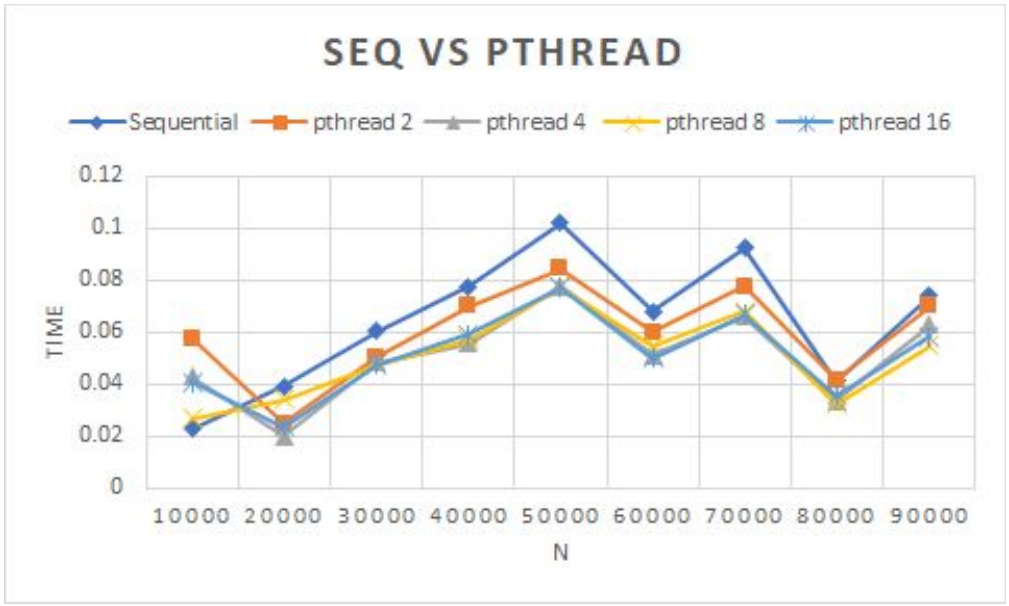
General design decisions:

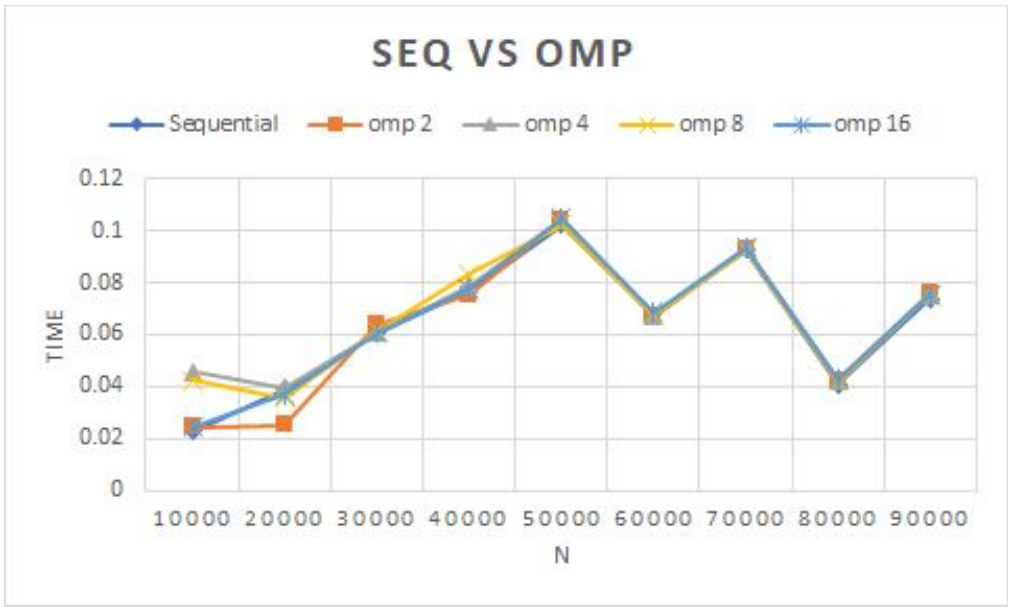
- The k-means iterations terminate when less than a certain number of points change the cluster they belong to. This is a better termination algorithm than a fixed number of iterations, in general, because we might do unnecessary number of iterations or may stop much before convergence.
- Minimalistic book-keeping is required to check if termination criteria is met. (If any point changes its cluster at an iteration, a variable, say delta, can be incremented by 1)
- The random initialisation is not seeded so that different runs of the programs can be compared

