
SC07 参加報告書

島 伸一郎

地球シミュレータセンター, 独立行政法人 海洋研究開発機構

平成 19 年 11 月 29 日

1 はじめに

2007/11/10-16 と Reno 市, Nevada 州, U.S. で開催された SC07 に参加してきた。そこで私が見聞きした事を報告する。私の考えた事が大体伝われば良いと思い, あまり確認を取らずに書いた事も多いため, 内容の正確さは保証できない。誤りをご指摘いただくと幸いである。

ちなみに, 私の興味は非線形現象全般に対する数理物理学的なアプローチにある。特に大学院では, 結合自励振動子系の協同現象を少し勉強した。現在, 複雑な現象をいかに効率良くシミュレーションできるかという事に興味がある。スパコンが必要な大規模計算を始めたのはここ 2 年程で, 雲の精密モデルを開発し地球シミュレータ上で走らせている。ハードウェアにはあまり強く無い。

1.1 SC07 概観

SC07 は high performance computing, networking, storage, and analysis に関する国際会議で今年で 20 回目の開催となる。企業展示・研究展示・各種口頭発表・各種表彰等が催されており, この分野に関する, 最新の技術, その適用, 企業の商品, 研究成果, 関連する米国の政策, 世界の動向を知る事ができる。参加者は 9000 人規模であり, 非常に大規模な会議であるが, 良く組織されている。

2 Keynote, “Programming Bits and Atoms,” Neil Gershenfeld

http://sc07.supercomp.org/schedule/event_detail.php?evid=11124

Massachusetts Institute of Technology (MIT) の Neil Gershenfeld による基調講演である。現在の computing technology がハードウェア, ソフトウェアとして縛られている種々の暗黙の仮定 (ドグマ) を見つめ直し, その制約を緩和する事で, computer の新しい方向性を探る話である。情報のデジタルな表現 (bits) とその情報を扱う物理的実体 (atoms) の関係を再考している様である。面白そうな話ではあったが, 考え方が新しすぎるせいか, どうもいまひとつピンと来なかった。

彼の提示した “digital dogma” は 5 つで, それぞれは以下の様に緩和された。

1. computers come in cases → computers come in rolls, buckets.
2. compilers optimize programs → optimization programs compilers.
3. bits can be 0 or 1 → bits can be between 0 or 1.
4. internetworking → interdevice internetworking
5. programs can describe things → programs can be things.

例えば, 1) は, コンピュータはケースに入っているが, 必ずしもその必要は無いのでは無いかと言う疑問提示の様である. 実際, 彼の学生であった Bill Butera *et al.* は “paintable computer” なる技術を開発. 2mm 四方の (無線で互いにやりとりする?) チップを沢山, ペンキの様に壁に塗る事で動作するコンピュータを作ったらしい. 他にも色々例を挙げ, Gershenfeld は “computers come in rolls, buckets” と言い換えた. 将来的に computer は壁紙のロールやペンキバケツで提供される, と言った感じか.

全体にこんな調子である. Gershenfeld は “FAB” や “When Things Start To Think” 等の本を出版しているので, 読めば理解の役に立つであろう.

また, 気泡で論理回路を作ったと言う論文は面白そうであった. “Microfluidic Bubble Logic,” Manu Prakash and Neil Gershenfeld, *Science* **315**, 832-835 (2007).

また, program (=thing?) の自己複製についても言及していた.

3 Plenary Talks

全ての参加者に向けられた総合講演が 4 つ催された.

3.1 “The American Competitiveness Initiative: Role of High End Computation,” Raymond Orbach

http://sc07.supercomp.org/schedule/event_detail.php?evid=11120

アメリカはスパコンの開発に力を注いでいるが, 最高性能のコンピュータを持っている事や作れる技術がある事が, どの様にアメリカの国益につながっているかと言った政治的な話, のはずである.

講演資料はこちらに掲載されている. http://www.science.doe.gov/News_Information/speeches/speeches/07/SC07/SC07.htm

3.2 “Cosmology’s Present and Future Computational Challenges,” George Smoot

http://sc07.supercomp.org/schedule/event_detail.php?evid=11121

最新の宇宙論の話. 講演者の George Smoot は天文学者で, 宇宙背景放射の異方性を発見した功績により, John Mather とともに 2006 年にノーベル物理学賞を受賞した. 本講演では, 宇宙にはまだまだ分からない事がたくさんあって, 例えば宇宙が加速膨張している理由を解明するためにも, スパコンが役に立つ, と言った話をしたはずである.

