

## Accelerated k-means Clustering on Multi-Core and GPGPU

*Wednesday, 11 November 2015 10:00 (30 minutes)*

Realistic simulations in plasma physics employ a large number of particles, usually beyond existing computer memory limits. To solve the problem, Particle-In-Cell (PIC) codes use the K-means clustering method to obtain a compact representation for so many particles by introducing mega-particles with "weight". However, the computational complexity of k-means is NP-hard to solve. Simpler heuristic algorithms such as Lloyd's exist, but their complexity still grows like  $O(n^2)$  with the number of particles. In addition, particle count is not constant during involved physical scattering processes, which may lead to exponential particle increase with simulation time.

The talk will present two main approaches for solving the performance bottleneck. First, combining KD-tree decomposition with a multi-core parallelization scheme based on OpenMP. Second, parallelize the algorithm on the GPGPU. Performance benchmarks, implementation issues and other remarks will be discussed.

The work is a joint effort performed at NBI during a PRACE project and ongoing activities.

**Primary authors:** Dr TRIER FREDERIKSEN, Jacob (Niels Bohr Institute); Mr MALÝ, Lukáš (VŠB-TUO: Technical University of Ostrava); Dr BURGDORFF KRISTENSEN, Mads Ruben (NBI); Mr BUTRASHVILY, Mordechai (Tel-Aviv University)

**Presenter:** Mr BUTRASHVILY, Mordechai (Tel-Aviv University)