CSCD 439/539 GPU Computing Lab 9

Tiled Matrix Multiplication Using Shared Memory

**Due Time:** May 7th 2014 9:50am. No Late Submissions are accepted. **Rules:** Your code must CUDA C Language. If your program shows a compilation error, you get a zero for this lab assignment.

**Submission:** Wrap up all your **source files and other data files** into a single zip file. Name your zip file as *FirstInitialYourLastName*Lab9.zip. For example, if your legal name is Will Smith, you should name your zip file as wsmithlab9.zip. A simple makefile has been provided in the zip file.

**Before you leave the laboratory, please show the TA or the instructor how your program works, they will give you a score for this Lab assignment.**

**For archive purpose, please also submit your single zip file on EWU Canvas by following CSCD439-01 Course 🡪Assignments🡪Lab9🡪 Submit Assignment to upload your single zip file.**

**Problem Description:**

Based on the lecture about tiled matrix-matrix multiplication on CUDA device, you are required to implement the following features and answer the questions.

In the provided lab package, you have two subfolders, **data** and **src**. The **data** folder contains all 2D matrices you will play with. The **src** folder has most of the source code you need to perform experiments.

1, Read the provided code in **src** folder, specifically read the main function in the source file matrix\_multiplication.cu.

2, Based upon the lecture notes, you have to implement the kernel \_\_global\_\_ void matrix\_multiply\_improved(float \*a, float \*b, float \*ab, size\_t width), which uses shared memory to perform a tiled matrix-matrix multiplication. You have to use dynamic arrays (**external** keyword) for tiles in the shared memory. (**hint** we have done this kernel already in lecture notes with minimal distinction.)

3, Change the main() function to invoke the kernel **matrix\_multiply\_improved()** you implemented, and time the kernel execution time. Then you have to compare the improved kernel with the provided simple kernel by filling out the following chart.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Input file | 1024.mat | 1024.mat | 2048.mat | 2048.mat | 2048.mat |
| Tile Dimensions | 8 X 8 | 16 X 16 | 8 X 8 | 16 X 16 | 32 X 32 |
| T1:time cost for Simple Kernel (ms) | .015337 | 0.09312 | 0.115066 | 0.077680 | 0.90740 |
| T2:time cost for Tiled Kernel (ms) | .008671 | .006759 | 0.069702 | 0.54264 | 0.47943 |
| Speedup = T1 / T2 | 1.7733 | 1.3777 | 1.650839 | 1.431512 | 1.892663 |