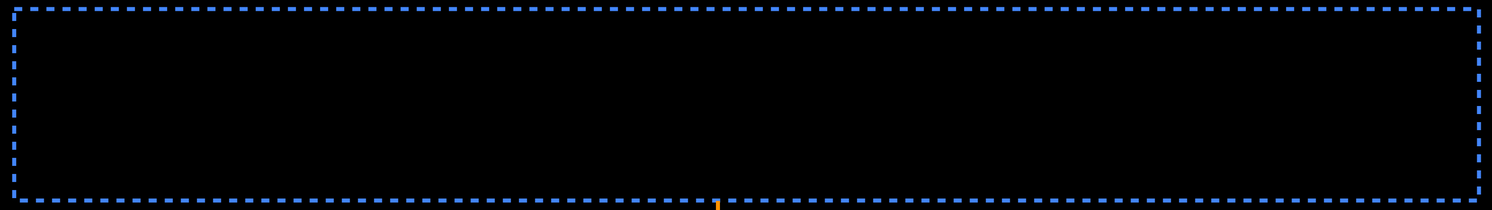
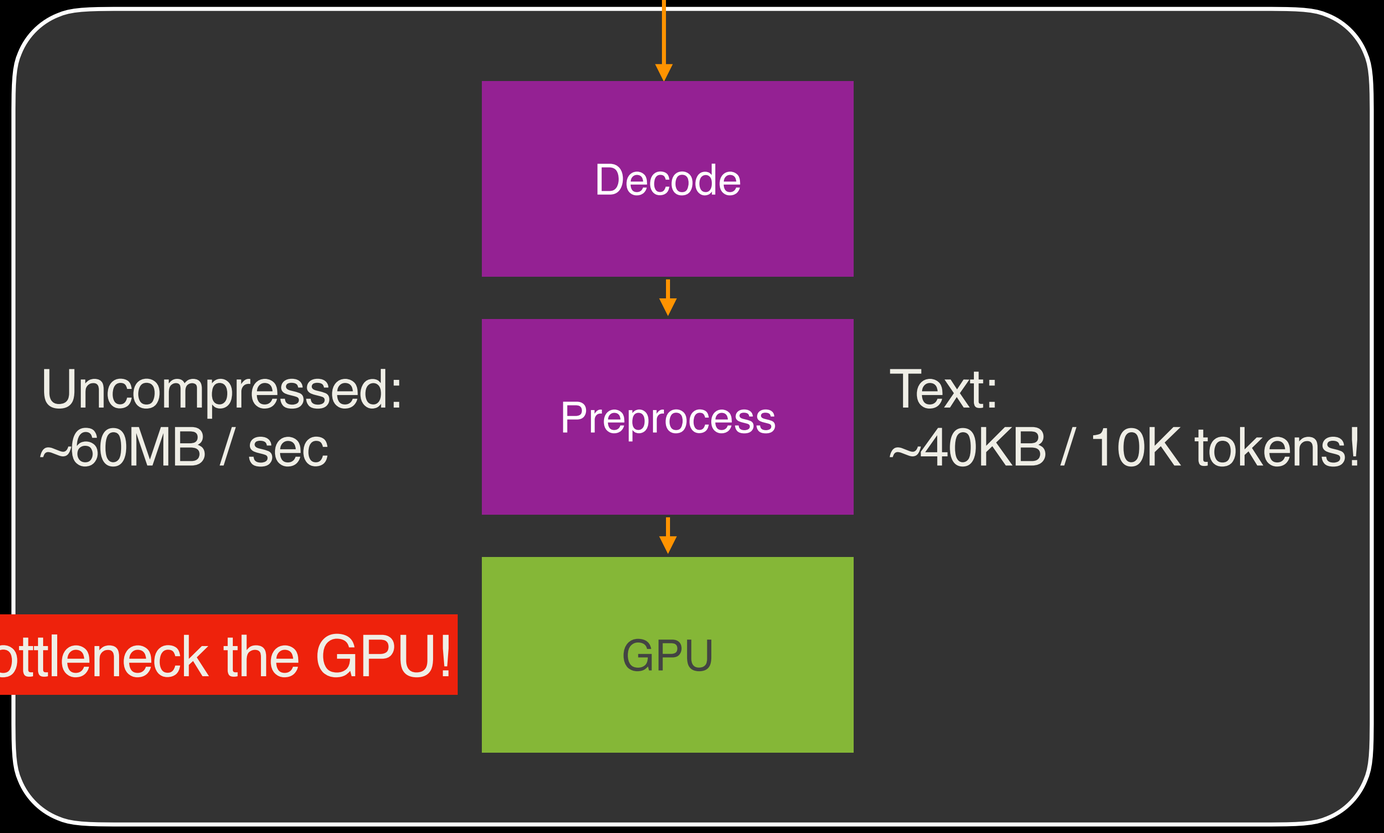


