

The following questions are based on the paper at http://www.intel.com/technology/itj/q12001/articles/art_2.htm and http://www.ece.lsu.edu/ee4720/s/hinton_p4.pdf (password needed off campus, will be given in class). See Homework 4 (<http://www.ece.lsu.edu/ee4720/2002/hw04.pdf>) for an introduction to the paper.

Problem 1: What is the maximum IPC of the IA-32 (in μops)? Put another way, the Pentium 4 is an n -way superscalar processor, what is n ?

Problem 2: The Pentium 4 can decode no more than one IA-32 instruction per cycle. How then can it execute more than one IA-32 instruction per cycle (at least for small code fragments prepared by a friendly programmer)?

Problem 3: One problem with superscalar systems noted in class is the wasted instructions following the delay slot of a taken branch near the beginning of a fetch group. How does the Pentium 4 avoid this?

Problem 4: The fast ($2\times$) integer ALUs have three stages, an initiation interval of 1 fast cycle ($\frac{1}{2}$ processor cycle), and a latency of zero fast cycles. Why is this surprising (not the one half part)? How does it do it?

Problem 5: In describing store-to-load forwarding the paper describes a special case for which data could be forwarded (bypassed) but is not because it would be too costly. Using MIPS code (or IA-32 if you prefer) provide an example of this special case.

Problem 6: In Figure 8 the performance of a 1 GHz Pentium III is compared to a 1.5 GHz Pentium 4. Why is it reasonable for the Pentium 4 to be compared at a higher clock frequency?