HW 5 (2021-27764 안지수)

1. Warming up

(a) NVIDIA TITAN RTX

(b) GFLOPS = N (GHz) * M (Cores) * 2 FLOPS/cycle = 1.1770 * 4608 * 2 = 16.31232 TFLOPS

(c) 8192 * 8192 * 4bytes / 15.75GB/s = 0.0159 sec

2. Matrix Multiplication with OpenCL

(a) 두가지 포인트로 병렬화를 수행하였다.
 (1) 행렬 C의 한 원소를 계산하기 위해 K번 Local memory에 접근하는 부분을 float형 변수를 선언하여 이 변수에서 연산을 수행하고 모든 연산이 끝난 후에 행렬 C의 원소를 업데이트 하는 방식으로 변경하였다.
 (2) local work size를 NVIDIA TITAN RTX의 한SM의 cuda core(64개)의 배수에 맞추어 (warp 사이즈를 맞추기 위해) 변경해 보면서 최고성능을 얻을 수 있는 local work size를 찾았다.

(b) 본 실험에서는 위에서 언급한 두가지 포인트로 병렬화를 수행했는데, (1)의 적용 유무와 (2)의 local work size를 조절해가며 성능을 조사하였다. (단, M=N=K= 4096)
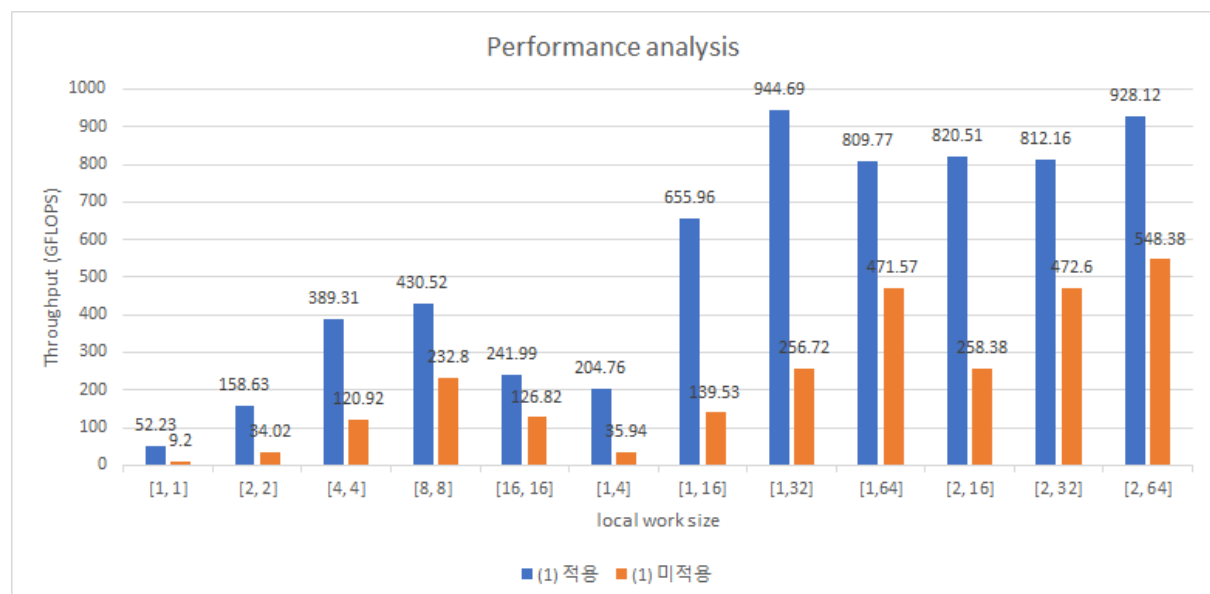
(c) 모두 Valid

```
  Problem size: M = 896, N = 1536, K = 640
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.010534 sec
Validating...
Result: VALID
Reference time: 0.052526 sec
Reference throughput: 33.537961 GFLOPS
Your Avg. time: 0.010534 sec
Your Avg. throughput: 167.235858 GFLOPS
```

```
  Problem size: M = 1408, N = 512, K = 1792
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.011782 sec
Validating...
Result: VALID
Reference time: 0.065624 sec
Reference throughput: 39.371340 GFLOPS
Your Avg. time: 0.011782 sec
Your Avg. throughput: 219.295978 GFLOPS
```

```
  Problem size: M = 640, N = 512, K = 2048
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.007440 sec
Validating...
Result: VALID
Reference time: 0.072318 sec
Reference throughput: 18.559335 GFLOPS
Your Avg. time: 0.007440 sec
Your Avg. throughput: 180.398329 GFLOPS
```

```
  Problem size: M = 512, N = 384, K = 1664
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.004532 sec
Validating...
Result: VALID
Reference time: 0.338610 sec
Reference throughput: 1.932344 GFLOPS
Your Avg. time: 0.004532 sec
Your Avg. throughput: 144.379529 GFLOPS
```

```
  Problem size: M = 1152, N = 768, K = 896
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.008840 sec
Validating...
Result: VALID
Reference time: 0.141425 sec
Reference throughput: 11.210542 GFLOPS
Your Avg. time: 0.008840 sec
Your Avg. throughput: 179.340637 GFLOPS
```

```
  Problem size: M = 1024, N = 384, K = 1664
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.007320 sec
Validating...
Result: VALID
Reference time: 0.098309 sec
Reference throughput: 13.311274 GFLOPS
Your Avg. time: 0.007320 sec
Your Avg. throughput: 178.781111 GFLOPS
```

```
  Problem size: M = 1792, N = 640, K = 2048
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.017776 sec
Validating...
Result: VALID
Reference time: 0.065665 sec
Reference throughput: 71.539717 GFLOPS
Your Avg. time: 0.017776 sec
Your Avg. throughput: 264.267281 GFLOPS
```

```
  Problem size: M = 384, N = 896, K = 640
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.003511 sec
Validating...
Result: VALID
Reference time: 0.182518 sec
Reference throughput: 2.412928 GFLOPS
Your Avg. time: 0.003511 sec
Your Avg. throughput: 125.426715 GFLOPS
```

```
  Problem size: M = 256, N = 1024, K = 1920
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.006793 sec
Validating...
Result: VALID
Reference time: 0.073033 sec
Reference throughput: 13.783342 GFLOPS
Your Avg. time: 0.006793 sec
Your Avg. throughput: 148.197276 GFLOPS
```

```
  Problem size: M = 1792, N = 768, K = 1152
  Number of iterations: 1
  Number of warmup iterations: 0
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Calculating...(iter=0)
Calculating done!(iter=0): 0.014586 sec
Validating...
Result: VALID
Reference time: 0.065740 sec
Reference throughput: 48.233768 GFLOPS
Your Avg. time: 0.014586 sec
Your Avg. throughput: 217.393996 GFLOPS
```

