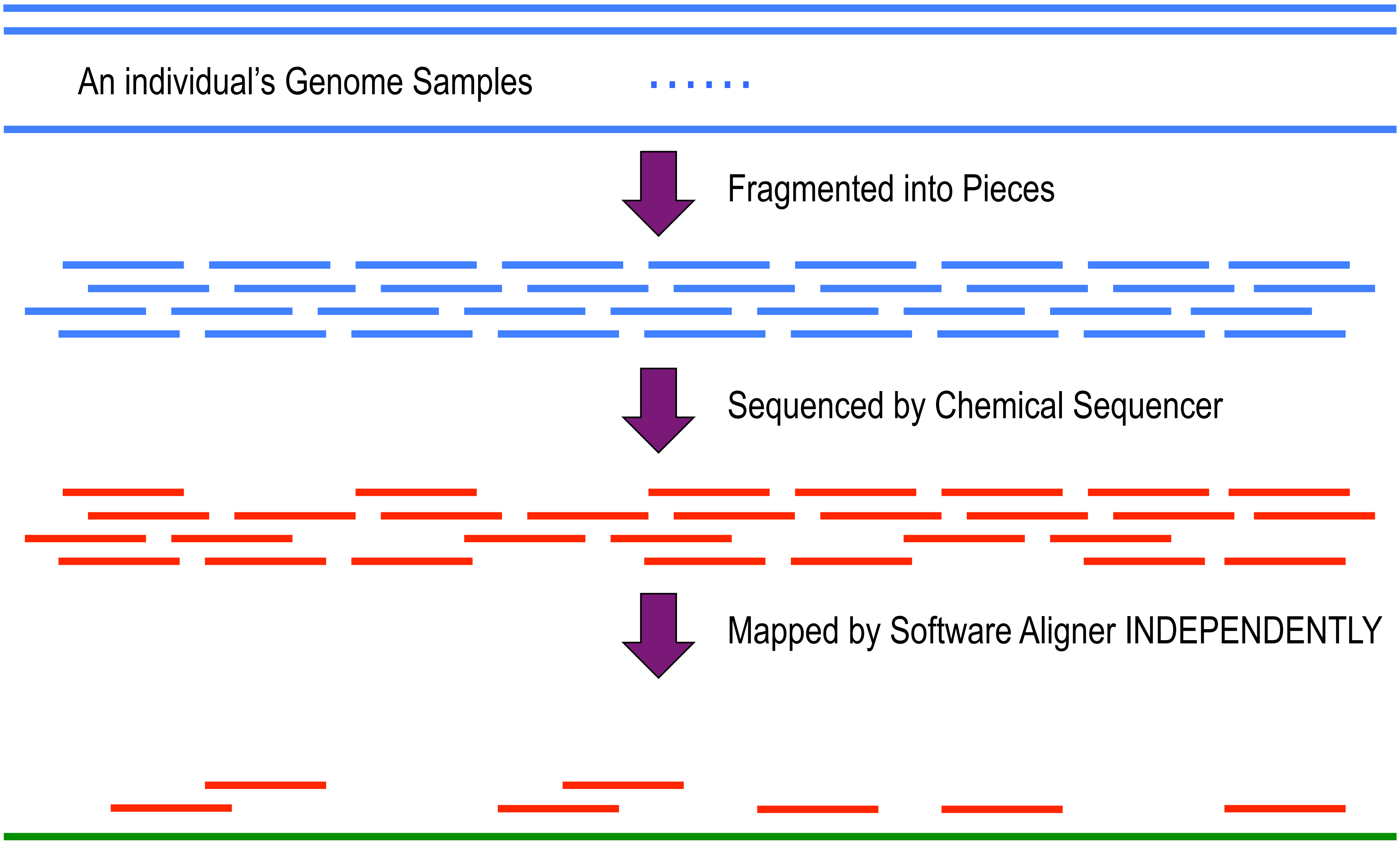
UCLA COMPUTER SCIENCE DEPARTMENT

Center for Domain Specific Computing

A Novel High-Throughput Acceleration Engine for Read Alignment

Yu-ting Chen, Jason Cong, Jie Lei, **Peng Wei**

Read Alignment: An Approximate Matching Problem



An individual's Genome Samples ……

Fragmented into Pieces

Sequenced by Chemical Sequencer

Mapped by Software Aligner INDEPENDENTLY

