Faculty of Engineering
Cairo University

Parallel Computing

#CMP4005

Project Proposal

Submitted to:

Eng. Mohamed Abdullah

Submitted By:

NAME	SEC	BN	ID
Sarah Mohamed Hossam	1	30	9202618
Basma Hatem Elhoseny	1	16	9202381

Parallel Computing
Big assignment proposal

NAME	SEC	BN	WORKLOAD
Sarah Mohamed Hossam	1	30	Data points assignment to clusters kernel
Basma Hatem Elhoseny	1	16	Centroids update kernel

Idea:

A CUDA implementation of the K-Means clustering algorithm,where the main computational workload is divided into two kernels:

- The first kernel computes the distances between each data point and all centroids simultaneously, assigning each point to the nearest centroid. This step utilizes parallelism across CUDA threads, with each thread computing distances for one data point against all centroids.
- The second kernel updates the centroid positions based on the new assignments, with each thread responsible for computing the mean position of all data points assigned to a particular centroid.

These kernels iterate in an alternating fashion until convergence is achieved. This approach maximizes parallelism, leveraging the GPU's computational power to accelerate the clustering process.

Diagram :