3.3 “HPC in Academia and Industry - Synergy at Work,” Michael Resch

http://sc07.supercomp.org/schedule/event_detail.php?evid=11122

スパコンを有効に利用するためには学术界と産業界の連携が不可欠であるが, 実際にどの様なビジネスプランが考えられるかと言った提案の話.

3.4 “Toward Millisecond-scale Molecular Dynamics Simulations of Proteins,” David E. Shaw

http://sc07.supercomp.org/schedule/event_detail.php?evid=11123

今回の会議の中で一番興味深かった講演である。

講演者の D. E. Shaw は computer scientist であり, D. E. Shaw & Co. という評価の高いヘッジファンドの創設者である。2001 年から D.E. Shaw Research, LLC (DESRES) の chief scientist として専ら科学的研究に従事し, 現在は投資や資産運用には多くの時間を割いてはいない。

DESRES の方向性は, 研究期間が長期に渡り, リスクも高いが, 科学的に金銭的に社会的に大きな見返りが期待できる基礎研究を進める事にある。現在特に, たんぱく質を初めとする生体高分子の分子動力学シミュレーション (MD) の研究を活発に進めている。グループは, シミュレーション生化学, 計算科学, 応用数学, コンピュータ工学, の専門家集団から成る。今回の講演は, DESRES の研究状況の報告と言えよう。

研究の興味は, たんぱく質の構造や振る舞いを分子動力学シミュレーションにより明らかにする事で, 生物に関する我々の理解を深めるとともに, 癌の分子的挙動からの理解や, 製薬に役立てようと言う物である。

さて, 現在の課題は, 現状の技術ではまだまだ MD 計算が遅いと言う事である。彼らが目標とするのは, 25000~50000 原子から成るたんぱく質と, 水分子を含めた溶媒から成る系を 1 ms (= 0.001 秒) シミュレーションする事である。1 ms とは, たんぱく質の折り畳みなど, 生物学的に興味深い動的現象がおこる程度の時間スケールである。そのためには少なくとも 500 倍~10000 倍程度まで現状の計算速度を向上させなくてはいけない。彼らの戦略は, 新しいアルゴリズムと新しいアーキテクチャを作る事である。

ソフトウェアは, 既に開発が終わっている様で, “Desmond” という名前である [1-4]。高速な並列 MD パッケージで, 並列化効率が今までのどのコードよりも良く, 高精度で, 市販のクラスタマシンでも動作する。特に Neutral Territory Method という方法が新しいらしい。講演では, 図 1 の様にその性能が示された。右の数値は一日かけて計算する事で何ナノ秒 (ns) のシミュレーションが出来るか比較した表。Desmond on Cluster マシンが, Blue/Gene や MDGRAPE-3 を大きく上回って見取れる。現在, “Desmond” は大学や非商用目的には無料配付しているそうである。

ハードウェアは, “Anton” という専用機を作っている。設計は終わっている様であり, 2008 年の終りには完成するとか。図 1 の下 2 列にその性能が見積もられているが, Desmond on Anton の速さにはびっくりである。仕組みは良く分からないが, “almost never need to access off-chip memory” との事である。詳しくは [5]。

既に Desmond を使って科学的な成果が上がっている [6]。今後の動向は要注目である。

参考文献

1. D. E. Shaw, “A Fast, Scalable Method for the Parallel Evaluation of Distance-Limited Pairwise Particle Interactions”, *J. Comp. Chem.* 26, 1318 (2005).
2. K. J. Bowers, R. O. Dror, and D. E. Shaw, “The midpoint method for parallelization of particle simulations”, *J. Chem. Phys.* 124, 184109 (2006).
3. K. J. Bowers *et al.*, “Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters”, *Proc. SC2006* (2006).

4. K. J. Bowers, R. O. Dror, and D. E. Shaw, “Zonal methods for the parallel execution of range-limited N-body simulations”, J. Comp. Phys. 221, 303 (2007).
5. D. E. Shaw *et al.*, “Anton, a Special-Purpose Machine for Molecular Dynamics Simulation”, Proc. ISCA2007 (2007).
6. I. T. Arkin *et al.*, “Mechanism of Na⁺/H⁺ Antiporting”, Science 317, 799 (2007).

