

## 과제 #2

M1522.006700 확장형 고성능 컴퓨팅 (001)  
M3239.005400 데이터사이언스를 위한 컴퓨팅 2 (001)

박찬정

서울대학교 전기정보공학부  
2023-24013

### 1 Theoretical Peak Performance of CPU

Fig. 와 같이 계산 노드의 CPU 정보를 얻었다.

- (a) Intel(R) Xeon(R) Silver 4216 CPU @ 2.10GHz
- (b) 2개의 CPU가 장착되어 있다.
- (c) Base clock frequency는 2.10GHz이고, boost clock frequency는 3.20 GHz이다. Base clock과 boost clock은 각각 CPU가 동작할 수 있는 최소, 최대의 clock이며, CPU는 현재 수행중인 workload가 필요로 하는 연산 성능을 판단하여 해당 범위 내에서 clock speed를 조정한다. CPU가 boost clock frequency로 동작하기 위한 조건으로는, 현재 수행중인 workload가 높은 성능을 필요로 하고, 공급 되는 전력이 충분하며, CPU의 온도가 너무 높지 않아야 한다.
- (d) Physical core는 32개, logical core는 64개이다. Theoretical peak performance 계산을 위해서는 physical core 개수인 32를 사용해야 하며, 그 이유는 logical core는 하나의 physical core가 동시에 여러 thread를 연산할 수 있도록 하는 기술(i.e. Hyper-Threading, SMT 등)을 적용하여 얻어진 가상의 개념이므로, 실제 연산의 성능은 결국 physical core에 달려 있기 때문이다.
- (e) Intel(R) Xeon(R) Silver 4216 CPU 는 1개의 AVX512 FMA 유닛을 갖고 있다. 이를 바탕으로 한 clock cycle당 실행되는 AVX512 instruction의 수는 1개라고 가정하자.  
AVX512 instruction이 하나당 수행할 수 있는 FP32 연산의 수는  $512/32 = 16$ 개이다.  
따라서 하나의 core가 한 clock cycle에 수행할 수 있는 FP32 연산의 수는 16개이다.
- (f) Theoretical peak performance를 다음과 같이 계산할 수 있다.

$$\begin{aligned} R_{peak} &= \frac{(\# \text{ of FP32 operation})}{(\# \text{ of core})(\# \text{ of clock})} \times (\text{clock frequency}) \times (\# \text{ of core}) \\ &= 1 \times 3.2 \times 32 \\ &= 102.4 \text{ (GFLOPS)} \end{aligned}$$

### 2 Matrix Multiplication using Pthread

•