# CMPT 295 Assignment 8 Solutions (2%)

1. [6 marks] *Latency and Throughput Bounds*

   (a) [2 marks] To be able to issue every clock cycle makes sense if the function unit is pipelined. In other words, each multiplication is broken into a sequence of 2 stages, and because of pipelining, there can be 2 multiplications in progress at a time, each at a different stage.

   In the case of `fmul`, where $I = L = 4$, the function unit is not pipelined, i.e., it has one stage with a very long propagation delay. That's the worst possible case. The most time you ever need to pay is the propagation delay, i.e., once the last computation has finished, the next one can immediately begin. Therefore, $I > L$ is not possible.

   (b) [2 marks] The latency bound is simply the latency of each operation, i.e., 2 CPI for `mul` and 4 CPI for `fmul`. These translate into 1.0 GIPS for `mul` and 0.5 GIPS for `fmul`.

   The throughput bound is the issue divided by the capacity, i.e., 0.5 CPI for `mul` and 2.0 CPI for `fmul`. Or 4.0 GIPS for `mul` and 1.0 GIPS for `fmul`.

   (c) [1 mark] When there are no data dependencies, the CPU can reach its throughput bound, i.e., 4.0 GIPS for `mul` and 1.0 GIPS for `fmul`.

   (d) [1 mark] When the calculations are sequential, the CPU can do no better than its latency bound, i.e., 1.0 GIPS for `mul` and 0.5 GIPS for `fmul`.