Sequential vs Parallel Processing Complexity

Sequential Processing: O(n)

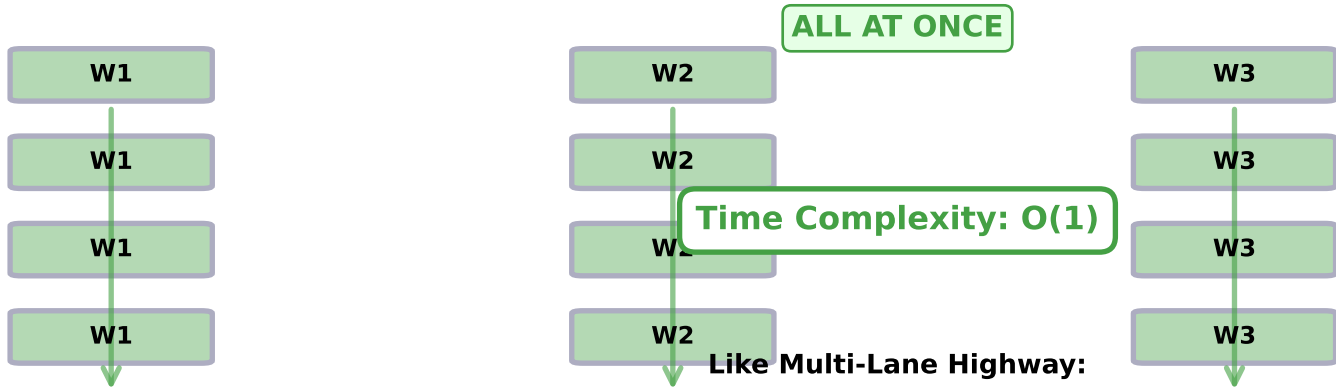| W1 | → | W2 | → | W3 | → | W4 | → | W5 |

WAIT   WAIT   WAIT   WAIT
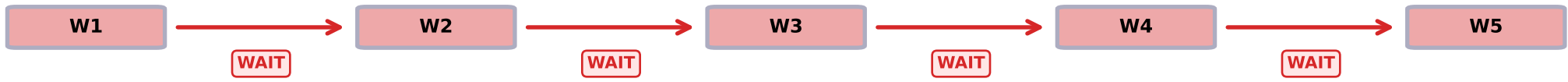
Time Complexity: O(n)

Like Single-Lane Highway:

Parallel Processing: O(1)

ALL AT ONCE

W1	W2	W3
W1	W2	W3
W1	W2	W3
W1	W2	W3
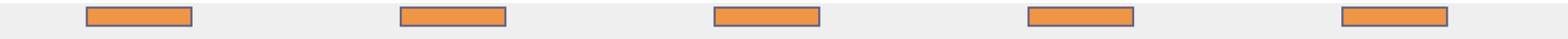
Time Complexity: O(1)

Like Multi-Lane Highway:

GPU Core Utilization

100x Speedup Achieved!

100%
50%
0%

1%
RNN
(Sequential)

5%
RNN+Attention
(Hybrid)

90%
Transformer
(Parallel)