

CSE 520: Computer Architecture 2

Assignment 3: Report

➤ **AUTHOR:**

Name: Viraj Savaliya

ASU ID: [REDACTED]

➤ **OUTPUT TABLE:**

NOTE: 1. All measurements output values of Execution time are in seconds.

2. The Reference C execution time is the average value of the 6 times it ran.

3. All the readings and model information were taken on host machine:

“en6169317-217.etslabs.dhcp.asu.edu” of BYENG 217 Lab.

Matrix size	Reference C implementation (single thread)	OpenCL on CPU			OpenCL on GPU		
		Normal kernel	Tiled kernel (tile_size=8)	Tiled kernel (tile_size=16)	Normal kernel	Tiled kernel (tile_size=8)	Tiled kernel (tile_size=16)
512x512	0.231985167	0.00217012	0.0135882	0.00141325	0.000369696	0.000376864	0.000249248
1024x1024	2.910065	0.0328663	0.0864493	0.00865014	0.00282243	0.00289094	0.00182947
2048x2048	50.60116667	0.285005	0.686938	0.0978163	0.0231036	0.0229905	0.0145656

OBSERVATION:

As mentioned by Prof.Lee during the lecture about the increase in the execution time when we increase the matrix size in GPU for a fixed tile size then the ratio is approximately 8 as $O(n) = 2^3$.

For e.g.: a) For tile size 8 in GPU, $0.0229905 / 0.00289094 = 7.9526$

b) For tile size 8 in GPU $0.00289094 / 0.000376864 = 7.6710$

➤ **FILES INCLUDED:**

1. main.cpp
2. Makefile
3. matrix_mul.cl
4. mm_all.sh
5. README
6. Report.pdf

➤ **CPU MODEL INFO:**

1. Platform Name	Intel(R) CPU Runtime for OpenCL(TM) Applications
2. Number of devices	1
3. Device Name	Intel(R) Core(TM) i9-9820X CPU @ 3.30GHz
4. Device Vendor	Intel(R) Corporation
5. Device Vendor ID	0x8086
6. Device Version	OpenCL 2.1 (Build 0)
7. Driver Version	18.1.0.0920
8. Device OpenCL C Version	OpenCL C 2.0
9. Device Type	CPU
10. Max compute units	20
11. Max clock frequency	3300MHz
12. Device Partition	(core)
a. Max number of sub-devices	20
b. Supported partition types	by counts, equally, by names (Intel)
13. Max work item dimensions	3
14. Max work item sizes	8192x8192x8192
15. Max work group size	8192
16. Preferred work group size multiple	128
17. Max sub-groups per work group	1
18. Preferred / native vector sizes	
a. char	1 / 32
b. short	1 / 16
c. int	1 / 8
d. long	1 / 4
e. half	0 / 0 (n/a)
f. float	1 / 8
g. double	1 / 4 (cl_khr_fp64)
19. Global memory size	67239464960 (62.62GiB)
20. Max memory allocation	16809866240 (15.66GiB)
21. Unified memory for Host and Device	Yes
22. Minimum alignment for any data type	128 bytes
23. Alignment of base address	1024 bits (128 bytes)
24. Max size for global variable	65536 (64KiB)
25. Preferred total size of global vars	65536 (64KiB)
26. Global Memory cache type	Read/Write
27. Global Memory cache size	262144 (256KiB)
28. Global Memory cache line size	64 bytes
29. Max number of pipe args	16
30. Max active pipe reservations	13107
31. Max pipe packet size	1024
32. Local memory type	Global
33. Local memory size	32768 (32KiB)
34. Max number of constant args	480
35. Max constant buffer size	131072 (128KiB)
36. Max size of kernel argument	3840 (3.75KiB)