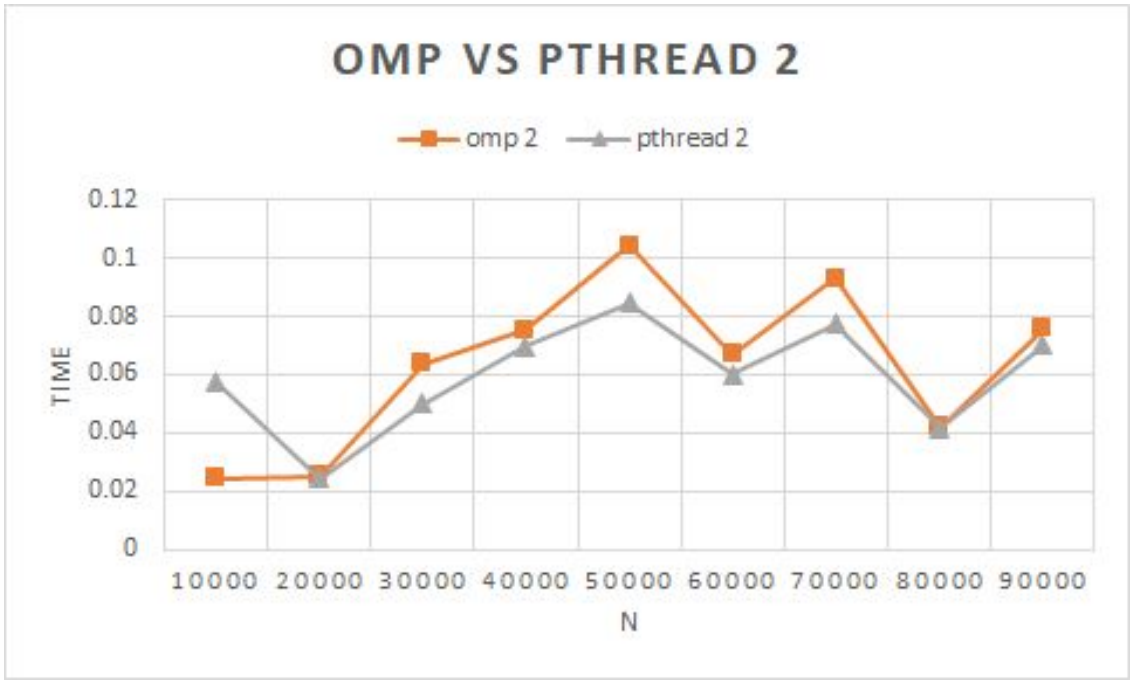
Parallelisation strategy:

