Project #6 – OpenCL Linear Regression

Hyun-Taek Oh
ohhyun@oregonstate.edu
2 June 2024

1. What machine you ran this on?

The program was conducted on a Predator HELIOS 300 (2022).

CPU:	12th Gen Intel(R) Core™ i9-12900H
Motherboard:	Mainboard PH315-55 Intel Ci912900H GN20-E6
Memory:	16GB DDR5 (8GB * 2)
Server:	rabbit.engr.oregonstate.edu

2. Show the table and graphs

m	b	DataSize	LocalSize	MegaXYsProcessedPerSecond
5	6.99	4096	8	29.21
5	6.99	4096	32	58.76
5	6.99	4096	64	61.82
5	6.99	4096	128	35.9
5	6.99	4096	256	44.59
5	7	16384	8	165.05
5	7	16384	32	184.74
5	7	16384	64	187.35
5	7	16384	128	190.41
5	7	16384	256	141.97
5	7	65536	8	231.59
5	7	65536	32	229.44
5	7	65536	64	205.88
5	7	65536	128	158.81
5	7	65536	256	253.5
5	7	262144	8	337.37
5	7	262144	32	371.25
5	7	262144	64	493.86
5	7	262144	128	583.33
5	7	262144	256	359.79
5	7	1048576	8	954.97
5	7	1048576	32	1464.49
5	7	1048576	64	1199.94
5	7	1048576	128	1620.61
5	7	1048576	256	1401.12
5.01	7	4194304	8	1829.25
5.01	7	4194304	32	3075.77
5.01	7	4194304	64	3629.15
5.01	7	4194304	128	2871.85
5.01	7	4194304	256	3684.92

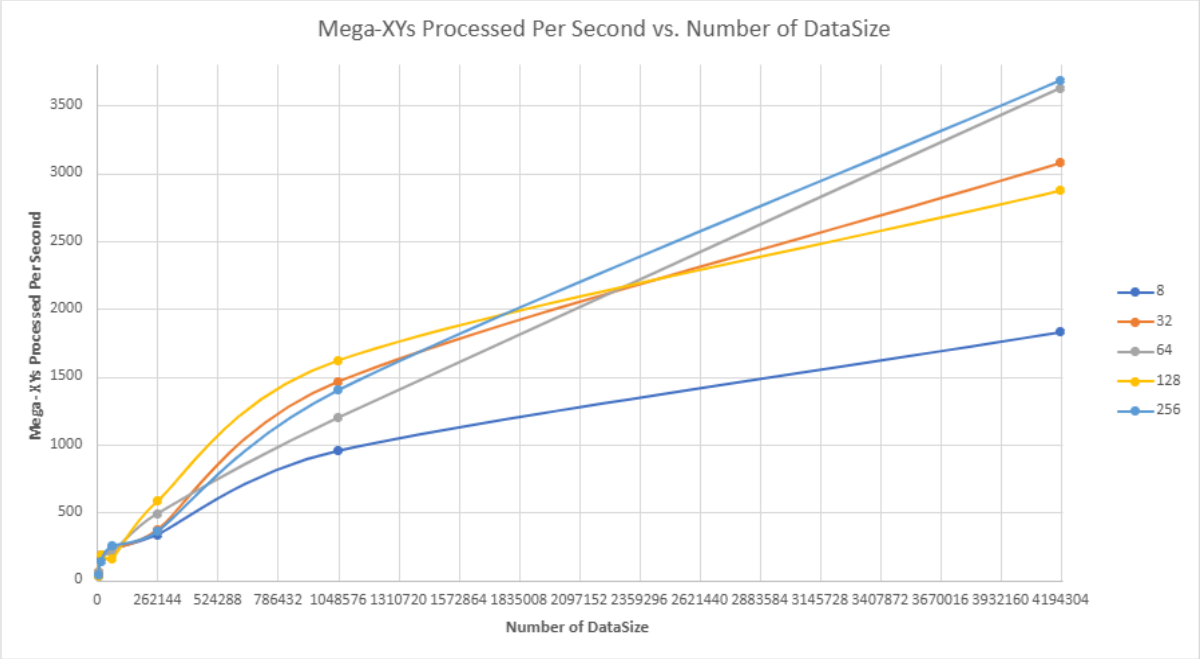
Datasize \ LocalSize	8	32	64	128	256
4096	29.21	58.76	61.82	35.9	44.59
16384	165.05	184.74	187.35	190.41	141.97
65536	231.59	229.44	205.88	158.81	253.5
262144	337.37	371.25	493.86	583.33	359.79
1048576	954.97	1464.49	1199.94	1620.61	1401.12
4194304	1829.25	3075.77	3629.15	2871.85	3684.92

