

計算科学の現状と今後の展望

奥田洋司 矢川元基

東京大学



1. はじめに

計算科学のこれまでの発展において、その最大の原動力はコンピュータの急速な進歩であった。表1は商用スーパーコンピュータの概ねの発展を示したものである。ピーク性能について見れば、黎明期から現在まで5年ごとに10倍ずつ演算性能が向上していることがわかる。コンピュータ性能の向上は単に解析性能をスケラブルに押し上げただけではなく、解析アルゴリズムにも変革をもたらした。例えば、データフロー計算機や並列計算機なくしては、非同期アルゴリズムや領域分割解法は研究者の興味をそれほど惹くものではなかったであろう。

計算科学が今後の様々な技術開発にいかにか寄与し得るかを考えるには、コンパイラなどハードウェアに強く依存するソフトウェアすなわちアプリケーション開発の支援環境も含め、ハードウェアの発展あるいはその利用形態の動向を見極めることが重要である。

2. 大規模計算科学

2.1 大規模計算科学の課題

スーパーコンピュータを駆使して大規模シミュレーションを実行するうえで、一般的に課題となっている問題について述べる。当面、並列計算機（分散メモリまたは分散共有メモリ）の趨勢は今後も変わらず、以下に挙げる項目は、現時点あるいは近い将来においても当てはまる問題である。

(1) 大規模分散データ処理

計算全体の中でメッシュ生成や領域分割（パーティショニング）が以前にも増して重要になりつつある。大規模データはもはや1領域データとして保持することも、ハードディスクやテープに出力することも非常に困難となる。また、領域分割は通信量、負荷分散、解析アルゴリズムの収束性などにも大きな影響をおよぼす。さらに、物理構造データ、CADデータ、シミュレーション用メッシュデータ、可視化用データのように多様なデータ形態を一貫して取り扱えるようなデータモデリングが必要となる。

年代	機種名	性能 (flops)
1975 ~ 1980	CRAY-1, CYBER 205	O (0.1G)
1980 ~ 1985	X-MP, CRAY-2, S-810, VP-200, SX-2	O (1G)
1985 ~ 1990	Y-MP, ETA-10, iPSC860, S-820, VP-400	O (1G ~ 10G)
1990 ~ 1995	Y-MP C90, T3D, CM-5, nCUBE2, Paragon XP/S, S-3800, SR2001, VP-2600, VPP500, AP1000, SX-3, Cenju3	O (10G ~ 100G)
1995 ~ 2000	T90, T3D, T3E, VPP700, AP3000, SP2, SX-4, SR2201	O (100G ~ 1T)
2000 ~	SR-8000, SX-5, SX-6, VPP5000, ORIGIN 2000	O (1T ~ 10T)

表1 代表的な商用スーパーコンピュータ

(2) 大規模並列演算技術（水平型並列）

これまで差分法や有限要素法による多くの大規模計算の経験が蓄積されているが、近年可能となっている分散大規模データ、例えば数億自由度にもものぼる計算はまだ世界的にも例が少ない。シミュレーションのためのアルゴリズムはもとより、モデリングから可視化までを超並列計算機上で on-memory で実行できる環境を構築しなければならない。マルチフィジクス（連成）シミュレーション技術もこの範疇に含まれる。

(3) マルチスケール (垂直型並列)

古くて新しい問題であるが、材料強度、乱流、気象、固体地球、バイオなど様々な分野で、時空間の広いスケールにまたがる事象の解明がシミュレーションに期待されている。例えば、合金において添加物濃度のわずかな差がマクロな物性に大きな影響を与える場合など、大きなスケールでの現象がマイクロな世界によって支配されている場合である。しかしながら、最小粒度の解像度で一様な時空間の離散化を行うことは原理的に不可能であり、何らかの均質化 (粗粒化、アップスケーリング、ネスティングなどとも呼ばれる) の導入が必要である。したがって、均質化に際してのモデルの妥当性、そのモデルを用いたマルチスケール解析、の 2 段階評価がシミュレーションに要求される。

