# P2Z Scaling Tracks and Bsize

Tres Reid

5/12/20

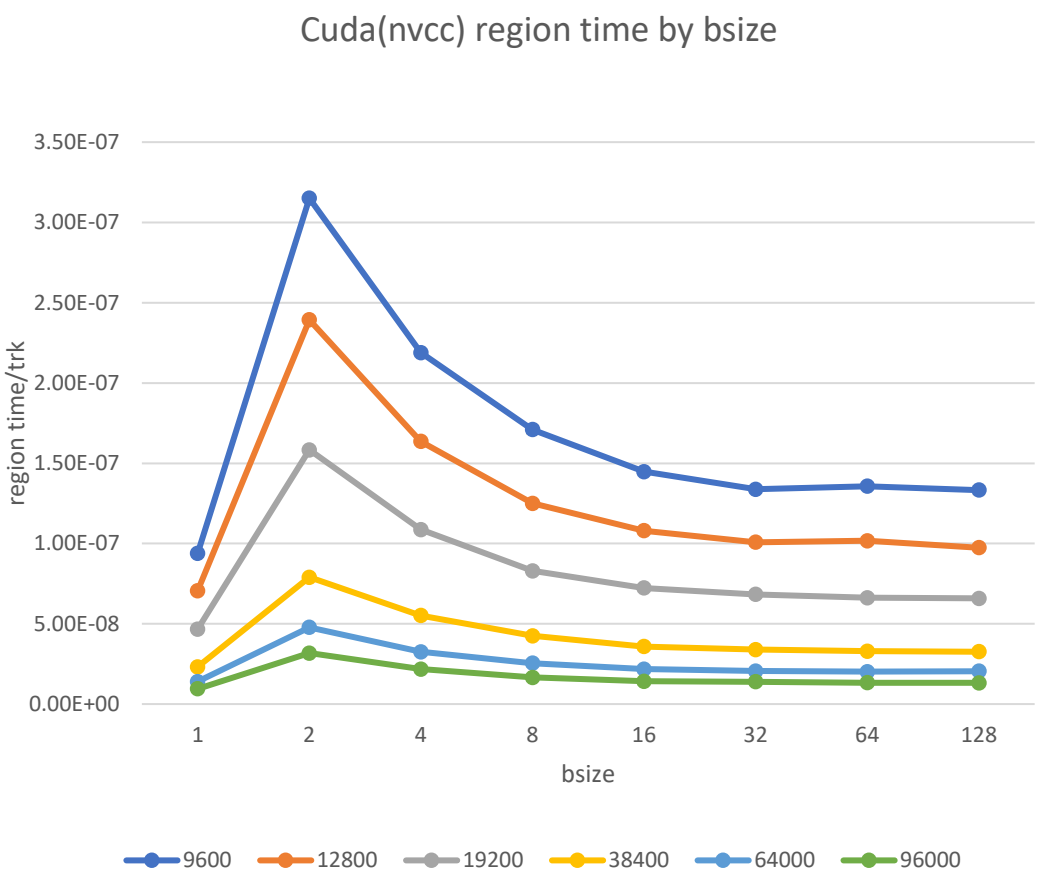
# Scaling

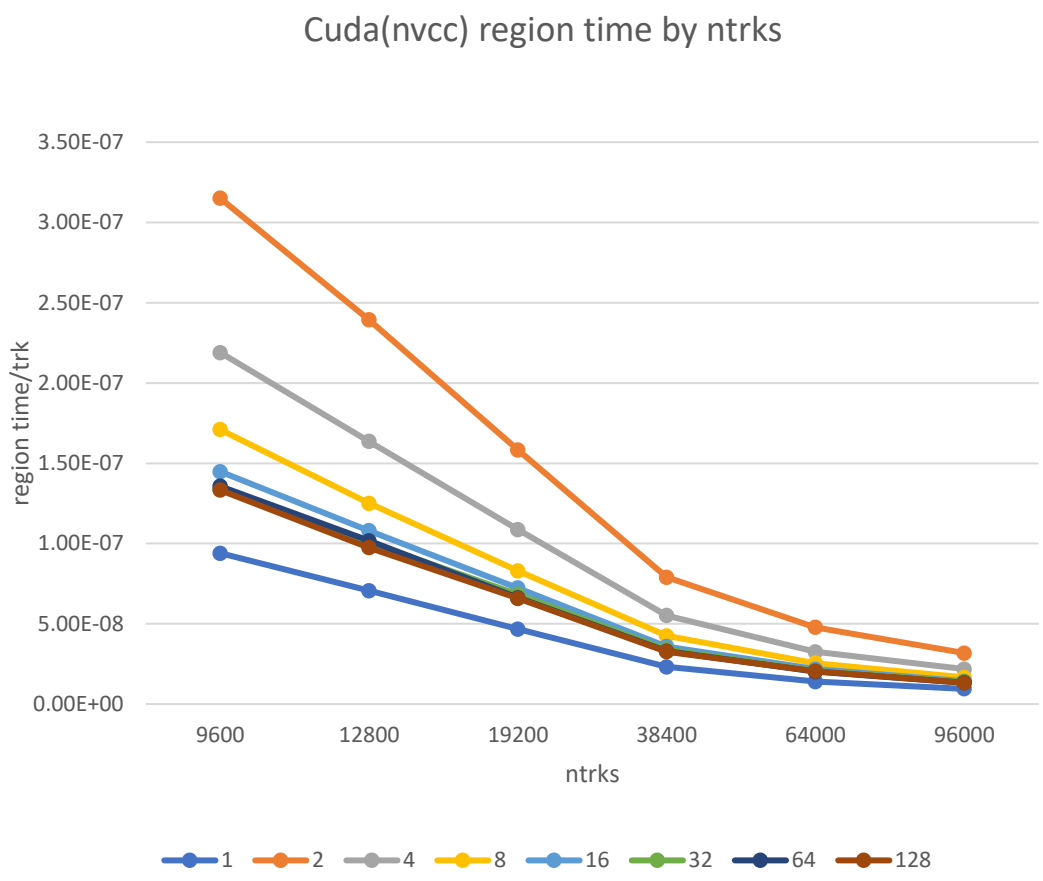
- Made a new tuning script and updated to Makefile and source codes to allow the number of tracks and bsize to be changed when making a particular version
  - Defaults set when to defined in Makefile
  - Points are found from taking the average over 5 runs
- Scaling number of tracks
  - Values of {9600,12800,19200,38400,64000,96000}
    - = 128*{75,100,150,300,500,750}
      - Values chosen to be divisible by 128 (so nb is a whole number through the range of bsize: nb= ntrks/bsize)
      - Quick scan across an order of magnitude starting from the previously used value
        - Gap could be filled in later if necessary
  - Cuda is the only mode that is affected by the number of tracks
    - More tracks -> faster runtime per track
    - Affects both memory transfer time and compute time for cuda
    - Does not affect ACC memory transfer or compute time
- Scaling the size of bsize (size of simd operations)
  - Powers of 2: {1,2,4,8,16,32,64,128}
    - I stopped at 128 since 9600 isn't divisible by 256.
    - I think CUDA doesn't have enough shared memory space after 128 (eigen looses memory at 32)
  - GPU
    - Memory transfer time doesn't depend on bsize.
    - Clear trend that higher bsize decrease computation time by orders of magnitude
      - Flattens out a bit around 32-> warp size
  - CPU
    - Less clear, but there is still a trend for higher bsize to speedup region time/trk
      - Factor of ~2 speedup going from 1 to 128 in most cases
      - Expected to flatten ~16 but this doesn't seem to be the case.
    - Not true for eigen

TODO

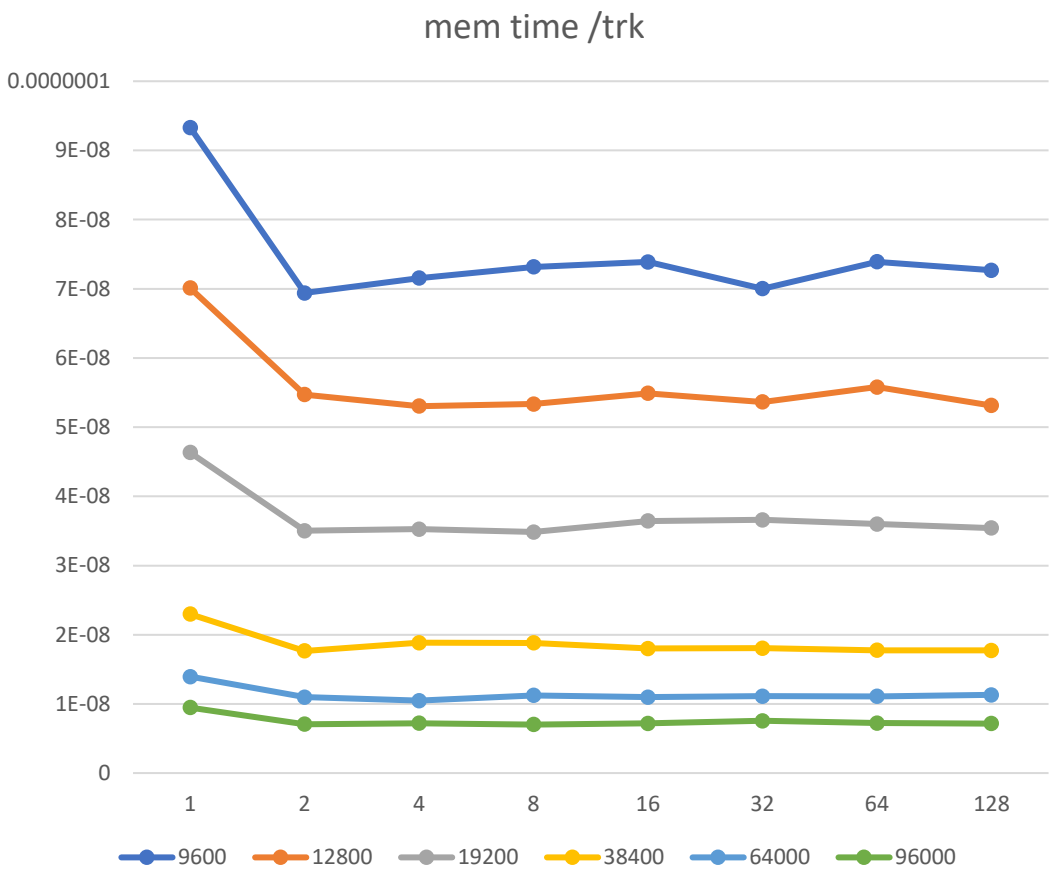
- Tune CUDA streams, blocks/grid, threads/block
- Tune tbb block size
- Try higher values of bsize?
  - Find a cutoff for all modes
  - Investigate shared memory cutoff for CUDA
  - Double check bsize dependence on linear scale?
- Other remaining tasks
  - Fix CUDA memory transfer by stream
  - Redo eigen implementation
  - Finish alpaca GPU implementation