Comparative Performance of Anton
(Dihydrofolate reductase benchmark; 23,558 atoms)

Platform	Size	ns/day
GROMACS on single processor	1 processor core	0.4
MDGRAPE-3	12 ASICs	3.3
IBM Blue Matter on Blue Gene/L	512 processor cores	4.8
IBM Blue Matter on Blue Gene/L	512 ASICs	9.5
NAMD on cluster	256 processor cores*	13.7
Desmond on cluster (benchmark params)	512 processor cores	61.7
Desmond on cluster (production params)	512 processor cores	173.4
Anton (benchmark params)	512 ASICs	4,547.4
Anton (production params)	512 ASICs	14,115.2

Ref: Shaw, et al., Proc. ISCA07, 2007

図 1: SC07 における David E. Shaw の発表より転載。一日かけて計算する事で何ナノ秒 (ns) のシミュレーションが出来るか比較した表。Desmond on Cluster も十分に速いが、Desmond on Anton の速さは驚異的である!!

4 Awards

4.1 The 2007 HPC Challenge Awards

http://sc07.supercomp.org/schedule/event_detail.php?evid=11293

詳しくは<http://www.hpcchallenge.org>を見る事。与えられたベンチマーク問題をいかに解決したかが評価される様である。賞は2種類に分類され、Class 1は、4つ与えられたベンチマーク問題それぞれについて、最も performance の良かった計算に対して。Class 2はベンチマーク問題の最も“エレガント”な実装に対して。

Class 2は以下の2件に対して与えられた。

Award	Recipient	Affiliation	Language
Most Productive Research Implementation	Vijay Saraswat	IBM	X10
Most Productive Commercial Implementation	Sudarshan Raghunathan	Interactive Supercomputing	Python/Star-P

X10 (http://domino.research.ibm.com/comm/research_projects.nsf/pages/x10.index.html) は IBM が開発する新しいプログラム言語らしい。特徴は、マルチコアプロセッサやクラスタマシン向けで、object 指向で java を理解するとか吐き出すとか、内部的に (ユーザーがいちいち指定しなくても?) 並列化やデータを分散させるとか、Partitioned Global Address Space (PGAS) な言語の一種だとか。スパコンと言うのは性能を引き出すためには面倒なプログラムのチューニングが避けられないが、その手間を省き容易に High Performance Computing を可能とするユーザーに優しい言語らしい。発表資料はこちら。 <http://www.hpcchallenge.org/presentations/sc2007/hpcc-awards2-2007-x10.ppt>

Star-P (<http://www.interactivesupercomputing.com>) もプログラムを並列化しスパコンの性能を引き出す助けとなるソフトの様である。例えば Python, MATLAB, R と言った PC 用のソフトをスパコン上で効率よく動かす事ができるらしい。発表資料はこちら。 <http://www.hpcchallenge.org/presentations/sc2007/hpcc-awards2-2007-star-p.pdf>

これらの様な並列化を支援する新しいプログラミング言語は多数開発が進んでいる様で、今回 SC07 では他にも Unified Parallel C (UPC), Co-Array FORTRAN, Titanium, Chapel と言う名前を聞いた。日本では次世代スパコンの開発を進めているわけだが、プログラミング言語についてはどの様な議論がなされているのだろうか?

4.2 Gordon Bell Prize

High Performance Computing がいかに科学の問題へ適用されたかと言う点が評価される。今年、Peak Performance 賞と Price/Performance 賞と特別賞の 3 つが用意されていた。ただし、Peak Performance 賞は必ず選ばれるらしいが、他は該当無しもあり得る。(詳しくは <http://sc07.supercomputing.org/?pg=gordonbell.html>.)

最終選考には 4 件が残り、次の 1 件が受賞した。計算の詳細は各ページより pdf で配布されている。

4.2.1 Winner, “Extending Stability Beyond CPU-Millennium: Micron-Scale Atomistic Simulation of Kelvin-Helmholtz Instability,” James N. Glosli *et al.*

http://sc07.supercomp.org/schedule/event_detail.php?evid=11133

これが受賞した計算である。

溶融した銅とアルミの Kelvin-Helmholtz (KH) 不安定性を Molecular Dynamics (MD) Simulation した。KH 不安定性は、流体の境界面がせん断流によって不安定化し波や渦が発生するマクロな現象である。これをミクロの原子のシミュレーションから再現した所が科学的に新しい。

200 億個の原子を BlueGene/L (BG/L) で計算し、103.9TFlops の実効性能が出た。理論ピーク性能が 596.4TFlops なのでピーク性能比は 17.4%である。なお、BG/L の増強に伴い、計算性能も申請時より上回っている。

