(1) (4) As minimum feature sizes decrease, what happens to wire resistance? Why? Are wire delays or transistors more likely to be the most significant limit on clock frequency in the future?

(2) (2 pts) Clock rates have grown by a factor of 1000 while power consumed has only grown by a factor of 30. How was this accomplished?

(3) (2) Which metric should the architect worry about, power or energy? Why?

(4) (2) How does reducing the minimum feature size affect power density?

(5) (2) What is a "name" hazard/dependence? What is a "true" hazard/dependence?

(6) (3) What is the "threshold voltage"?

(7) (2) When dealing with control hazards, it is not enough to predict the branch direction - what else must we provide?

(8) (2) What is one of the simplest ways to exploit parallelism among instructions?

(9) (3) What is "Yield"? Is it important? Why or why not?

(10) (5) What is the difference between static and dynamic power consumption? Which one was the dominant contributor to overall power consumption 20 years ago? What is the situation now, and why has it changed?

(11) (4) What is a benchmark program? What is the **perfect** benchmark? What were the 4 types of benchmarks we discussed in class?

(12) (4) Register allocation is a very effective and useful compiler optimization. However, it cannot be used in certain situations for correctness reasons - explain why.

(13) (4) What are the 3 pipeline hazards? Which one can be solved by throwing money at it?

(14) (2 pts) What is a "balanced" pipeline?

(15) (5) In an MOS device, there is a gate, drain, and source. Briefly explain how this device works. Drawing a picture is recommended.

(16) (3) Why do most current chips feature multiple cores on the same die?

(17) (6) Why is branch prediction important? What performance enhancing techniques have made it so? List 3 examples of existing Static Branch Prediction strategies in order of (average) increasing effectiveness.

(18) (6) Suppose floating point instructions are responsible for 51% of the execution time of a particular benchmark, and the floating point square root (FPSQ) by itself is responsible for 21% of the execution time. You have a choice between speeding up the FPSQ instruction by a factor of 11, or making all floating point instructions run faster by a factor of 2. Which should you do? (Show how you set up the equations, although you do not need to actually solve them.)

(19) (4) Choosing the number of operands explicitly specified per instruction has an impact on the overall instruction size. Give an example of the smallest and the largest instructions, and what each is called.

(20) (4 pts) Processor A requires 540 instructions to execute a given program, uses 1 cycles per instruction, and has a cycle time of 10 ns. Processor B requires 3 cycle per instruction, but only requires 200 instructions to do the same program. What must the cycle time of Processor B be in order to give the same CPU time as Processor A?

(21) (4) Stack machines get their operands from the stack which is in memory, but it's possible to keep the top of the stack in registers or in the L1 cache. So in spite of the fact that the operands can be fetched in a single cycle, stack machines are still not popular. Why?

(22) (3) Why do control flow changing instructions use PC-relative addressing?

(23) (3) What is difference between caller saved and callee saved? Which one is usually used by conservative compilers?

(24) (4) Briefly describe the phase-ordering problem.

(25) (3) The book lists a number of things that the architect can do to help the compiler writer. Give three of them.

(26) (3) The Worst Case Path fixes the clock rate for a pipelined processor. What factors must be included when calculating the Worst Case Path?

(27) (3) What is forwarding, and why is it useful?

(28) (3 pts) Is it possible to have a WAW dependency in our 5-stage pipeline? Why or why not?

(29) (5) The standard MIPS has a 5-stage pipeline, and uses a load delay slot. If the machine is redesigned to be an 8-stage pipeline, with the following stages:

F D1 RR E1 E2 M1 M2 WB (where RR stands for Register Read)

Assuming the machine has bypass/forwarding logic, how many load delay slots will this new design require if the memory returns the value at the end of M2? At the end of M1?