

Parallelizing TSP Supplemental Info

Tables

GPU vs CPU Speedup

Dataset Name	Dataset Size	CPU time(ms)	GPU Time(ms)	CPU Result	GPU Result	Speedup	Dataset Source
Ulysses16	16	0.15	54.04	75.1654	75.1654	0.003	TSPLIB
Att48	48	31	190.06	35579.4	35579.42	0.163	TSPLIB
KroA100	100	603.91	732.28	23294.31	23294.35	0.825	TSPLIB
a280	280	1967.17	334.48	2749.78	2749.78	5.881	TSPLIB
gr202	202	3242.83	492.63	513.77	513.77	6.583	TSPLIB
pcb442	442	91507.3	2411.9	94.17	97.94	37.94	TSPLIB
gr666	666	331424.05	5423.07	3353.81	3353.81	61.114	TSPLIB
uy734	734	1381234	24482.07	87106.21	87106.21	56.418	WATERLOO
mu1979	1979	77032000	2004000	NA	91303	inf	WATERLOO

GPU Scalability

Devices	Time(ms)	CPU Time(ms)	Speedup
1	93201.24	1381234	14.82
2	48105.81	1381234	28.71
3	33096.63	1381234	41.73
4	24482.07	1381234	56.42

GPU Profiling

Dataset Name	Dataset Size	Kernel Time Percent	Comm Overhead Percent	Kernel Time-Overhead Ratio	Avg Occupancy
Ulysses16	16	82.42	17.58	4.69	0.75
Att48	48	83.75	16.25	5.15	4
KroA100	100	85.29	14.71	5.8	4
a280	280	90.95	9.05	10.05	17.5
gr202	202	88.28	11.72	7.53	35.5
pcb442	442	94.25	5.75	16.39	70
gr666	666	96.45	3.55	27.17	79
uy734	734	97.19	2.81	34.59	84
mu1979	1979	98.9	1.1	89.91	96