計算に際し、アルゴリズムの最適化や、データの入出力の最適化も行っているが、一番面白いのは、hardware error をプログラム上で修復している点である。あまりに沢山 CPU があるので、5 時間程でキャッシュエラーが生じてしまう。エラーが生じた事はハードウェアが検知するが、修正はしないらしい。と言ってプログラムを再スタートしてしまうと、効率が悪い。そこで、ソフトウェア上でパリティチェックをかける、様な事をするらしい。

面白いとは思いますが、KH 不安定性を原子から計算する意義が良く分からなかった。マクロな流体計算で十分 KH 不安定性を再現できるならば、原子から計算するのは無駄である。そこら辺の議論が無かった様に思われる。

ちなみに CPU-Millennium とは、のべ CPU 時間で 1000 年と言う意味。BG/L だと 212992 processor core があるので、1 CPU-Millennium は、現実の計算時間では $1000 \times 365 \text{ 日} / 212992 \approx 1.7 \text{ 日}$ に相当する。

4.2.2 Finalist, “A 281 Tflops Calculation for X-ray Protein Structure Analysis with the Special-Purpose Computer MDGRAPE-3,” Yousuke Ohno *et al.*

http://sc07.supercomp.org/schedule/event_detail.php?evid=11131

X 線散乱のデータを元に、たんぱく質の 3 次元構造を解析した。

分子動力学計算用の専用機である MDGRAPE-3 を使って計算し、281.6TFlops の実効性能が出た。理論ピーク性能が 302.8TFlops なのでピーク性能比は 93.0% である。

Non-equispaced Discrete Fourier Transformation の計算に MDGRAPE-3 を使ったと言う点が興味をひく。

実効性能は前節の物を上回り、科学的にも重要な結果だと感じられるが、受賞に至らなかったのは、専用機であったためであろうか？ 授賞式には出席しなかったので、詳しいことは分からない。

4.2.3 Finalist, “First-Principles Calculations of Large-Scale Semiconductor Systems on the Earth Simulator,” Takahisa Ohno *et al.*

http://sc07.supercomp.org/schedule/event_detail.php?evid=11132

年々集積度の上がっていく半導体素子だが、それに伴い CMOS トランジスタは 10nm 程度のサイズに近づきつつある。これは原子が 10 個並ぶ程度の素子と言う事であり、トンネル電流など量子力学的現象が顕在化する。(現在 Intel では 45nm の配線幅の CPU を出荷を始めている。) この計算は、不純物添加半導体の量子力学的振る舞いを、密度汎関数理論に基づく第一原理計算をしたと言う話。

系は原子 10648 個から成る一辺 6nm の半導体。(周期境界条件?) 地球シミュレータ (ES) を 512 ノード使って計算し、16.2TFlops の実効性能が出た。理論ピーク性能が 32.77TFlops なのでピーク性能比は 49.4% である。

計算プログラムは、「戦略的革新シミュレーションソフトウェアの研究開発」プロジェクトで開発が行われている、PHASE (http://www.rss21.iis.u-tokyo.ac.jp/theme/multi/material/material_softwareinfo.html) を ES 用にチューニングして使った。

計算結果は実験結果とも良く一致したとか。(基底状態のエネルギーが?)

4.2.4 Finalist, “WRF Nature Run,” John Michalakes *et al.*

http://sc07.supercomp.org/schedule/event_detail.php?evid=11134

ペタスケールマシンの登場を見越して、最新の気象モデルを IBM BlueGene/L や Cray XT4 で走らせたと言う話。

WRF (<http://www.wrf-model.org/index.php>) は大気現象の研究・現業数値予報の両方に役立つ事を目的とし、米国で開発されている気象シミュレーションモデルである。領域的な、カーテシアン座標系の、非静力学的で (= 静水圧近似していない)、完全圧縮な、モデル。

地形の考慮されていない理想半球を、水平 5 km, 6 sec の時間ステップで計算。乾燥大気のみを扱った、つまり水蒸気は考慮していない。90 日間の振る舞いをシミュレーションする、(予定だったがまだ走っていない?)

