

Assignment 1

Abhishek Chandratre
achand12@uncc.edu

Microbenchmark: Maximum Flops and Iops

1. Expectation of Performance

a. Max flops expected from the machine

- i. Number of Cores = $8 * 2 = 16$
- ii. Number of Instruction per FMA = 2
- iii. Number of FMA = 2
- iv. Clock = 3.2GHZ
- v. Number of Floats per instruction = $250/32 = 8$

1. Result = 1.638 TeraFlops

b. Max iops expected from the machine

- i. Number of Cores = $8 * 2 = 16$ (VECTOR INSTRUCTIONS)
- ii. Number of Instruction = 1
- iii. Number of Vector Int = 3
- iv. Clock = 3.2 GHz
- v. Number of Int per instruction = 8

1. 1.228

- vi. Number of Cores = $8 * 2 = 16$ (INTEGER ALU)
- vii. Number of Instruction = 1
- viii. Number of Integer ALU = 1
- ix. Clock = 3.2 GHZ

1. 51.2

2. Total = 1.228 + 51.2 = 1.279 Teralops

2. Realization

a. Write code to get peak flops

- i. Code attached to submission

b. Write code to get peak iops

- i. Code attached to submission

3. Measurement

a. How many flops code achieved?

- i. **1.698 TeraFlops**

b. How many iops code achieved?

- i. **0.968 Teralops**

c. Does that match expectation? Where does this discrepancy come from?

- i. TeraFlops match expectation, Teralops do not match the expectation.

- ii. Regarding integers we have 3 ports which can do vector integer operations and one port which can do integer / branch operation. While calculating theoretical value, we calculate using 3 vector instruction and 1 scalar instruction. In reality, we need a for loop which has two scalar operations namely, conditional (branch) and increment(scalar integer operation), thus we cannot fully utilize all the three vector instruction ports at a given time.
- iii. If we try to load all three vector unit and ignore two instruction needed for for loop, then I got **0.7 Teralops**, which is worst that performing two vector and two for instruction which gives me **0.9 Teralops**
- d. Can you do better?
 - i. If we add more port we can achieve more Flops or lops. For lops we can if instead of having port 5 do vector and integer alu in Haswell, If these operations were are separate port then we definitely could have reach near the theoretical value.

4. Reporting

- a. Write short report that explain your technique, explaining and finding
 - i. Technique for finding Flops.
 - 1. Used two FMA in a loop
 - 2. Used asm("") to make sure no loop optimization occurs
 - 3. Printed values after the loop for same reason
 - 4. Also used omp parallel.
 - ii. Technique for finding lops.
 - 1. Used two vector instruction additions in a loop
 - 2. Used asm("") and printing of data to avoid loop optimization.
 - 3. Also used omp parallel
- b. Submit archive of report and code.
 - i. Attached to submission