NOTE: It is my policy to give a failing grade in the course to any student who either gives or receives aid on any exam or quiz.

INSTRUCTIONS: Circle the letter of the one best answer for each question. Answer all questions: there is no penalty for guessing. Unless otherwise indicated, all questions count equally.

1. What is the average number of memory accesses per instruction if 20% of the instructions are *lw*, 10% are *sw*, and the remainder are a mix of arithmetic and branch instructions?
   A. 0.1
   B. 0.2
   C. 1.1
   D. 1.2
   E. 1.3

2. What is the ALU used for in *lw* and *sw* instructions?
   A. To calculate the effective address.
   B. To calculate the hit ratio.
   C. To calculate the miss ratio.
   D. To calculate the difference.
   E. It is not used because they don't do arithmetic.

3. A 256M × 32 memory system is built using 64M × 16 memory chips. How many memory chips will be needed, and what will be their organization?
   A. Two chips, next to each other.
   B. Three chips, on top of each other.
   C. Four chips, in a square.
   D. Six chips, in a hexagon.
   E. Eight chips, in a 4 × 2 matrix.

4. Using the same design as Question 3, how many address lines go into the memory system, and how many go into each memory chip?
   A. 28 go into the memory system, and 26 go into each memory chip.
   B. 26 go into the memory system, and 28 go into each memory chip.
   C. 2 go into the memory system, and 8 go into each memory chip.
   D. 6 go into the memory system, and 8 go into each memory chip.
   E. 28 go into the memory system, and they all go into each memory chip.

5. Still using the same design as Question 3, how many data lines go into the memory system, and how many go into each memory chip?
   A. 32 go into the memory system, and each memory chip is connected to 16 of them.
   B. 16 go into the memory system, and each memory chip is connected to all of them.
   C. 28 go into the memory system, and each memory chip is connected to 32 of them.
   D. 32 go into the memory system, and each memory chip is connected to all of them.
   E. No data lines go into the memory system; they all go into the decoder.

6. Still using the same design as Question 3, what are the name and structure of the device that determines which row of memory chips to enable for a particular operation?
   A. It is a 3 × 2 multiplexer.
   B. It is a 2 × 3 multiplexer.
   C. It is a 4 × 2 decoder.
   D. It is a 2 × 4 decoder.
   E. It is a 5 × 7 glossy.

7. What is the name of the input on a memory chip that is used to control whether it takes part in a memory operation or not?
   A. Do-op
   B. No-op
   C. Pre-op
   D. Pre-nup
   E. Chip Select

8. What are the two operations that the CPU can request a memory system to perform?
   A. Hit and miss
   B. Read and write
   C. Add and subtract
   D. Remember and forget
   E. Appendectomy and tooth extraction

9. What is a *balanced pipeline*?
   A. A pipeline with a different number of stages for different classes of instructions.
   B. A pipeline with no hazards.
   C. A pipeline with an instruction occupying each stage.
   D. A pipeline with the same propagation delays for each stage.
   E. A pipeline that corrects its number of stages if the clock speed drifts from its nominal value.

10. A CPU that is controlled by a state machine where the sequence and number of states is determined by the op code of the instruction being executed is:
   A. A single cycle design
   B. A multi-cycle design
   C. A pipelined design
   D. A cyclical cycle design
   E. A bicycle design

11. A single cycle design with a clock speed of 1 GHz is converted to a pipeline design with 5 stages. Ideally, the new clock speed will be:
   A. 0.2 GHz
   B. 0.5 GHz
   C. 1.0 GHz
   D. 2.0 GHz
   E. 5.0 GHz

12. What is the average CPI (Clocks Per Instruction completed) for a single-cycle CPU design?
   A. 0.9
   B. 1.0
   C. 1.1
   D. 1.2
   E. *lw* and *sw* instructions have a larger CPI than other instructions in the single-cycle design.

13. What is the ideal CPI for a 5 stage pipelined CPU design after the pipeline is filled?
   A. 1.0
   B. 2.0
   C. 3.0
   D. 4.0
   E. 5.0

14. How might a *data hazard* occur in a single cycle design?
   A. If the program branches before the result is needed.
   B. If the program branches after the result is needed.
   C. If there is a *lw* instruction followed by a *sw* instruction
   D. If there is a cache miss.
   E. There can't be a data hazard in a single cycle design because each instruction is processed completely before the next one is started.

15. What is a *data hazard*?
    A. When an instruction needs the result produced by a previous instruction, but the previous instruction hasn't actually produced the result yet.
    B. When two instructions produce different answers.
    C. When one instruction produces the same answer.
    D. When a branch invalidates a result.
    E. When a jump invalidates a branch.

16. How long does it take to fetch and process one instruction in a five-stage pipeline? (*Note:* this is *not* the same as Question 13.)
    A. One clock cycle
    B. Five clock cycles
    C. One millisecond
    D. 2.5 clock cycles
    E. It depends on the op code of the next instruction after it.

17. What is the maximum number of instructions that are completed each clock cycle in a CPU with a five stage pipeline? (*Note:* this is essentially the same as Question 13.)
    A. One
    B. Two
    C. Three
    D. Four
    E. It depends on the clock speed.

18. What is *register forwarding*?
    A. A technique for increasing the clock speed in the single-cycle design.
    B. A technique for decreasing the clock speed in the multi-cycle design.
    C. A technique for increasing the number of clock cycles in the pipelined design.
    D. Obtaining an operand from a pipeline register instead of waiting for it to be written to the register file.
    E. Obtaining an operand from the register file instead of waiting for it to be written to a pipeline register.