BG/L の 15360CPU を使って計算し、実効性能約 3.35 TFlops が出た模様。この CPU 数だと理論ピーク性能は 43.0TFlops で、ピーク性能比は 7.8%。一方、XT4 では 12288CPU を使って、実効性能 8.76 TFlops が出たらしい。この CPU 数だと理論ピーク性能は 63.7 TFlops で、ピーク性能比は 13.7%。事前に提出された論文には、増強される予定の IBM BlueGene/L で走らせて、64 TFlops 以上の性能を出す見込みだと書いてある。間に合わなかったのか、失敗したのか。怪しげな報告であった。

地球上の大気は、低波数では 2 次元乱流的な -3 乗のエネルギーカスケードを示し、高波数では 3 次元乱流的な $-5/3$ 乗のエネルギーカスケードを示す事が観測されている。これは、大きな空間スケールで見れば 2 次元的だが、小さな空間スケールで見れば 3 次元的であるためと解釈できよう。高解像度の計算をする事で、 -3 乗から $-5/3$ 乗へエネルギースペクトルが遷移する様子が再現出来た事が、この計算の科学的成果らしい。

4.3 The Seymour Cray Award

http://sc07.supercomp.org/schedule/event_detail.php?evid=11150

Seymour Cray 賞は High Performance Computing System, つまりハードウェア, に対して深く貢献した個人に対して与えられる。今年は, Kenneth Batcher (Kent State University) に贈られた。受賞理由は, “for the fundamental theoretical and practical contributions to massively parallel computation, including parallel sorting algorithms, interconnection networks, and pioneering designs of the STARAN and MPP computers.”

彼は, STARAN と MPP という名前の, 2 つの SIMD 並列プロセッサを作った。また, odd-even mergesort と bitonic mergesort という名の 2 つの並列ソーティングアルゴリズムを開発したらしい。受賞講演は http://www.cs.kent.edu/~batcher/SC07_Batcher.pdf に公開されている。

4.4 The Sidney Fernbach Award

http://sc07.supercomp.org/schedule/event_detail.php?evid=11150

Sidney Fernbach 賞は High Performance Computer を革新的なアプローチにより活用する事に貢献した個人に贈られる。今年は David Keyes (Columbia University) が受賞した。理由は, “Outstanding contributions to the development of scalable numerical algorithms for the solution of nonlinear partial differential equations (PDEs) and for his exceptional leadership in high-performance computation.”

受賞講演や詳しい解説は本人のページから入手出来る. <http://www.columbia.edu/~kd2112/fernbach.html>, http://www.columbia.edu/~kd2112/Fernbach_2007.pdf. Keyes は, 偏微分方程式の陰的数値解法と並列計算の専門家で, 特に Newton-Krylov-Schwarz 法と言うのが有名ならしい.

陽解法だと, 現象の時間スケールがあらわに効いてくるため, 時間スケールが混在すると確かに計算が困難になる. 陰解法はその問題を避ける事が出来る. しかし, 陰解法は並列化が難しい, と私は思っていたのだが, そうでも無い様である. 少し勉強してみようかと思う. <http://www.math.odu.edu/~keyes/papers.html> に文献が掲載されている. D. E. Keyes, 1995, "Aerodynamic Applications of Newton-Krylov-Schwarz Solvers", in Proceedings of the 14th International Conference on Numerical Methods in Fluid Dynamics, Springer, New York, pp. 1-20. の解説が詳しいかもしれない.

5 Papers

いくつかピックアップして紹介する. 各ページから論文を pdf で入手できる.

5.1 Computational Biology, "A Preliminary Investigation of a Neocortex Model Implementation on the Cray XD1," Kenneth L. Rice, Christopher N. Vutsinas, Tarek M. Taha

http://sc07.supercomp.org/schedule/event_detail.php?evid=11062

人間の脳新皮質 (neocortex) は知覚, 学習や認知をつかさどる部分である. 最近, George and Hawkins model と言う, 新しい脳新皮質の数値モデルが提唱された. このモデルは hierarchical Bayesian network model の枠組みに含む事ができ, 要は Bayes の定理を基に, 因果関係を推論する新しい人工知能のモデルと言う所であろうか. これを色々工夫し並列化効率を上げ, Cray の XD1 と言うスパコン上に実装し, 画像のパターンマッチングのテストも行ったと言う話.

スパコンの性能を生かして計算できたと言う点が新しく, まだ科学的には新しい結果が出ている訳では無さそうである. モデル自体も注目に値する物の様であり, 今後どの様な結果が出てくるのか興味が湧く. そのうち, スパコンが脳のように自律的に自分で考え学習したりする様に出来ないのか妄想してしまう. ちなみに, 新皮質は脳の中でも構造が均質な?ため並列化がしやすい部分の様である.

