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: Mark the best answer for each question on your answer sheet. All questions count equally; the number of points given is a relative weight, not an absolute number.

---

1.   (5 Points) A non-pipelined processor has a 250 MHz clock. If the processor is redesigned to use a perfectly balanced pipeline with depth 6, what would be the clock speed of the new processor?
   A. Approximately 400 MHz.
   B. Approximately 1.00 GHz
   C. Approximately 1.25 GHz
   D. Approximately 1.50 GHz
   E. Approximately 1.75 GHz

2.   (5 Points) Continuing from the previous question, what would be the expected speedup due to pipelining?
   A. 4.0
   B. 5.0
   C. 6.0
   D. 7.0
   E. 8.0

3.   (5 Points) What would be the expected speedup expressed as a percentage?
   A. 300%
   B. 400%
   C. 500%
   D. 600%
   E. 700%

4.   (5 Points) What will be the relationship of the CPI without pipelining to the CPI with pipelining? (Don't ask me what CPI stands for; that's what I'm testing!)
   A. No change
   B. The pipleined CPI will be 6 times the unpipelined CPI.
   C. The unpipelined CPI will be 6 times the pipelined CPI.
   D. It depends on the hit ratio.
   E. It depends on the miss ratio.

5.   (5 Points) What will happen to the instruction *latency* as a result of introducing the pipeline?
   A. The latency will remain unchanged.
   B. The latency will decrease from 6 cycles to one cycle as a result of using the pipeline.
   C. The latency will increase from one cycle to 6 cycles as a result of using the pipeline.
   D. The latency will depend on the opcode of the instruction.
   E. The latency will depend on the throughput.

6.   (5 Points) What will happen to the instruction *throughput* as a result of introducing the pipeline?
   A. The throughput will remain unchanged.
   B. The throughput will increase by a factor of 6 as a result of using the pipeline.
   C. The throughput will decrease by a factor of 6 as a result of using the pipeline.
   D. The throughput will depend on the number of instructions in the program.
   E. The throughput will depend on the latency.

7.   (5 Points) What is a *data* hazard?
   A.   When there are not enough functional units to meet the demands of all the instructions in a pipeline.
   B.   When an instruction in a pipeline is a branch and partially executed instructions following it in the pipeline aren't supposed to be executed.
   C.   When an instruction needs the result of a previous instruction, but the previous instruction is still in the pipeline and its result hasn't been written back to the registers yet.
   D.   When a pipeline result is invalid because of arithmetic overflow.
   E.   When an arithmetic calculation produces a carry out of the leftmost position.

8.   (5 Points) What is a *control* hazard?
   A.   When there are not enough functional units to meet the demands of all the instructions in a pipeline.
   B.   When an instruction in a pipeline is a branch and partially executed instructions following it in the pipeline aren't supposed to be executed.
   C.   When an instruction needs the result of a previous instruction, but the previous instruction is still in the pipeline and its result hasn't been written back to the registers yet.
   D.   When two instructions try to read from the same register at the same time.
   E.   When a single instruction tries to read from two different registers at the same time.

9.   (5 Points) What is a *structural* hazard?
   A.   When there are not enough functional units to meet the demands of all the instructions in a pipeline.
   B.   When an instruction in a pipeline is a branch and partially executed instructions following it in the pipeline aren't supposed to be executed.
   C.   When an instruction needs the result of a previous instruction, but the previous instruction is still in the pipeline and its result hasn't been written back to the registers yet.
   D.   When a program has illegally nested loops that the pipeline can't handle.
   E.   When the program has illegally nested loops that the compiler has to fix.

10.   (5 Points) What is a pipeline *stall*?
   A.   When instructions are not allowed to move through the pipeline for one or more clock cycles due to a data or structural hazard.
   B.   When instructions enter the pipeline too fast.
   C.   When instructions exit the pipeline too fast.
   D.   When the same instruction enters and exits the pipeline in a single clock cycle.
   E.   When an instruction later in the pipeline causes an earlier instruction to produce the wrong answer.

11.   (5 Points) What is the relationship between a pipeline *bubble* and a pipeline *stall*?
   A.   A bubble is a stall that doesn't cause a delay.
   B.   A stall is a bubble that doesn't cause a delay.
   C.   A bubble can only occur if there is a control hazard.
   D.   A stall can only occur if there is a control hazard.
   E.   They are the same thing.

12.   (5 Points) What is *result forwarding*?
   A.   A mechanism for dealing with control hazards in which one branch instruction nullifies another branch instruction.
   B.   A mechanism for dealing with structural hazards in which one branch instruction nullifies another branch instruction.
   C.   A mechanism for dealing with data hazards by obtaining the result of a previous instruction from an internal buffer instead of waiting for it to be written back to a register.
   D.   A mechanism for dealing with structural hazards in which an arithmetic instruction nullifies a branch instruction that has not yet left the pipeline.
   E.   A mechanism for dealing with structural hazards in which an arithmetic instruction nullifies a branch instruction that has already left the pipeline.

