

コモディティな並列処理の OR における可能性 ---クラスタと Grid コンピューティングの動向---

松岡 聡 (情報理工学研究科数理・計算科学専攻 / JST)

matsu@is.titech.ac.jp, <http://matsu-www.is.titech.ac.jp>

1. はじめに

本講演では、高性能コモディティハードウェアの普及と高速ネットワークによる計算インフラの急激なパラダイムシフトと、それによってもたらされる莫大な計算パワーによる OR の計算数学的可能性、および OR が計算インフラのテクノロジーに及ぼしうる新たな技術革新について述べる。

従来高性能計算では主に物理化学系、特に流体や構造系の PDE 解法が主体であったが、計算機技術の進歩により高性能計算が普遍的になり、他の科学技術の分野でも容易に莫大な計算力を手にすることが可能になりつつある。特に、(コモディティ)クラスタ計算機技術、および Grid 計算技術により、「普通の」計算科学者および計算数学者は現在 10 年前と比較して数十万倍の計算パワーを手中にすることが可能になりつつあり、それによって OR で従来は困難だと思われていた NP 問題が一部解かれつつあるなど、そのパワーを生かす研究が試みられている。本講演によって、計算数学と、計算機科学(特に High-Performance Computing(高性能計算)の分野)との相互の絆が深まり、新しい科学的あるいは工学的な展開につながれば大変幸いである。

2. コモディティクラスタと Grid 計算---新しい高性能計算基盤のパラダイム。

筆者の研究室では高性能並列計算のソフトウェアの研究を行っており、特に近年では以下のソフトウェア技術を中心として研究開発を行っている:

(コモディティ)クラスタコンピューティングのソフトウェア技術: クラスタ計算機は複数の独立した計算機を超高性能 LAN で結合し、単一システムのイメージを提供する並列計算技術である。特に、一般の PC 技術を活用するコモディティクラスタでは、近年の PC の高性能化と低価格化により、従来のスパコンの数十倍から数百倍のコストパフォーマンスを実現する。現状の世界最大のクラスタは米国 Albuquerque 市の Sandia 国立研究所の CPlant クラスタであり、2000 年 8 月現在では 1120 ノードの Compaq Alpha プロセッサを擁し、毎秒 1TeraFlops あまりのピーク性能を誇る。また、NSF Pittsburgh Supercomputing Center では、2001 年春には 2728 プロセッサ、ピーク性能が 6Teraflops のクラスタが導入され、その時点で世界二位の速度を有する計算機となる。(ちなみに、高性能計算機の速度は<http://www.top500.org> にその詳細と現状の 500 位までのリストがあり、半年に一度更新されている。我が国の最速の計算機は東京大学計算機センターの日立 SR8000 であり、1Teraflops 程度。)

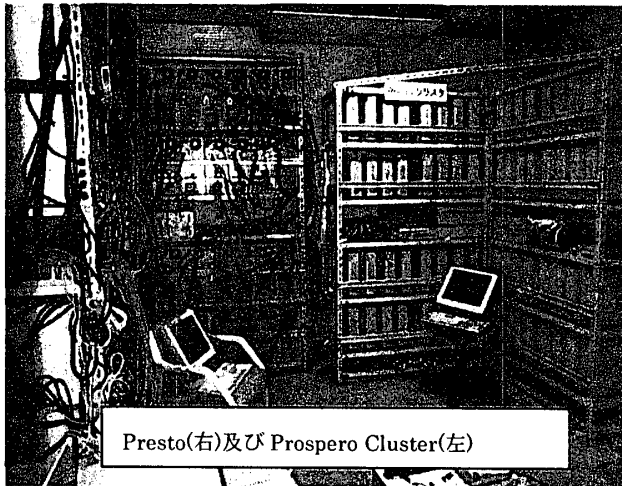
さらに、今後は半導体技術の指数的かつ急激な性能向上、および巨大クラスタの構築技術の進歩により、数年後には、ピーク性能数 TeraFlops~数十 TeraFlops、メモリ数 Tera バイト、および Peta バイト級の二次記憶を備

えるクラスタ計算機の構築が可能となる。また、小規模では数百万円で数十~数百 GigaFlops の並列計算機が構築可能になり、今後の並列計算機の爆発的な普及をもたらすとみなされている。後にのべるが、OR の諸問題もクラスタ計算機により、従来は困難だった問題が現実的な時間内で解けるようになる可能性がある。

