CS 4370 - Parallel Programming

Bella Brickler, Jahcorian Ivery, Renee Paxson

Professor Meilin Liu

18 October 2024

Project 2: Tiled Matrix Multiplication

Compiling Instructions

1. `sftp` into a server with access to a GPU and a CUDA compiler and run `put Brickler-Ivery-Paxson_Project2.cu`
2. `ssh` into the server in another terminal
3. run `srun -p a100 --gres=gpu:1 --pty bash` to claim a node
4. run `nvcc Brickler-Ivery-Paxson_Project2.cu -o Brickler-Ivery-Paxson_Project2` to compile the code
5. run `./Brickler-Ivery-Paxson_Project2` to execute the code
6. To change the size of the matrices and thread block, change `#define MATRIX_WIDTH`, `#define BLOCK_SIZE`, and `#define TILE_WIDTH` on lines 13-15 of `Brickler-Ivery-Paxson_Project2.cu`, then restart from step 1

Run Notes

- The program will print full matrices if they are 8*8 or smaller. Any larger matrices will only have the first line printed.
- For this program to run properly, `BLOCK_SIZE` and `TILE_WIDTH` must be the same
- The program will print `TEST PASSED` to the terminal if the GPU addition or multiplication matches the CPU addition or multiplication. If it prints `TEST FAILED` the two matrices do not match, and the code may have been tampered with in some way.