

Wikipedia-1M, quantized, 32 threads