5.2 Computational Biology, "Anatomy of a Cortical Simulator," Rajagopal Ananthanarayanan, Dharmendra Modha

http://sc07.supercomp.org/schedule/event_detail.php?evid=11063

C2 と言う脳皮質のモデルを, 解法やメモリや通信などをチューニングして IBM の BlueGene/L に乗せ, ラット (大型のネズミ?) 程度の脳皮質に相等する大規模計算を実現したと言う話.

研究の興味は前節の物と同じであるが, モデルはもっと生理学的に忠実な物である. 脳は, その基本素子である神経細胞のネットワークとみなせるが, これを, 神経細胞 (neuron) の発火モデル, 電気信号を運ぶ軸索 (axon) のモデル, 伝わってきた信号を接触する神経細胞に伝達するシナ

プス (synapse) のモデル, の集団としてモデル化している. 発火モデルは Izhikevich モデルを使っている様である. 現象論的な低自由度のモデルながら, Hodgkin-Huxley model の様に多彩に振る舞い, Integrate-and-Fire model の様に計算効率の良いモデルだそう (Izhikevich EM (2003) Simple Model of Spiking Neurons. IEEE Transactions on Neural Networks, 14:1569-1572). モデルとしては Spike-Timing Dependent Plasticity (STDP) を導入した所が新しいらしい.

計算の工夫は色々している様だが, “allow overlapping communication” (MPI の非同期通信を利用?) という部分に興味を引かれた.

人間の脳皮質はラットの 500 倍位である. 彼らは long way to go と言っているが, 私にはわりと直ぐに実現できそうな数字に感じられた. なお, Stephen Hawking の “.. if very complicated chemical molecules can operate in humans to make them intelligent then equally complicated electronic circuits can also make computers act in an intelligent way.” という言葉を引用していたが, 全くその通りだと思う. 脳の HPC 分野の今後が楽しみな気がする. 脳その物が多数の神経細胞のネットワークからなるため, vector 並列型スパコンよりも scaler 並列型スパコンの方が構造的に近い気がする.

また, IBM では同じ様な方向性で BlueBrain プロジェクトと言うのも別で走っているらしい.

5.3 Computational Biology, “Large-scale Maximum Likelihood-based Phylogenetic Analysis on the IBM BlueGene/L,” Michael Ott *et al.*

http://sc07.supercomp.org/schedule/event_detail.php?evid=11064

Best Paper (BP) Finalist に残ったが BP には選ばれなかった論文の一つである.

各生物種の遺伝子情報を比較する事で進化系統樹を構成しようと言うのが研究の興味である. 具体的には, (幾つかの種からなる) ある生物集団の全遺伝子情報が与えられたとき, 考えうる進化系統樹の中から, 最尤法の意味で最も可能性の高い系統樹を, 膨大な遺伝子情報を基に見つけると言う問題設定である. 最尤な系統樹を見つけると言う決定問題は NP 困難であり, 多項式時間の決定論的アルゴリズムは現存しないため, 経験的方法でこれを近似的に推定する. この論文では, Randomized Axelerated Maximum Likelihood (RAxML) 法と言う, 正確で速くてメモリも使わない最新手法を, さらにチューニングする事で IBM BlueGene/L 上に移植し, かつてない大量のデータ (各 566470 塩基対からなる 212 組の DNA) を効率良く解析したと言う話.

実際に進化系統樹を作るのはまだ無理で, さらに改良を進めている様である. 系統樹に限らず, 大量のデータの構造を解析する, 分類するのは重要なテーマだと思う.

5.4 System Architecture, “The Cray BlackWidow: A Highly Scalable Vector Multiprocessor,” Dennis Abts

http://sc07.supercomp.org/schedule/event_detail.php?evid=11053

この論文が Best Paper Award を受賞した.

Cray 社は XT5h と言うハイブリッドスーパーコンピュータをリリースした. これは, 複数のタイプのプロセッサを自在に組み合わせる事で構成する事ができる, 革新的なスパコンである. 3 種類のアーキテクチャが blade と言う形で提供され, AMD Opeteron をベースとする “XT5 blade” と言う scaler processing blade, “X2” と言う vector processing blade, “XR1” と言う reconfigurable

processing blade, がある。最後の“XR1”は Field Programmable Gate Array (FPGA) という技術が使われており, 論理回路をユーザーが再定義できるらしい。

