

Projet final ACP

Programmation parallèle en python appliquée à la bio-informatique

Dans ce projet on s'intéresse à la parallélisation d'algorithmes de l'état de l'art en bio-informatique. Parmi ces algorithmes on s'intéresse aux algorithmes suivants : Smith-Waterman pour l'alignement local de deux séquences d'ADN, Needleman-Wunch alignement global de deux séquences, et BLAST pour la recherche de séquences dans une base de données. Les algorithmes de Smith Waterman et Needleman-Wunch sont des algorithmes exacts (permettent de trouver la solution optimale) basés sur la programmation dynamique. BLAST est un algorithme et outil très utilisé pour la recherche de séquences dans une BDD et il s'agit d'une heuristique.

Le présent projet consiste à implémenter un de ces algorithmes (au choix) en utilisant des bibliothèques de programmation parallèle sous python.

Le compte rendu du travail réalisé doit contenir les étapes suivantes :

- 1-Présentation de l'algorithme (ses étapes, pseudo code et explication)
- 2-Etude de la Complexité temporelle et spatiale de l'algorithme.
- 3- Utilisation d'un outil de profiling pour analyser les goulots d'étranglements dans l'algorithme (quelles sont les instructions ou tâche, qui prennent le plus de temps) et qu'on devrait par la suite paralléliser.
- 4-Analyse de l'algorithme et proposition d'un modèle de parallélisation sur une architecture multicœurs. Justifier le choix de votre modèle.
- 5-Implémentation du modèle parallèle en utilisant une des techniques citées dans ce cours.
- 6-Tests de performances avec mesure du temps (calculer accélération et efficacité) entre la version parallèle et version séquentielle.

Références :

[1] des implémentations en Python (séquentiel) de ces trois algorithmes peuvent être trouvées dans le livre : « BioInformatics Algorithms design and implémentation in Python »

[2] « Python High Performance Programming » livre de Gabrielo Lanaro , Open source.