However…it is not as easy as we thought

- In theory

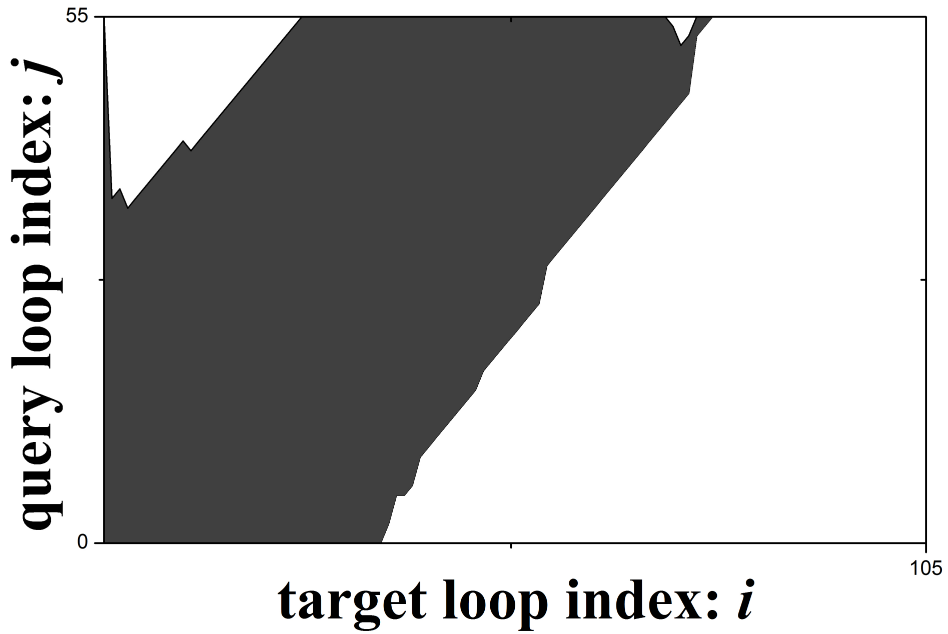
$$H(i,j) = \max \begin{Bmatrix} 0 & \\ H(i-1,j-1) + w(a_i,b_j) & \text{Match/Mismatch} \\ H(i-1,j) + w(a_i,-) & \text{Deletion} \\ H(i,j-1) + w(-,b_j) & \text{Insertion} \end{Bmatrix}, 1 \le i \le m, 1 \le j \le n$$

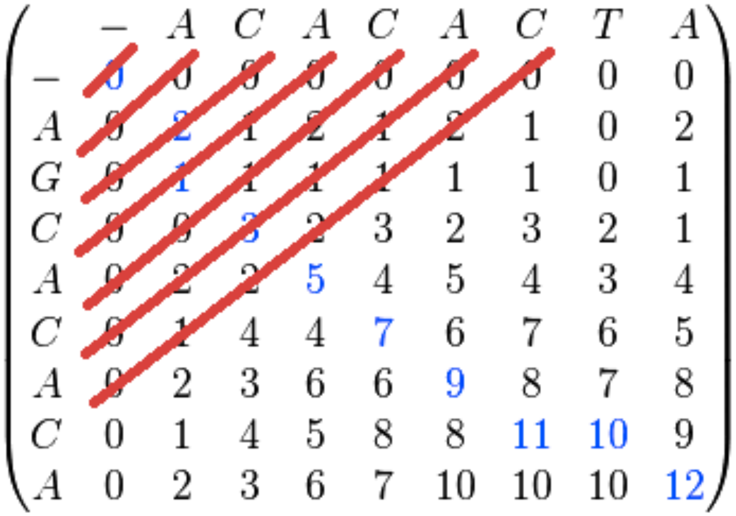
~ 30 Lines of C Code, Simple, Well-shaped, Anti-diagonal Parallelism, Well Studied