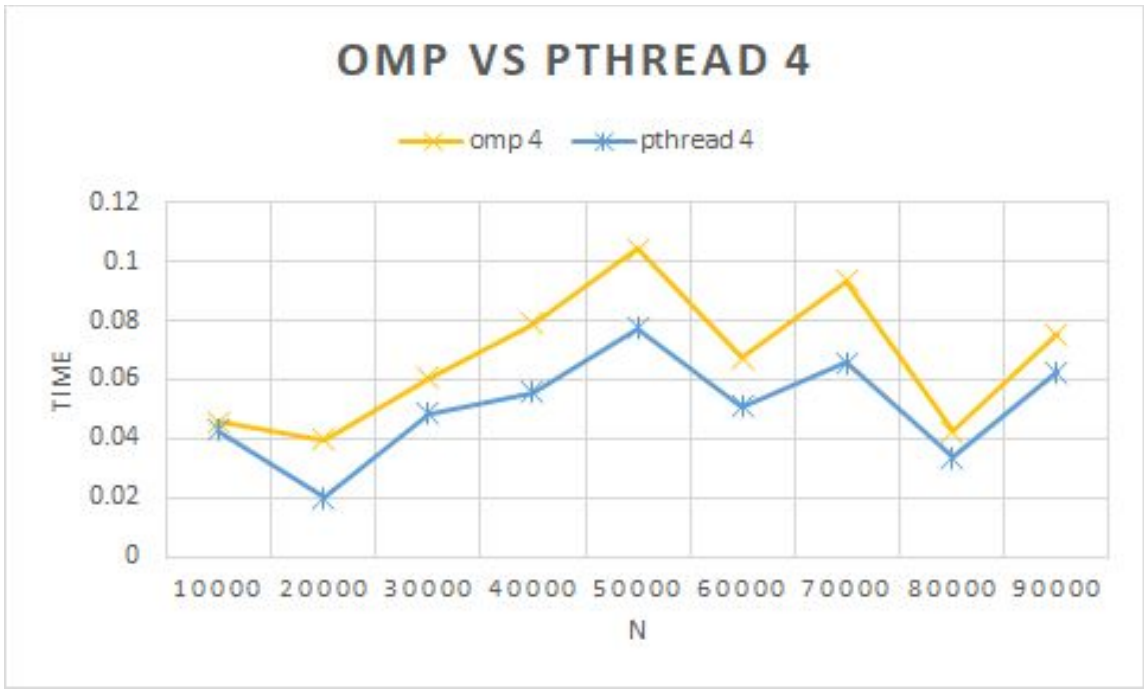
- Computation intensive step is the calculation of closest centroid for each point in the dataset (in the algorithm provided for reference). The computation for each point is independent and hence can be parallelised.
- Note: Due to the chosen termination condition, the variable delta has to be made atomic as it is a shared variable)
- No load balancing was done due to the assumption that the points are random.

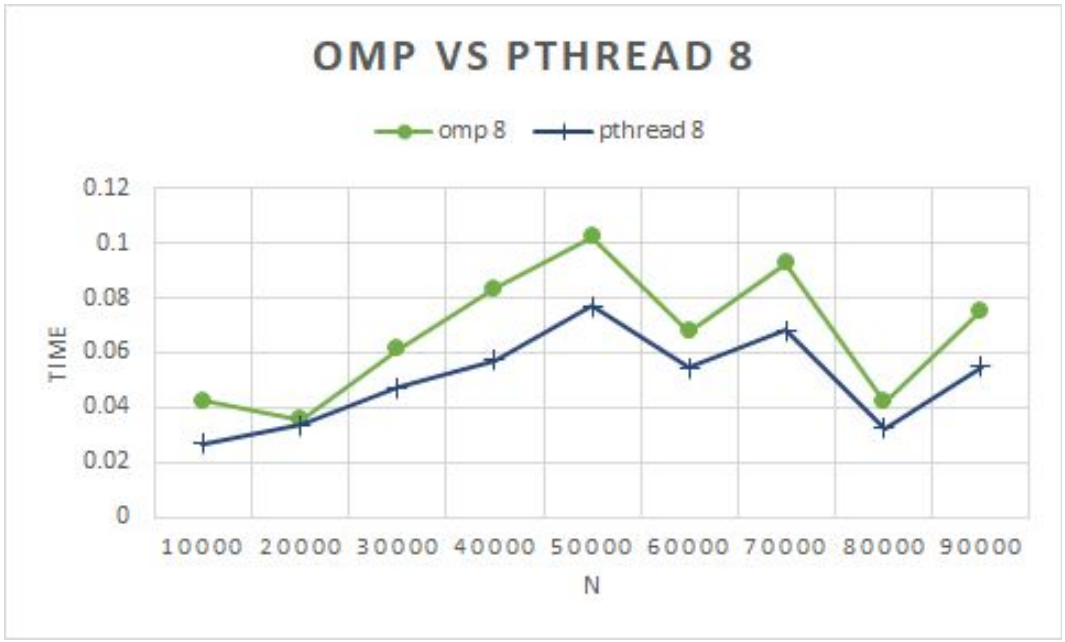
For any implementation, we can not compare across different values of n because the number of iterations depends heavily on the generated dataset. If the random initialisations in the program match that of the dataset, lesser number of iterations are required (random generation of centroid values is unseeded)