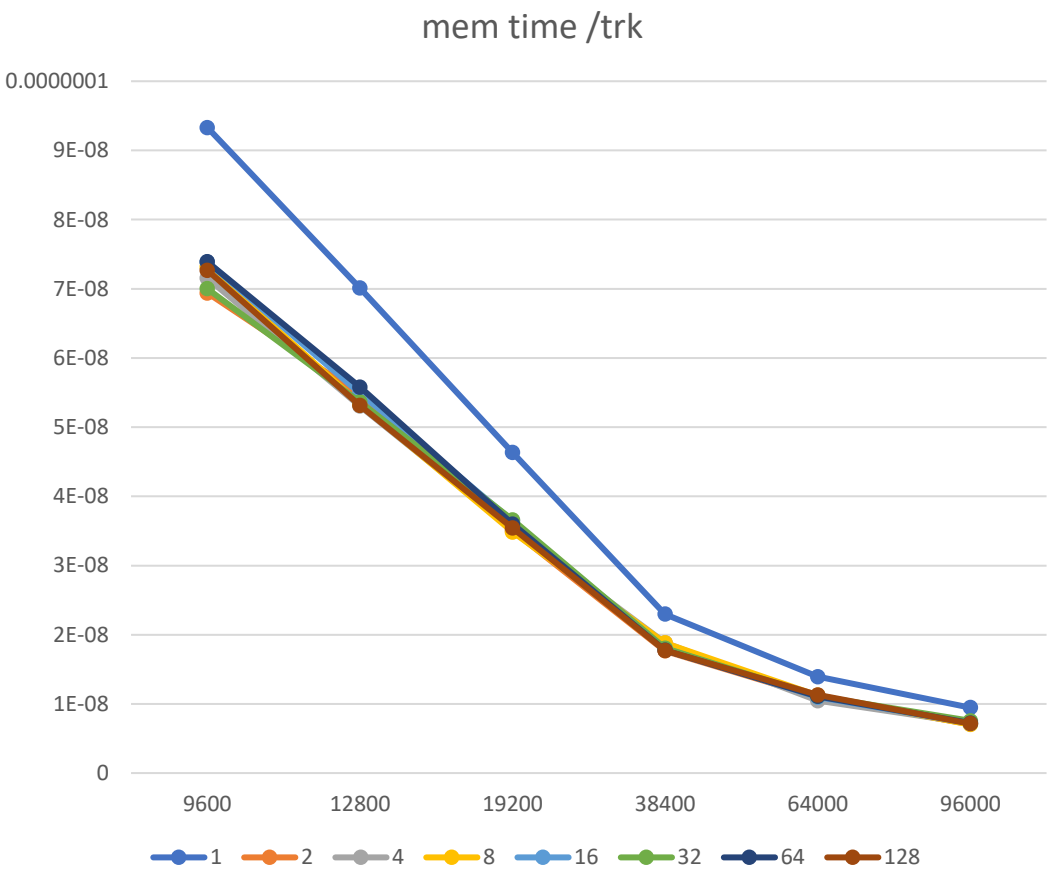
# Cuda nvcc


Cuda(nvcc) region time by bsize
3.50E-07
3.00E-07
2.50E-07
2.00E-07
1.50E-07
1.00E-07
5.00E-08
0.00E+00
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

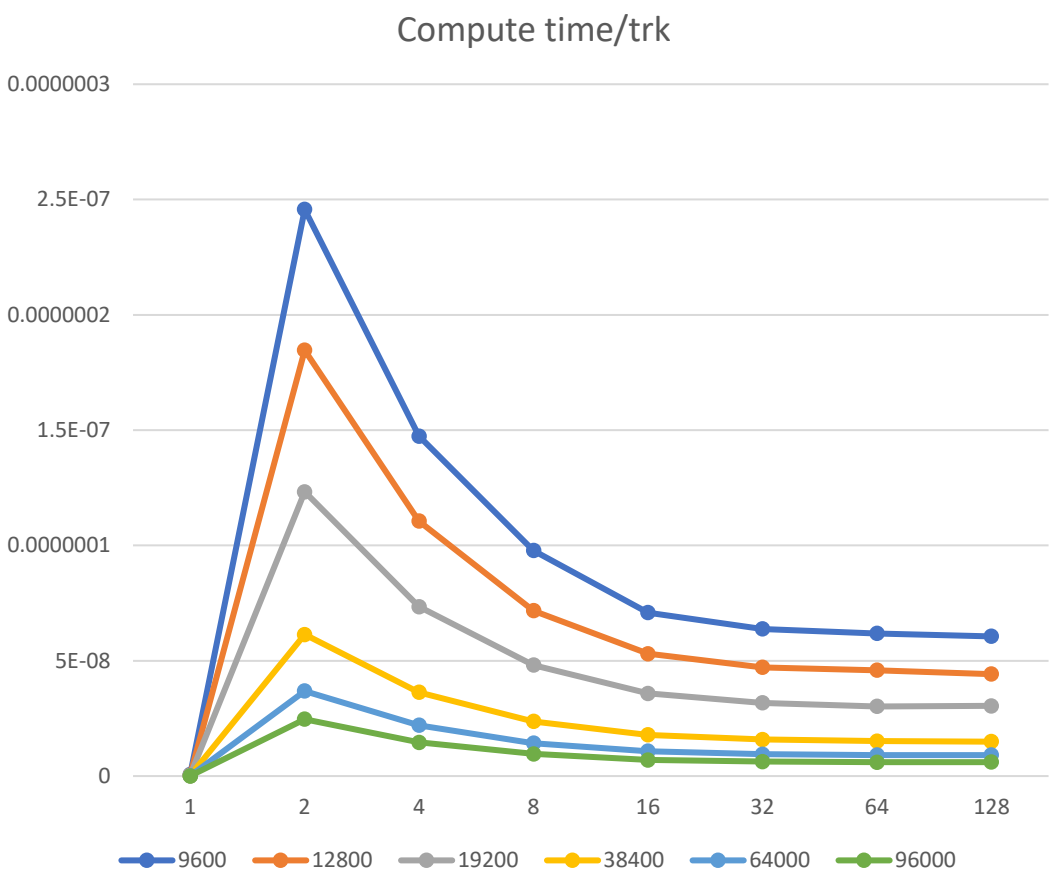

Cuda(nvcc) region time by ntrks
3.50E-07
3.00E-07
2.50E-07
2.00E-07
1.50E-07
1.00E-07
5.00E-08
0.00E+00
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