(4) 大規模非構造格子データの可視化

可視化の処理はもはやポストプロセッシングとして解析から分けるよりも、項目 (1) でも触れたように、必然的に解析部の一部として考えるべきである。そして、大規模データの入出力を高速に行い、計算途中の結果を随時表示するリアルタイムな可視化手法や、データの間引き/抽出技法、並列画像処理技術などが、とくに有限要素法などの非構造格子に対して求められる。

2.2 大規模計算科学プロジェクト

地球シミュレータ計画 (後出) が 2001 年度完成に向けて進行中である。地球環境の諸問題に対するシミュレーションソフトウェアの多くが、科学技術振興調整費「高精度の地球変動予測のための並列ソフトウェア開発に関する研究 (平成 10 年度～14 年度)」の中で、地球科学と情報科学の研究者による協力体制のもとで開発が進められている。また、日本学術振興会未来開拓学術研究推進事業「計算科学 (平成 9 年度～13 年度)」では、現在の Tflops (T (テラ)= 10^{12} , flops (フロップス): 1 秒間あたりの浮動小数点演算の数) 級計算機が今後も 5 年間に 1～2 桁高速化されること、大規模工学設計用ソフトウェアが開発されること、実験的なアプローチでは解明が不可能な地球規模解析および物質解析が実現すること、を目指している。理化学研究所では「生体力学シミュレーション研究プロジェクト (平成 11 年度～15 年度)」を推進している。生体の構造と機能について厳密な力学モデルを構築し、その複雑な大規模方程式を高速で

解くことにより、高度な診断、新しい治療手段の開発、スポーツ工学やリハビリテーションへの寄与などを目指している。同志社大学では、文部科学省学術フロンティア推進事業「知能情報科学とその応用 (平成 12 年度～16 年度)」が発足した。本プロジェクトでは、生物が持つインテリジェントな方法の工学への応用を目的として、知的エンジニアリングデザインの手法の構築、知的情報処理およびメディア情報処理、知的化のための基礎理論の構築を推進している。

3. 大規模計算科学のための計算機環境

3.1 計算機の高速度化技術

演算性能の向上は主にプロセッサのクロックの高速化、アーキテクチャにおける並列処理によって推進されてきた。並列処理には、ベクトル計算、命令の並列実行、マルチプロセッサ化などの処理があり、クロックの高速化以上に大きな寄与を果たしている。

(1) ベクトル計算

ベクトルデータを同時に処理できるベクトルレジスタとパイプライン (セグメント) 化されたベクトル演算器との組み合わせによって高速演算が実現される。複数のベクトル演算機を装備することにより、さらに高速化が図られる。これらの技術は、いわゆるベクトル型スーパーコンピュータで用いられている。

(2) RISC

RISC (縮小命令セット方式) と呼ばれている方式では、使用頻度が高く実行時間の短い単純な命令セットのみを、メモリよりもアクセス速度の速いキャッシュメモリとのアクセスにより高速処理している。また、スーパースカラと呼ばれる計算方式では、複数の命令列のうち可能なものから並列実行される。この方式のプロセッサはパソコンに広く用いられている。また、次項の並列計算機の要素プロセッサとしても用いられている。

(3) マルチプロセッサ

アーキテクチャ面での並列化のレベルをもう一段階上げたものがいわゆる並列計算機である。複数のプロセッサにメモリをどう割り当てるかによって、共有メモリ型、分散メモリ型、共有・分散メモリ型の 3 つの方式がある。共有メモリ型は、全プロセッサから等距離の 1 つの大容量メモリにアクセスするもので、多くのベクトル型スーパー

コンピュータはこの方式をとっている。分散メモリ型とは、プロセッサとメモリの組を1つの単位（ノード）とし、それらを様々なトポロジーを有するネットワークで接続したものである。共有・分散メモリ型は、共有メモリ型マルチプロセッサを1ノードとし、それらをネットワーク接続したアーキテクチャで、いわば共有メモリ型と分散メモリ型を組み合わせた方式である。SX4、VPP700、SR8000、地球シミュレータなど、最近のハイエンド機の多くはこれに分類される。

