CSE327: Parallel Programming Spring 2015

Course Objective

In this class, you will learn to program NVIDIA's graphics processing units (GPUs) using CUDA (Compute Unified Device Architecture). According to General-Purpose computation on Graphics Processing Units (GPGPU) paradigm, video cards with massive parallelism capabilities, certainly available to essentially anyone nowadays, can be used to enhance the speed of countless computationally intensive algorithms. We study how CUDA and the GPU capabilities can be subjugated to deploy fast and effective parallel implementations for many computational parallel patterns.

Course Schedule

The calendar will be available before the start of the second semester on February 2015. To navigate the course lectures and assignments, just click on the name of the topic or the assignment.

1. Lectures

Week	Topics
1	CUDA Introduction
2	CUDA Thread Execution
3	CUDA Debugging Performance Analysis
4	Matrix Multiplication
5	<u>Tree-based Reduction</u>
6	Advanced Parallel Reduction
7	Prefix Sum (Scan)
8	<u>Histogram</u>
9	Host-Device Data Transfer and Streams

Department of Computers and Systems Engineering Faculty of Engineering University of Minia

2. Assignments

Week	Assignments
1	<u>Vector Addition</u>
2	Basic Matrix Multiplication
3	Matrix Multiplication using Shared Memory
4	Image Convolution
5	<u>Dot Product</u>
6	Prefix Sum for Large Data Sets
7	Statistics using Concurrent Kernels
8	

Grading

All assignment solutions must be delivered duo date as instructed. Solutions duo the hard deadline will get a %20 penalty. The assignment is evaluated as follows:

- Correctness (%75)
- Performance (%25)

The correctness stage is essential and if you followed the instructions carefully in each assignment and completed the objectives successfully, you will pass this stage. For performance assessment, your program timing or the execution time will be compared with the average of our implementation's timing and the best timing achieved by your class mates.