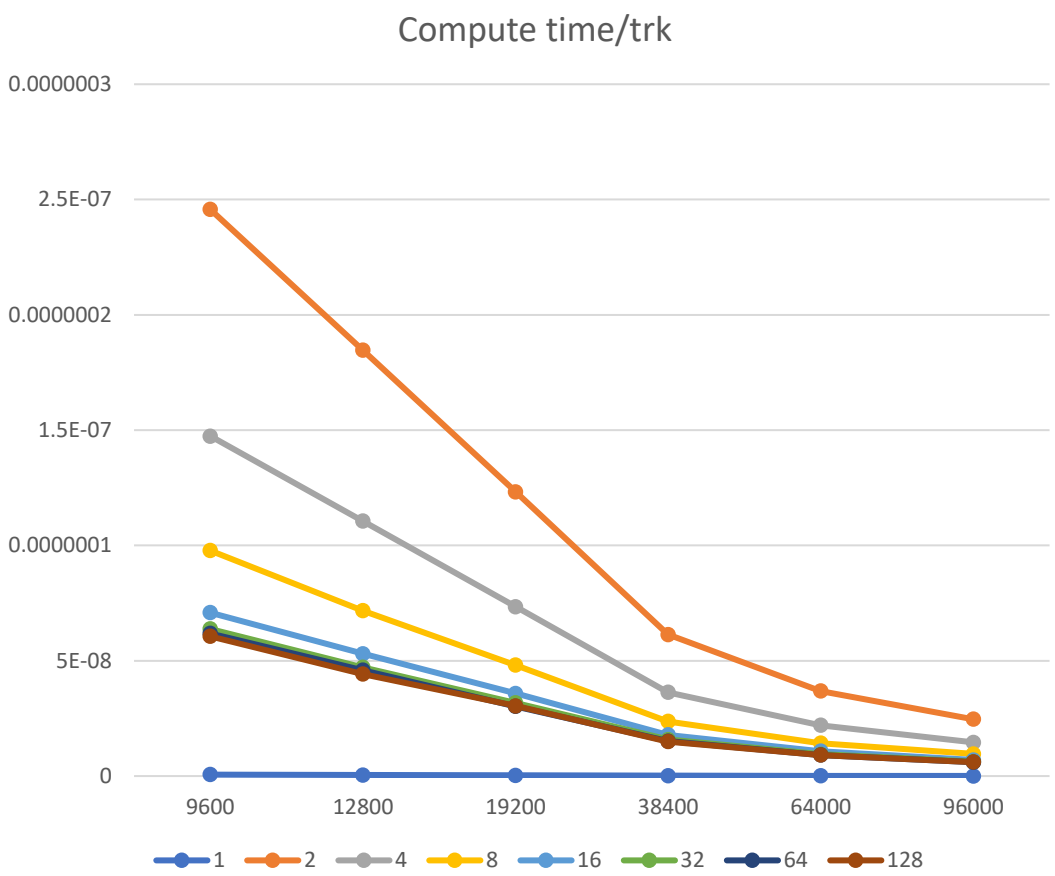
# Cuda mem transfer time



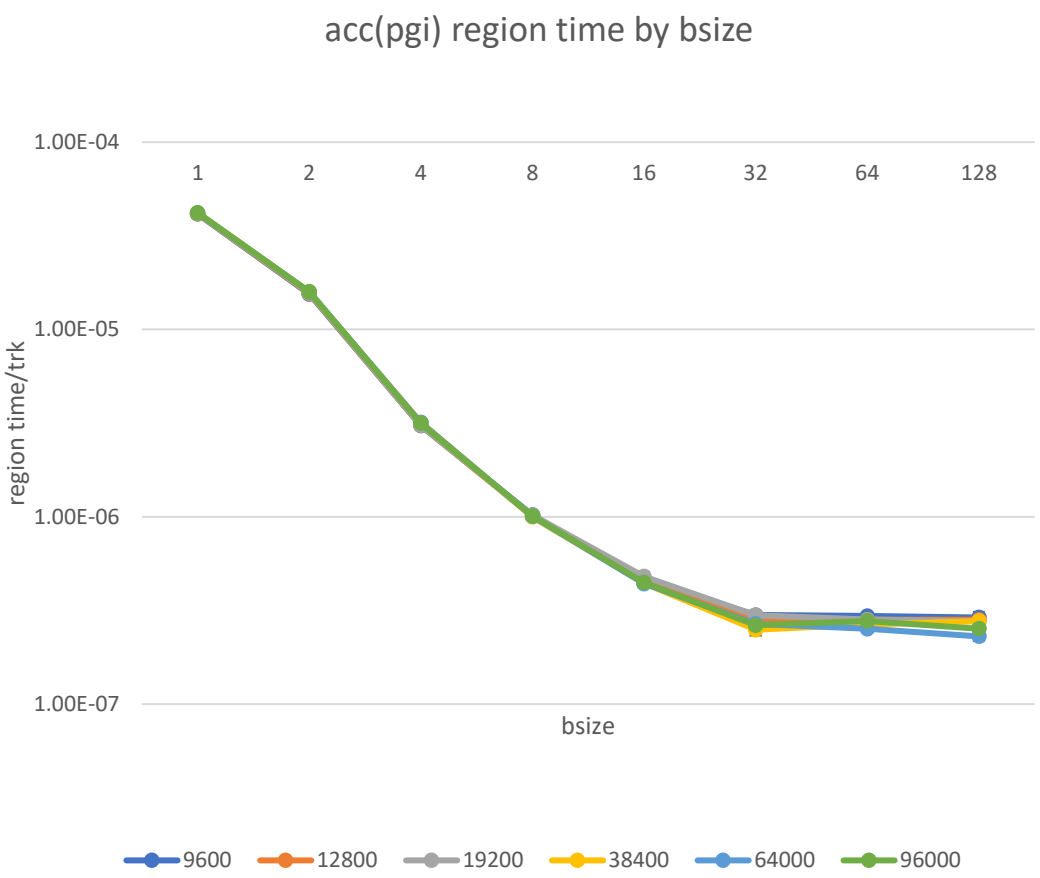



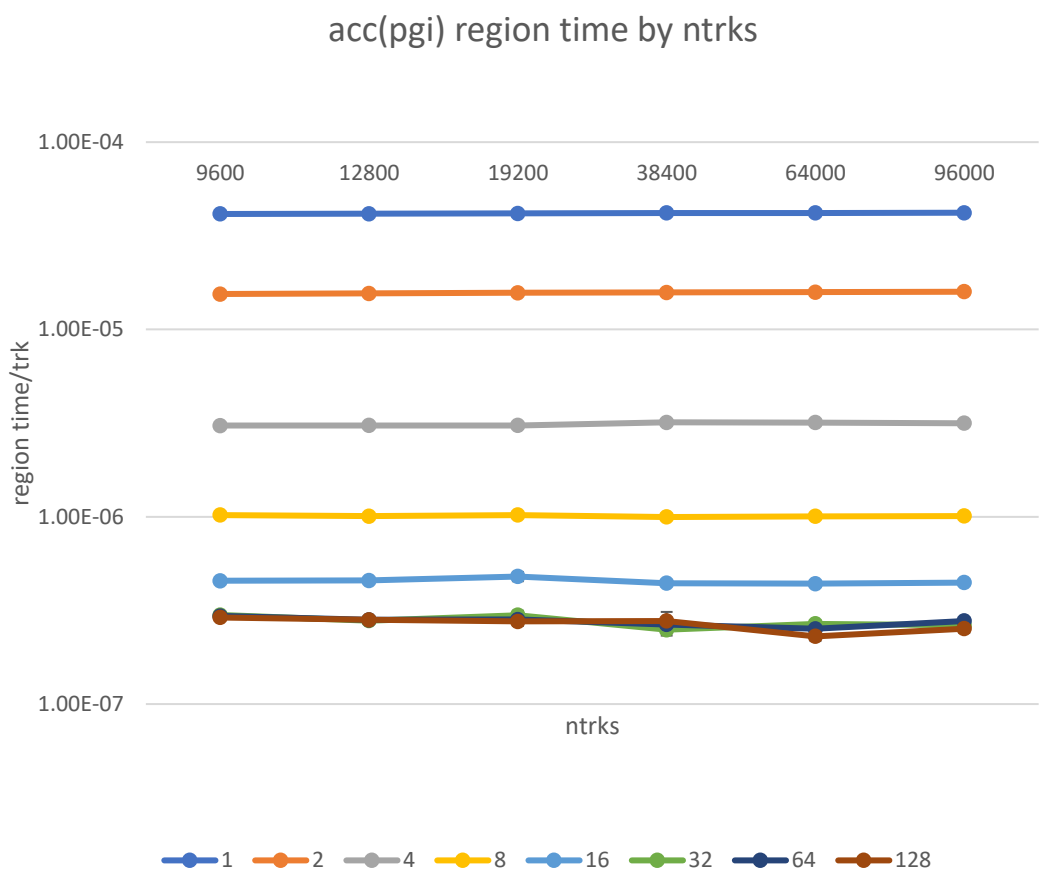
# Cuda computation time



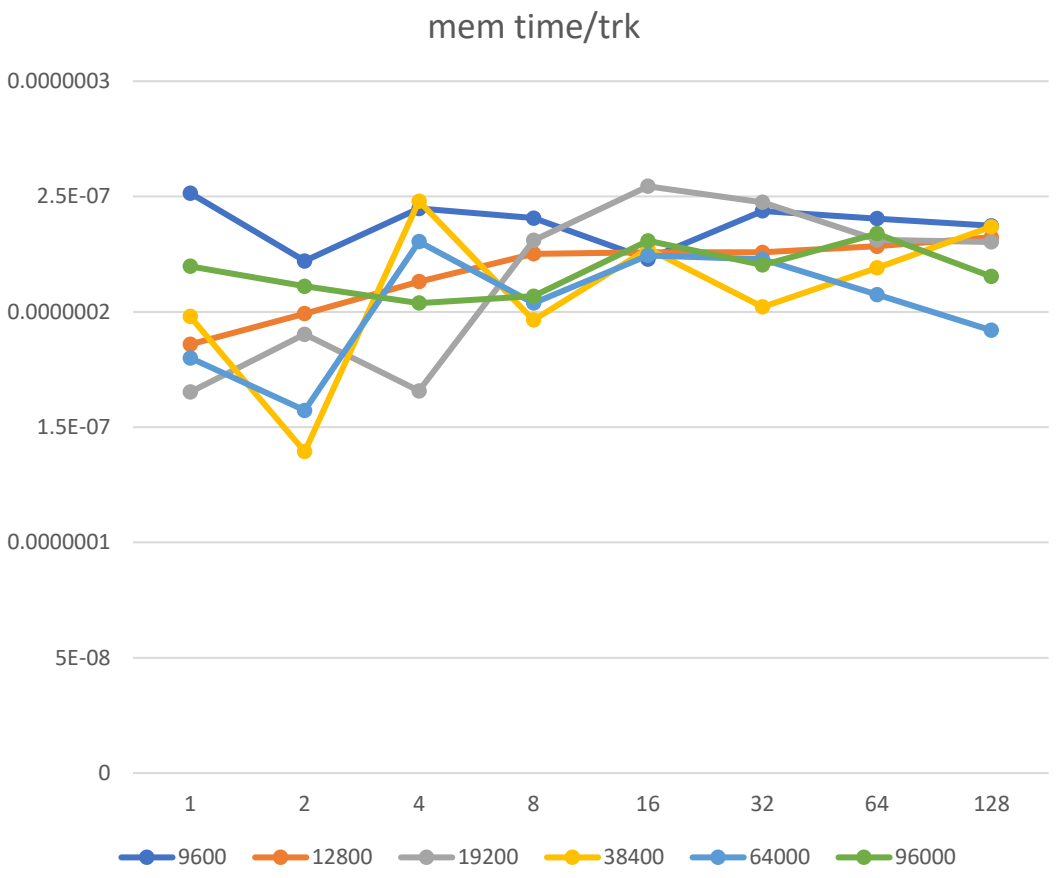



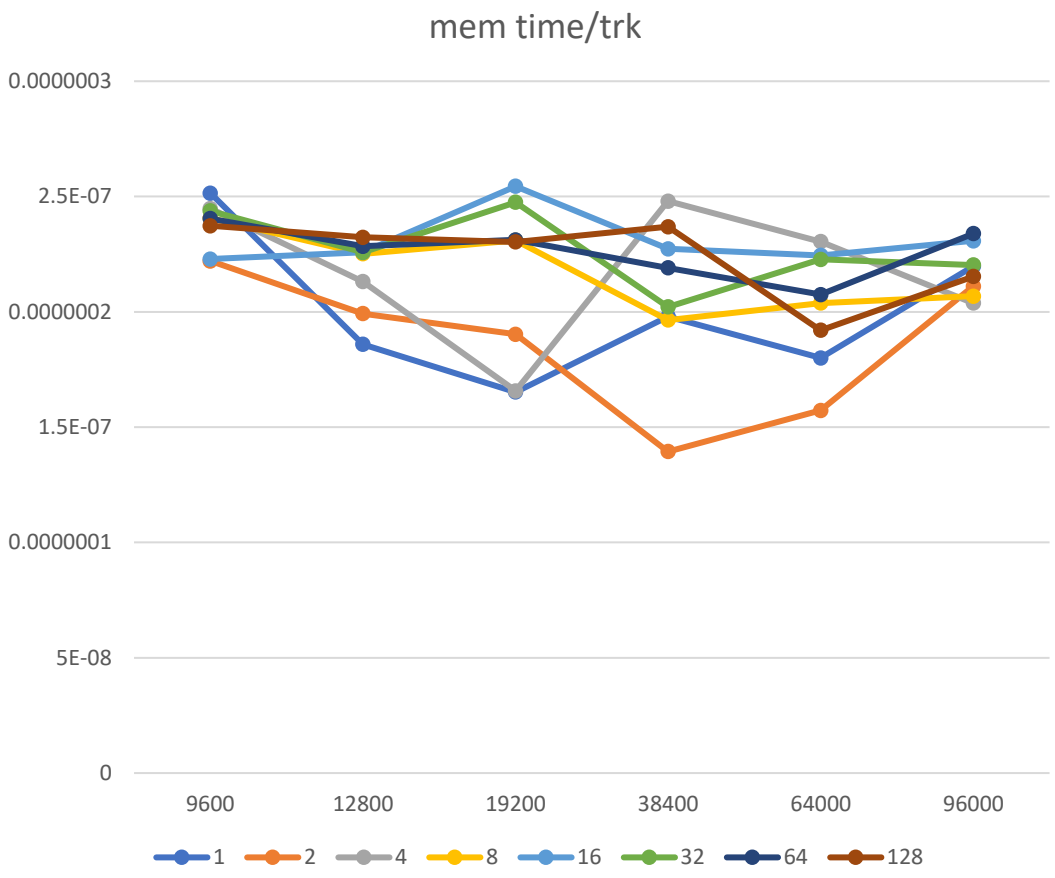
# Acc pgi


acc(pgi) region time by bsize
1.00E-04
1.00E-05
1.00E-06
1.00E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

