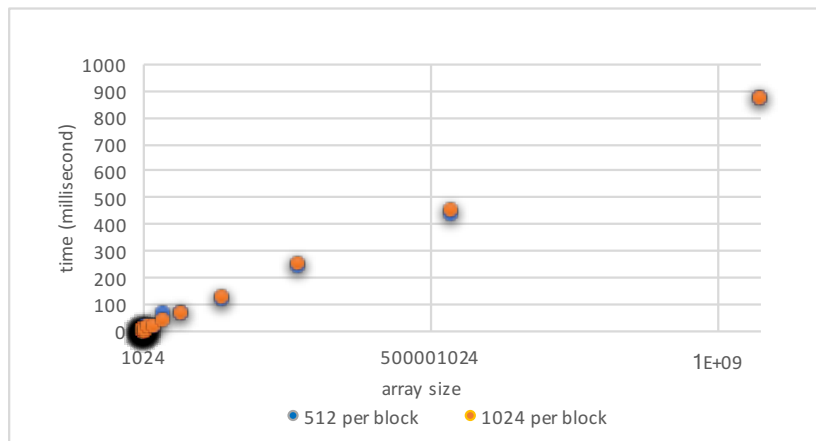
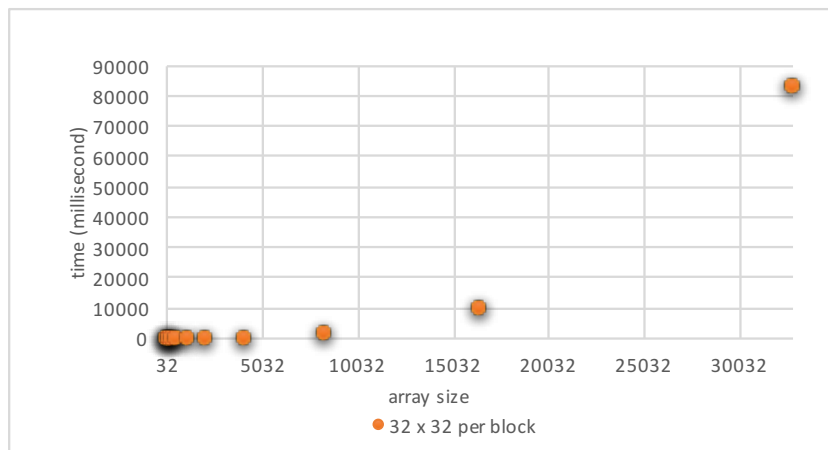


1. c)



2. c)



d)

dimension	32 x 32 per block	16 x 16 per block	8 x 8 per block
milliseconds	82836.8	142385	490744

As above result shows, when block\_dim is 32, it gets the best performance.

e) Under the same input size and threads per block, the runtime of naïve GPU approach is 828490 which is worse than tiled GPU approach, since the latter utilizes shared memory to reduce latency.

f) Under the same input size, the runtime of CPU approach is beyond Euler time limit which is worse than both GPU approaches, since the latter replaces some loops by threads to allow parallelism.