➤ **GPU MODEL INFO:**

1. Platform Name	NVIDIA CUDA
2. Number of devices	1
3. Device Name	Quadro RTX 5000
4. Device Vendor	NVIDIA Corporation
5. Device Vendor ID	0x10de
6. Device Version	OpenCL 1.2 CUDA
7. Driver Version	410.104
8. Device OpenCL C Version	OpenCL C 1.2
9. Device Type	GPU
10. Device Topology (NV)	PCI-E, 65:00.0
11. Max compute units	48
12. Max clock frequency	1815MHz
13. Compute Capability (NV)	7.5
14. Device Partition	(core)
a. Max number of sub-devices	1
b. Supported partition types	None
15. Max work item dimensions	3
16. Max work item sizes	1024x1024x64
17. Max work group size	1024
18. Preferred work group size multiple	32
19. Warp size (NV)	32
20. Preferred / native vector sizes	
a. char	1 / 1
b. short	1 / 1
c. int	1 / 1
d. long	1 / 1
e. half	0 / 0 (n/a)
f. float	1 / 1
g. double	1 / 1 (cl_khr_fp64)
21. Global memory size	16868376576 (15.71GiB)
22. Max memory allocation	4217094144 (3.927GiB)
23. Unified memory for Host and Device	No
24. Integrated memory (NV)	No
25. Minimum alignment for any data type	128 bytes
26. Alignment of base address	4096 bits (512 bytes)
27. Global Memory cache type	Read/Write
28. Global Memory cache size	786432 (768KiB)
29. Global Memory cache line size	128 bytes
30. Image support	Yes
31. Local memory type	Local
32. Local memory size	49152 (48KiB)
33. Registers per block (NV)	65536
34. Max number of constant args	9
35. Max constant buffer size	65536 (64KiB)
36. Max size of kernel argument	4352 (4.25KiB)