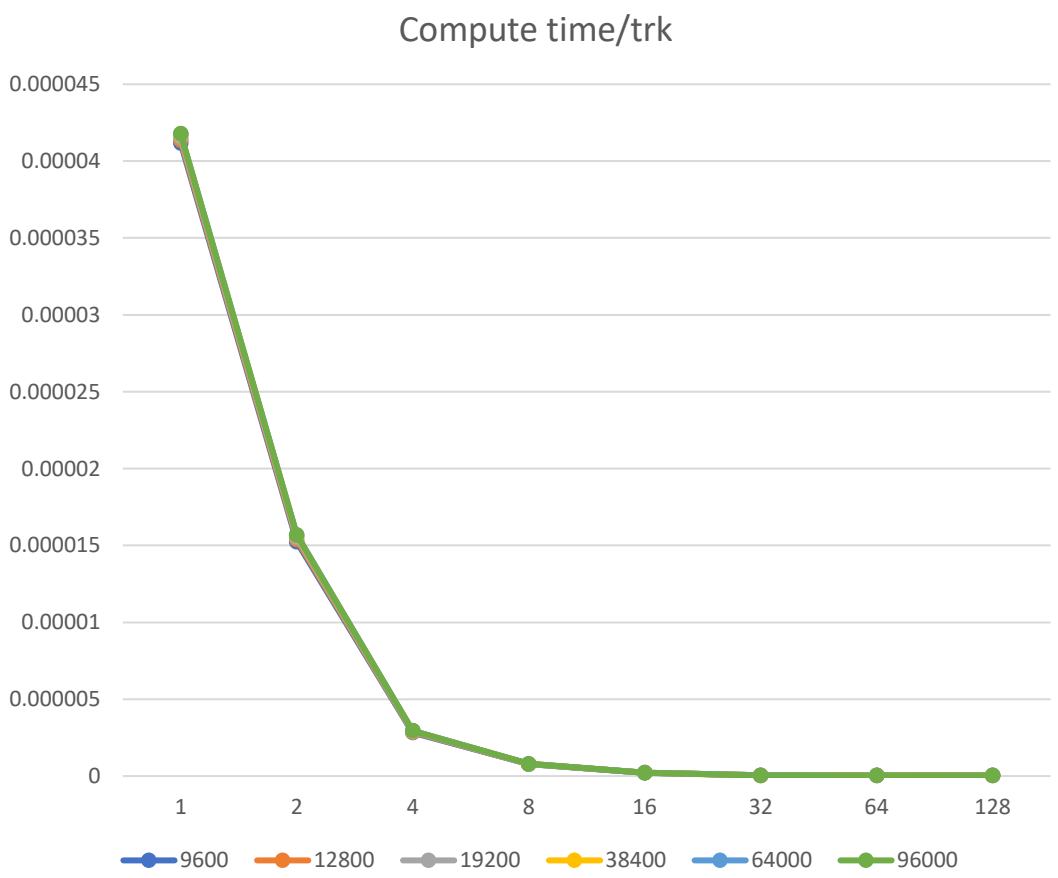

acc(pgi) region time by ntrks
1.00E-04
1.00E-05
1.00E-06
1.00E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