- In practice…

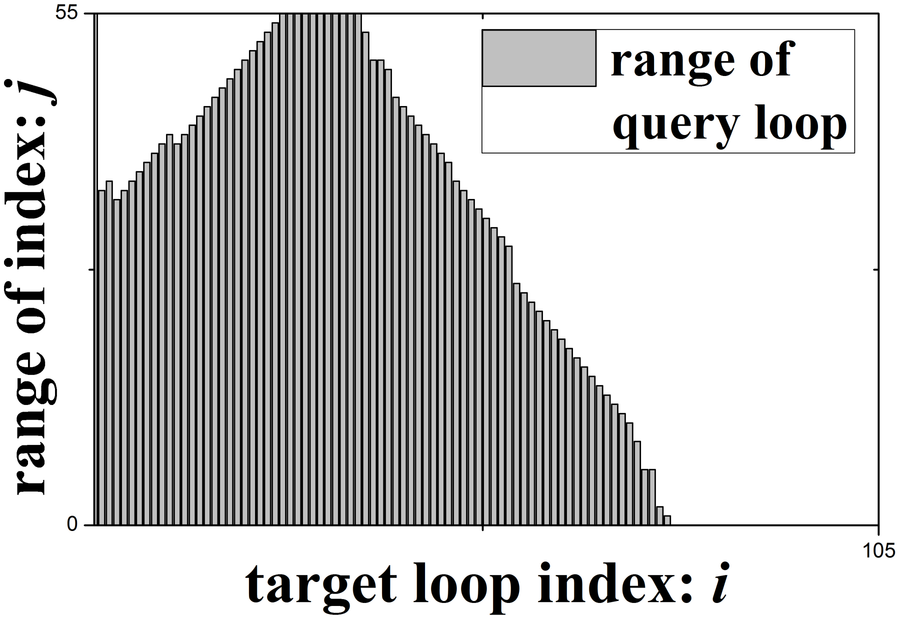
~ 150 Lines of C Code, Complicated, Full of Dependencies, No Parallelism, Existing work can not even achieve the same results