しかしながら、単純なノード計算機の結合のみでは高性能・高信頼性・使いやすさの達成は困難であり、高度なソフトウェア技術とクラスタ計算に適したアルゴリズムの開発が望まれている。海外ではクラスタに関して種々の研究が活発に行なわれており、また VLinux などのクラスタのインテグレータが急増しており、その高いコストパフォーマンスと改善したユーザビリティが注目を集めている;しかし、残念なことにわが国の大型計算機メーカーはスパコン至上主義であり、クラスタへの対応は遅れている[1]。(通産省の RWC (<http://www.rwcp.or.jp>) プロジェクトだけは独自の高性能クラスタ技術の研究開発で内外で高い評価を受けている。)

- Grid コンピューティング: インターネットの普及と急激な高速化に伴い、TFLOPS 級の計算機、および数百 T ~ ペタバイト級のストレージを数 Gigabit から数百 Gigabit の WAN で結合することによって高性能計算およびそのデータセットを広範に提供するための計算インフラおよびパラダイムである[2]。大規模計算だけではなく、超広帯域マルチメディアの連携による遠隔協調計算、多数のセンサのリアルタイム処理、および Digital Library のテラバイト級のデータベース処理、なども含み、単に高性能計算機を remote login で遠隔で使うという話ではない。米国ではこの 2-3 年間で急速に研究開発が進み、The Next Generation Internet (NGI)および Internet2 などの超広帯域次世代インターネットと呼応して、急激にハード、ソフト、DB、および人的なインフラの整備が進行している。代表的なソフトウェアとしては Globus[3], Legion[4], Condor[5], Netsolve[6]などがあり、プログラムとしても国立研の Alliance, DOE の DOE2000, NSF の NPACI, NASA の IPG, DOE/DOD の ASCI などがあり、莫大な研究資金が注ぎ込まれつつある。最近では IETF(インターネットの標準化委員会)にならない、Grid 技術の研究および標準化を目指す The Grid Forum (<http://www.gridforum.org/>)も設立された。一方わが国のメーカー、大学、および研究所はこの流れからは取り残されているが、来年度から幾つかのプロジェクトが立ち上る可能性がある。

当研究室では大規模コモディティクラスタを複数構築し、システムソフトウェアおよびテストアプリケーションを構築するとともに、通産省電子総合研究所らと共同研究で



Presto(右)及び Prospero Cluster(左)

Grid コンピューティングのプロジェクトである Ninfl[7]を遂行している。写真は科学技術庁の「さきがけ 21」の研究プログラムにおいて当研究室で構築された Presto クラスタと Prospero クラスタである。中でも Prospero は 900Mhz の Intel Pentium III の PC をベースとして、64 ノード/128CPU/40Gbyte Memory/2TeraByte Disk を持ち、さらに本年度では倍の規模に増強され、ピーク性能 230GFlops に達する予定であり、研究室全体としては 4 台のクラスタで 400GFlops を得る。これは、東工大の計算機センターのスパコン(NEC SX-5+SGI Origin2000)と同程度の性能だが、ハードウェアのコスト的には数十分の 1 である。また、高速ネットワークと Grid のソフトウェアを通じ、電総研の 256GFlops のクラスタを連携しての使用が可能となる。

3. 莫大な汎用の計算力と OR への適用の可能性

このように、すでに大学の一研究室で大学の計算センターに匹敵する計算力を活用することが可能である。一般的には応用分野として、OR 系の問題の適用はまだまだではあるが、幾つかの研究プロジェクトが遂行しており、すでに問題サイズで世界記録を樹立した等の報告がある。

米国 Argonne 国立研究所では、NEOS[8]、及び MetaNEOS/MW[9]を研究開発している。NEOS は種々の最適化問題のソルバをパッケージ化し、WWW や mail などのインターネット上のインターフェースを付与するシステムである。ユーザは問題とそのソルバをブラウザとをなどで指定すると、バックエンドの計算機に適切に送付され、結果が返送・表示される。ユーザは高速な計算機の準備・インストールや管理の手間・データフォーマットなどから解放される。現状で、NEOS には内点法や非線形の種々のソルバが登録され、世界中から利用可能である。