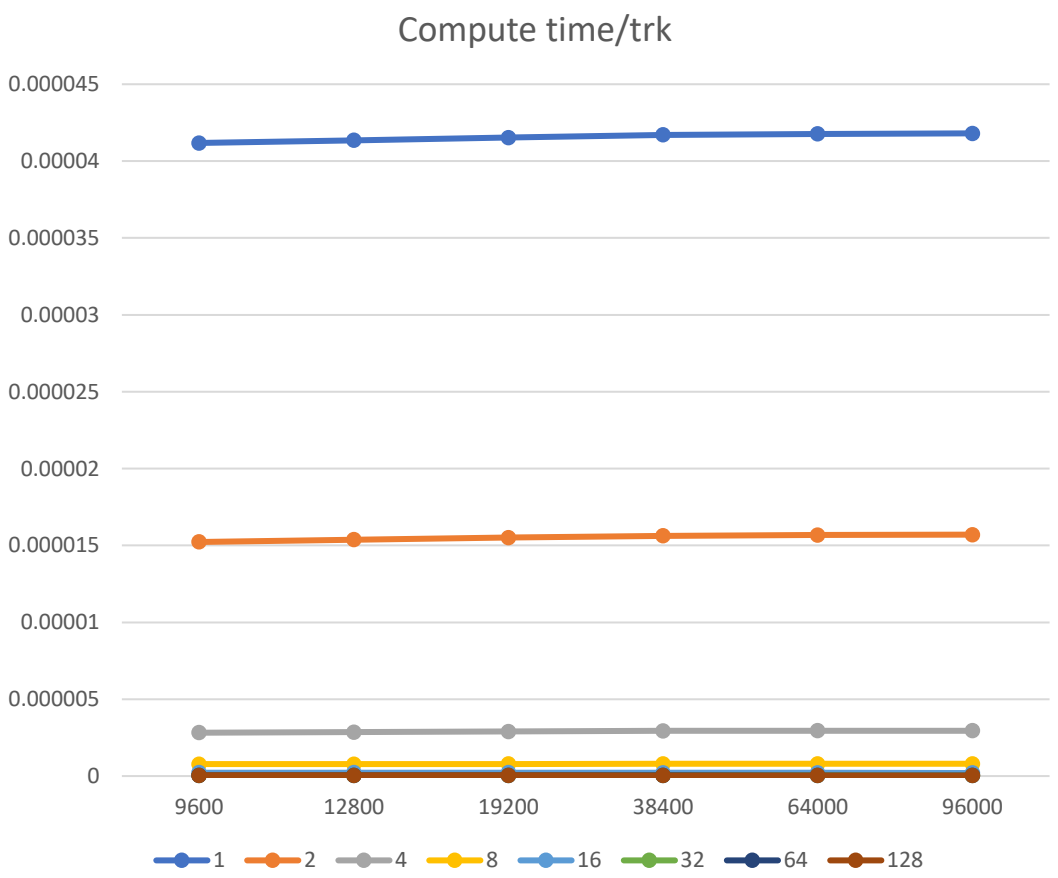
# Acc mem transfer time



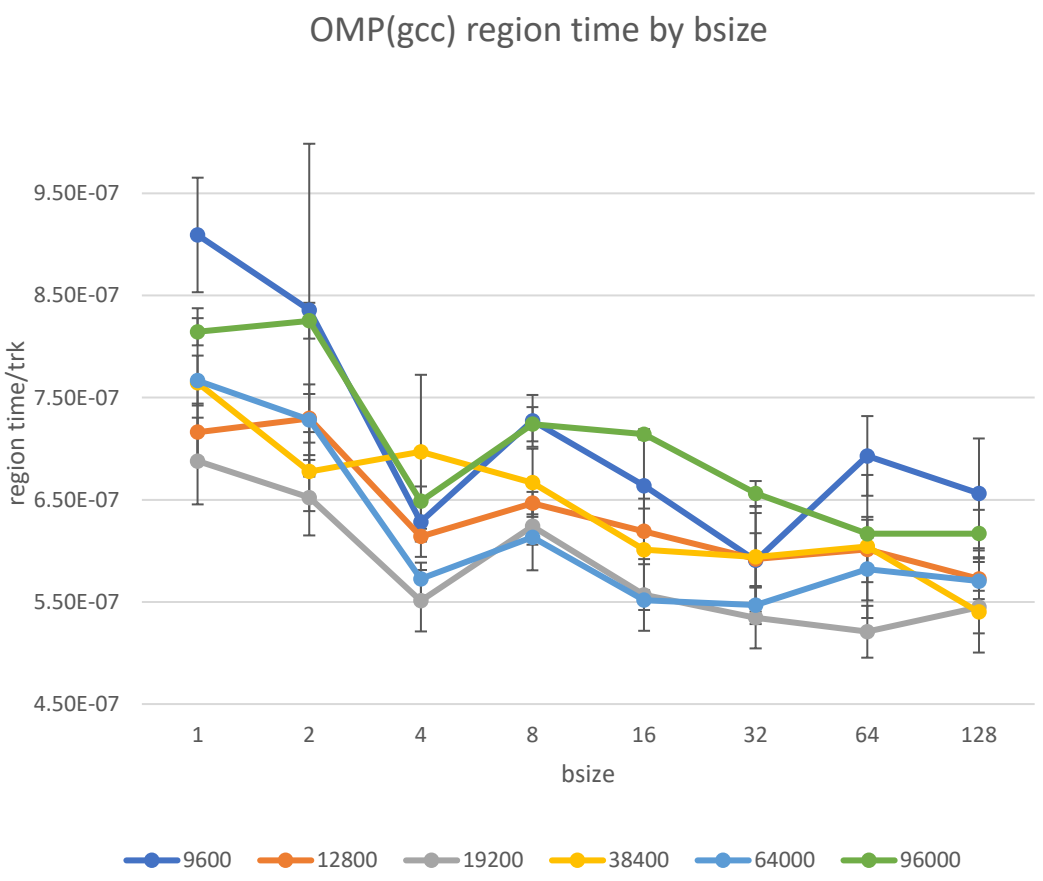



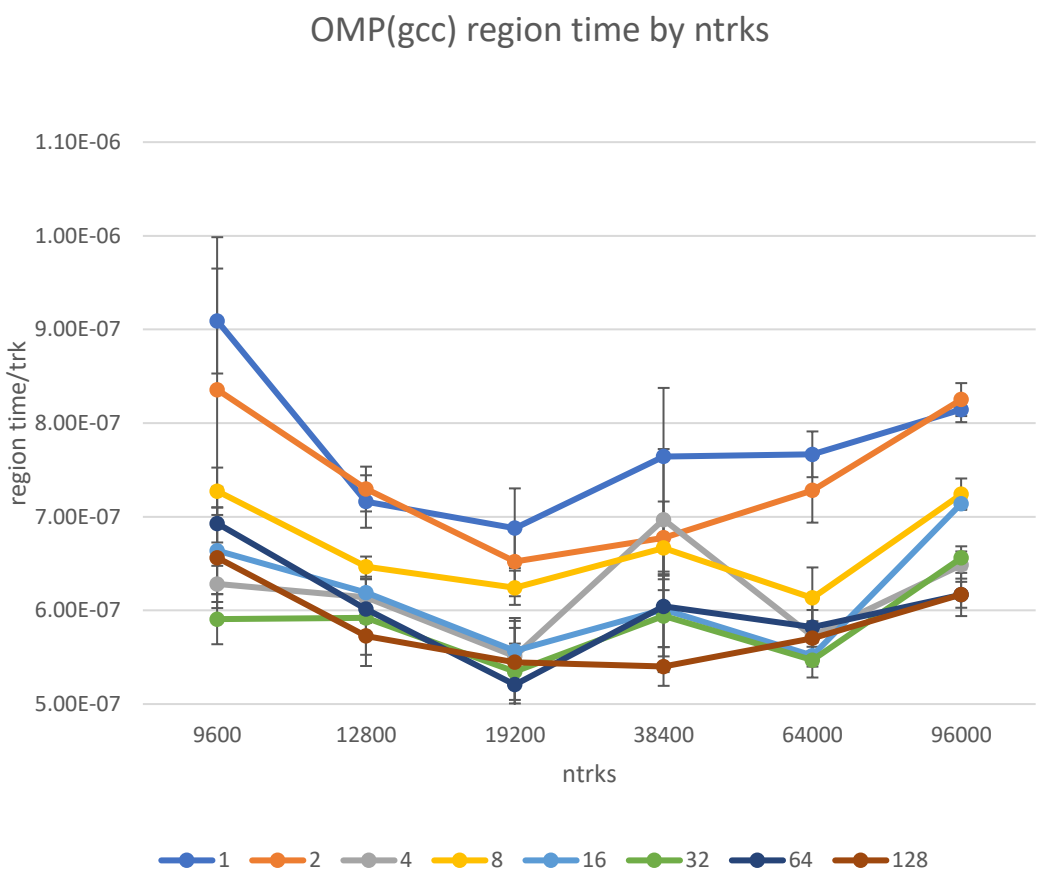
# Acc computation time



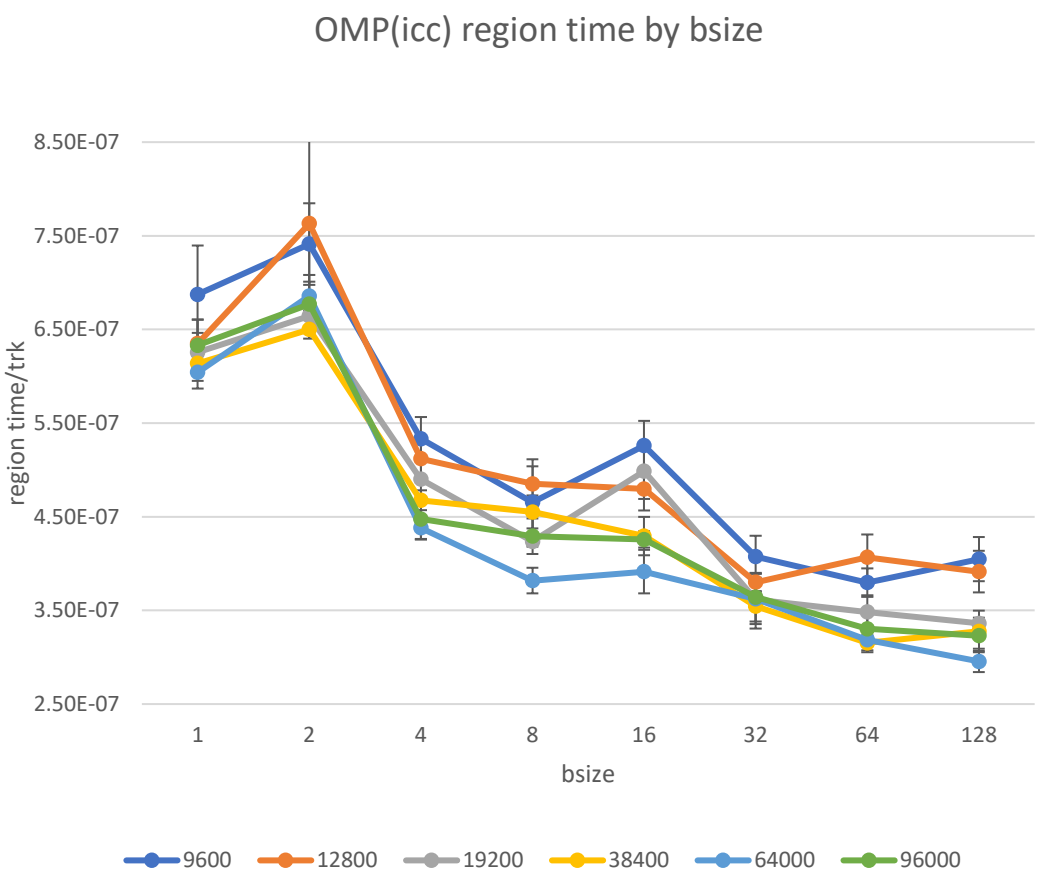



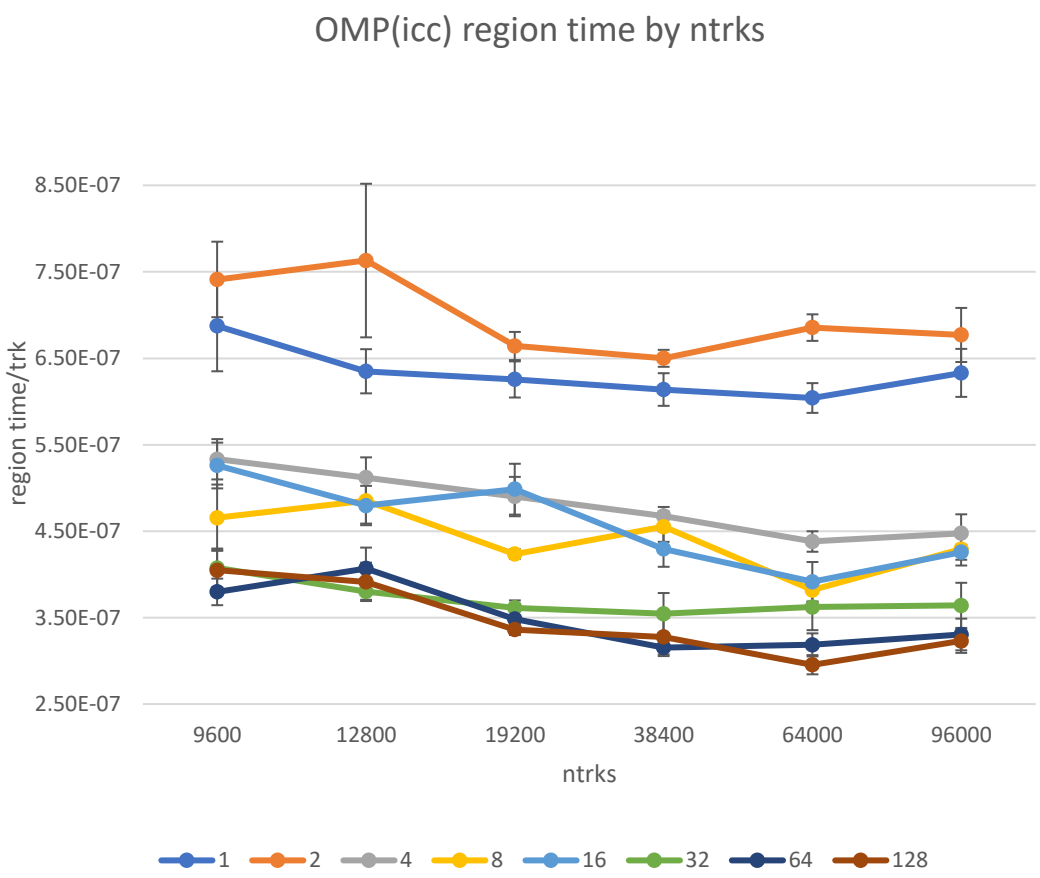
# OMP gcc


OMP(gcc) region time by bsize
9.50E-07
8.50E-07
7.50E-07
6.50E-07
5.50E-07
4.50E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

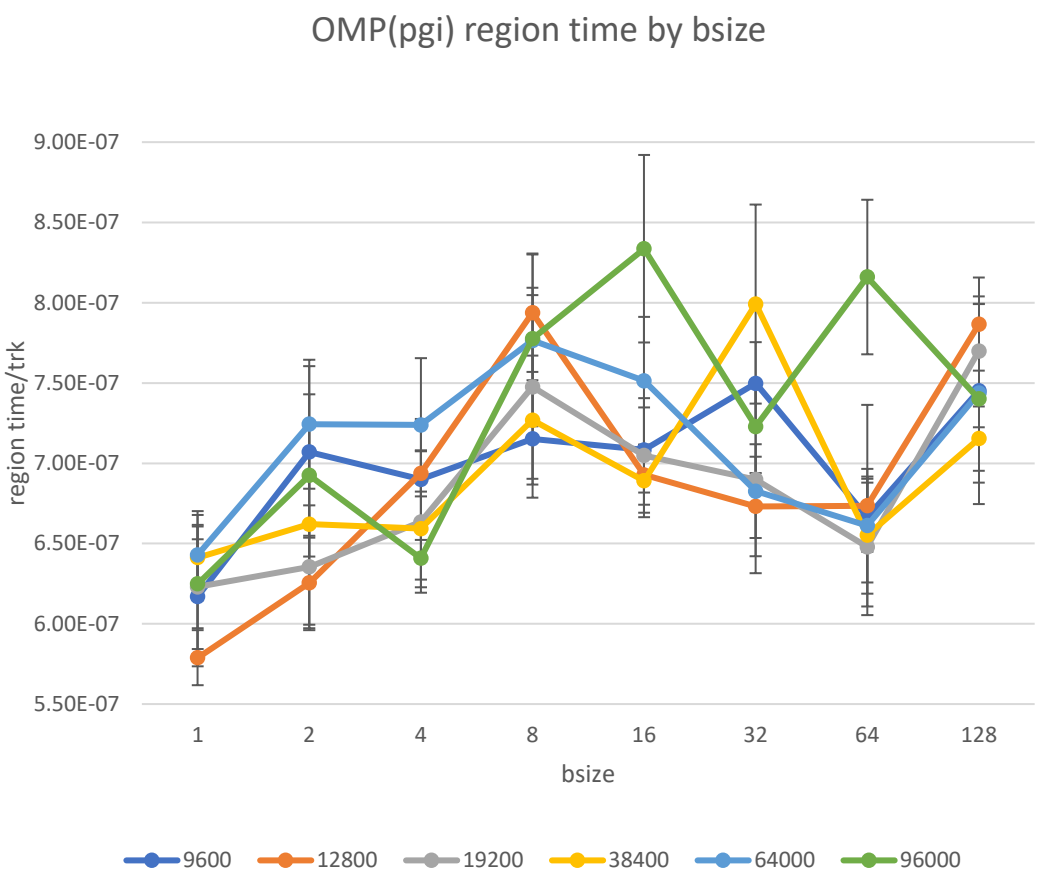

OMP(gcc) region time by ntrks
1.10E-06
1.00E-06
9.00E-07
8.00E-07
7.00E-07
6.00E-07
5.00E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

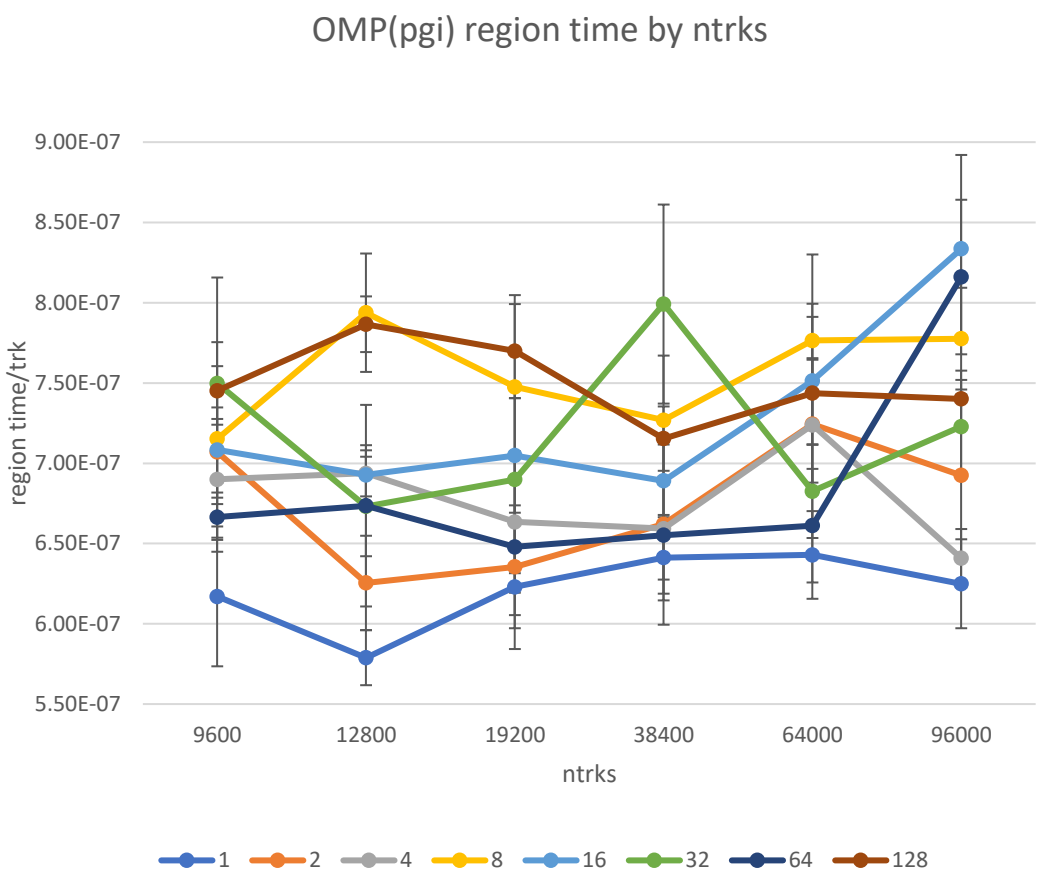
# OMP icc


OMP(icc) region time by bsize
8.50E-07
7.50E-07
6.50E-07
5.50E-07
4.50E-07
3.50E-07
2.50E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

