

block_k = 64×fp16

LDS 32 banks