➤ CONSOLE SCREENSHOTS:

```
vsavaliy@en6169317-217: ~  
vsavaliy@en6169317-217:~$ make  
g++ -O3 -std=c++11 main.cpp -o matrix_mul -lOpenCL  
vsavaliy@en6169317-217:~$ bash mm_all.sh  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 1  
SIZE is 512  
TILE_SIZE is 16  
Reference C matrix multiplication: 0.255458 sec  
Starting KERNEL on CPU!  
Starting NAIVE KERNEL on CPU!  
OpenCL matrix multiplication: 0.00217012 sec  
Matrix Multiply for NAIVE KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 1  
SIZE is 1024  
TILE_SIZE is 16  
Reference C matrix multiplication: 2.86661 sec  
Starting KERNEL on CPU!  
Starting NAIVE KERNEL on CPU!  
OpenCL matrix multiplication: 0.0328663 sec  
Matrix Multiply for NAIVE KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 1  
SIZE is 2048  
TILE_SIZE is 16  
Reference C matrix multiplication: 50.6442 sec  
Starting KERNEL on CPU!  
Starting NAIVE KERNEL on CPU!  
OpenCL matrix multiplication: 0.285005 sec  
Matrix Multiply for NAIVE KERNEL on CPU done!  
*****
```

```
vsavaliy@en6169317-217: ~  
Matrix Multiply for NAIVE KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 2  
SIZE is 512  
TILE_SIZE is 8  
Reference C matrix multiplication: 0.224356 sec  
Starting KERNEL on CPU!  
Starting TILED KERNEL on CPU!  
OpenCL matrix multiplication: 0.0135882 sec  
Matrix Multiply for TILED KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 2  
SIZE is 1024  
TILE_SIZE is 8  
Reference C matrix multiplication: 2.95861 sec  
Starting KERNEL on CPU!  
Starting TILED KERNEL on CPU!  
OpenCL matrix multiplication: 0.0864493 sec  
Matrix Multiply for TILED KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 2  
SIZE is 2048  
TILE_SIZE is 8  
Reference C matrix multiplication: 50.4613 sec  
Starting KERNEL on CPU!  
Starting TILED KERNEL on CPU!  
OpenCL matrix multiplication: 0.686938 sec  
Matrix Multiply for TILED KERNEL on CPU done!  
*****
```

```
vsavaliy@en6169317-217: ~  
Matrix Multiply for TILED KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 2  
SIZE is 512  
TILE_SIZE is 16  
Reference C matrix multiplication: 0.224504 sec  
Starting KERNEL on CPU!  
Starting TILED KERNEL on CPU!  
OpenCL matrix multiplication: 0.00141325 sec  
Matrix Multiply for TILED KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 2  
SIZE is 1024  
TILE_SIZE is 16  
Reference C matrix multiplication: 2.86427 sec  
Starting KERNEL on CPU!  
Starting TILED KERNEL on CPU!  
OpenCL matrix multiplication: 0.00865014 sec  
Matrix Multiply for TILED KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 1  
MM_KERNEL is 2  
SIZE is 2048  
TILE_SIZE is 16  
Reference C matrix multiplication: 50.9348 sec  
Starting KERNEL on CPU!  
Starting TILED KERNEL on CPU!  
OpenCL matrix multiplication: 0.0978163 sec  
Matrix Multiply for TILED KERNEL on CPU done!  
*****
```

```
vsavaliy@en6169317-217: ~  
Matrix Multiply for TILED KERNEL on CPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 1  
SIZE is 512  
TILE_SIZE is 16  
Reference C matrix multiplication: 0.228349 sec  
Starting KERNEL on GPU!  
Starting NAIVE KERNEL on GPU!  
OpenCL matrix multiplication: 0.000369696 sec  
Matrix Multiply for NAIVE KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 1  
SIZE is 1024  
TILE_SIZE is 16  
Reference C matrix multiplication: 2.92088 sec  
Starting KERNEL on GPU!  
Starting NAIVE KERNEL on GPU!  
OpenCL matrix multiplication: 0.00282243 sec  
Matrix Multiply for NAIVE KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 1  
SIZE is 2048  
TILE_SIZE is 16  
Reference C matrix multiplication: 54.3838 sec  
Starting KERNEL on GPU!  
Starting NAIVE KERNEL on GPU!  
OpenCL matrix multiplication: 0.0231036 sec  
Matrix Multiply for NAIVE KERNEL on GPU done!  
*****
```

```
vsavaliy@en6169317-217: ~  
Matrix Multiply for NAIVE KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 2  
SIZE is 512  
TILE_SIZE is 8  
Reference C matrix multiplication: 0.227707 sec  
Starting KERNEL on GPU!  
Starting TILED KERNEL on GPU!  
OpenCL matrix multiplication: 0.000376864 sec  
Matrix Multiply for TILED KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 2  
SIZE is 1024  
TILE_SIZE is 8  
Reference C matrix multiplication: 2.9719 sec  
Starting KERNEL on GPU!  
Starting TILED KERNEL on GPU!  
OpenCL matrix multiplication: 0.00289094 sec  
Matrix Multiply for TILED KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 2  
SIZE is 2048  
TILE_SIZE is 8  
Reference C matrix multiplication: 48.602 sec  
Starting KERNEL on GPU!  
Starting TILED KERNEL on GPU!  
OpenCL matrix multiplication: 0.0229905 sec  
Matrix Multiply for TILED KERNEL on GPU done!  
*****
```

```
vsavaliy@en6169317-217: ~  
Matrix Multiply for TILED KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 2  
SIZE is 512  
TILE_SIZE is 16  
Reference C matrix multiplication: 0.231537 sec  
Starting KERNEL on GPU!  
Starting TILED KERNEL on GPU!  
OpenCL matrix multiplication: 0.000249248 sec  
Matrix Multiply for TILED KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 2  
SIZE is 1024  
TILE_SIZE is 16  
Reference C matrix multiplication: 2.87812 sec  
Starting KERNEL on GPU!  
Starting TILED KERNEL on GPU!  
OpenCL matrix multiplication: 0.00182947 sec  
Matrix Multiply for TILED KERNEL on GPU done!  
*****  
Starting the CSE520 Assgn3 runs!  
MM_CPU is 2  
MM_KERNEL is 2  
SIZE is 2048  
TILE_SIZE is 16  
Reference C matrix multiplication: 48.5809 sec  
Starting KERNEL on GPU!  
Starting TILED KERNEL on GPU!  
OpenCL matrix multiplication: 0.0145656 sec  
Matrix Multiply for TILED KERNEL on GPU done!  
*****  
vsavaliy@en6169317-217:~$
```