この論文で詳細が報告されている“BlackWidow”はこの vector processing blade “X2” の開発コードネームである。

詳しい事は良く理解できなかったが, 題名にもある通り, highly scalable なのが売りの様である。大規模な構成でもよく性能を引き出す事が出来ると言う事であろう。NEC も最近 SX-9 という新しい vector processor を発表したので, どの様な違いがあるのか (Cray のブースで) 聞いてみたところ, Cray の物の方が highly scalable だと言っていたがどうなんだろうか。印象に残ったのは, 将来的に MPI に取って代わるかもしれない, 先進的な global address space language である, Co-Array Fortran や Unified Parallel C をサポートしている点。

5.5 Performance Optimization, “Application Development on Hybrid Systems,” Roger D. Chamberlain *et al.*

http://sc07.supercomp.org/schedule/event_detail.php?evid=11105

複数のアーキテクチャが混在するハイブリッドマシンは HPC の新しいターゲットである。しかし, その様な環境で, ユーザがマシンの性能を引き出すには大変な手間がかかる事が見越される。そこで, プログラムの最適化を支援する様な新しい開発環境を作っていると言う論文。

6 Masterworks

産業界で High Performance Computing (HPC) がどの様に使われているかと言うのがテーマである。車などの製品の設計に使うと言う話などがあつた。生産ライン上で飛び散らない様にするため, プリングルズ (ポテトチップス) の周りの流体解析をしていると言う話もあつて驚いた。最近では映画の作成にもスパコンを使うそうである。撮影の後, もっと海の色を鮮やかにする, 草原を春らしくする, 登場人物が目立つようにするといった, 色合いやコントラスト等々の調整に使うらしい。

7 Birds of a Feather (BOF)

7.1 TOP500 Supercomputers

http://sc07.supercomp.org/schedule/event_detail.php?evid=11303

第 30 回 TOP500 リストが発表された。世界のスパコンの性能を Linpack benchmark で評価する。

最速は引き続き, IBM と the Department of Energy’s (DOE) が共同開発する BlueGene/L System である。ノードを大幅に増設した結果, 前回の 280.6 TFlop/s から比べパフォーマンスは大きく向上し 478.2 TFlop/s となった。その他 highlight はこちらに掲載されている。 <http://www.top500.org/lists/2007/11/highlights>

スパコンはその計算に対するエネルギー効率でも評価すべきとの話があつた。Green500 と言われる物であろう。「パフォーマンス/使用電力量」を評価の指標としようと言う話であつた。

今後の展開についても話があった。これからは沢山のコアからなるプロセッサ (many-core processor) が増えるだろうと言う話があった。例えば Intel は 80 コアプロセッサを開発している。そうになると、その性能を使いこなすためには既存のソフトウェアを大幅に書き直さなくてはいけなくなるとの指摘。

次回のリストが発表されるのは 2008 年 6 月であるが、その頃には IBM が開発する Roadrunner(<http://www.lanl.gov/roadrunner/>) というスパコンが完成しており、注目である。これは、Cell B.E. プロセッサと、AMD Opteron プロセッサからなるハイブリッドな構成となる。Cell B.E. プロセッサは、IBM、ソニー、東芝によって共同開発された物で、ゲーム機である PlayStation に採用された技術である。Cell プロセッサもマルチコアプロセッサである。

7.2 “Bayesian Network Awareness,” Bill Rutherford and Loki Jorgenson

http://sc07.supercomp.org/schedule/event_detail.php?evid=11310

“Bayesian Network” によりスパコンの “Network Awareness” を実現しようと言う話。

最近のスパコンは、沢山のノードのネットワークにより構成されている。状況に応じて負荷の大きい部分もあれば、小さい部分もあって、それも時々刻々と変化する。もしも、ネットワーク自身が、現在の状態を知覚し、状況に応じて動的に適応すると、スパコンは効率良く機能するだろう。つまりスパコン内のネットワークの自律的な制御である。このような概念を “Network Awareness” と呼ぶらしい。その実現には、学習や予測が不可欠であるが、それに “Bayesian Network” が使えそうだという提案である。

研究の方向性は面白そうだが、まだ具体性にかける気がした。最終的に、スパコン上で走る各ユーザーのジョブプログラムと連動して、動的負荷分散もうまい事自律的に調整してくれると素晴らしい。