Challenge 1: Bioinformaticians have done something we dislike



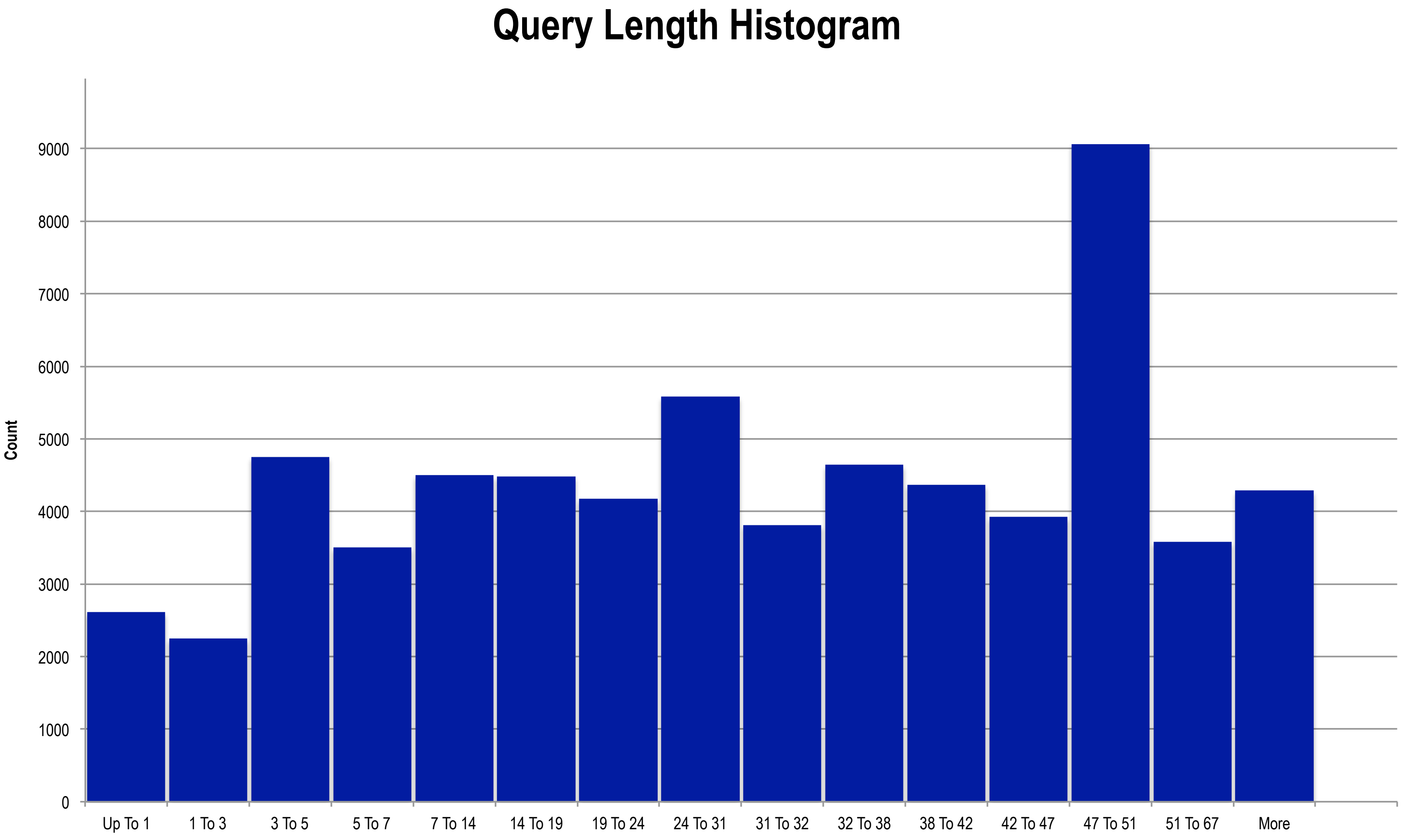


$$\left(\begin{array}{c|ccccccccc} & - & A & C & A & C & A & C & T & A \\ \hline - & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} & 0 & 0 \\ A & \cancel{0} & \cancel{2} & \cancel{1} & \cancel{2} & \cancel{1} & \cancel{2} & 1 & 0 & 2 \\ G & \cancel{0} & \cancel{1} & \cancel{1} & \cancel{1} & \cancel{1} & 1 & 1 & 0 & 1 \\ C & \cancel{0} & \cancel{0} & \cancel{3} & \cancel{2} & 3 & 2 & 3 & 2 & 1 \\ A & \cancel{0} & \cancel{2} & \cancel{2} & 5 & 4 & 5 & 4 & 3 & 4 \\ C & \cancel{0} & \cancel{1} & 4 & 4 & 7 & 6 & 7 & 6 & 5 \\ A & \cancel{0} & 2 & 3 & 6 & 6 & 9 & 8 & 7 & 8 \\ C & 0 & 1 & 4 & 5 & 8 & 8 & 11 & 10 & 9 \\ A & 0 & 2 & 3 & 6 & 7 & 10 & 10 & 10 & 12 \end{array}\right)$$