13. (5 Points) What is the name for the technique in which both forks following a branch are executed in parallel, and the unused fork is discarded once the result of the branch condition has been determined?
    A. Delayed branch
    B. Speculative execution
    C. Twin forks
    D. Negative forks
    E. Positive forks

14. (5 Points) How many pipeline registers are used in an n stage pipeline?
    A. 0
    B. 1
    C. $n - 1$
    D. n
    E. $n + 1$

15. (5 Points) How many bits must each pipeline register hold?
    A. 32
    B. Enough to control all the parts of the datapath in the next stage of the pipeline.
    C. Enough to control all the parts of the datapath in the next stage of the pipeline plus enough to control further stages in the pipeline, such as the destination register number.
    D. Enough to hold the result of the ALU operation in the previous stage of the pipeline.
    E. Enough to hold the opcode and register numbers of the instruction being executed.

16. (5 Points) Most pipelined processors improve performance by adjusting the number of pipeline stages depending on the opcode of the instruction being executed.
    A. True
    B. False

17. (5 Points) What is the number of pipeline stages in the MIPS processor design of Chapter 6?
    A. 0
    B. 1
    C. 5
    D. 20
    E. 50

18. (5 Points) What is the approximate number of pipeline stages in Intel Pentium IV processors?
    A. 0
    B. 1
    C. 5
    D. 20
    E. 50

19. (5 Points) Are you having fun yet?
    A. Yes, Dr. Vickery.

20. (5 Points) What is the purpose of using a memory hierarchy?
    A. It is a way to manage the tradeoffs between memory capacity, cost per bit, and access time.
    B. It is a way to optimize memory capacity without regard to cost or access time.
    C. It is a way to optimize memory cost without regard to capacity or access time.
    D. It is a way to optimize memory access time without regard to cost or capacity.
    E. It is a way of extending the non-volatile nature of disks to main memory and the registers.

21. (5 Points) What is the essential difference between SRAM and DRAM with regard to their roles in a memory hierarchy?
    A. SRAM has a larger capacity than DRAM with the same cost.
    B. DRAM is faster than SRAM with the same cost.
    C. SRAM is slower than DRAM with the same capacity
    D. DRAM is more expensive than SRAM with the same speed.
    E. SRAM is faster, smaller, and costs more per bit than DRAM.

22.   (5 Points) Which type of device would be most appropriate for a cache memory?
   A.   SRAM
   B.   DRAM
   C.   Flash
   D.   Disk
   E.   Tape

23.   (5 Points) Why is disk so much slower than RAM?
   A.   Because it is non-volatile (doesn't lose information when turned off).
   B.   Because of network delays.
   C.   Because of bus delays.
   D.   Because of mechanical delays.
   E.   It's not slower, it's faster.

For the following questions, assume a processor uses byte addressing, has a word size of 4 bytes, a virtual address space of 4GB bytes, and a physical address space of 1 GB. *Help:* All questions asking about cache refer to physical addresses, not virtual addresses. For all questions, the cache can hold 64K bytes, and each cache block can hold 16 bytes. Page size is 1024 words.

24.   (5 Points) How wide is a virtual address?
   A.   30 bits
   B.   32 bits
   C. 2^{30} bits
   D. 2^{32} bits
   E. 2^{32} bytes

25.   (5 Points) How wide is a physical address?
   A.   30 bits
   B.   32 bits
   C. 2^{30} bits
   D. 2^{32} bits
   E. 2^{30} bytes

26.   (5 Points) How many bits in the byte offset field of an address? (The answer is the same whether the address is physical or virtual.)
   A.   0
   B.   1
   C.   2
   D.   4
   E.   8

27.   (5 Points) How many cache blocks are there?
   A.   1K
   B.   2K
   C.   4K
   D.   8K
   E.   16K

28.   (5 Points) Assume the cache is 4-way set associative. How wide are the tag, cache index (set number), block-offset, and byte-offset fields, in that order from left to right?
   A.   16, 10, 2, 2
   B.   14, 12, 2, 2
   C.   12, 12, 4, 2
   D.   12, 12, 2, 4
   E.   8, 8, 8, 8

29.   (5 Points) Assume the cache is fully associative. How wide is the cache index field?
   A.   0 bits
   B.   10 bits
   C.   12 bits
   D.   14 bits
   E.   16 bits

30.   (5 Points) How many virtual pages are there?
   A.   10
   B. 2^{10}
   C.   20
   D. 2^{20}
   E. 2^{32}

31.   (5 Points) How many physical pages are there?
   A.   10
   B. 2^{10}
   C.   20
   D. 2^{20}
   E. 2^{18}

32.   (5 Points) How many entries are there in a page table?
   A.   10
   B. 2^{10}
   C.   20
   D. 2^{20}
   E.   18

33.   (5 Points) Under what conditions would a "valid" bit be false?
   A.   When there was overflow.
   B.   When there was an interrupt.
   C.   When a page or block has not yet been loaded into this level of the hierarchy.
   D.   Whenever there is a cache hit.
   E.   When the network is down.