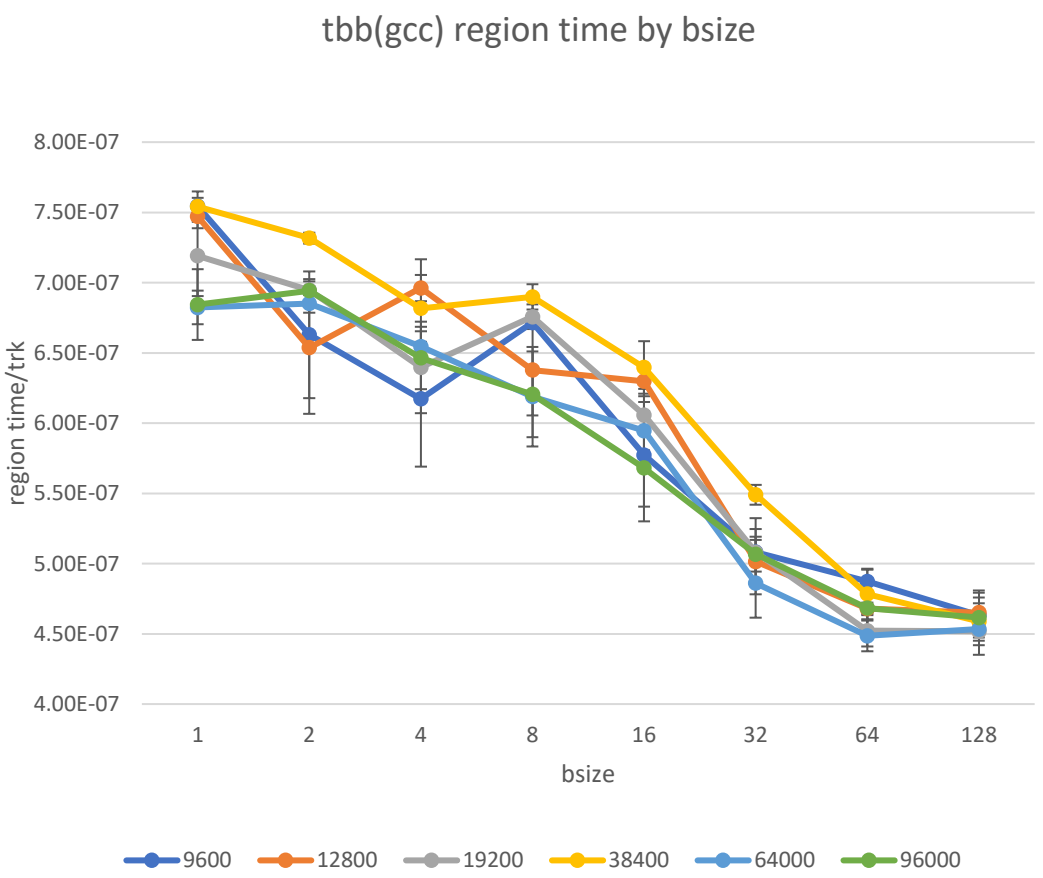

OMP(icc) region time by ntrks
8.50E-07
7.50E-07
6.50E-07
5.50E-07
4.50E-07
3.50E-07
2.50E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

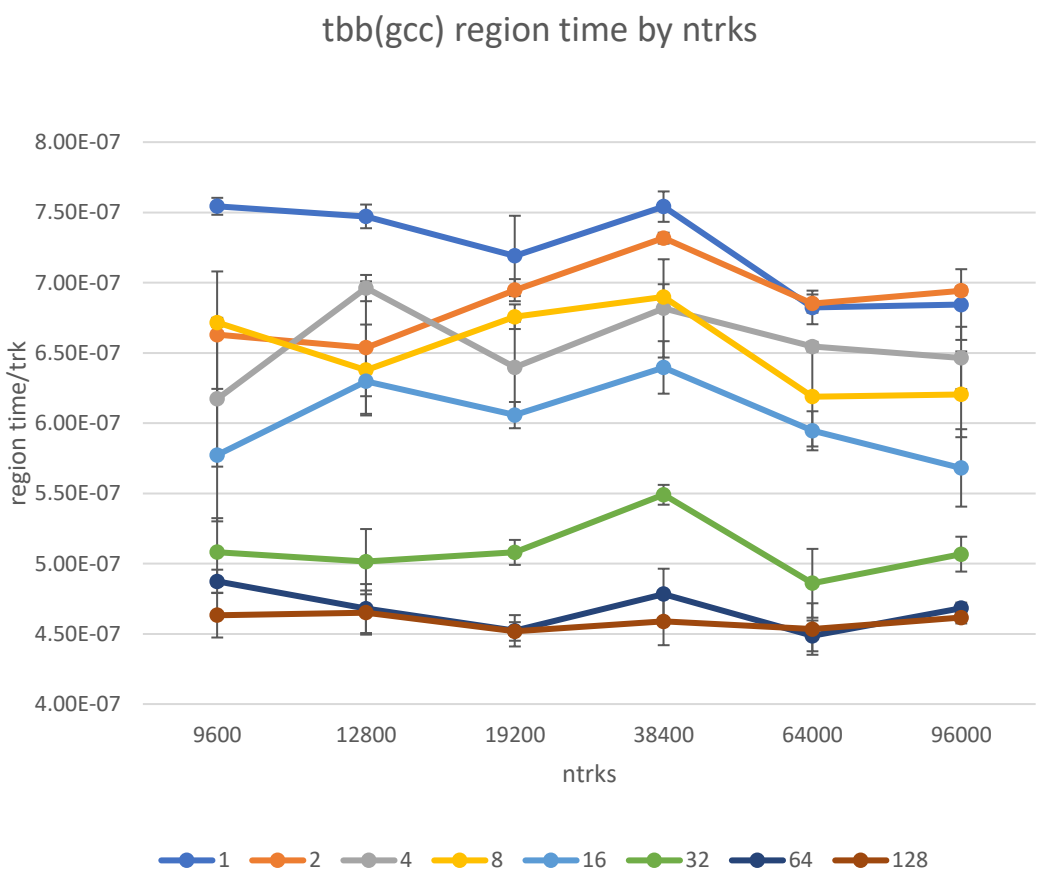
# OMP pgi


OMP(pgi) region time by bsize
9.00E-07
8.50E-07
8.00E-07
7.50E-07
7.00E-07
6.50E-07
6.00E-07
5.50E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

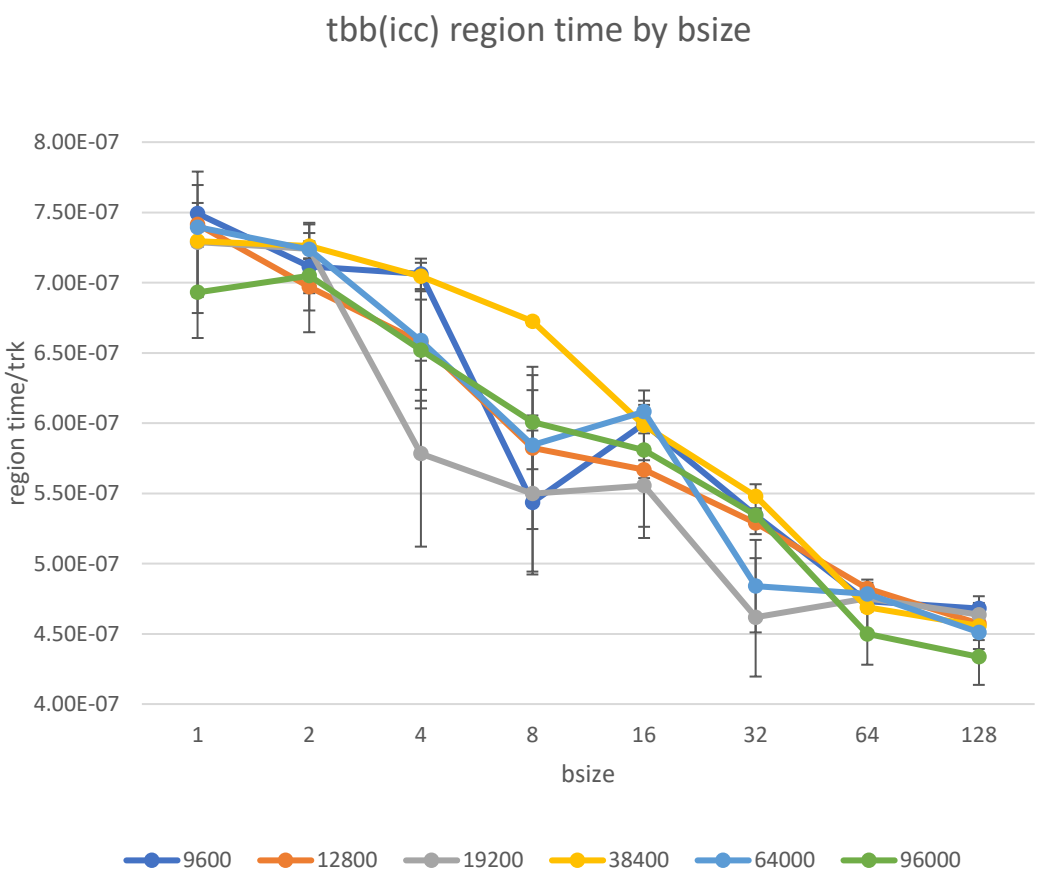

OMP(pgi) region time by ntrks
9.00E-07
8.50E-07
8.00E-07
7.50E-07
7.00E-07
6.50E-07
6.00E-07
5.50E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