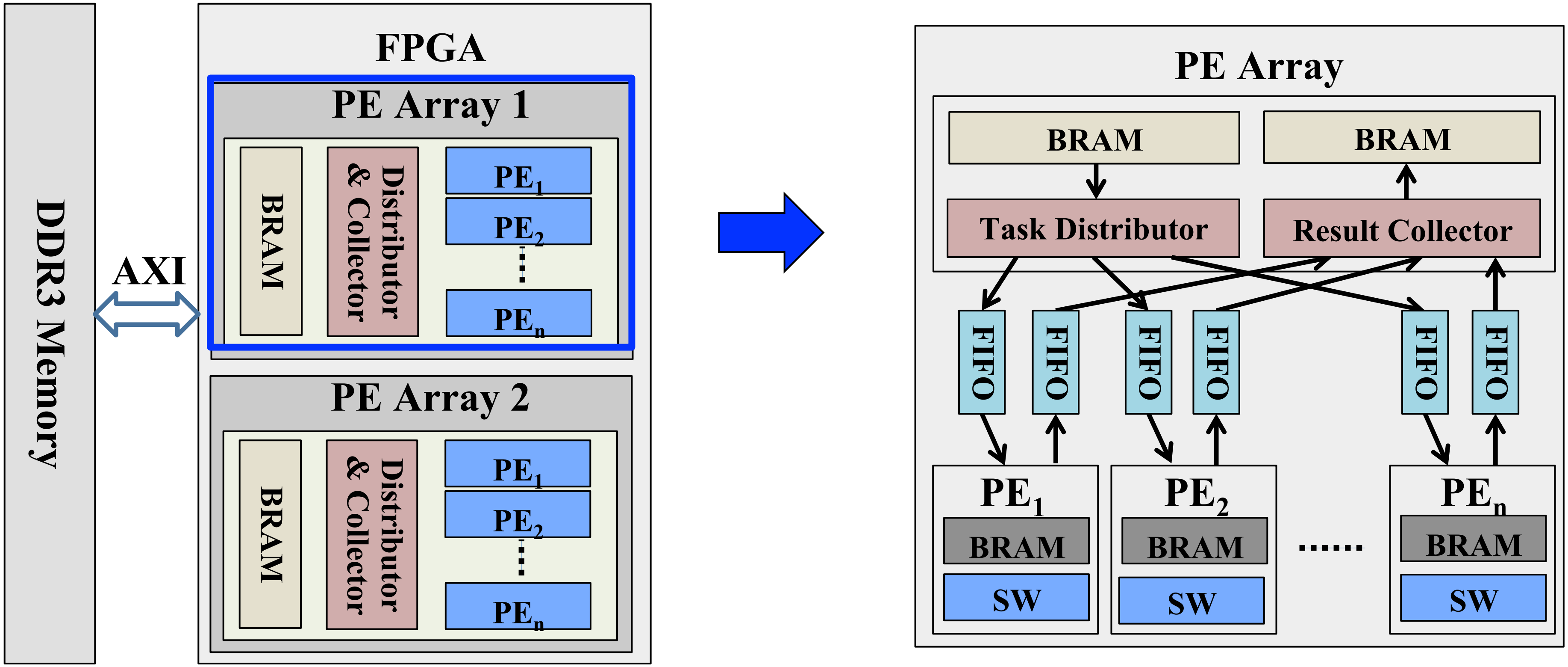


- Calculate ~50% of Matrix[i][j]-s
- Anti-diagonal parallelism disappears.
- It changes the results, raising a correctness issue.

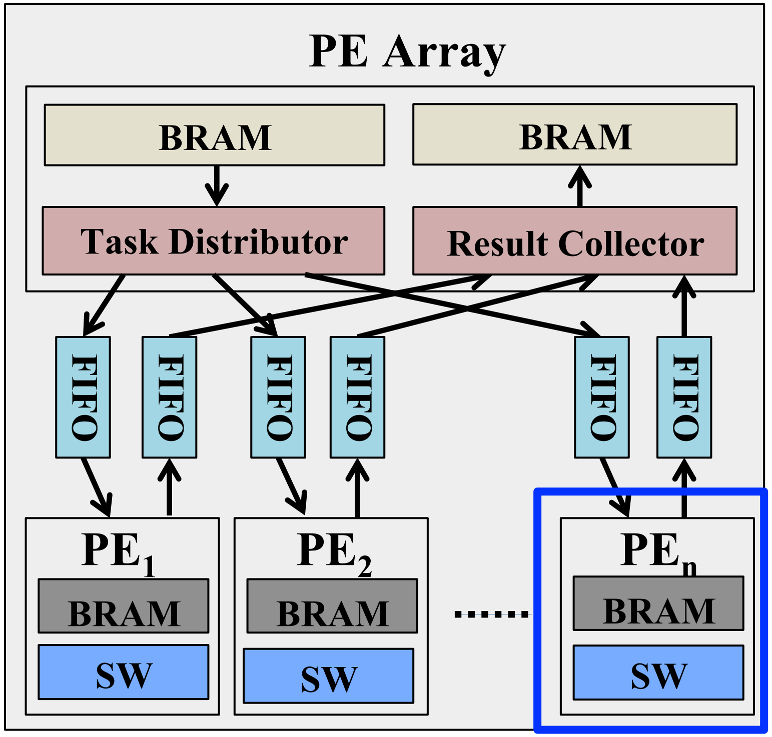
Challenge 2: Input Size Varies Dramatically