34.   (5 Points) Under what condition would a "dirty" bit be true?
   A.   When a block or page has been modified but not written back to the next lower level of the
        hierarchy yet.
   B.   When a block or page is causing a pipeline stall.
   C.   When the TLB is empty.
   D.   When the TLB is full.
   E.   When there is no TLB available.

35.   (5 Points) What is a TLB?
   A.   A mechanism for translating disk addresses into cache block numbers.
   B.   A cache that holds part of a page table.
   C.   A page table that holds part of a cache.
   D.   A part of the kernel that is executed when there is an interrupt.
   E.   The smallest amount of information that can be read from or written to a disk.

36.   (5 Points) Which statement is true about a page table?
   A.   All programs share a single page table.
   B.   A page table is part of the CPU.
   C.   The kernel is responsible for putting information in the page table and the CPU uses the
        information to translate virtual addresses to physical addresses.
   D.   The CPU is responsible for putting information in the page table and the kernel uses the
        information to decide which process to schedule.
   E.   The page table is managed completely by the CPU, but the kernel can store a copy on disk when
        there is a mistake.

37.   (5 Points) If a hit takes one cycle to complete and the cache miss penalty is 100 clock cycles, what is the average number of clock cycles per memory access for a processor with a hit ratio of 0.9? (Pick the closest answer.)
   A.   100.9
   B.   10.90
   C.   1.900
   D.   9.100
   E.   9.010

38.   (5 Points) What is the advantage of an interleaved memory design?
   A.   The processor doesn't have to wait for the kernel to decide which page to load.
   B.   The kernel doesn't have to do a context switch when there is an interrupt.
   C.   The bandwidth between the processor and memory is higher.
   D.   The bandwidth between the memory and the disk is lower.
   E.   Multiple banks of memory can be accessed in parallel, which is faster.

39.   (5 Points) What is the bandwidth of a bus?
   A.   The sum of the number of address wires and the number of data wires.
   B.   The number of address wires.
   C.   The number of data wires.
   D.   The number of data wires times the clock speed of the bus.
   E.   The log base 2 of (the number of address wires plus the number of data wires).

40.   (5 Points) Which of the following statements is true?
   A.   A sector is the same thing as a page.
   B.   The sector size is always an integer multiple of the page size.
   C.   The page size is always an integer multiple of the sector size.
   D.   There is no relationship between the size of a page and the size of a sector.
   E.   Sectors are parts of physical memory, but pages are in cache.

41.   (5 Points) What is the purpose of reference bits?
   A.   They are used to tell whether a disk drive will have rotational delay or not.
   B.   They are used to tell whether a page needs to be written back to disk or not.
   C.   They are used to tell whether a cache block needs to be written back to RAM or not.
   D.   They tell whether a page is in SRAM or DRAM.
   E.   They are used in implementing Least Recently Used replacement algorithms.

42.   (5 Points) What is the difference between write back and write through?
   A.   With write back, main memory is updated immediately, but with write through, the memory is updated only when the block is replaced.
   B.   With write through, main memory is updated immediately, but with write back, the memory is updated only when the block is replaced.
   C.   With write through, the block is replaced when it is written, but with write back, it is replaced when it is read.
   D.   With write through, the block is replaced when it is read, but with write back, it is replaced when it is written.
   E.   They are two names for the same thing.

43.   (5 Points) What is a split cache design?
   A.   A design in which part of the cache resides in main memory.
   B.   A design in which part of the page table resides on disk.
   C.   A design with separate instruction and data caches.
   D.   A design with separate caches for disk and main memory.
   E.   A design with multiple cycles per cache access.

44.   (5 Points) A cache miss caused by the first access to a block that has never been in the cache is called:

    F.   A compulsory or cold start miss.
    G.   A capacity miss.
    H.   A conflict or collision miss.
    I.   A conflict of interest miss.
    J.   A missed opportunity.

45.   (5 Points) A cache miss that occurs in a set-associative or direct-mapped cache when multiple block compete for the same set and that are eliminated in a fully associative cache of the same size is called:

    A.   A compulsory or cold start miss.
    B.   A capacity miss.
    C.   A conflict or collision miss.
    D.   A conflict of interest miss.
    E.   A missed opportunity.

46.   (5 Points) Which statement describes page faults most accurately?

    A.   The kernel detects page faults, and signals the CPU so it can obtain the correct page from disk.
    B.   The CPU detects page faults, and signals the kernel so it can obtain the correct page from disk.
    C.   The disk detects page faults, and signals the kernel so it can obtain the correct page from the CPU.
    D.   The disk detects page faults, and signals the CPU so it can obtain the correct page from the kernel.
    E.   Page faults are ignored whenever performance is important.