Screenshots

Nvprof results of Ulysses16 and Att48

```

-- vi --/bash_profile -- .08@cuda5:~/gpu/project -- ssh ppo208@access.cims.nyu.edu -- ssh ppo208@cuda5:~/ -- ssh ppo208@access.cims.nyu.edu

==36842== NVPROF is profiling process 36842, command: ./gpu_tsp 16 data/a16 data/a16_optimum
Time taken = 661.745668 milliseconds
Global minimum value is 75.166413
==36842== Profiling application: ./gpu_tsp 16 data/a16 data/a16_optimum
==36842== Profiling result:
Type      Time      Time      Calls      Avg      Min      Max      Name
GPU activities: 82.77%  2.648ms  112  22.682us  22.847us  23.744us  two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
12.72%  398.34us  224  1.7439us  1.4888us  6.9120us  [CUDA memcopy DtoH]
4.52%  138.62us  116  1.1950us  1.8888us  1.8888us  [CUDA memcopy HtoD]
API calls: 95.63%  658.33ms  16  41.146ms  4.9778us  192.86ms  cudaMalloc
2.39%  16.423ms  32  513.22us  3.2840us  13.284ms  cudaStreamCreate
1.18%  7.5762ms  336  22.548us  6.1598us  57.882us  cudaMemcopyAsync
0.32%  2.1928ms  388  5.6490us  153ns  230.37us  cuDeviceGetAttribute
0.26%  1.7918ms  112  15.991us  11.484us  77.207us  cudaLaunchKernel
0.15%  1.8538ms  4  265.45us  256.93us  275.49us  cuDeviceTotalMem
0.05%  350.04us  13  26.926us  6.5768us  221.52us  cudaFree
0.04%  296.66us  112  2.6480us  1.9258us  9.2430us  cudaStreamSynchronize
0.02%  177.92us  4  44.480us  48.584us  52.759us  cuDeviceGetName
0.02%  189.85us  4  27.443us  22.267us  33.232us  cudaMemcopy
0.01%  62.602us  8  7.8250us  2.9590us  16.413us  cudaSetDevice
0.00%  21.513us  4  5.3780us  2.5268us  12.470us  cuDeviceGetPCIBusId
0.00%  3.6270us  1  3.6270us  3.6270us  3.6270us  cuDeviceGetDeviceCount
0.00%  3.2390us  8  404ns  198ns  764ns  cuDeviceGet
0.00%  2.3140us  3  771ns  152ns  1.3930us  cuDeviceGetDeviceCount
0.00%  894ns  4  223ns  176ns  314ns  cuDeviceGetUuid

[ppo208@cuda5 project]$ nvprof ./gpu_tsp 48 data/a48 data/a48_optimum
==13305== NVPROF is profiling process 13305, command: ./gpu_tsp 48 data/a48 data/a48_optimum
Time taken = 663.428867 milliseconds
Global minimum value is 35379.429688
==13305== Profiling application: ./gpu_tsp 48 data/a48 data/a48_optimum
==13305== Profiling result:
Type      Time      Time      Calls      Avg      Min      Max      Name
GPU activities: 83.72%  58.238ms  2448  23.784us  22.816us  25.376us  two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
12.04%  8.3762ms  4896  1.7180us  1.4888us  11.200us  [CUDA memcopy DtoH]
4.23%  2.9481ms  2452  1.1990us  1.1190us  11.232us  [CUDA memcopy HtoD]
74.75%  612.62ms  16  38.289ms  5.2940us  159.79ms  cudaMalloc
19.54%  169.28ms  7344  21.855us  5.7778us  2.8518ms  cudaMemcopyAsync
3.88%  31.819ms  2448  12.998us  10.457us  1.5893ms  cudaLaunchKernel
0.78%  6.3765ms  2448  2.6840us  1.6778us  1.3697ms  cudaStreamSynchronize
0.51%  4.2808ms  32  131.17us  2.6230us  996.67us  cudaStreamCreate
0.29%  2.3715ms  388  6.1110us  168ns  370.58us  cuDeviceGetAttribute
0.13%  1.0381ms  4  259.53us  257.23us  261.57us  cuDeviceTotalMem
0.05%  442.68us  13  34.052us  7.8749us  291.13us  cudaFree
0.02%  193.15us  4  48.207us  42.197us  66.378us  cuDeviceGetName
0.02%  129.02us  4  32.253us  27.766us  40.051us  cudaMemcopy
0.01%  69.229us  8  8.6538us  3.1710us  16.213us  cudaSetDevice
0.00%  20.286us  4  5.0710us  2.9880us  10.591us  cuDeviceGetPCIBusId
0.00%  3.5570us  1  3.5570us  3.5570us  3.5570us  cudaGetDeviceCount
0.00%  3.1300us  8  391ns  199ns  817ns  cuDeviceGet
0.00%  2.4230us  3  807ns  156ns  1.7390us  cuDeviceGetDeviceCount
0.00%  1.2380us  4  307ns  214ns  489ns  cuDeviceGetUuid

[ppo208@cuda5 project]$