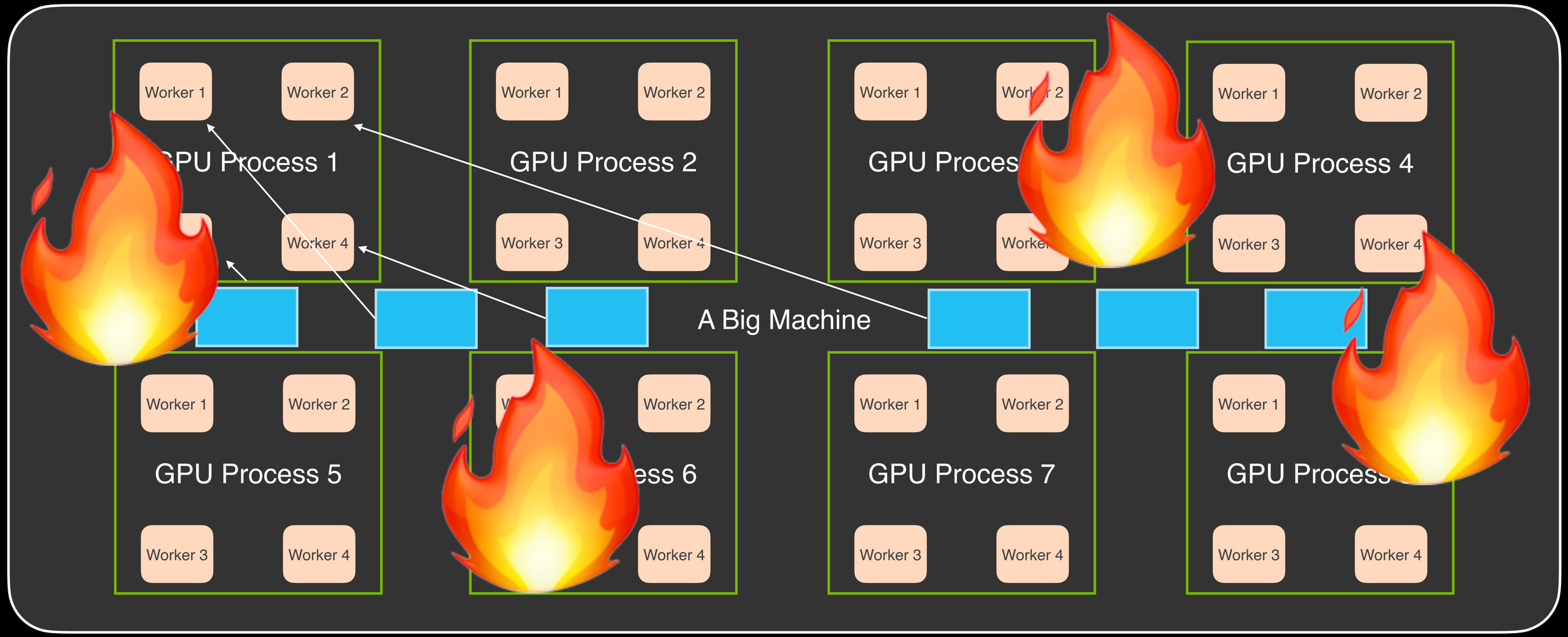
Compressed data size:
O(10KB) / 1080p frame
O(1MB) / sec



Uncompressed:
~60MB / sec

Text:
~40KB / 10K tokens!

Don’t Bottleneck the GPU!


Worker 1
Worker 2
GPU Process 1
Worker 3
Worker 4
GPU Process 2
GPU Process 4
A Big Machine
GPU Process 5
GPU Process 7