Mikhail Andreev

EC527 Lab 9

Hpcl-10 (2.53 GHz)

**Assignment 9**

**Part I: MMM using global memory only**

Timing:

1K elements:

GPU outer loop time: 71.321854 (msec)  
GPU inner loop time: 66.807617 (msec)  
CPU time: 9097.080078(msec)

2K elements:

GPU outer loop time: 490.746613 (msec)  
GPU inner loop time: 476.247742 (msec)  
CPU time: 77356.093750(msec)

Differences between CPU and GPU:

Max difference = 422212465065984.000000  
Min difference = 0.000000  
Average difference = 32833141735424.000000  
Average = 1159456227666161565696.000000

2K elements:

Max difference = 844424930131968.000000  
Min difference = 0.000000  
Average difference = 65723879653376.000000  
Average = 2353149382649547063296.000000

Tolerance as percentage:

1K elements:

Max Tolerance = 0.000036%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.000003%

2K elements:

Max Tolerance = 0.000036%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.000003%

Array Creation code:



**Part II: MMM using shared memory**

Timing:

1K elements:

GPU outer loop time: 26.674944 (msec)  
GPU inner loop time: 22.116961 (msec)  
CPU time: 9483.327148(msec)

2K elements:

GPU outer loop time: 177.982727 (msec)  
GPU inner loop time: 162.597473 (msec)  
CPU time: 77534.179688(msec)

Differences between CPU and GPU:

1K elements:

Max difference = 179158470832611655680.000000  
Min difference = 1000000.000000  
Average difference = 9443112734200692736.000000  
Average = 1156209132334827438080.000000

2K elements:

Max difference = 844424930131968.000000  
Min difference = 0.000000  
Average difference = 65723879653376.000000  
Average = 2353149382649547063296.000000

Tolerance as percentage:

1K elements:

Max Tolerance = 15.495335%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.816730%

2K elements:

Max Tolerance = 0.000036%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.000003%

**Part III: MMM using shared memory and loop unrolling**

Timing:

1K elements:

GPU outer loop time: 26.811424 (msec)  
GPU inner loop time: 22.321184 (msec)  
CPU time: 9489.315430(msec)

2K elements:

GPU outer loop time: 179.556671 (msec)  
GPU inner loop time: 165.040131 (msec)  
CPU time: 77353.265625(msec)

Differences between CPU and GPU:

1K elements:

Max difference = 179158822676332544000.000000  
Min difference = 1000000.000000  
Average difference = 9443345830665781248.000000  
Average = 1156208287909897306112.000000

2K elements:

Max difference = 8022036836253696.000000  
Min difference = 0.000000  
Average difference = 1155057097637888.000000  
Average = 2353149382649547063296.000000

Tolerance as percentage:

1K elements:

Max Tolerance = 15.495377%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.816751%

2K elements:

Max Tolerance = 0.000341%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.000049%

**Extra Credit I: Different ThreadBlock Sizes**

Timing:

ThreadBlock = 4:

GPU outer loop time: 163.819427 (msec)  
GPU inner loop time: 159.275391 (msec)  
CPU time: 9573.115234(msec)

ThreadBlock = 8:

GPU outer loop time: 43.526176 (msec)  
GPU inner loop time: 38.943871 (msec)  
CPU time: 9522.835938(msec)

ThreadBlock = 16:

GPU outer loop time: 26.857281 (msec)  
GPU inner loop time: 22.299776 (msec)  
CPU time: 9646.983398(msec)

ThreadBlock = 32:

Error while running Kernel

ThreadBlock = 64:

Could not run.

Differences between CPU and GPU:

ThreadBlock = 4:

Max difference = 2603643534573568.000000  
Min difference = 0.000000  
Average difference = 411123832586240.000000  
Average = 1160537373051707195392.000000

ThreadBlock = 8:

Max difference = 2603643534573568.000000  
Min difference = 0.000000  
Average difference = 411123832586240.000000  
Average = 1160537373051707195392.000000

ThreadBlock = 16:

Max difference = 179158822676332544000.000000  
Min difference = 1000000.000000  
Average difference = 9441774628549689344.000000  
Average = 1156209624916036681728.000000

ThreadBlock = 32:

Error while running Kernel

ThreadBlock = 64:

Could not run.

Tolerance as percentage:

ThreadBlock = 4:

Max Tolerance = 0.000224%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.000035%

ThreadBlock = 8:

Max Tolerance = 0.000224%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.000035%

ThreadBlock = 16:

Max Tolerance = 15.495358%  
Min Tolerance = 0.000000%  
Average Tolerance = 0.816614%

ThreadBlock = 32:

Error while running Kernel

ThreadBlock = 64:

Could not run.

**Extra Credit II: Different Matrix Sizes**

Timing:

.5K elements:

1.5K elements:

2.5K elements:

3K elements:

Differences between CPU and GPU:

.5K elements:

1.5K elements:

2.5K elements:

3K elements:

Tolerance as percentage:

.5K elements:

1.5K elements:

2.5K elements:

3K elements:

**Extra Credit III: Breaking Matrix Size**

**Extra Credit IV: Web GPU Code**