



GEMM-M=1024, GEMM-N=4096 (BERT-large)