

クラスタ計算と広域計算に適応した並列 GA モデルの構築 ~ 研究紹介 ~

Parallel Genetic Algorithm's Model Applied to Cluster Computing and Global Computing

谷村 勇輔

Yusuke TANIMURA

Abstract: This resume is introduction to my research. My research has three keywords: Cluster Computing, Global Computing, Genetic Algorithms. The former, two words indicate my computing environments. Today, Seamless science computing portals that are set of remote resources, are needed. We study Parallel Genetic Algorithm's model as application software on such computing environments. This study will be important in the future.

1 はじめに

本発表ではこれまでの研究成果と今後の研究目標について説明を行う。本研究では Cluster Computing 技術を導入して PC クラスタを構築し、PC クラスタ上でのアプリケーションの 1 つとして並列 GA のモデルを考えてきた。その結果、並列 GA のモデルとして分散 GA が有効であることがわかった。今後はさらにそれを発展させ、傾斜法との組み合わせや 1 島 2 個体にした特別な分散 GA モデルの研究を行う予定である。また計算環境についても、遠隔地にある複数のクラスタを利用した Global Computing 環境に発展させ、その環境においての並列 GA のモデルについても研究を行う予定である。

2 Cluster Computing

近年のコンピュータ技術・文化は急激な進化を遂げている。これにより今日では、性能の良いハードウェアを低価格で入手できるようになり、Linux のように安定して動作する OS や MPICH といった通信ソフトウェアなどのミドルウェアを無料でかつ自由に利用できるようになっている。そしてこれらが、1994 年頃より始まった「日常において一般的に利用されるハードウェアを組み合わせることで手作りの並列計算機を作る」ための研究プロジェクト (Beowulf プロジェクト) と結びつき、今日では多くの研究機関において COTS (commodity and off the shelf) なハードウェアとオープンソースな PC UNIX を利用した PC クラスタが構築されている。

PC クラスタの魅力はそのコストパフォーマンスの高さにある。市販の専用並列計算機を購入した場合に比べて、数十分の 1、数百分の 1 のコストで自家製の並列計算機を手に入れることができるのである。最近では PC クラスタを構築するためのドキュメントが整備され、SCore や FAI (Fully Automatic Installation)、RedHat6.2 の clustering パッケージなどクラスタに必要なソフトウェアのインストールツールも整備されつつあ

る。PC クラスタ技術は、かなりの部分で実用段階にきていると思われる。

PC クラスタの構築は、本研究の基盤ではあるが主目的ではない。しかし、本研究の計算機環境は PC クラスタであり、性能の良い並列プログラムを書くために、クラスタの構築・運用技術は必須であると思われる。

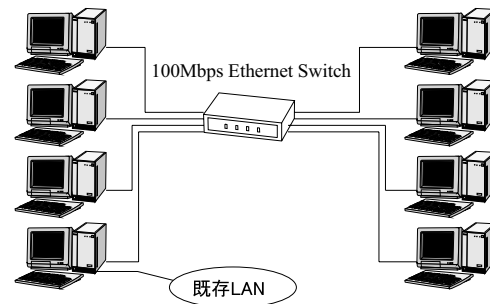


Fig. 1 PC クラスタ型並列計算機概念図

3 Global Computing

3.1 Grid

インターネットの普及とともに高速なコンピュータネットワークが整備されつつあり、地球上の任意の地域に置かれた計算機に対して遠隔的に計算させることが可能となっている。そこで、それに基づいて超高性能広域計算を行なう仮想的な「グローバルコンピュータ」を実現するための研究が米国を中心に行われている。これにはコンピュータ、ネットワーク、データベース、手法、人間といった各リソースを強く結びつけ、ユビキタスな分散サービスを一般的な形で提供することが必要であり、そのための技術は一般に Grid 技術と呼ばれている。Grid 技術の最終目標は「サイエンス・ポータル」の提供である。

これまでに開発されている Grid 技術としてはクライ

アント・サーバ型の Ninf, Netsolve, ミドルウェアとしての globus, legion, Nimrod/G, あるいは Java ベースのものがある。

3.2 シームレスな計算環境

Cluster Computing と Global Computing はこれまで、別々の研究分野であった。しかし昨年より、2つのコミュニティが結びつき、遠隔地にある計算拠点に Cluster Computing の技術を利用し、遠隔地を結びつけるのに Global Computing の技術を利用しようという動きが出てきた。そしてユーザには、グローバルコンピュータ上の計算サーバに Cluster が使われていることを意識させないシームレスな計算環境を実現することが目指されている。

4 並列遺伝的アルゴリズム

これまでには、次のような並列 GA のモデルについて研究を行っている。

- (1) Coarse grained モデルと Fine grained モデルの比較
- (2) 傾斜法 (準ニュートン法) を組み合わせたハイブリッド GA
- (3) 2 個体 GA (昨年度, 知的システムデザイン研究室在籍の浜崎による研究)

Coarse grained モデルとは、一般的に島モデルと呼ばれる分散 GA のことである。Fine grained モデルとは GA の計算において、ホットスポットとなる評価計算の部分を並列化するモデルのことである。すなわち、評価以外の GA の各遺伝的操作は 1 プロセッサで逐次的に行い、評価の部分については各スレーブノードのプロセッサ数だけ並列に計算を行う。(1) に示した 2 つのモデルの比較検討を行った結果、Coarse grained モデルの方が並列化によるオーバーヘッドを抑えることができ、かつ GA の解の探索効率も上昇することがわかった。

(2) の研究は現在進行中である。これは DGA の各世代のエリートに対して傾斜法を適用し、解の探索性能を上げようという新しいハイブリッド DGA のモデルである。DGA により大域探索を行い、傾斜法により近傍探索を行うことで通常の DGA に比べて解の探索効率が向上する可能性がある。

(3) の研究については、今後さらに深く研究を進める必要があると思われる。しかし 2 個体 GA のモデルは、並列分散モデルを自由につくることができ、Global Computing を行う上で興味深いモデルであると思われる。

5 まとめ

本研究は Cluster Computing, Global Computing の技術を学び、それを利用して常に新しく高性能な計算機

環境を実現しながら、その上で動くアプリケーションを考える。アプリケーションの 1 つとして並列分散 GA は興味深い題材であり、計算環境に応じた適切なモデルを考えていく。研究を遂行するにあたり必要とされる技術は特でない。ただし特徴として、コンピュータに触れている時間が長いことや英語のドキュメントを読む機会が多いことが挙げられるかもしれない。

参考文献

- 1) Grid forum : <http://www.gridforum.org>
- 2) The Globus Grid Toolkit : <http://www.globus.org/>
- 3) IEEE Task Force on Cluster Computing : <http://www.dgs.monash.edu.au/rajkumar/tfcc/index.html>
- 4) SCore : <http://pdswww.rwcp.or.jp/>
- 5) Extreme Linux : <http://www.extremelinux.org/>
- 6) RedHat : <http://www.redhat.com>
- 7) MPICH : <http://www-unix.mcs.anl.gov/mpi/mpich/>
- 8) Enrique Alba, Jose M. Troya, "A Survey of Parallel Distributed Genetic Algorithms", Complexity vol.4, No.4(1999).
- 9) D.E., Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading(1998).
- 10) D., Levine, A Parallel Genetic Algorithm for the Set Partitioning Problem, MCS-P458-0894(1994).