

week6实验记录

zxp

April 13, 2024

1 environment

实验1 cpu:Inter i5-12400f (2.5 GHz)‘

gpu:rtx4060ti

System:Ubuntu 22.04.1

Compiler:gcc 12.3

实验2 cpu:Intel(R) Xeon(R) Gold 6330 CPU @ 2.00GHz (使用时申请了56个核心并且使用独占指令)

System:CentOS7

Compiler:9.5 和icpc

2 code

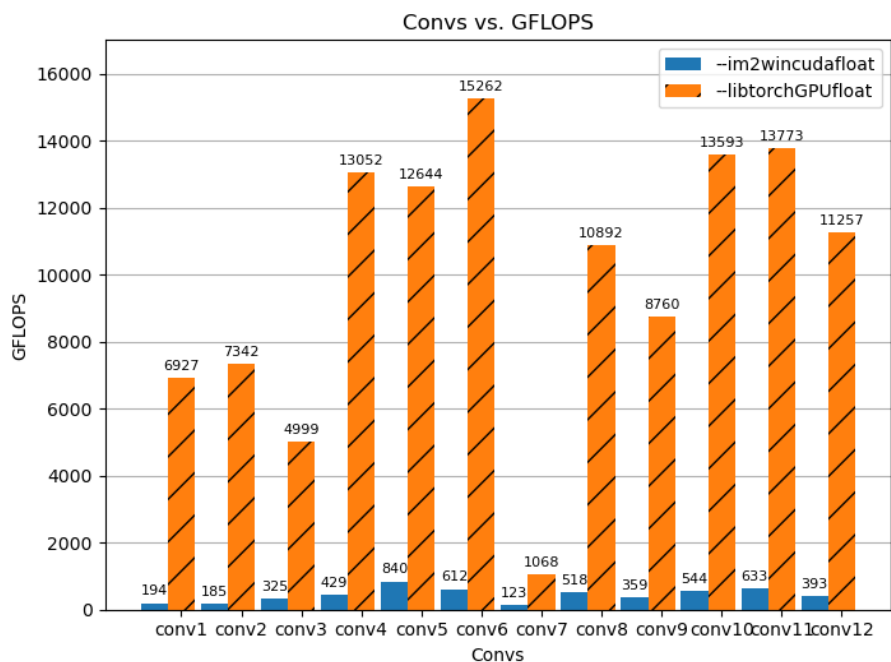


Figure 1: GPU

3 Experiment

为了有参照，将卢帅师兄调用libtorch中在GPU上实现卷积的代码跑起来了。卢帅师兄写的优化im2win和调用cudnn的代码还未跑起来，还有报错。

3.1 Analysis

libtorch比未优化的im2win快10多倍，这是自然的，毕竟还没有优化im2win，接下来需要在GPU上对im2win优化，不过应该直接在服务器上跑优化的实验，不然会出现优化CPU的那种本地跑起来效果不错但服务器上效果很差的情况。

4 Experiment2

openblas不能像mkl一样设置变量就指定使用avx2还是avx512，而是需要编译不同版本，编译时添加NO_AVX512=1可以关闭avx512。之前发现openmp和openblas冲突，然后这周写了用pthread实现并行，在自己机器上效果还行，但在服务器上效果不好。不过后面发现不需要自己写pthread，openblas官方考虑了这个问题，自己的程序使用openmp并行化时编译的时候需要加上USE_OPENMP=1。编译单线程版本的openblas需要在编译时加上USE_THREAD=0 USE_LOCKING=1（USE_LOCKING=1是为了解决和openmp冲突的，不加会报错）。然后在使用的时候链接需要的版本。结果看马吉祥的图。

4.1 Analysis

openblas并行版本比MKL低不少，单线程版本和MKL差不多。这周本来想在服务器上编译libtorch，结果编译报错，怀疑是环境问题。