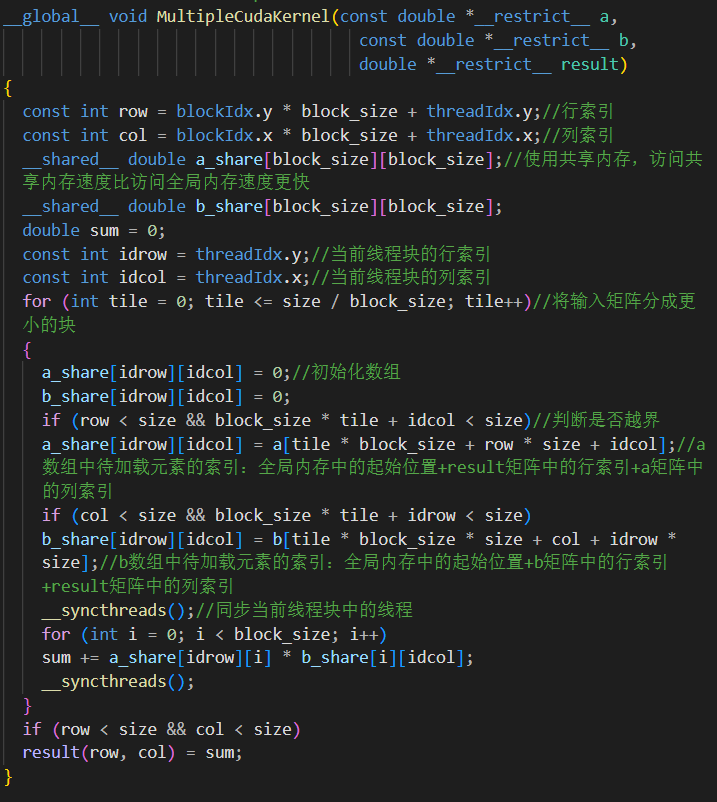
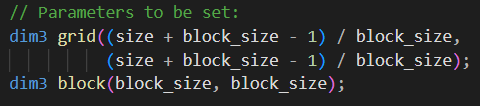
**实验三：GEMM 通用矩阵乘法实验报告**

**3220105775王傲哲**

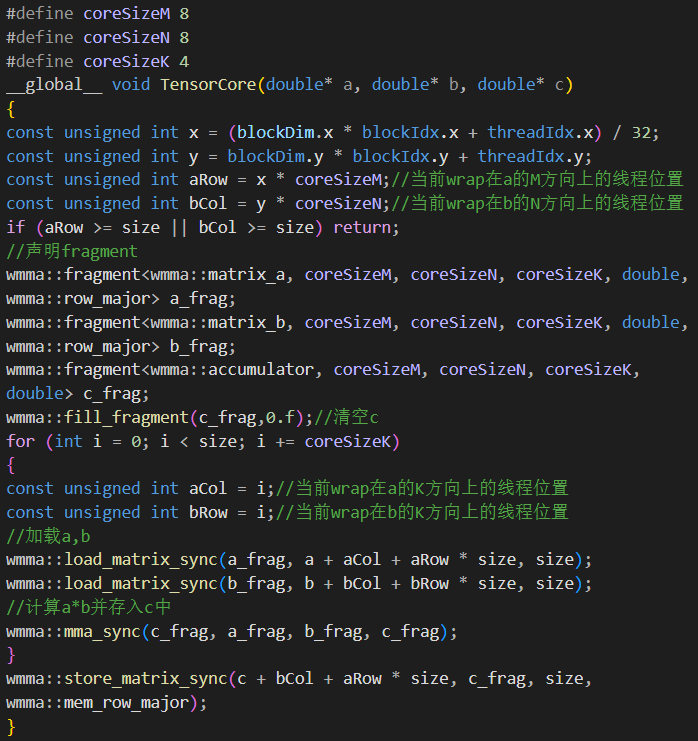
1. 源代码及思路

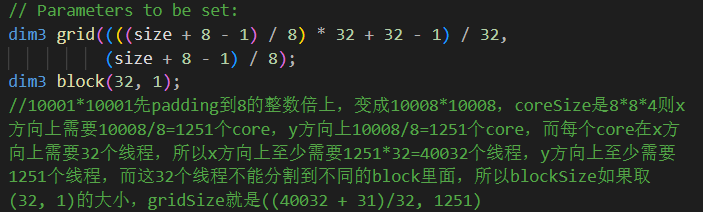
基础部分(见<baseline.cu>)：

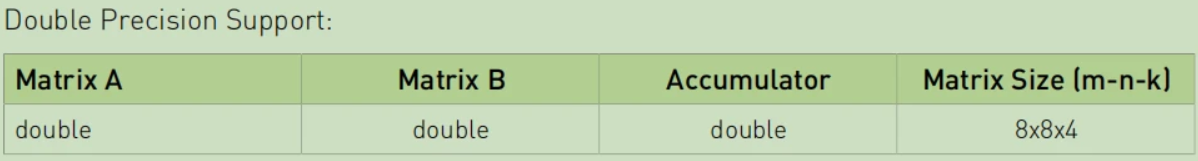




Tensor Core部分(见<baseline2.cu>):







1. 编译参数

在baseline.cu同目录下创建Makefile文件，终端输入make指令进行编译

基础部分：

nvcc baseline.cu -o gemm -lcublas -O3 -cudart=shared -Xcompiler -fopenmp -arch=sm\_75

Tensor Core部分：

nvcc baseline.cu -o gemm -lcublas -O3 -cudart=shared -Xcompiler -fopenmp -arch=sm\_80

1. 运行参数

基础部分：

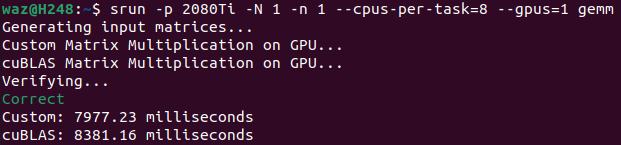
srun -p 2080Ti -N 1 -n 1 --cpus-per-task=8 --gpus=1 gemm

Tensor Core部分：

srun -p A100 -N 1 -n 1 --cpus-per-task=8 --gpus=1 gemm

1. 相对性能

基础部分：



Relative Performance = 0.9518

Tensor Core部分：



很遗憾在截止日期前未成功使用Tensor Core通过测试，经查询可能是可能是因为指针没有与处理器所需的边界对齐造成的，虽然暂时不理解为什么报错处是CUDA\_CALL(cudaDeviceSynchronize())，但我会在空余时间里继续调试直至成功。