CSE 5441 SP2021 (8386)

Lab 4 CUDA

Mengfan Zhu

04/12/2021

1 Final runtimes

Run all data files using the serial program and parallel programs. Table 1 shows the "real" time from time(1).

Table 1: Overall runtimes of serial, arbitrary order parallel and SIMD order parallel programs

Data File	Serial	Arbitrary Parallel	SIMD Parallel
PC_data_x1	0m03.506s	5m38.352s	5m14.070s
PC_data_t00100	0m22.443s	5m33.709s	3m48.913s
PC_data_t01000	3m26.711s	15m1.467s	6m25.927s
PC_data_t05000	17m35.247s	20m16.154s	7m55.423s
PC_data_t10000	34m51.188s	27m23.559s	9m44.418s
PC_data_t50000	246m38.747s	115m51.851s	41m25.817s

Table 2 shows the total producer runtime from time(2).

Table 2: Producer runtimes of serial, arbitrary order parallel and SIMD order parallel programs

Data File	Serial	Arbitrary Parallel	SIMD Parallel
PC_data_x1	2s	2.82min	2.62min
PC_data_t00100	12s	2.78min	1.92min
PC_data_t01000	1.72min	7.51min	3.22min
PC_data_t05000	8.72min	10.13min	3.97min
PC_data_t10000	17.32min	13.70min	4.87min
PC_data_t50000	123.23min	57.92min	20.72min

Table 3 shows the total consumer runtime from time(2).

Table 3: Consumer runtimes of serial, arbitrary order parallel and SIMD order parallel programs

Data File	Serial	Arbitrary Parallel	SIMD Parallel
PC_data_x1	2s	2.82min	2.62min
PC_data_t00100	10s	2.77min	1.90min
PC_data_t01000	1.73min	7.50min	3.22min
PC_data_t05000	8.87min	10.13min	3.95min
PC_data_t10000	17.53min	13.68min	4.87min
PC_data_t50000	123.42min	57.93min	20.72min

2 The max and min values for transformed keys

The max and min values for transformed keys for each data file are shown in Table 4.

Table 4: Max and min values for transformed keys

Data File	**max**	**min**
PC_data_x1	639	7
PC_data_t00100	1020	3
PC_data_t01000	1023	1
PC_data_t05000	1023	0
PC_data_t10000	1023	0
PC_data_t50000	511	6

3 Results summary

From Table 1, the overall runtimes for CUDA parallel programs are larger than the serial program when the size of data is small, and are much less than the serial program when the size of data is large.That shows CUDA parallel programs are more suitable for the cases when we have a huge number of data.

Besides, the runtimes for SIMD order parallel program are always less than the arbitrary parallel program. That shows reordering the input data for gpu can improve the performance.