

**CS 491/521 Parallel Programming**  
**Homework 5: CUDA & Paper Reviews**  
**Due: 12/01, Thursday, 11:59 pm**

1. Matrix multiplication on GPU (50 points):

- (1) Write a C program for matrix multiplication:  $Z_{cpu} = X * Y$ .  
 $X$ ,  $Y$ , and  $Z_{cpu}$  are three  $N * N$  matrixes. Generate  $X$  and  $Y$  with random numbers, and then perform the multiplication and store the result in  $Z_{cpu}$ . The program should allow the user to specify the value of  $N$ .
- (2) Extend the program in (1) by adding a CUDA kernel to perform the same matrix multiplication (i.e.,  $X * Y$ ). The kernel should not use any shared memory. Store the result in another matrix called  $Z_{gpu}$ . The program should verify the correctness of the kernel by comparing the result of the kernel (i.e.,  $Z_{gpu}$ ) with the result in (1) (i.e.,  $Z_{cpu}$ ). Also, the program should measure the execution time of the kernel.
- (3) Extend the program in (2) by adding another CUDA kernel which uses shared memory to perform the matrix multiplication. Store the result in a matrix called  $Z_{gpu-s}$ . The program should verify the correctness of the kernel by comparing the result of the kernel (i.e.,  $Z_{gpu-s}$ ) with the result in (1) (i.e.,  $Z_{cpu}$ ). Also, the program should measure the execution time of the kernel.
- (4) Let  $N = 1024$ , compare the execution time of the kernels in (2) and (3). Explain why the performance are similar or different.
- (5) Change  $N$  to another value, and compare and explain the results again.
- (6) Write a report about your design of the kernels as well as the experimental results. Submit your report as well as your code to Canvas. Your report should include instructions for the TA to compile and execute your code on a CS departmental machine.

2. Review the following two papers, and submit the reviews to Canvas (50 points):

(1) Paper#1:

**(Big)Data in a Virtualized World: Volume, Velocity, and Variety in Cloud Datacenters (FAST'14)**

[https://www.usenix.org/system/files/conference/fast14/fast14-paper\\_birke.pdf](https://www.usenix.org/system/files/conference/fast14/fast14-paper_birke.pdf)

(2) Paper#2:

**Ceph: A Scalable, High-Performance Distributed File System (OSDI'06)**

<http://dl.acm.org/citation.cfm?id=1298485>

Tips:

The basic format of the review includes 5 paragraphs:

- (a) summary
- (b) strengths
- (c) weaknesses
- (d) unsolved problem(s)/opportunities for future work
- (e) questions for the authors

A good reviewer should think critically:

- \* What's the paper really about (not what the paper says it's about)
- \* Dig deeply to understand strengths & weaknesses (again, you should have your own judgement instead of simply agreeing with the paper)
- \* Assume reader of your review has read the paper. Just convey to me your own understanding/opinions/ideas.