Practical Session 4 : K-means parallelism

Objective

The objective of this lab is to analyse some sequential problem and apply parallelism methods using the *OpenMP* library. You should write your code using the *C++* language.

The goal of this exercise is to create the parallel program that is efficient. The problem that is solve in the proposed program is the k-means algorithm

The provided sequential code use a standard approach to k-means clustering without any form of optimization. The data structures used to hold the data are STL containers.

- Study the performances of the sequential program by modifying the number of clusters, the size of the problems, the threshold.
- Identify the dependencies (time and space) between the different parts of the program and between the data.
- Write a parallel version of the program using some approaches mentioned during the lectures. Several strategies are possible.
- Study the performances of your code by changing the number of processors and the size of the problem.