コモディティプロセッサをネットワーク結合したものを一般にクラスタ計算機と呼ぶが、これも立派な並列計算機である。100Base-Tなど比較的安価で低速なLANで結合されたものはBeowulf計算機と称され、それと差別化してMyrinetなど高速ネットワーク結合されたものをスーパークラスタと呼ぶことがある。近年では世界のスーパーコンピュータの性能一覧であるTOP500[1]の上位にこのタイプが多く見られるようになった。後述のASCI計画はこの戦略をとっている。

3.2 ASCI計画

1995年にスタートした米国のASCI (Accelerated Strategic Computing Initiative、戦略的コンピューティング加速構想) [2]では、DOE (エネルギー省) 傘下の3研究所SNL (サンディア国立研究所)、LLNL (ローレンスリバモア)、LANL (ロスアラモス) がIntel、IBM、SGI、DEC、SUNを支援しつつ、2004年までにピーク性能100Tflopsを目指している。ASCIの特徴のひとつは、コモディティ技術のスケールアップによって大規模なコンピュータを構築しようとしていることである。これによって、プロジェクトの成果を産業界へスムーズに展開することが可能である。約10年かけて、1T、3T、10T、30T、100Tと5段階で推進する計画であり、これまでほぼ予定通りに開発が進んでいる。これらASCIの一連のコンピュータはTOP500[1]の常にトップを占めている。2000年6月からは、3段階目にあたる10Tマシン (ニックネームWhite) のLLNLへの納入が始まった。まもなく4段階目の30Tマシン (ニックネームQ) が運用開始とされており、さらに2004年頃にはピーク性能100Tflops (ニックネーム不明) のコンピュータが登場する。

3.3 地球シミュレータ計画

地球シミュレータ計画[3]とは、地球変動予測を目指し

て仮想地球をコンピュータ上に再現する「地球シミュレータ」を開発する、我が国のプロジェクトである。1997年度より5ヵ年計画、400億円の予算規模で進行中である。具体的には、超高速並列計算機システム「地球シミュレータ」の開発、大気・海洋分野のモデリングと高解像度シミュレーション、固体地球分野のモデリングとシミュレーション、地球シミュレータ用大規模並列ソフトウェアの開発、が進められている。

地球シミュレータはピーク性能40Tflops、総主記憶10Tbyteの分散共有メモリ型並列計算機であるが、このスペックは、気候・気象・海洋分野のシミュレーションにおいて経度方向4,000点、緯度方向2,000点、鉛直方向150点、メッシュあたりの倍精度変数200~400個が扱えるメモリ、実効速度を(地球シミュレータ計画立案時点の)1,000倍である5Tflopsを実現するものとしている。2002年の完成時には汎用機としては世界最高速のコンピュータとなるはずである。ハードウェアの詳細については文献[3]に譲るが、8PEでメモリを共有したノードを構成し、全体は640ノードからなる。各PEはベクトルプロセッサである。建屋は冷却系や外部記憶、ネットワークのスペースも含めると体育館のような大きさとなり、現在、横浜近郊に建設中である。

3.4 GRIDとITBL

次世代HPCインフラとして日米欧で精力的に研究開発が進められているインターネット利用技術がGRIDである[4]。ネットワーク接続された情報資源に対して、必要なときに、必要なサービスを即座に簡便に享受できるというコンセプトを実現するために、多様なOS、計算機、超高速ネットワーク、データベースを統一的なツールやインターフェイスで取り扱うためのミドルウェア、プロトコル、等の開発が行われている。これらの整備により、世界のスパコンの同時利用(メタコンピューティング)、世界の遊休パソコン資源の利用(ボランティアコンピューティング)、遠隔地での共同研究環境の提供(アクセスグリッド)、のほか、遠隔地の大規模データ利用計算科学(Data Intensive Computing)、WEB上でのアプリケーション提供(Grid ASP)などが期待されている。

