dependencies that are unknown at compile time, such as those arising from memory references or cache misses.
   (D) Some of the techniques use Register Renaming to mitigate Write-After-Write (WAW) and Write-After-Read (WAR) hazards.
   (E) It simplifies the hardware control logic compared to static scheduling, leading to lower power consumption.

8. (10%) Which of the following statements are true for load-store architecture?
   (A) Arithmetic instructions can directly use memory addresses as operands to reduce the number of instructions.
   (B) Only specific instructions (typically Load and Store) are permitted to access memory.
   (C) Arithmetic and logical instructions should have their operands in registers.
   (D) It typically uses variable-length instructions to minimize the total size of the program in memory.
   (E) It requires fewer registers than a Register-Memory architecture because data stays in memory longer.

9. (10%) What is correlated branch prediction? Give me an example of a (m, n) Predictor with m=3, n=2.

10. (10%) Consider the following code sequence. F1- F10 are registers. Assume the following execution time for different operations: MUL takes 8 cycles, DIV takes 20 cycles, and ADD and SUB both take 1 cycle to complete. The instructions are issued into the pipeline in order, but out of order execution and completion is allowed. What hazards will be encountered when executing this code sequence?

MUL F6, F5, F6
SUB F2, F7, F8
DIV F10, F6, F5
ADD F5, F8, F2