```

Nvprof results of KroA100 and a280

```
==42214== NVPROF is profiling process 42214, command: ./gpu_tsp 100 data/kroA100 data/kroA100_optimum
Time taken = 824.868584 milliseconds
Global minimum value is 23294.351562
==42214== Profiling application: ./gpu_tsp 100 data/kroA100 data/kroA100_optimum
==42214== Profiling result:
Type      Time(%)    Time      Calls      Avg      Min      Max      Name
GPU activities: 85.25%  298.50ms  10700  27.149us  25.439us  28.880us  two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
                10.85%  36.98ms  21400  1.7280us  1.5990us  12.880us  [CUDA memcpy DtoH]
                3.90%  13.28ms  10704  1.2418us  1.1510us  11.528us  [CUDA memcpy HtoD]
API calls:      47.31%  693.70ms  32100  21.610us  5.3550us  1.8337ms  cudaMemcpyAsync
                41.39%  686.88ms  16      37.938ms  5.6140us  159.36ms  cudaMalloc
                9.17%  134.43ms  10700  12.563us  9.9560us  1.9588ms  cudaLaunchKernel
                1.60%  23.46ms  10700  2.1920us  1.6880us  583.89us  cudaStreamSynchronize
                0.27%  3.9354ms  32      122.98us  2.9230us  906.96us  cudaStreamCreate
                0.14%  2.1225ms  388      5.4700us  152ns  214.09us  cuDeviceGetAttribute
                0.07%  996.18us  4      249.04us  242.24us  253.79us  cuDeviceTotalMem
                0.03%  449.38us  13      34.567us  6.2880us  269.35us  cudaFree
                0.01%  177.26us  4      44.315us  38.963us  48.257us  cudaMemcpy
                0.01%  177.11us  4      44.277us  41.057us  52.357us  cuDeviceGetName
                0.00%  48.779us  8      6.0970us  2.6190us  12.354us  cudaSetDevice
                0.00%  18.205us  4      4.6510us  1.7260us  10.441us  cuDeviceGetPCIBusId
                0.00%  3.5970us  1      3.5970us  3.5970us  3.5970us  cuDeviceGetCount
                0.00%  2.9610us  8      378ns  183ns  678ns  cuDeviceGet
                0.00%  2.1320us  3      710ns  244ns  1.2530us  cuDeviceGetCount
                0.00%  902ns  4      225ns  172ns  325ns  cuDeviceGetUuid
[ppo200@cuda5 project]$ nvprof ./gpu_tsp 280 data/a280 data/a280_optimum
==1197== NVPROF is profiling process 1197, command: ./gpu_tsp 280 data/a280 data/a280_optimum
Time taken = 725.388286 milliseconds
Global minimum value is 2749.783203
==1197== Profiling application: ./gpu_tsp 280 data/a280 data/a280_optimum
==1197== Profiling result:
Type      Time(%)    Time      Calls      Avg      Min      Max      Name
GPU activities: 90.95%  217.10ms  4480  48.459us  46.399us  53.439us  two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
                6.46%  15.415ms  8960  1.7200us  1.6310us  11.232us  [CUDA memcpy DtoH]
                2.59%  6.1803ms  4480  1.3780us  1.2790us  38.392us  [CUDA memcpy HtoD]
API calls:      56.59%  686.47ms  16      37.842ms  4.4920us  185.14ms  cudaMalloc
                36.50%  390.53ms  13440  29.057us  6.1340us  1.5529ms  cudaMemcpyAsync
                5.27%  56.364ms  4480  12.581us  10.244us  1.4250ms  cudaLaunchKernel
                0.85%  9.0929ms  4480  2.0290us  1.7280us  9.9730us  cudaStreamSynchronize
                0.34%  3.6260ms  32      113.32us  2.8330us  900.37us  cudaStreamCreate
                0.23%  2.4476ms  388      6.3800us  152ns  347.89us  cuDeviceGetAttribute
                0.10%  1.0462ms  4      261.55us  249.28us  284.12us  cuDeviceTotalMem
                0.05%  532.99us  4      133.25us  112.80us  175.55us  cudaMemcpy
                0.04%  450.89us  13      34.683us  6.0430us  292.27us  cudaFree
                0.02%  183.22us  4      45.821us  42.125us  52.276us  cuDeviceGetName
                0.01%  61.835us  8      7.7290us  2.7180us  19.373us  cudaSetDevice
                0.00%  21.423us  4      5.3550us  2.1520us  13.183us  cuDeviceGetPCIBusId
                0.00%  3.9340us  1      3.9340us  3.9340us  3.9340us  cuDeviceGetCount
                0.00%  3.5350us  8      441ns  180ns  784ns  cuDeviceGet
                0.00%  2.1610us  3      728ns  159ns  1.3650us  cuDeviceGetCount
                0.00%  1.2480us  4      312ns  202ns  419ns  cuDeviceGetUuid
[ppo200@cuda5 project]$
```