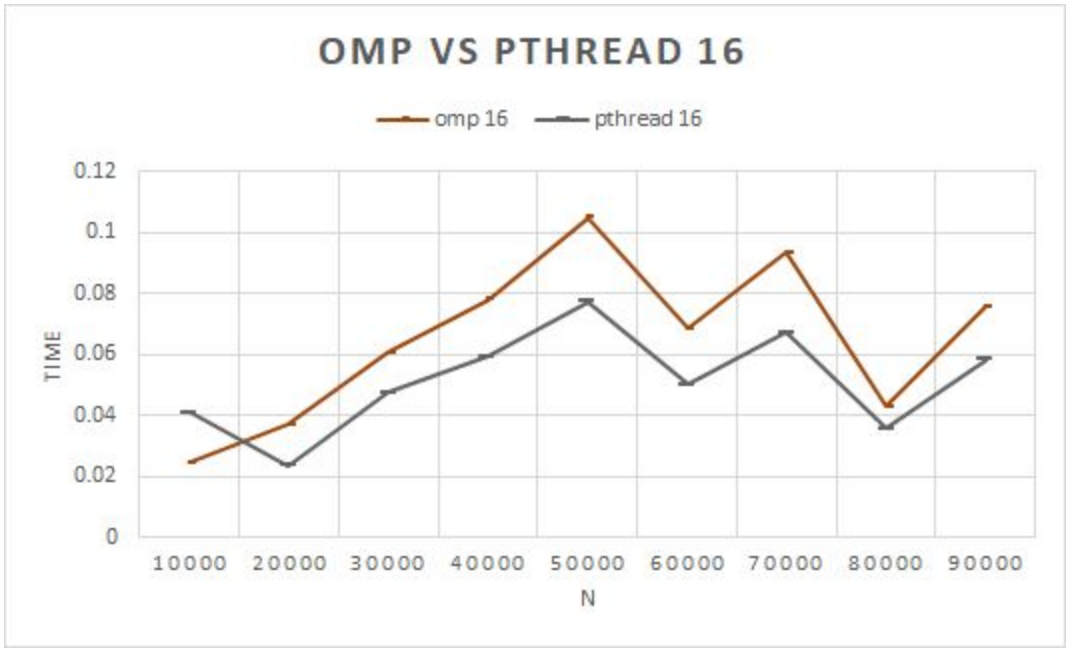

SEQ VS PTHREAD
Sequential
pthread 2
pthread 4
pthread 8
pthread 16
0.12
0.1
0.08
0.06
0.04
0.02
0
TIME
10000 20000 30000 40000 50000 60000 70000 80000 90000
N


SEQ VS OMP
Sequential
omp 2
omp 4
omp 8
omp 16
0.12
0.1
0.08
0.06
0.04
0.02
0
TIME
10000 20000 30000 40000 50000 60000 70000 80000 90000
N


OMP VS PTHREAD 2
omp 2
pthread 2
0.12
0.1
0.08
0.06
0.04
0.02
0
TIME
10000 20000 30000 40000 50000 60000 70000 80000 90000
N


OMP VS PTHREAD 4
omp 4
pthread 4
0.12
0.1
0.08
0.06
0.04
0.02
0
TIME
10000 20000 30000 40000 50000 60000 70000 80000 90000
N


OMP VS PTHREAD 8
omp 8
pthread 8
0.12
0.1
0.08
0.06
0.04
0.02
0
TIME
10000 20000 30000 40000 50000 60000 70000 80000 90000
N


OMP VS PTHREAD 16
omp 16
pthread 16
0.12
0.1
0.08
0.06
0.04
0.02
0
TIME
10000 20000 30000 40000 50000 60000 70000 80000 90000
N