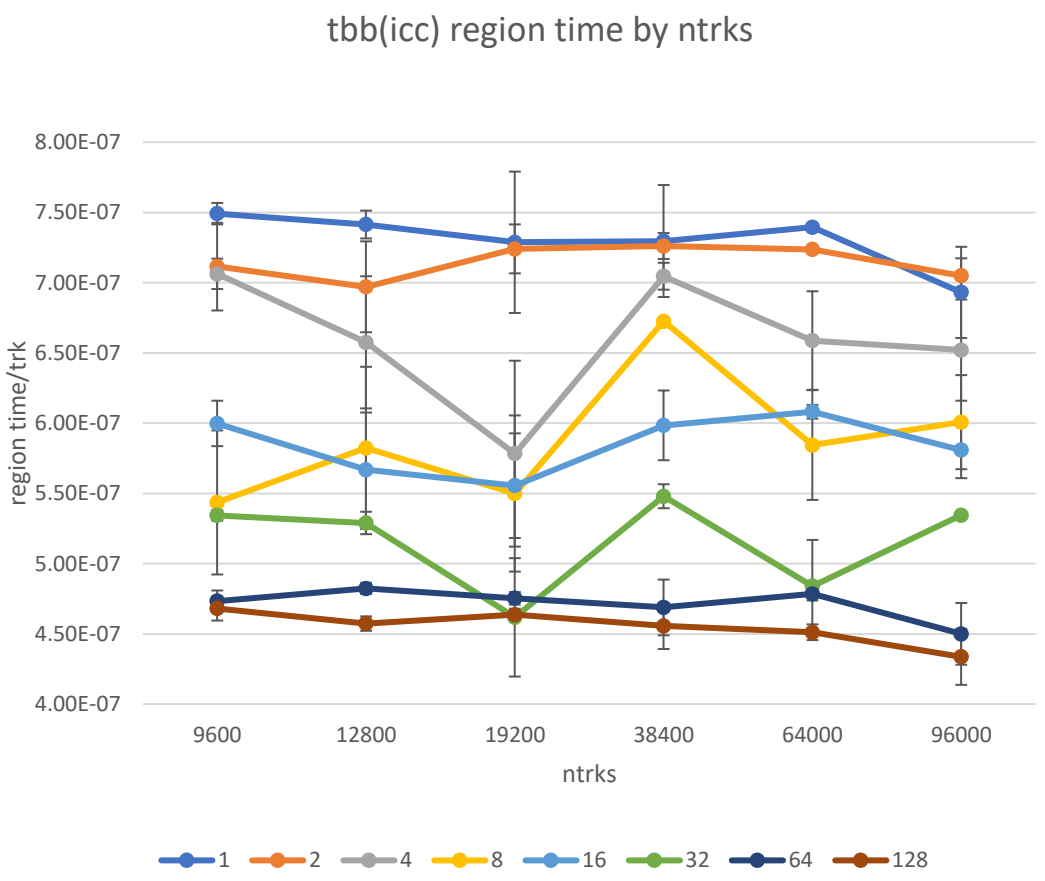
# Tbb gcc


tbb(gcc) region time by bsize
8.00E-07
7.50E-07
7.00E-07
6.50E-07
6.00E-07
5.50E-07
5.00E-07
4.50E-07
4.00E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

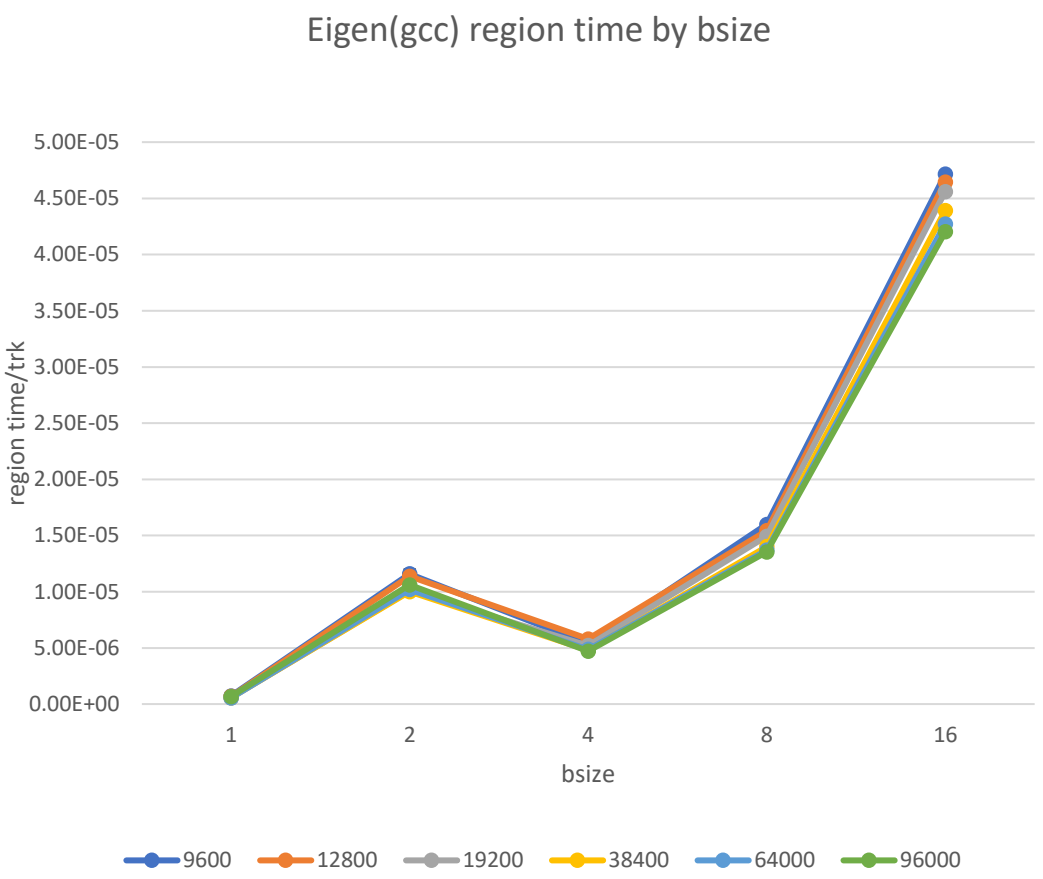

tbb(gcc) region time by ntrks
8.00E-07
7.50E-07
7.00E-07
6.50E-07
6.00E-07
5.50E-07
5.00E-07
4.50E-07
4.00E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

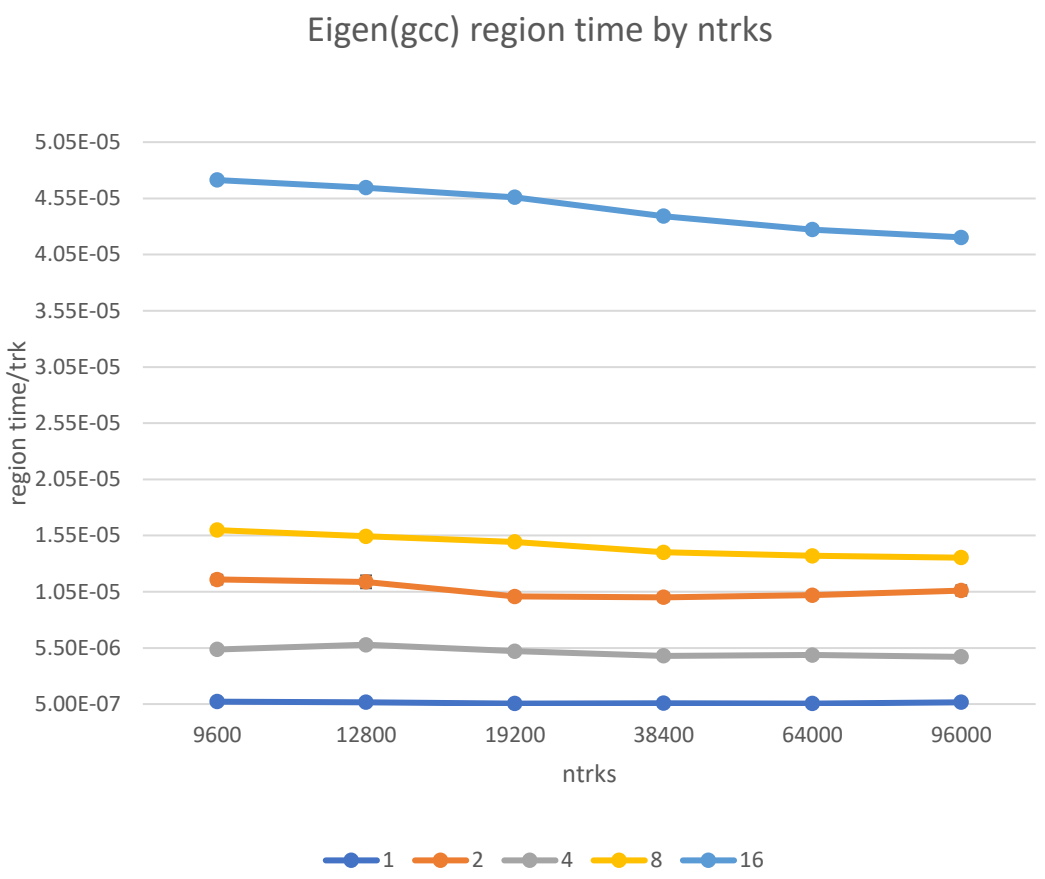
# Tbb icc


tbb(icc) region time by bsize
8.00E-07
7.50E-07
7.00E-07
6.50E-07
6.00E-07
5.50E-07
5.00E-07
4.50E-07
4.00E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000

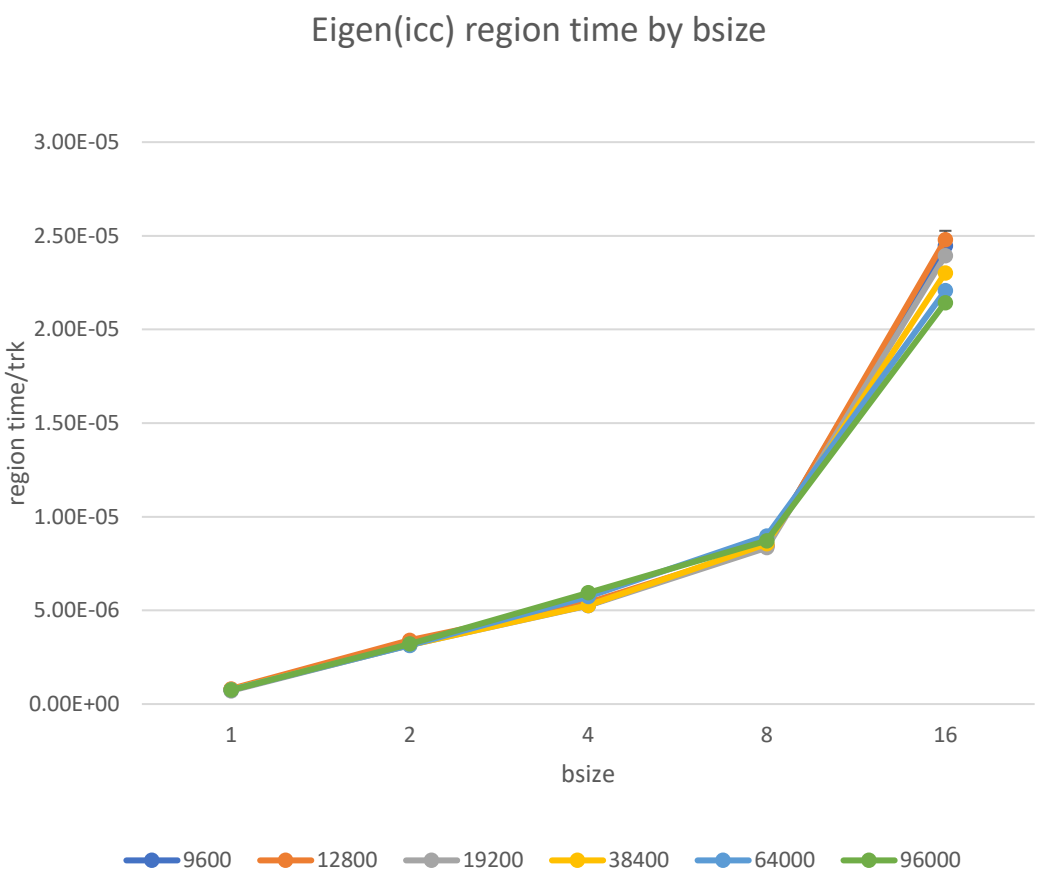

tbb(icc) region time by ntrks
8.00E-07
7.50E-07
7.00E-07
6.50E-07
6.00E-07
5.50E-07
5.00E-07
4.50E-07
4.00E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128

