Consider the following code segment, with the values MAX = 2^{20} and WSIZE = 5 and the following data:

- Read from x takes 10 cycles (accounts for lw add)
- Read from w takes 10 cycles (accounts for lw add)
- Write to a register takes 1 cycle.
- Multiply takes 10 cycles (accounts for writing to the register)
- Each addition takes 5 cycles (accounts for writing to the register)
- Write to y takes 10 cycles
- Each comparison takes 5 cycles
- L1 cache hits require 5 cycles
- Cache miss penalty is 10 cycles

```
for( i = 0; i < MAX; ++i ){
        t = 0;
        for( j = 0; j < WSIZE; ++j ){
                t += x[i+j]*w[j];
                y[i] = t;
        }
}
```

a) Without any modifications, how many cycles does the code segment take?

for i: i=0 ⇒ Once, i < MAX ⇒ 5×2^{20}, ++i ⇒ 5×2^{20}, t=0 ⇒ 1×2^{20} = $11\times2^{20}+1$ cycles

for j: j=0 ⇒ Once ⇒ 2^{20}, j < WSIZE ⇒ $5\times5\times2^{20}$, ++j ⇒ $5\times5\times2^{20}$ ⇒ 51×2^{20} cycles

x[i+j] ⇒ i+j ⇒ $5\times5\times2^{20}$, w[i+j] ⇒ $10\times5\times2^{20}$, w[i] = $10\times5\times2^{20}$,

Multiply: $10\times5\times2^{20}$, Add: $5\times5\times10^{20}$

y: $10\times8\times2^{20}$

$(11+51+25+50+50+50+25+50)\times2^{20}+1$

$312\times2^{20}+1$ cycles

Consider the following code segment, with the values MAX = 2^{20} and WSIZE = 5 and the following data:

- Read from x takes 10 cycles (accounts for lw add)
- Read from w takes 10 cycles (accounts for lw add)
- Write to a register takes 1 cycle.
- Multiply takes 10 cycles (accounts for writing to the register)
- Each addition takes 5 cycles (accounts for writing to the register)
- Write to y takes 10 cycles
- Each comparison takes 5 cycles
- L1 cache hits require 5 cycles
- Cache miss penalty is 10 cycles

```
for( i = 0; i < MAX; ++i ){
        t = 0;
        for( j = 0; j < WSIZE; ++j ){
                t += x[i+j]*w[j];
                y[i] = t;
        }
}
```

Load w(0) from w(4) into registers

50 cycles total

b) Discuss the opportunities for reduction of cycle times and determine the improvement in performance when implemented in SIMD with 8 processors → Assume strong scaling

for i → Still $10 \times 2^{20} + 1$ cycles

$t = 0 \rightarrow 1 \times 2^{20}$

Load x(i) → x(i+4) into cache → Reduces load from 10 to 5

$t += x[i] \times w0 \rightarrow (5 + 10) \times 2^{20} = 15 \times 2^{20}$ cycles

Other 4: $t += x[i+\alpha] \cdot w\alpha \Rightarrow (5 + 5 + 10 + 5) \times 2^{20} = 25 \times 2^{20}$

i+α, load x, multiply, + t+=

$4 \times 25 \times 2^{20} = 100 \times 2^{20}$ cycles

Cache miss, every cycle: 10×2^{20}

$y(i) = t \Rightarrow 10 \times 2^{20}$ cycles

Sum: $(10 + 1 + 15 + 100 + 10 + 10) \times 2^{20} + 50 + 1 \Rightarrow 146 \times 2^{20} + 51$

Non-parallel improvement

$$\frac{312}{146} \approx 2.14\times$$

Parallel w/ Strong Scaling

$$\frac{312}{(146/8)} = 17.09\times$$