2.2. Superscalar Processors (10 points)

Modern CPUs contain multiple cores and each core has multiple copies of the same physical resources, such as execution units, issue control or data paths. This allows the execution of more than 1 instruction per core and CPU cycle. For example, the new Intel(R) Core(TM) i9-7900X CPU contains 10 cores, where each core allows the execution of 2 simultaneous threads. Its CPU Core Pipeline functionality is based on the Skylake Microarchitecture with a 6-wide superscalar design per thread able to execute 6 floating point operations per cycle . Please give a brief explanation of each answer.

1. Calculate the theoretical peak performance (in FLOPS – floating point operations per second) of an Intel(R) Core(TM) i9-7900X processor, having a clock frequency of 3.3 GHz.
2. What would be the theoretical execution time (in seconds), on the machine described previously, of calculating the dot products of two full nxn matrices, one having 10 4 x 10 4 elements, represented as floating point numbers with single precision?
3. Consider this processor running at 3.3 GHz. Assume that the average CPI (clock cycles per instruction) is 1. Assume that 10% of all instructions are stores, and that each store writes 8 bytes of data. How many processors will a 800-GB/s bus be able to support without becoming saturated?

By Jiahui Tang

(1)

- 1 cpu contains 10 cores;
- 1 core allows 2 threads;
- 1 thread 6 flops per cycle;
- Clock Frequency 3.3 GHZ/second;

Thus we have $3.3 * 10 * 2 * 6 = 396$ GFLOPS/sec as the theoretical peak performance of this intel(R) Core(TM) i9-7900X processor.

(2) dot products of two full $n * n$ matrics:

For each element in the new $n * n$ matrix, it requires n product + $(n - 1)$ addition

Thus in total it requires computations of

$$n^2 * (2n - 1) = 2n^3 - n^2$$

and we have

$$n = 10^4$$

It requires approximately $= 2n^3 - n^2 = 2 * 10^{12} - 10^8$ operations

Thus, in total, the theoretical execution time is

$$\frac{2 * 10^{12} - 10^{8}}{396 * 10^{9}} = 5.0503$$

seconds

(3) For the bus, the problem clearly stated it stores 8 bytes/instruction (regardless of it's AVX store or normal store)

Thus we have

$$\frac{800}{8} = 100G$$

store instructions/sec

We could have

$$\frac{100}{10\%} = 1000G$$

instruction/sec at maximum for not becoming saturated.

As the i9 processor has

$$3.3ghz * 10 * 2 = 66G$$

instructions/sec

(Note: doesn't need to multiply 6 here, as 6 is floating point operation but not instruction)

We could support at maxium

$$1000G/66G = 15.15$$

--> we can only support 15 Intel(R) Core(TM) i9-7900X CPU processors but not 16 for not being saturated

In [ ]: