

ゲノム情報学アプリケーションの 自己適応的並列化／最適化に関する研究

●福田 晃¹⁾ ◆國枝 義敏²⁾ (H13-14) ◆中西 恒夫³⁾(H14) ◆南里 豪志³⁾(H14) ◆小坂 隆浩¹⁾(H14) ◆北須賀 輝明³⁾(H14)

1) 九州大学大学院システム情報科学研究院 2) 和歌山大学システム工学部 (現 立命館大学情報理工学部)
3) 九州大学情報基盤センター 4) 大阪産業大学工学部 (現 同志社大学工学部)

〈研究の目的と進め方〉

ゲノム情報学アプリケーションを対象として、PCクラスタ群などの分散システム環境上で、ユーザに使い勝手のよいプログラミング環境を提供するとともに、対象計算機システムにあわせて自己適応的に並列／分散化、最適化する手法、ならびに汎用性に優れたソフトウェア基盤を構築し、ゲノム情報学アプリケーションの高速化を図ることを目的とする。

〈研究開始時の研究計画〉

- ・PCクラスタ環境上で、ゲノム情報学アプリケーションの高速化を図る。
- ・ソフトウェア基盤を構築する。

〈研究期間の成果〉

- ・ゲノム情報学アプリケーションの分析
分析の結果、並列処理としては、データベースとのマッチングをとるデータ並列処理が大部分を占めることが分かった。
- ・共通並列プログラミングモデルの設計と実装
共通プログラミングモデルの一つである分散共有メモリシステムの設計を行い、PCクラスタおよびグリッド環境に実装した。
- ・自己適応的並列化／最適化手法の研究
ヘテロな分散環境において、ノード性能、ネットワーク性能などを考慮したスケジューリング手法を提案し、本スケジューリング方式を実装した。
- ・ユーザインタフェースの設計と実装
ゲノムアプリケーションを高速実行（並列／分散実行）するサービス環境をユーザに提供するWebサービス環境を設計し、構築した。
- ・ゲノムアプリケーションの高速化
Globusに対応するように、HMMerを書き換え、HMMerをヘテロ分散環境に実装した。さらに、上記のスケジューリング方式、共通プログラミングモデルなどを統合して、分散環境に実装して、この環境上で、ゲノムアプリケーション（HMMer）を実行することにより、高速実行の有効性を実証した。

〈国内外での位置づけ〉

ゲノム情報学アプリケーションをPCクラスタ上で分散実行する研究は、研究遂行当時は、わずかならが行われていたものの、ハードウェア環境に自己適応させる研究は見当たらなかった。この後、種々の研究が開始されはじめた。本研究は、その契機になったと、考えている。

〈主な発表論文〉

- ・T.Koita, Y.Kofune, Y.Inoue, and A.Fukuda: Implementation and Evaluation of Resource Allocation for a Genomic Application Program on the Grid, Proc. AI2003, pp.524-528, 2003.
- ・Y.Inoue, T.Koita, and A.Fukuda: Performance

Evaluation of Scheduling Schemes for a Genomic Application on OBIGrid, Proc. CIIT 2003, pp. 699-703, 2003.

- ・小坂, 小舟, 井上, 福田: Grid環境におけるゲノムアプリケーションInterProScan実行方式の検討, 情報処理学会 ハイパフォーマンスコンピューティングと計算科学シンポジウム (HPCS2003), pp.9-12, 2003.
- ・井上, 小坂, 福田: ゲノムアプリケーション InterProScan のファイル管理機能, 情報処理学会 先進的計算基盤シンポジウム (SACIS2003), pp.183-184, 2003.
- ・小坂, 小舟, 井上, 福田: Grid環境におけるゲノムアプリケーション実行支援機能, 人工知能学会第24回JSAI SIG-MBI 分子生物情報研究会, 2003.
- ・小舟, 東原, 小坂, 福田: Grid環境におけるMPI-HMMERの性能評価, 第2回情報科学技術フォーラム講演論文集 (FIT 2003), pp.59-60, 2003.
- ・井上, 小坂, 福田: Grid環境におけるInterProScanスケジューリング機能の実装, 第2回情報科学技術フォーラム講演論文集 (FIT2003), pp.61-62, 2003.
- ・S.Ikebe, T.Nakanishi, and A.Fukuda: Redundant Scheduling: An Algorithm for Scheduling Parameter Sweep Applications in Grid Environment, Proc. PDPTA2002, Vol.II, pp.693-698, 2002.
- ・奥野, 横手, 齋藤, 上原, 國枝, 福田: 分散共有メモリシステムにおける WaveFront 法を用いたホモロジー解析の並列化, 情報処理学会研究会, 2002-OS-89/2002-EVA-2, pp.63-70, 2002.
- ・池辺, 中西, 福田: メタコンピューティング環境におけるポストゲノムアプリケーション HMMer のスケジューリング, 情報処理学会研究会, 2002-ARC-147/2002-HPC-89, pp.19-24, 2002.
- ・池辺, 中西, 福田: メタコンピューティング環境を用いたポストゲノムシーケンスライブラリ HMMer の高速化に向けて, 九州大学情報基盤センター年報, Vol.2 (2002), pp.117-122, 2002.
- ・S.Saito, S.Yokote, T.Uehara, and Y.Kunieda: The Implementation of a Compiler Controlled Software Distributed Shared Memory System "Fagus" as a Runtime Support System for Automatic Parallelizing Compiler, Proc.PDPTA2001, pp.1186-1192, 2001.