Lab 7 report

6314061 – Quang Tuan Le

March 2023

1 Task 1: A simple test of your GPU hardward and CUDA programming environment

- To do this task, I login to onyx server via VS code, and using these steps to get information to answer the questions: cd /usr/local/cuda-11.8/extras/demo_suite and ./deviceQuery.

```
qle010@onyx:~ 101% cd /usr/local/cuda-11.8/extras/demo_suite
qle010@onyx:/usr/local/cuda-11.8/extras/demo_suite 102% ./deviceQuery
./deviceQuery Starting...

 CUDA Device Query (Runtime API) version (CUDART static linking)

Detected 8 CUDA Capable device(s)

Device 0: "NVIDIA A100 80GB PCIe"
  CUDA Driver Version / Runtime Version          11.8 / 11.8
  CUDA Capability Major/Minor version number:    8.0
  Total amount of global memory:                 81100 MBytes (85039775744 bytes)
  (108) Multiprocessors, ( 64) CUDA Cores/MP:     6912 CUDA Cores
  GPU Max Clock rate:                            1410 MHz (1.41 GHz)
  Memory Clock rate:                             1512 Mhz
  Memory Bus Width:                              5120-bit
  L2 Cache Size:                                 41943040 bytes
  Maximum Texture Dimension Size (x,y,z)         1D=(131072), 2D=(131072, 65536), 3D=(16384, 16384, 16384)
  Maximum Layered 1D Texture Size, (num) layers  1D=(32768), 2048 layers
  Maximum Layered 2D Texture Size, (num) layers  2D=(32768, 32768), 2048 layers
  Total amount of constant memory:               65536 bytes
  Total amount of shared memory per block:       49152 bytes
  Total number of registers available per block: 65536
  Warp size:                                     32
  Maximum number of threads per multiprocessor:  2048
  Maximum number of threads per block:           1024
  Max dimension size of a thread block (x,y,z): (1024, 1024, 64)
  Max dimension size of a grid size    (x,y,z): (2147483647, 65535, 65535)
  Maximum memory pitch:                          2147483647 bytes
  Texture alignment:                             512 bytes
  Concurrent copy and kernel execution:          Yes with 3 copy engine(s)
  Run time limit on kernels:                     No
  Integrated GPU sharing Host Memory:            No
  Support host page-locked memory mapping:       Yes
  Alignment requirement for Surfaces:            Yes
  Device has ECC support:                        Enabled
  Device supports Unified Addressing (UVA):      Yes
  Device supports Compute Preemption:            Yes
  Supports Cooperative Kernel Launch:            Yes
  Supports MultiDevice Co-op Kernel Launch:      Yes
  Device PCI Domain ID / Bus ID / location ID:   0 / 26 / 0
```

- Here are the answers:
  - There are 8 GPU devices in my machine.
  - The maximum amount of shared memory per thread block is 49152 bytes.
  - The maximum dimension size of a thread block is (1024, 1024, 64).

- The maximum number of registers available per thread block is 65536.
- The global memory size is 81100 MBytes.

2 Task 2: Find the maximum

- The code file and Makefile is attached, and here is my result:

```
qle010@onyx:/ 109% ls
a/    bin@   depot/  disk/  home/   lib@    local@       media/  mnt/  netdepot/  proc/           root/  sbin@  sys/  usr/
aul/  boot/  dev/    etc/   homes/  lib64@  lost+found/  misc/   net/  opt/       restoresymtable  run/   srv/   tmp/  var/
qle010@onyx:/ 110% cd /a/buffalo.cs.fiu.edu./disk/jccl-001/homes/qle010/Courses/COP4520/Lab7
qle010@onyx:~/Courses/COP4520/Lab7 111% /usr/local/cuda-11.8/bin/nvcc findMax.cu -o findMax
qle010@onyx:~/Courses/COP4520/Lab7 112% ./findMax
GPU max: 99 | GPU time: 0.004096
CPU max: 99 | CPU time: 0.008192
qle010@onyx:~/Courses/COP4520/Lab7 113% make
/usr/local/cuda-11.8/bin/nvcc findMax.cu -o findMax
qle010@onyx:~/Courses/COP4520/Lab7 114% ./findMax
GPU max: 99 | GPU time: 0.003072
CPU max: 99 | CPU time: 0.009216
qle010@onyx:~/Courses/COP4520/Lab7 115% make clean
rm -f *.o
qle010@onyx:~/Courses/COP4520/Lab7 116% ls
findMax*  findMax.cu  Makefile
qle010@onyx:~/Courses/COP4520/Lab7 117%
```