8 Panels

パネリストがあるテーマについて公開討論するセッションである。

8.1 “Fifty Years Of Fortran,” David Padua, Henry M. Tufo, John Levesque, Richard Hanson

http://sc07.supercomp.org/schedule/event_detail.php?evid=11126

Fortran の公開 50 周年を記念した討論会である。当初の苦労話やその設計の理念、現状と課題、それから将来について。

そもそも Fortran は科学計算用に特化した、シンプルでいて性能が出る事が売りのプログラミング言語であった。開発と普及には苦労があった。近年ハードの発展と複雑化・多様化に伴い、Fortran は昔よりもはるかに複雑になっている。懸念としては、C/C++ に似てきていて、そのアイデンティティーを失いつつある事とか、Fortran を教育する場も減っている事とか。現在、Fortran の拡張に関し、明確な方針があるわけでは無い様だが、将来が暗い訳ではない。

などなど話題になっていた。現在新たな 50 年に向けての転換期と言う印象である。

8.2 “Progress in Quantum Computing,” Wim Van Dam, Mark Heiligman, Geordie Rose, Will Oliver, Eli Yablonovich

http://sc07.supercomp.org/schedule/event_detail.php?evid=11130

量子コンピュータは、状態の重ね合わせが可能と言う、量子力学特有の振る舞いを利用することで、超並列計算の実現が期待される、開発途上の技術である。このパネルではその進展状況について討論された。

現在のコンピュータは、bit に情報を保持する。bit は 0 か 1 のどちらかの状態のみしか表現する事ができない。一方、量子コンピュータでは、qubit に情報を保持する。qubit は、0 と 1 の 2 つの状態の線形な重ね合わせ状態を表現できる。この qubit に作用して、新しい 0 と 1 の重ね合わせ状態を作る量子論理ゲートを用意する事で量子コンピュータが実現する。沢山の状態の重ね合わせを表現し操作できる事から、明らかに古典コンピュータ以上の可能性を秘めるが、まだまだ実用化にはほど遠いのが現状である。(15 = 3 × 5 の因数分解ができる位にはなっている, IBM 2001.)

実用化に向けては多数の困難がある。まず、ハードウェア、つまり qubit を沢山保持し演算する様な機械、を作るのが難しい。特に、純粋な重ね合わせ状態を長時間保持するのが難しい。この点に関しては少しづつ進展が見られていて、様々な物理的実装の方法が提案されているそうである。Ion traps, Optical Lattices, Single photonics, Spin centers in solids, Superconducting circuits, etc. まだ決定的な方法が見つかったわけでは無く、また、エラー訂正の問題が深刻だとか。

また、量子的演算が実行可能な装置を作る事が出来たとしても、古典的コンピュータよりも効率よく計算できる具体的な問題とその解法・アルゴリズムを見つけなければ量子コンピュータの存在価値が無い。1994 年に、暗号解読とも密接に関連する、素因数分解を効率良く量子計算する Shor のアルゴリズムが報告されたが、その後しばらく進展が無かった様である。しかし、2002 年から新しい問題とアルゴリズムの発見が少しずつ進んでいる様である。

そんな中、最近のニュースは、D-Wave 社が 28-qubit から成る世界初の商用量子コンピュータを作ったと言う話。ただし、この話には懐疑的な意見が多い様である [1]。今回その D-Wave 社の Geordie Rose もパネリストとして参加している [2]。また D-Wave 社は展示会場にブースを設け、実演もして見せたとか。この D-Wave マシンだが、パネルでは「何か重要な側面を捕えているが、量子コンピュータとして機能するかどうかは明らかでない」と評されていた。少し関連資料を探してみた [1-3] が、量子コンピュータとして機能している証拠はまだ不十分の様である。Rose によると、より詳しい報告は間もなく学術論文として公表するとの事である。さらに、来年には数 100-qubit からなるマシンを作るそうであり、そのぐらいの規模になれば、もし本物ならば、量子コンピュータの特徴が顕著になるはずである。

参考文献

1. “Quantum leap of faith,” Nature 446, 245 (2007), <http://www.nature.com/nature/journal/v446/n7133/full/446245a.html>
2. Geordie Rose, “Image matching with a 28-qubit superconducting quantum computer,” http://dwave.files.wordpress.com/2007/11/20071110_sc07_panel.ppt
3. M.H.S. Amin, “Adiabatic Quantum Computation with Noisy Qubits,” http://dwave.files.wordpress.com/2007/11/20071119_mit_d-wave_1.ppt

9 おわりに

SC07を通してHPC分野の最新情勢に触れる事ができました。この貴重な機会を提供しお世話して下さった理化学研究所の皆様へ感謝致します。また行動を共にしたレポーターのお二人にも感謝致します。大変楽しい旅になりました。