Nvprof results of gr202 and pcb442

```
==36340== NVPROF is profiling process 36340, command: ./gpu_tsp 202 data/gr202 data/gr_202_optimum
Time taken = 955.965181 milliseconds
Global minimum value is 513.771362
==36340== Profiling application: ./gpu_tsp 202 data/gr202 data/gr_202_optimum
==36340== Profiling result:
Type Time(%) Time Calls Avg Min Max Name
GPU activities: 88.28% 504.11ms 14140 35.651us 33.920us 39.008us two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
8.49% 48.469ms 28280 1.7130us 1.5990us 13.824us [CUDA memcpy DtoH]
3.23% 18.424ms 14144 1.3020us 1.2150us 28.928us [CUDA memcpy HtoD]
API calls: 55.74% 1.06364s 42420 25.074us 5.4390us 1.9981ms cudaMemcpyAsync
32.63% 622.67ms 16 38.917ms 4.5210us 161.09ms cudaMalloc
9.46% 180.50ms 14140 12.765us 10.175us 662.26us cudaLaunchKernel
1.73% 33.074ms 14140 2.3390us 1.6880us 1.7439ms cudaStreamSynchronize
0.21% 4.0542ms 32 126.69us 3.0170us 937.37us cudaStreamCreate
0.11% 2.1404ms 388 5.5160us 164ns 213.77us cuDeviceGetAttribute
0.05% 1.0475ms 4 261.87us 257.12us 271.76us cuDeviceTotalMem
0.03% 634.61us 13 48.816us 6.6260us 436.33us cudaFree
0.02% 290.83us 4 72.708us 66.772us 87.770us cudaMemcpy
0.01% 174.62us 4 43.655us 42.027us 48.085us cuDeviceGetName
0.00% 81.446us 8 10.180us 3.2330us 40.409us cudaSetDevice
0.00% 19.039us 4 4.7590us 1.8510us 11.499us cuDeviceGetPCIBusId
0.00% 3.1810us 1 3.1810us 3.1810us 3.1810us cudaGetDeviceCount
0.00% 3.1450us 8 393ns 176ns 653ns cuDeviceGet
0.00% 1.7360us 3 578ns 160ns 1.0050us cuDeviceGetCount
0.00% 977ns 4 244ns 207ns 327ns cuDeviceGetUuid
[ppo208@cuda5 project]$ nvprof ./gpu_tsp 442 data/pcb442 data/pcb442_optimum
==5888== NVPROF is profiling process 5888, command: ./gpu_tsp 442 data/pcb442 data/pcb442_optimum
Time taken = 3373.702306 milliseconds
Global minimum value is 97.941277
==5888== Profiling application: ./gpu_tsp 442 data/pcb442 data/pcb442_optimum
==5888== Profiling result:
Type Time(%) Time Calls Avg Min Max Name
GPU activities: 94.26% 5.52363s 80670 68.471us 56.383us 101.41us two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
4.00% 234.28ms 161340 1.4520us 1.3120us 11.584us [CUDA memcpy DtoH]
1.75% 102.52ms 80674 1.2700us 1.1510us 106.78us [CUDA memcpy HtoD]
API calls: 82.67% 9.11184s 242010 37.650us 6.1250us 10.258ms cudaMemcpyAsync
10.03% 1.10606s 80670 13.710us 10.257us 3.7853ms cudaLaunchKernel
5.55% 611.90ms 16 38.244ms 4.0210us 165.34ms cudaMalloc
1.65% 181.54ms 80670 2.2500us 1.6990us 1.6411ms cudaStreamSynchronize
0.04% 3.9181ms 32 122.44us 3.0200us 932.53us cudaStreamCreate
0.02% 2.5898ms 13 199.22us 7.6790us 1.8545ms cudaFree
0.02% 2.0742ms 388 5.3450us 151ns 200.69us cuDeviceGetAttribute
0.01% 1.0460ms 4 261.50us 239.76us 279.21us cudaMemcpy
0.01% 997.60us 4 249.40us 234.67us 271.82us cuDeviceTotalMem
0.00% 189.00us 4 47.269us 41.449us 56.351us cuDeviceGetName
0.00% 41.265us 8 5.1580us 3.1020us 10.925us cudaSetDevice
0.00% 18.096us 4 4.5240us 2.0270us 9.7460us cuDeviceGetPCIBusId
0.00% 5.5520us 8 694ns 187ns 2.9890us cuDeviceGet
0.00% 2.8860us 1 2.8860us 2.8860us 2.8860us cudaGetDeviceCount
0.00% 2.0790us 3 693ns 186ns 1.1310us cuDeviceGetCount
0.00% 1.0390us 4 259ns 166ns 364ns cuDeviceGetUuid
[ppo208@cuda5 project]$
```

