

CIS6930: Programming on Massive Parallel Systems

GPU Project 2: A better version of SDH computing program

Submitted by – Ghalib Saleem

I have decided to implement the algorithm 3 from the paper, Algorithms and Framework for Computing 2-body Statistics on GPUs by Napath Pitaksiranan, Zhila Nouri, and Yi-Cheng Tu for the implementation of this project. The paper describes a series of techniques for achieving high performance in dealing with a group of problems that share similar computing patterns as the SDH problem.

Algorithm 3: SDH with Output Privatization

Local Var: t (Thread id), b (Block id)
Global Var: B (Block size), M (total number of blocks)

- 1: $SHMOut \leftarrow$ Initialize shared memory to zero
- 2: $reg \leftarrow$ the t -th datum of b -th input data block
- 3: **for** $i = b + 1$ to M **do**
- 4: $R \leftarrow$ the i -th input data block loaded to cache
- 5: syncthreads()
- 6: **for** $j = 0$ to B **do**
- 7: $d \leftarrow \text{DisFunction}(reg, R[j])$
- 8: atomicAdd($SHMOut[d]$, 1)
- 9: **end for**
- 10: **end for**
- 11: $L \leftarrow$ the b -th input data block loaded to cache
- 12: syncthreads()
- 13: **for** $i = t + 1$ to B **do**
- 14: $d \leftarrow \text{DisFunction}(reg, L[i])$
- 15: atomicAdd($SHMOut[d]$, 1)
- 16: **end for**
- 17: syncthreads()
- 18: $Output[b][t] \leftarrow SHMOut[t]$

The reason I chose the Algorithm 3 is the use of shared memory for the values of distance function. As we all know shared memory is faster than the local memory and Global memory and has lower latency. The privatization algorithm creates a private copy of the output and stores it so it tends to be utilized by the subset of threads.

After completion and execution of my code, the minimum total running time of the kernel for three execution with different input that I was able to achieve is:

Note: - GPU Old is Brute force and GPU Algo 3 is result of privatization.

Output 1

Number of Samples: 10000	Bucket Width: 1000	Block Size: 64
Kernel time: 93.46391 ms		

```
(base) ghalibsaleem@Ghalib-linux:~/Documents/USF/Cuda/proj2-ghalibsaleem$ nvcc proj2-ghalibsaleem.cu -o test
(base) ghalibsaleem@Ghalib-linux:~/Documents/USF/Cuda/proj2-ghalibsaleem$ ./test 10000 1000 64

***** GPU Old *****
Running time for GPU Old version: 0.000010

00:      16288 |      108733 |      275806 |      501165 |      767139 |
05:     1061461 |     1367080 |     1678079 |     1979292 |     2259514 |
10:     2510674 |     2730531 |     2906757 |     3039978 |     3120712 |
15:     3146650 |     3112938 |     3027899 |     2882944 |     2685905 |
20:     2426248 |     2110178 |     1743897 |     1356172 |     1021851 |
25:      749569 |      530258 |      358306 |      230657 |      138568 |
30:       76914 |       39502 |       19368 |        8666 |        3661 |
35:       1254 |        326 |         50 |          10 |           0 |
T:49995000

***** GPU Algo 3 *****
Running time for GPU new version: 0.000016

00:      16288 |      108733 |      275806 |      501165 |      767139 |
05:     1061461 |     1367080 |     1678079 |     1979292 |     2259514 |
10:     2510674 |     2730531 |     2906757 |     3039978 |     3120712 |
15:     3146650 |     3112938 |     3027899 |     2882944 |     2685905 |
20:     2426248 |     2110178 |     1743897 |     1356172 |     1021851 |
25:      749569 |      530258 |      358306 |      230657 |      138568 |
30:       76914 |       39502 |       19368 |        8666 |        3661 |
35:       1254 |        326 |         50 |          10 |           0 |
T:49995000

*****Time to generate:: 93.46391 ms*****
```