(d) 1035.36 GFLOPS의 결과를 얻었다.

```
  Problem size: M = 8192, N = 8192, K = 8192
  Number of iterations: 1
  Number of warmup iterations: 1
  Print matrix: off
  Validation: on
Initializing matrices...
Initializing matrices done!
Initializing...
OpenCL platform name: NVIDIA CUDA
OpenCL device name: NVIDIA TITAN RTX
Initializing done!
Warming up...
Warming up done!: 1.152979 sec
Calculating...(iter=0)
Calculating done!(iter=0): 1.061680 sec
Validating...
Result: VALID
Reference time: 9.658535 sec
Reference throughput: 113.838344 GFLOPS
Your Avg. time: 1.061680 sec
Your Avg. throughput: 1035.634185 GFLOPS
```

3. Multi-GPU

본 실험에서는 자료실에 올려놓은 vector addtion code중 vec_add_normal_io를 multi-gpu에서 동작할 수 있도록 수정한 뒤 진행하였다.
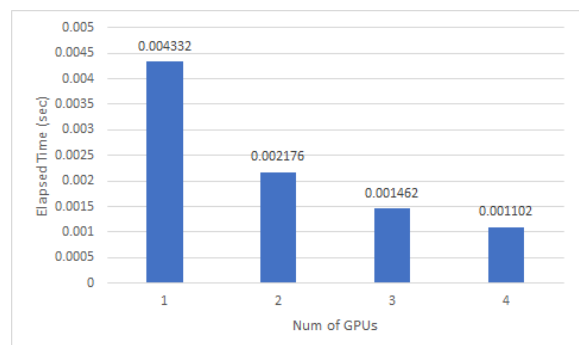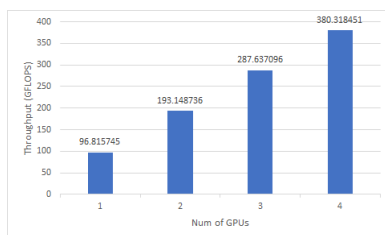
실험에서는 0.8GB의 float형 자료형 array의 vector addtion 을 수행하여 위 차트와 같은 실행시간과 Throughput을 얻을 수 있었다.