Nvprof results of gr666 and uy734

```
==2187== NVPROF is profiling process 2187, command: ./gpu_tsp 666 data/a666 data/a666_optimum
==2187== Warning: Profiling results might be incorrect with current version of nvcc compiler used to compile cuda app. Compile with nvcc compiler 9.0 or later version to get correct profiling results. Ignore this warning if code is already compiled with the recommended nvcc version
Time taken = 6779.433248 milliseconds
Global minimum value is 9353.886641
==2187== Error: Internal profiling error 4055:34.
===== Profiling result:
Type      Time(%)    Time      Calls      Avg      Min      Max      Name
GPU activities: 96.45%  15.2382s  130536  116.74us  103.26us  172.61us  two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
                2.41%   380.24ms  261072  1.4560us  1.3110us  15.679us  [CUDA memcpy DtoH]
                1.14%   180.34ms  130540  1.3810us  1.2480us  650.26us  [CUDA memcpy HtoD]
API calls: 88.06%  20.7577s  391608  53.006us  6.3750us  49.628ms  cudaMemcpyAsync
                7.84%   1.8482s  130536  14.158us  11.440us  713.44us  cudaLaunchKernel
                2.83%   665.96ms  16      41.622ms  4.8000us  186.45ms  cudaMalloc
                1.22%   287.49ms  130536  2.2020us  1.6730us  581.82us  cudaStreamSynchronize
                0.02%   3.9071ms  4       976.79us  589.22us  1.4642ms  cudaMemcpy
                0.01%   2.6345ms  384      6.8600us  160ns    649.31us  cuDeviceGetAttribute
                0.01%   2.4961ms  32      78.003us  2.7980us  660.00us  cudaStreamCreate
                0.01%   1.5053ms  13      115.80us  6.7250us  284.22us  cudaFree
                0.00%   1.0370ms  4       259.25us  251.06us  267.46us  cuDeviceTotalMem
                0.00%   103.83us  4       45.956us  42.133us  51.460us  cuDeviceGetName
                0.00%   76.386us  8       9.5480us  3.3630us  23.556us  cudaSetDevice
                0.00%   14.807us  4       3.7010us  1.8640us  7.8140us  cuDeviceGetPCIBusId
                0.00%   4.6660us  1       4.6660us  4.6660us  4.6660us  cudaGetDeviceCount
                0.00%   3.4340us  8         429ns  165ns    851ns    cuDeviceGet
                0.00%   1.6590us  3         553ns  181ns    948ns    cuDeviceGetCount
                0.00%   1.3180us  4         329ns  235ns    581ns    cuDeviceGetUuid
===== Error: CUDA profiling error.
[ppo2080@cuda5 project]$ nvprof ./gpu_tsp 734 data/a734 data/a734_optimum
Error in opening the file:===== Error: Application returned non-zero code 1
[ppo2080@cuda5 project]$ nvprof ./gpu_tsp 734 data/uy734 data/uy734_optimum
==41117== NVPROF is profiling process 41117, command: ./gpu_tsp 734 data/uy734 data/uy734_optimum
==41117== Warning: Profiling results might be incorrect with current version of nvcc compiler used to compile cuda app. Compile with nvcc compiler 9.0 or later version to get correct profiling results. Ignore this warning if code is already compiled with the recommended nvcc version
Time taken = 29185.559096 milliseconds
Global minimum value is 87106.218750
==41117== Error: Internal profiling error 4055:34.
===== Profiling result:
Type      Time(%)    Time      Calls      Avg      Min      Max      Name
GPU activities: 97.19%  77.9910s  495450  157.41us  123.17us  231.87us  two_opt(unsigned int*, float*, unsigned int, float*, unsigned int*)
                1.94%   1.5555s  990900  1.5690us  1.3120us  20.576us  [CUDA memcpy DtoH]
                0.87%   700.31ms  495454  1.4130us  1.2800us  653.24us  [CUDA memcpy HtoD]
API calls: 91.92%  98.8050s  1486350  66.474us  6.7190us  56.488ms  cudaMemcpyAsync
                6.48%   6.96729s  495450  14.062us  11.590us  12.618ms  cudaLaunchKernel
                0.90%   1.05794s  495450  2.1350us  1.6640us  654.80us  cudaStreamSynchronize
                0.61%   652.63ms  16      40.789ms  6.3900us  182.70ms  cudaMalloc
                0.00%   3.1455ms  4       786.37us  620.12us  1.1688ms  cudaMemcpy
                0.00%   2.5762ms  32      80.506us  2.7210us  582.86us  cudaStreamCreate
                0.00%   2.1395ms  384      5.5710us  150ns    218.25us  cuDeviceGetAttribute
                0.00%   1.8030ms  13      138.69us  7.7850us  324.64us  cudaFree
                0.00%   1.0127ms  4       253.17us  249.28us  264.45us  cuDeviceTotalMem
                0.00%   237.91us  4       59.476us  41.481us  113.10us  cuDeviceGetName
                0.00%   53.976us  8        6.7470us  3.1520us  13.780us  cudaSetDevice
```