19. What information is stored in the IF/ID pipeline register?
    A. The operands, the op code, and the ALU
    B. The instruction, the next instruction, and the one after that.
    C. The instruction and the address of the next instruction.
    D. The word from data memory and the branch target address.
    E. The jump, the branch, and the set less than.

20. What information is stored in the MEM/WB pipeline register?
    A. The word from memory, the output word from the ALU, the destination register number, and the controls to tell which word, if either, is to be written to the register file.
    B. The register file, the ALU, and the branch target address.
    C. The instruction memory, the data memory, and the jump address.
    D. The ALU control bits, the operands, and the number of the register to receive the result.
    E. All of the above, and more bits too tedious to mention.

21. Which choice lists only items in the CPU that are connected to the clock?
    A. The ALU, the multiplexers, and the memory.
    B. The register file, the PC, and the pipeline registers.
    C. The ALU and the register file.
    D. The Multiplexers and the pipeline registers.
    E. The adders, the sign extend unit, and the multiplexers.

22. Which design(s) have different execution times for different types of instructions?
    A. Single-cycle, multi-cycle, and pipeline.
    B. Single-cycle and pipeline.
    C. Single-cycle and multi-cycle.
    D. Only multi-cycle.
    E. Only single-cycle.

23. Which of the following is an example of good *locality of reference*?
    A. When a program branches to a subroutine.
    B. When a program executes a short loop.
    C. When a program adds the elements in a vector.
    D. All of the above.
    E. Choices B and C.

24. A memory system has a cache with 512K (524,288) lines with 64 bytes per line. Main memory is 4 GB. How many *blocks* of main memory are there?
    A. 512 K
    B. 512 K * 64
    C. 4 G / 64
    D. 4 G * 64
    E. 4 G

25. Using the same design as Question 24, what proportion of main memory can be in cache at one time?
    A. 64 / 4 G
    B. (512 K * 64) / 4 G
    C. 512 K / 4 G
    D. 512 K / (4 G * 8)
    E. 2.7184518459045

26. The number of bytes in a cache is 0.008 times the number of bytes of main memory, and the hit ratio is also 0.008. Which statement makes the most sense?
    A. The hit ratio is much higher than expected because the cache is so small relative to main memory.
    B. The hit ratio is just about what one would expect.
    C. The hit ratio is much lower than expected because locality of reference should cause values in the cache to be accessed multiple times before they are replaced.
    D. The hit ratio and the miss ratio must always be the same.
    E. The hit ratio should be 0.008 * 0.008 = 0.000064.

27. In a *direct mapped* cache design, a particular block of main memory:
    A. Can be loaded into any cache block.
    B. Can be loaded into any of four different cache blocks.
    C. Is always loaded into the same cache block.
    D. Is loaded into the cache block that will never be used again until the program finishes.
    E. Is loaded into each cache block on a rotating basis.

28. Using the same design as Question 24, how many bits are in the *offset* part of a main memory address?
    A. 6
    B. 19
    C. 7
    D. 25
    E. 32

29. Using the same design as Question 24, for a direct mapped cache, how many bits are in the *index* part of a main memory address?
    A. 6
    B. 19
    C. 7
    D. 25
    E. 32

30. Using the same design as Question 24, for a direct mapped cache, how many bits are in the tag part of a main memory address?
    A. 6
    B. 19
    C. 7
    D. 25
    E. 32

31. What are the *V* bits in a cache memory system used for?
    A. To tell whether a cache block has been written to or not.
    B. To tell how recently a cache block has been accessed.
    C. To determine the hit ratio of a cache block.
    D. To tell whether a cache block contains valid data or not.
    E. To establish the parity of a cache block.

32. If the cache access time is 100 psec, the main memory access time is 1 nsec, and the hit ratio is 0.95, what is the average memory access time?
    A. 100 psec
    B. 145 psec
    C. 950 psec
    D. 1.05 nsec
    E. 1.45 nsec

33. Why is *write back* better than *write through* for virtual memory systems?
    A. It's not; write through is better.
    B. It improves cache access time.
    C. Because disk access time is so much slower compared to memory access time.
    D. Because it is simpler to implement.
    E. Because the operating system is not involved in the write back algorithms.

34. Which of the following is a typical page size?
    A. 0.512 byte
    B. 512 bytes
    C. 512 kilobytes
    D. 512 megabytes
    E. 512 gigabytes

35. What part of a computer translates virtual addresses into physical addresses when a memory reference is made?
    A. The operating system, using the translation lookaside buffer
    B. The CPU, using the page table.
    C. The cache, using the *V* bit.
    D. The ALU, using the effective address algorithm.
    E. The web browser, using Google.

36. Compare the time penalty for a *page fault* compared to a *cache miss*:
    A. A page fault takes millions of times more time to handle than a cache miss.
    B. Page faults and cache misses are handled in exactly the same amounts of time.
    C. The time penalty for a page fault depends on the size of a cache block, but the time penalty for a cache miss is always exactly two clock cycles.
    D. Page faults are handled faster than cache misses because they occur at a higher point in the memory hierarchy.
    E. Cache misses cause multiple page faults when using a set associative virtual memory system.

37. How many clock cycles does the cache use to determine whether there is a hit or not?
    A. None, it uses combinational logic rather than sequential logic.
    B. One, because the comparator has one flip-flop for every tag bit in an instruction.
    C. One clock cycle per tag bit.
    D. One clock cycle per cache block.
    E. One clock cycle per bit in the EX/MEM pipeline register.