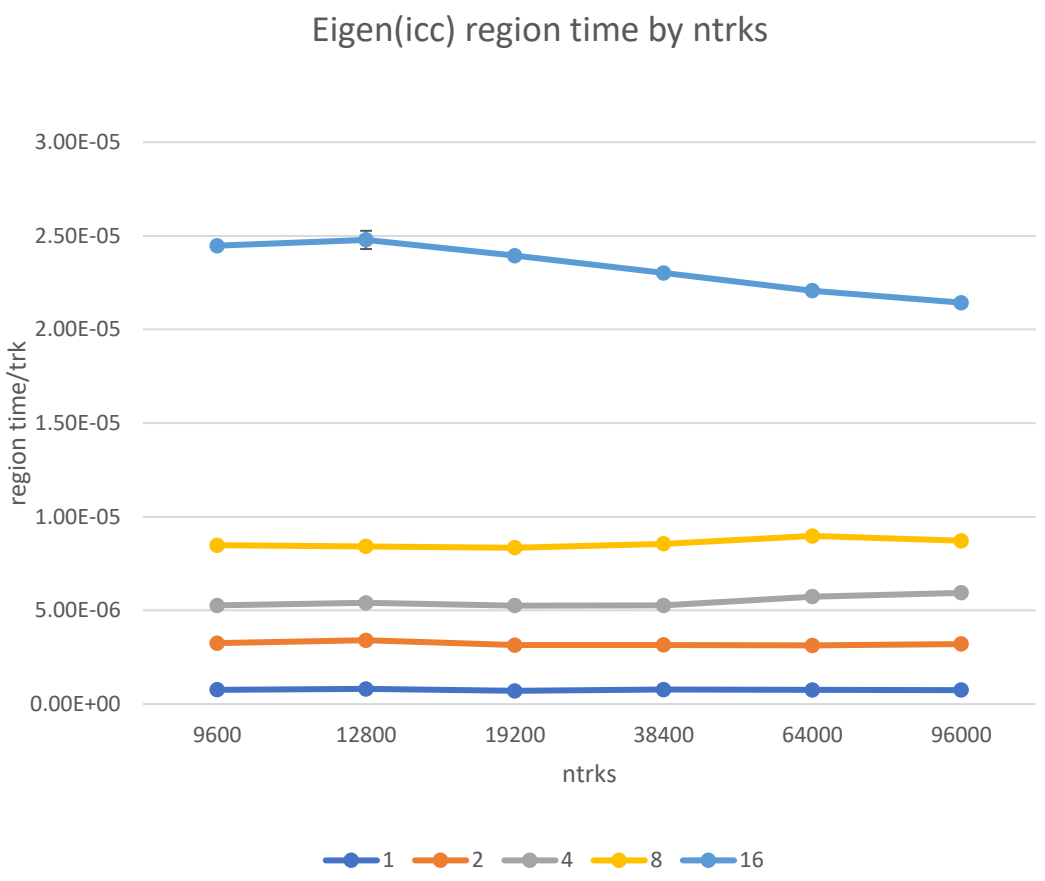
# Eigen gcc


Eigen(gcc) region time by bsize
5.00E-05
4.50E-05
4.00E-05
3.50E-05
3.00E-05
2.50E-05
2.00E-05
1.50E-05
1.00E-05
5.00E-06
0.00E+00
region time/trk
1
2
4
8
16
bsize
9600
12800
19200
38400
64000
96000

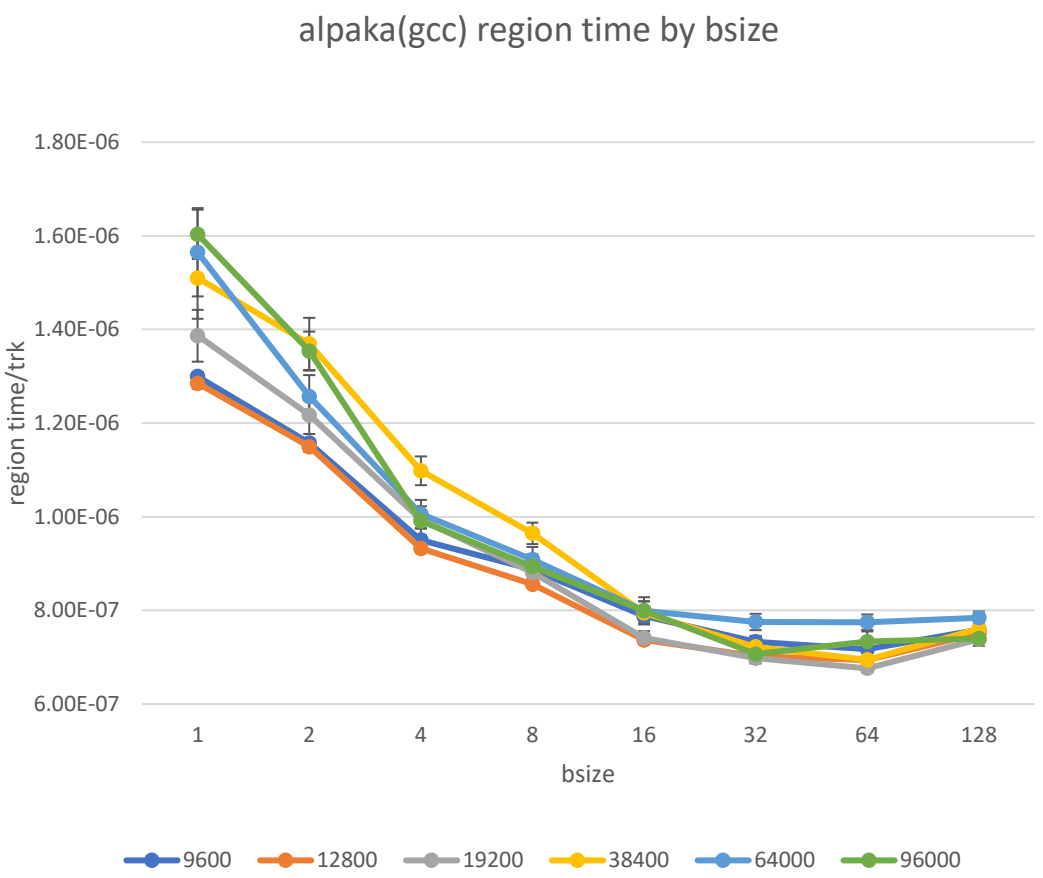

Eigen(gcc) region time by ntrks
5.05E-05
4.55E-05
4.05E-05
3.55E-05
3.05E-05
2.55E-05
2.05E-05
1.55E-05
1.05E-05
5.50E-06
5.00E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16

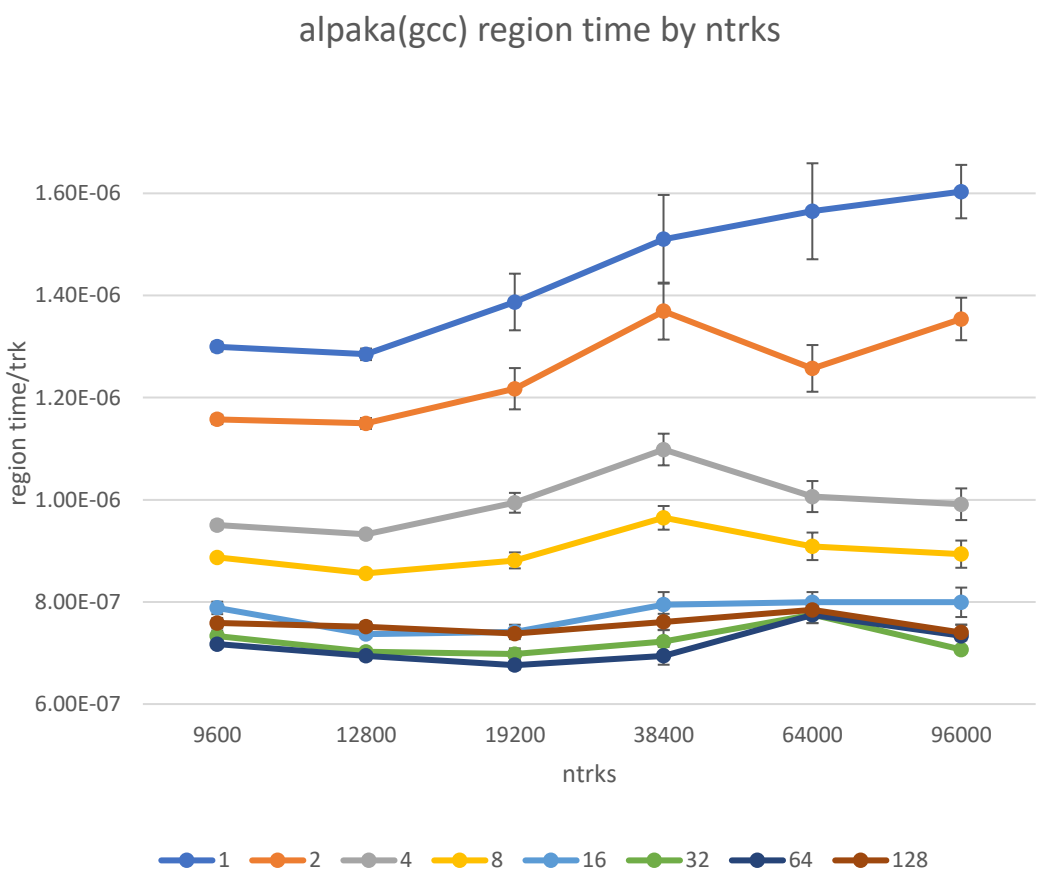
# Eigen icc


Eigen(icc) region time by bsize
3.00E-05
2.50E-05
2.00E-05
1.50E-05
1.00E-05
5.00E-06
0.00E+00
region time/trk
1
2
4
8
16
bsize
9600
12800
19200
38400
64000
96000


Eigen(icc) region time by ntrks
3.00E-05
2.50E-05
2.00E-05
1.50E-05
1.00E-05
5.00E-06
0.00E+00
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16

# Alpaka gcc


alpaka(gcc) region time by bsize
1.80E-06
1.60E-06
1.40E-06
1.20E-06
1.00E-06
8.00E-07
6.00E-07
region time/trk
1
2
4
8
16
32
64
128
bsize
9600
12800
19200
38400
64000
96000


alpaka(gcc) region time by ntrks
1.60E-06
1.40E-06
1.20E-06
1.00E-06
8.00E-07
6.00E-07
region time/trk
9600
12800
19200
38400
64000
96000
ntrks
1
2
4
8
16
32
64
128