Instructions

For the Breadth Exam, answer questions 1 through 4 and for the Depth Exam, answer questions 1 through 7. The questions are quite specific. If, however, some confusion should arise, be sure to state all your assumptions explicitly.

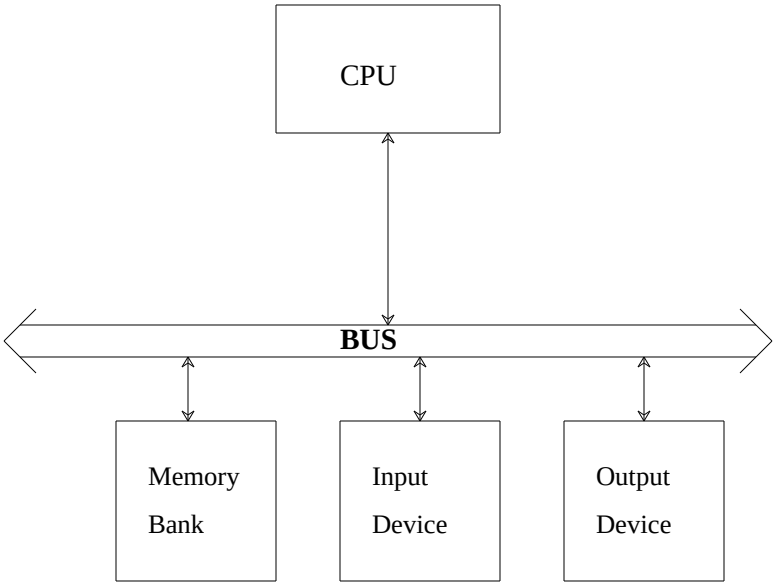
Breadth Exam

(1) What are ‘‘registers’’? What are ‘‘caches’’? What are the advantages and disadvantages of each? Why do most computers use both?

(2) The control logic for many CPUs is organized using microcode. What is the alternative to microcode? What are the advantages and disadvantages of using or not using microcode?

(3) You are designing a pipelined computer with multiple, pipelined functional units. The computer has a load/store, register-register instruction set. That is, all computation instructions have only register operands, while interactions with operands in memory are carried out with explicit load/store instructions. To generate memory addresses, the load/store instructions use a set of address registers that are distinct from the general-purpose data registers. The problem you are faced with is the choice of addressing modes for the load/store instructions. The choice of addressing modes you have is: i) absolute or direct mode, ii) register indirect mode, iii) displacement mode, iv) autoincrement mode, and v) autodecrement mode.

Which of the above addressing modes would hinder your ability to pipeline in the above architecture? For your answer, discuss in detail each mode and how it might/would not hinder a pipelined implementation of the above architecture.



(4) Consider a computer system with a CPU connected to a main memory bank, input devices and output devices through a bus, as shown in the figure above. The CPU is rated at 10 MIPS, *i.e.*, has a peak instruction execution rate of 10 MIPS. On the average, to fetch and execute each instruction, the CPU needs 40 bits from memory, 2 bits from an input device and sends 2 bits to an output device.

The peak bandwidths of the input and output devices are 3 Megabytes per second each. The bandwidth of the memory bank is 10 Megabytes per second, and the bandwidth of the bus is 40 Megabytes per second.

What is the peak instruction execution rate of the system as configured as above? What is the bottleneck in the system? Suggest *several ways* in which you might go about alleviating the bottleneck and improving the instruction execution rates.

Suggest one system organization where the CPU is likely to be the bottleneck.

Depth Exam

(5) A particular processor being designed is slated to execute instructions so fast that it is feared that one of the limits on speed will be simply incrementing the program counter (PC) in the time needed to issue an instruction. How can this problem be solved without compromising the speed of the processor?

(6) Since the early 1970s, most computers have had at least 3 levels in the memory hierarchy: i) a relatively small cache memory, ii) a somewhat larger main (or real or physical) memory, and iii) a large virtual memory.

Traditionally, the mapping from virtual to physical memory has been a fully-associative mapping, implemented with a page table. The mapping from physical to cache memory has been a set-associative one (with set associativities in the range 1-8), facilitated with a cache directory.

In recent years, thanks to advances in semiconductor memory technology, cache memories have become larger and have led to the consideration of caches with a very small degree of set-associativity, for example, direct-mapped caches. The main phenomenon that has led to this trend is that as caches become larger, the miss ratio is low enough and additional set-associativity will not be able to lower it further.

Today, main memories are also becoming quite large. Might we see a change in the virtual to physical mapping, i.e., the fully associative mapping being discarded in favor of a set-associative mapping? Why or why not?

(7) One reason many computers execute at less than their peak rate is the relatively slow execution of branches. Why are branches slow? Describe architectural and implementation techniques for mitigating the negative effect of branches.