Assignment #2: Programming with CUDA

Due date: May. 7th,   11:59pm

Solving a linear system in parallel using Gaussian elimination

The formal components of the assignment are listed below:

·   Write a CUDA program to solve a linear system using Gaussian elimination with partial pivoting.

·   A default configuration you must evaluate is to use a single thread for CPUs.

·   Optionally, you may evaluate a version with multiple threads for CPUs.

·   Write a document that describes how your programs work. This document should not include your programs, though it may include figures containing pseudo-code that sketch the key elements of your parallelization strategy for each implementation. Explain how your program partitions the data, work and exploits parallelism. Justify your implementation choices. Explain how the parallel work is synchronized.

Use problem size n = 8000 to evaluate the performance of your implementations. If your sequential running time is too long for the interactive queue, you may base your timing measurements on n=6000 or n=5000. Prepare a table that includes your timing measurements for the combination of the elimination and back substitution phases of your implementations on rc0.kaist.ac.kr.

It is recommended to evaluate various options for the number of threads and blocks etc. Show graphs for such evaluations.

Your submission should contain:

·   The code for your program and a makefile to build the code, and

·   A writeup about your programs in either Word or PDF format (PDF preferred).