Output 2

Number of Samples: 10000	Bucket Width: 500	Block Size: 32
Kernel time: 74.97648 ms		

```
(base) ghalibsaleem@Ghalib-linux:~/Documents/USF/Cuda/proj2-ghalibsaleem$ ./test 10000 500 32
```

```
***** GPU Old *****
```

```
Running time for GPU Old version: 0.000008
```

00:	2076	14212	37870	70863	113190
05:	162616	219792	281373	348231	418908
10:	492542	568919	643293	723787	799933
15:	878146	953657	1025635	1097551	1161963
20:	1225131	1285543	1340874	1389657	1435607
25:	1471150	1505326	1534652	1554499	1566213
30:	1574002	1572648	1562633	1550305	1527753
35:	1500146	1463004	1419940	1372128	1313777
40:	1249134	1177114	1097835	1012343	920271
45:	823626	723835	632337	548947	472904
50:	405296	344273	289347	240911	197652
55:	160654	129093	101564	78499	60069
60:	44360	32554	23054	16448	11533
65:	7835	5295	3371	2289	1372
70:	804	450	222	104	34
75:	16	8	2	0	0

```
T:49995000
```

```
***** GPU Algo 3*****
```

```
Running time for GPU new version: 0.000016
```

00:	2076	14212	37870	70863	113190
05:	162616	219792	281373	348231	418908
10:	492542	568919	643293	723787	799933
15:	878146	953657	1025635	1097551	1161963
20:	1225131	1285543	1340874	1389657	1435607
25:	1471150	1505326	1534652	1554499	1566213
30:	1574002	1572648	1562633	1550305	1527753
35:	1500146	1463004	1419940	1372128	1313777
40:	1249134	1177114	1097835	1012343	920271
45:	823626	723835	632337	548947	472904
50:	405296	344273	289347	240911	197652
55:	160654	129093	101564	78499	60069
60:	44360	32554	23054	16448	11533
65:	7835	5295	3371	2289	1372
70:	804	450	222	104	34
75:	16	8	2	0	0

```
T:49995000
```

```
*****Time to generate:: 74.97648 ms*****
```

Output 3

Number of Samples: 100000	Bucket Width: 1000	Block Size: 64
Kernel time: 6902.37646 ms		

```
(base) ghalibsaleem@Ghalib-linux:~/Documents/USF/Cuda/proj2-ghalibsaleem$ ./test 100000 1000 64

***** GPU Old *****
Running time for GPU Old version: 0.000010

00:      1640289 |      10843080 |      27544198 |      50026816 |      76708211 |
05:      106121559 |      136910632 |      167848146 |      197868797 |      225955796 |
10:      251328500 |      273244161 |      291068479 |      304327905 |      312470075 |
15:      315193417 |      312331390 |      303639589 |      289084180 |      268815791 |
20:      242887739 |      211519683 |      174891328 |      135536901 |      101644679 |
25:       74102047 |       52073269 |      34966533 |      22219570 |      13192634 |
30:       7249346 |       3683809 |      1759328 |       789152 |      317736 |
35:       109694 |       29703 |       5342 |       491 |       5
T:4999950000

***** GPU Algo 3 *****
Running time for GPU new version: 0.000050

00:      1640289 |      10843080 |      27544198 |      50026816 |      76708211 |
05:      106121559 |      136910632 |      167848146 |      197868797 |      225955796 |
10:      251328500 |      273244161 |      291068479 |      304327905 |      312470075 |
15:      315193417 |      312331390 |      303639589 |      289084180 |      268815791 |
20:      242887739 |      211519683 |      174891328 |      135536901 |      101644679 |
25:       74102047 |       52073269 |      34966533 |      22219570 |      13192634 |
30:       7249346 |       3683809 |      1759328 |       789152 |      317736 |
35:       109694 |       29703 |       5342 |       491 |       5
T:4999950000

*****Time to generate:: 6902.37646 ms*****
```