平成13年度より、日本原子力研究所計算科学技術推進センターが旧科学技術庁傘下機関と協力してITBL (IT-Based Laboratory) 計画を発足させた[5]。これは全国のスパコンを接続して計算資源、計算プログラム、データベース、などの資源の共有化を図り、仮想研究環境

を構築しようとするものである。

4. おわりに ペタへの道一

ASCI や地球シミュレータ計画の特徴は、ハードウェアだけではなくアプリケーションの開発も同時に重視されている点である。この点については、テラの上、ペタフロップス (Pflops) (P (ペタ)= 10^{15}) コンピュータの事前計画を見ても伺える[6]。米国では、Pflops コンピュータの可能性について、既に 1994 年の時点で、それ以後 20 年間の戦略を、デバイス技術、アーキテクチャ、システム・ソフトウェアとツール類、アプリケーションとアルゴリズム、などの面から検討を始めている。

アプリケーションへのニーズを最優先させ、特化されたコンピュータが専用機である。その例として、恒星系やタンパク質の分子動力学シミュレーションなどで必要となる多体系の数値計算のために開発されてきた専用機並列計算機 GRAPE や MDM などがある[7]。専用機ということもあって、これらは 2001 年の時点で 100T の演算速度を達成しようとしている。同時に、適用分野の拡大を図るべく、粒子法やメッシュレス法などへの適用も考えられている。

全てのアプリケーションに専用機が生まれるとは考えにくい。したがって、産業界における計算力学の役割を考えると、ハードのコストやその更新サイクルの速さから、ハードの汎用的利用の重要性も明らかである。ピーク性能を追求する中で次々と開発されるコンピュータアーキテクチャの多様性を吸収できるようなアプリケーション開発体制が必要である。

上記のように、最近あるいは近未来のハードウェア発展の動向を見たとき、当たり前なことではあるがアプリケーションへのニーズが非常に重要な位置を占めることがわかる。そして、専用機の適用分野拡大と、ソフトウェア側からのハードウェアの多様性の吸収、この相乗効果が計算力学の可能性をより押し広げて行くと思われる。

参考文献

[1] <http://www.netlib.org/benchmark/top500/top>

500.list.html

[2] <http://www.sandia.gov/ASCI/>

[3] 谷啓二、地球シミュレータ計画、数値流体力学、第 9 巻、第 1 号、pp.17-1、2000

[4] <http://www-cger.nies.go.jp/grid-j>

[5] 山口勇吉、武宮博、仮想研究環境 ITBL 基盤ソフトウェアの開発、数値流体力学、第 9 巻、第 3 号、pp.83-88、2001

[6] T.Sterling, P.Messina and P.H.Smith, 'Enabling Technologies for Petaflops Computing', MIT Press, 1995 (小林達、栗本武共訳、'ペタフロップスコンピュータ'、筑波出版会、1997)

[7] J.Makino and M.Taiji, 'Scientific Simulations with Special-Purpose Computers --- The GRAPE Systems', John Wiley and Sons, 1998

【略歴】

奥田洋司 (OKUDA Hiroshi)

東京大学大学院助教授。1962 年生まれ。1990 年東京大学大学院工学系研究科博士課程修了、工学博士。同大学講師、助教授、横浜国立大学助教授、を経て 2000 年より現職。並列有限要素法による流体・構造解析、固体地球シミュレーション、メッシュレス法、科学技術計算ミドルウェアなどの研究・教育に従事。(財) 高度情報科学技術研究機構特別招聘研究員を兼務。固体地球シミュレータ GeoFEM 開発グループリーダー

矢川元基 (YAGAWA Genki)

東京大学大学院教授。1942 年生まれ。1970 年東京大学大学院工学系研究科博士課程修了、工学博士。同大学講師、助教授、を経て 1984 年より現職。大規模計算科学、メッシュレス計算力学、ニューロによる逆問題解法開発、原子炉プラントの安全工学などの研究・教育に従事。日本工学アカデミー会員。米国機械学会 (ASME) フェロー。