MW は Master-Worker 並列プログラミングを主体とした Branch-and-Bound 組合せ最適化プログラミング用ライブラリである。ユーザは master-worker 流に問題記述をするのみで、システムが Grid 上の計算リソースを提供する Condor[5]システムに並列にタスクをスケジューリングして分配する。最新の報告では、Quadratic Assignment 問題に適用し、今までは大変困難とされていた $n = 30$ (nug30) の最適解を得るのに成功した。この際には、平均で数百プロセッサを一週間にわたり使用し、平均では毎秒あたり 100 万もの線形問題を解いている。

当研究室においても、当専攻の小島政和研究室らと共同し、クラスタ&Grid 環境で大規模な最適化問題を解く研究を行なっている。既に京都大学の藤澤氏が開発した SDPA

を用い、報告されているものより大規模な非線形問題を Ninfl システムを用いて Prospero クラスタで解いている。その際には問題によって 128 プロセッサで 100 倍程度の並列化効率を得ており、情報処理振興協会(IPA)の次世代アプリケーション開発事業「広域分散コンピューティングアプリケーションの開発」にてさらに Grid 環境でのテストが行なわれる予定である。また、同研究室の武田朗子氏は、代数方程式の根を求める際の Mixed Volume 問題において、Cyclic 14-roots problem を Prospero クラスタ上ではじめて解いたことを報告している(従来は Cyclic 13-roots が最高)。

逆に、OR の研究結果はクラスタ&Grid 計算に本質的に応用されていくことが期待されている。紙面の都合上詳しくは述べられないが、例えば、Grid 環境においては、莫大でかつ動的に変動しうる計算およびストレージの資源を適切にスケジューリングしなくてはならないが、このような問題はそもそも OR 分野のものである。

以上のように、OR の分野の研究者も、莫大な計算資源を活用することにより、物理化学に匹敵する新たな計算数学の領域の開拓が可能になると信ずる。無論、アルゴリズムやソフトウェアの研究開発は単純ではないが、計算機科学者らと共同で進行し、従来では考えられなかった規模の問題が次々と解かれ、それが有用な応用やあるいは新たな理論的発展につながる可能性がある。本講演がそれに向けた多少なりとも貢献できれば大変幸いである。

[参考文献]

- [1] 朴、渡辺、松岡、関口「インタラクティブエッセイ: これでもいいのか? 日本のスパコン」情報処理、Vol 41-07. (<http://www.ipsj.or.jp>に完全版掲載).
- [2] I. Foster and C. Kesselman (eds.). *The Grid: Blueprint for a New Computational Infrastructure*. Morgan Kaufmann, San Francisco, 1999.
- [3] I. Foster and C. Kesselman. "Globus: A Meta-computing Infrastructure Toolkit", *Intl J. Super-computer Applications*, 11(2):115-128, 1997. (<http://www.globus.org/>)
- [4] M. J. Lewis and A. Grimshaw. "The Core Legion Object Model", *Proc. IEEE Symp. on High Performance Distributed Computing (HPDC-5)*, Los Alamitos, California, Aug. 1996. (<http://legion.virginia.edu>)
- [5] M. Litzkow, M. Livny, and M. W. Mutka, "Condor - A Hunter of Idle Workstations", *Proc. 8th Int. Conf. of Distributed Computing Systems*, pp. 104-111, June, 1988. (<http://www.cs.wisc.edu/condor/>)
- [6] H. Casanova and J. Dongarra. "NetSolve: A Network Server for Solving Computational Science Problems", *Int. Journal of Supercomputer Applications and High Performance Computing*, 11(3), pp 212-223, Fall 1997. Also in *Proc. IEEE Supercomputing'96*, Pittsburgh. (<http://www.cs.utk.edu/netsolve/>).
- [7] A. Takefusa, S. Matsuoka, H. Ogawa, H. Nakada, H. Takagi, M. Sato, S. Sekiguchi, and U. Nagashima. "Multi-client LAN/WAN performance analysis of Ninfl: a High-Performance Global Computing System", In *Proc. IEEE Supercomputing '97*, 1997. (<http://ninf.etl.go.jp/>).
- [8] NEOS: Server for Optimization (Home page). <http://www.neos.mcs.anl.gov/> (2000年8月1日現在).
- [9] J.P. Goux, S. Kularni, J. Linderth, and M. Yoder. "An Enabling Framework for Master-Worker Application on the Computational Grid", *Proc. 9th IEEE Symp. on High-Performance Distributed Computing (HPDC-9)*, Pittsburgh, PA, Aug. 2000, pp. 43-50.