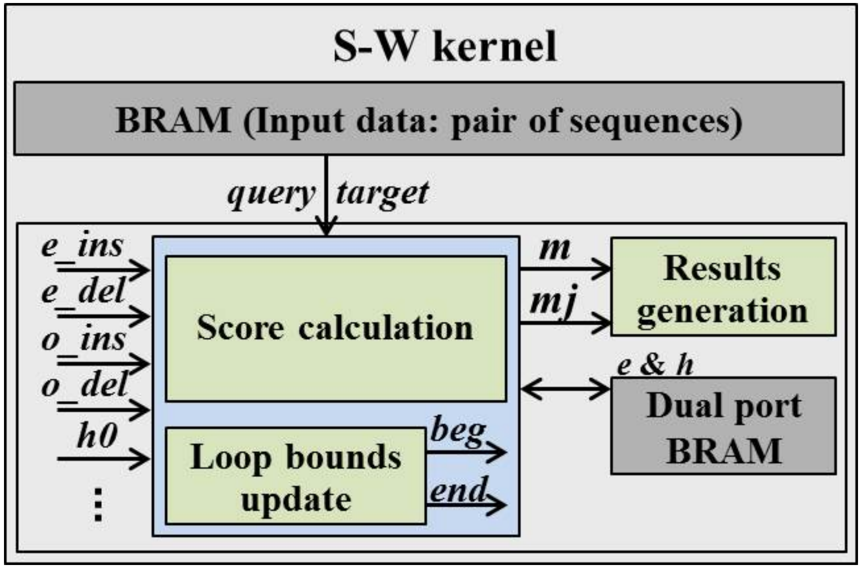


Accelerator Architecture



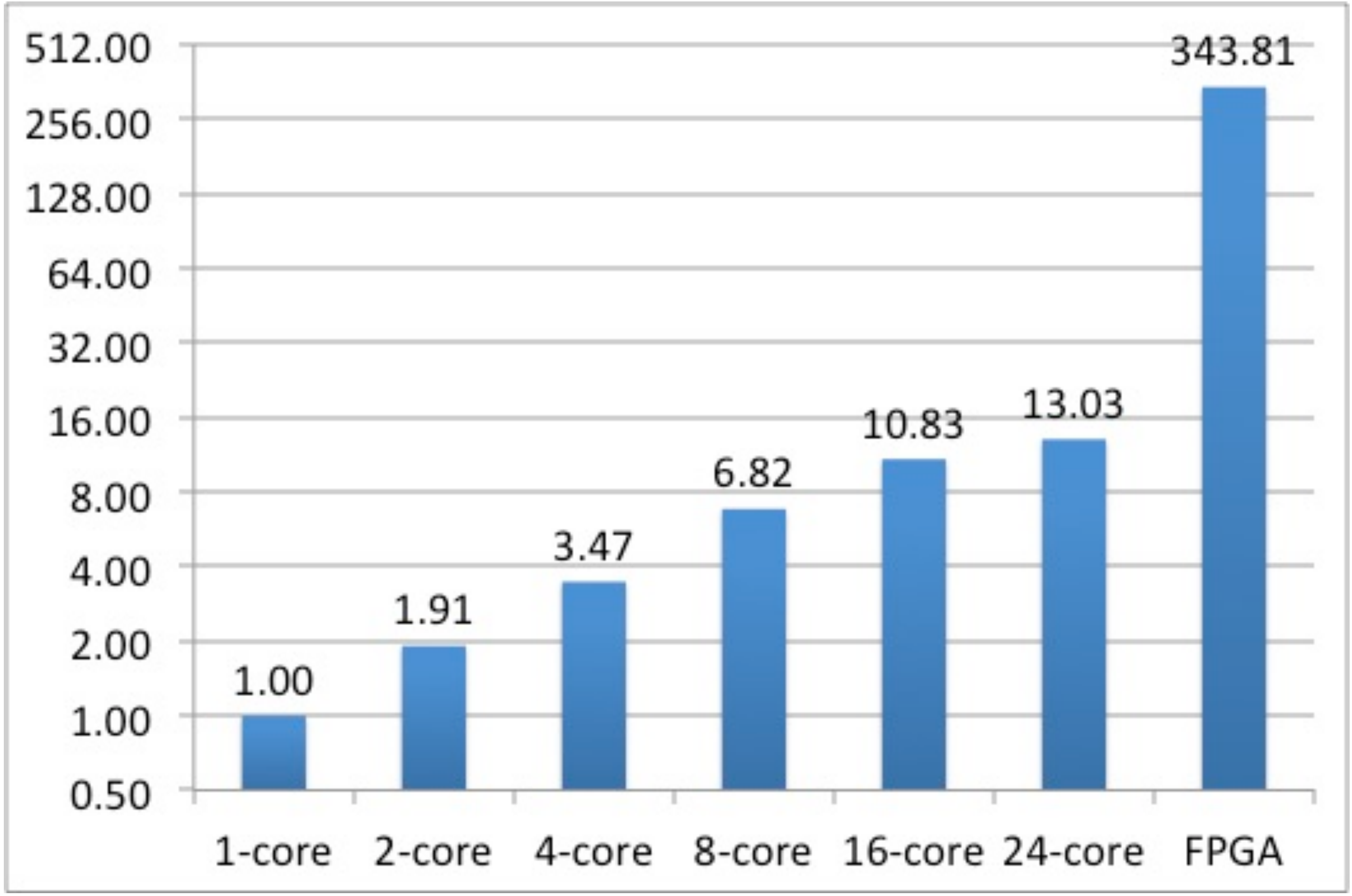
Accelerator Architecture





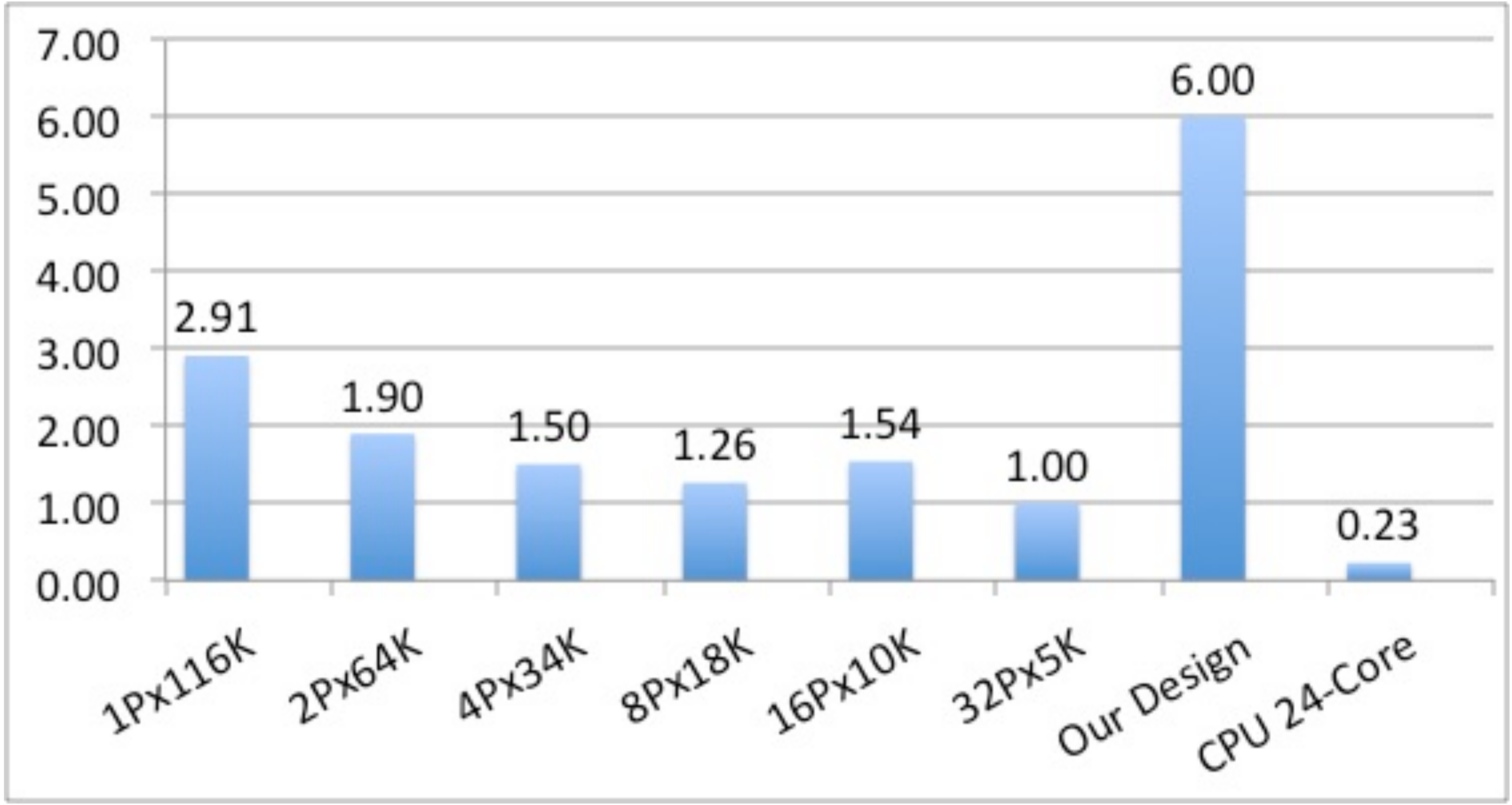
Experimental Results

- **Performance Improvement over CPU**



Experimental Results

- **Performance Improvement Over Wavefront-based Designs**



THANKS!