

2P14

入力ファイルの前処理とデータベースを用いた全解探索の分散処理： Charmm を用いた検討

矢部誠, ○玉城哲平, 上田一義, 武田穰, 小泉淳一
横浜国大院工(〒240-0067 神奈川県横浜市保土ヶ谷区常盤台 79-1)

1. 緒言

分子の再安定構造を求める計算では、位置のパラメータを少しずつ変化させながら総当たりに求めていく全解探索法が広く用いられている。このような膨大なエネルギー計算においては、Charmm に代表されるシミュレーションソフトウェアを用いて繰り返し計算を行う際に、CPU 並列化での計算効率向上が困難であること、計算過程の出力ファイルが巨大になり扱いづらいことが問題となっていた。そこで、入力ファイルの前処理とデータベース（以下、DB）を扱うプログラムを作成し、既存の計算ソフトウェアに手を加えず計算速度の向上とディスクスペースの有効利用を試みた。

2. 方法

Web サーバー (Apache + PHP + PostgreSQL) で動作する PHP アプリケーションを開発した。データを格納しジョブの進捗を管理するマシン (A) と、計算を行うマシン (B) とに役割が分かれる。

まず、用意した入力ファイルのループ処理の部分を (A) で分割し、ジョブリストを作成する。(B) はジョブリストを監視し、計算を実行する。結果が得られたら (A) へと結果を格納し、次々に分割された計算を行う。

本プログラムは計算ソフトウェアそのものには一切手を加えていないため、結果の信頼性を損なわずに分散処理を実現した。そして、投入したプロセッサ (コア) 数や計算マシンの数にほぼ比例した、計算の高速化が可能である。さらには、低スペックのマシンも計算資源として活用することが可能である。また、Web サーバーや DB サーバーという環境を採用したことで、各マシン間のデータ転送の実装が容易になり、開発時間を大幅に短縮することができた。巨大なファイルとなるログファイルは圧縮し、ホストマシンへアップロードして管理する。スクリプトに応じた解析プログラムを作成すれば、解析作業を容易に行うことができる。

3. 結果

分割プログラムの実施例として、 β -D-glucose と N-acetylglucosamine から成る二糖の最安定構造計算を行った。クロック周波数 2.3GHz, 24 コア (12 コア \times 2CPU) のコンピューター上で計算を行ったところ、60 日ほどかかっていた計算を 40 時間に短縮することができた。また、計算過程を出力したファイルが分割されることで内容参照が容易になり、圧縮により使用ディスクスペースを 1/30 にすることができた。

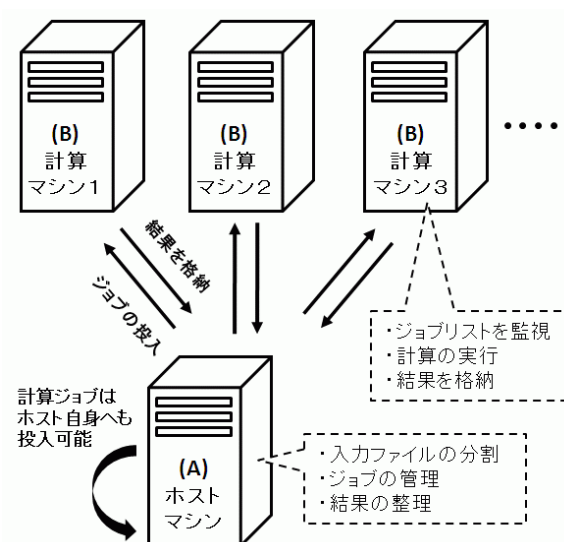


Figure 1 分散処理の仕組み