< **Table 1.** The tables of Result (Left) and Performance (Right) >

The program ran on rabbit server, and there are two tables: result and performance. In the result table, ‘m’ and ‘b’ are equal to 5 and 7 respectively. In the data file “p6.data”, the pair of X and Y imply that the value X is quite small, meaning

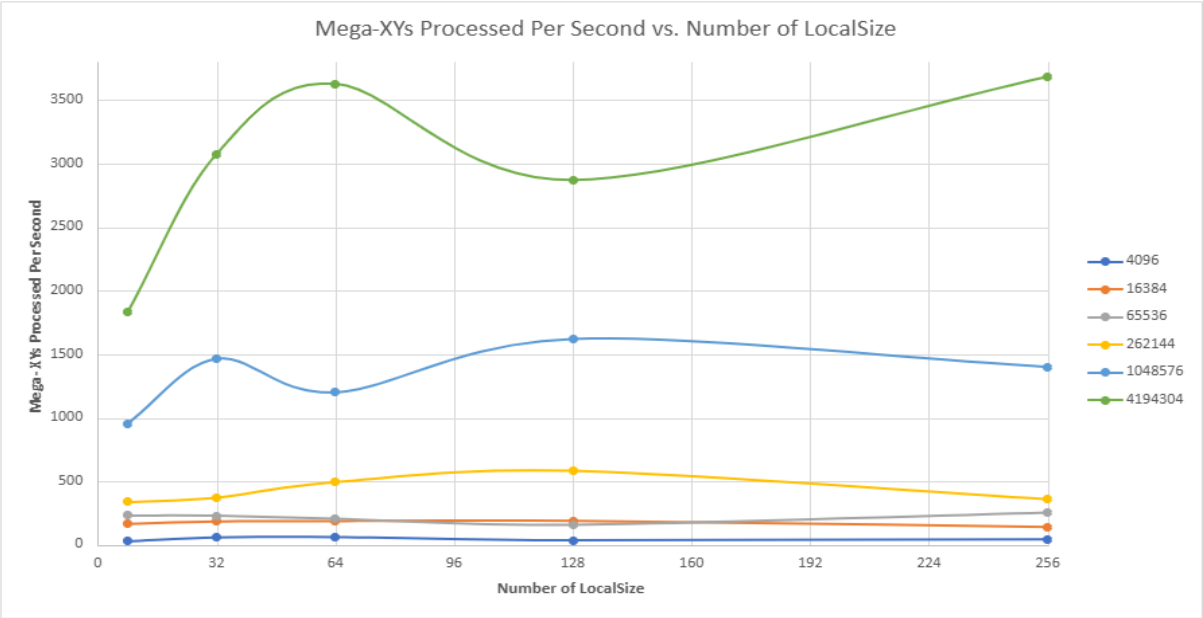
that X has a little influence on the result of the equation, and the value Y is fluctuated, meaning that Y is in the larger range than X. From the findings, we can guess that m and b in the result table hold the equation of "$y = mx + b$".

In the performance table, local size can significantly affect the performance when data size is large enough. However, when the data size is small enough, the huge number of local size does not necessarily guarantee higher performance than others since local size 64 has best performance when data size is 4096, not 256.



< **Figure 1.** The graph of performance versus DATASIZE, with LOCALSIZE >

As shown in Figure 1, the graph of performance shows that the number of local sizes can be affected by the number of data sizes. Most graphs gradually grows up except the local size 8. The local size 8 displays the worst performance when the number of data size is large enough.



< **Figure 2.** The graph of performance versus LOCALSIZE, with DATASIZE >

Figure 2 shows the performance graphs indicating that the correlation between the number of data size and local size. Like the previous graphs in Figure 1, the number of local size can considerably affect the performance when the data size is large enough. In the case of local size 256, the performance is best when the size of data is equal to 4194304.

3. **What patterns are you seeing in the performance curves? What difference does the size of data make? What difference does the size of each work-group make?**

In general, the larger the size of data and local are, the higher performance displays. The graphs in Figure 1 gradually grows up, depending on the size of data. Likely, the graphs in Figure 2 shows that the most local size has best performance. However, the performance is not necessarily guaranteed when the local size is large. For example, In the data size 4194304, the performance of local size 64 is better than

one of local size 128.

When the size of data is large enough, the performance depends on the number of each work-group. Particularly, the size of data is small enough, there is a little difference between graphs, and even the small number of work-group has better performance.

When the size of each work-group is large enough, the performance depends on data size. Data size affects work-group's doing since the data size is small enough, there would be some of idle work-group.

4. Why do you think the patterns look this way?

There are some factors affecting the performance of the program. Not only local and data size, but the server conditions also can affect it. Most trials show minor differences, but they produce different patterns of graphs. What's more, the communication between CPU and GPU may yield different patterns of graphs. Some code in Project #6 send data to GPU for computation, and the results from GPU go back to the CPU for next instruction.