

100 Tflops の分子動力学シミュレーション 専用計算機MDMの開発状況

成見 哲、薄田 竜太郎、古沢 秀明、川井 敦、
古石 貴裕、泰岡 顕治[†]、戎崎 俊一

理化学研究所情報基盤研究部計算科学技術推進室

[†] 慶應義塾大学理工学部機械工学科

我々は現在 100Tflops の分子動力学シミュレーション専用計算機 Molecular Dynamics Machine (MDM) を開発中である。MDM は、Ewald 法を用いたクーロン力および分子間力の計算のために専用に開発される計算機であり、タンパク質を含んだ系のようにクーロン力が支配的で大規模な分子動力学シミュレーションを加速することが出来る。MDM は WINE-2、MDGRAPE-2、ホストコンピュータで構成されており、それぞれクーロン力の波数空間部分、クーロン力の実空間部分および分子間力、結合力その他の計算を行う。2000 年末に完成予定であり、現在波数空間部分では 29Tflops の実効速度を得た。

Development Status of the MDM: 100 Tflops Special-purpose Computer for Molecular Dynamics Simulations

Tetsu Narumi, Ryutaro Susukita, Hideaki Furusawa, Atsushi Kawai,
Takahiro Koishi, Kenji Yasuoka[†], and Toshikazu Ebisuzaki

Computational Science Division, Advanced Computing Center,
RIKEN (The Institute of Physical and Chemical Research)

[†] Department of Mechanical Engineering, Keio University

We are now developing Molecular Dynamics Machine (MDM), a 100 Tflops special-purpose computer for classical molecular dynamics simulations. It can accelerate such as a large system with proteins by a specialized hardware for the Coulomb and Van der Waals force calculations. MDM is composed of WINE-2, MDGRAPE-2 and a host computer. WINE-2 accelerates the calculation for wavenumber-space part of the Coulomb force, while MDGRAPE-2 accelerates the calculation for real-space part of the Coulomb and van der Waals forces. The host computer performs other calculations. It will complete at the end of the year 2000. We got 29 Tflops of the effective performance for wavenumber-space part.

1 はじめに

分子動力学シミュレーションはさまざまな物質の物性を原子レベルで解析するために広く用いられている手法である。タンパク質を含んだ系など膨大な原子を含んだ系の分子動力学シミュレーションに対する要求は強い。分子動力学シミュレーションでは、原子は古典的な粒子と扱われ、クーロン力、分子間力、結合力で他の原子と相互作用する。最も計算の大変な部分は非結合力の部分であり、1万粒子を越

える系ではクーロン力がそのほとんどを占めることになる。Ewald 法 [1] を使うと、周期的境界条件のもとで計算量を有限(粒子数の 1.5 乗のオーダー)に出来るもののそれでも膨大である。

百万原子規模の大規模なシミュレーションを可能にするために、我々は Molecular Dynamics Machine (MDM) と名付けた専用計算機を開発中である。MDM は Ewald 法を使用することが出来る。WINE-2、MDGRAPE-2 およびホストコンピュータ

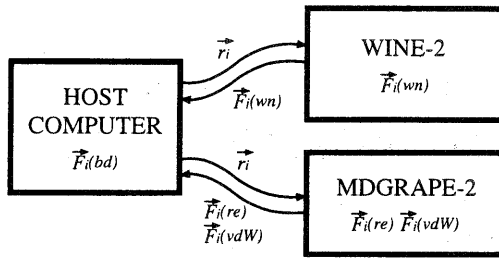


Figure 1: Molecular Dynamics Machine の構造

から構成されている (図 1)。WINE-2 はクーロン力の波数空間部分を、MDGRAPE-2 はクーロン力の実空間部分および分子間力を、HOST コンピュータは結合力や時間積分などその他の計算を行う。MDM のピーク速度は 100Tflops を予定しており、2000 年末に完成予定である [2, 3, 4]。このように大部分を占める力の計算箇所だけ専用のハードウェアを作るというアイデアは、既に GRAPE [5, 6] と呼ばれるハードウェア群によって有効性が示されている。本論文では、2000 年 7 月現在までに完成しているハードウェアの説明と、その上でのシミュレーションのパフォーマンスについて報告する。第 2 章では専用計算機での計算の概要を説明し、第 3 章では MDM のハードウェアについて述べる。第 4 章ではパフォーマンスについて述べ、第 5 章で今後の速度改善の方法について議論する。

2 Ewald 法によるクーロン力の計算方法

Ewald 法では粒子 i に働くクーロン力 $\vec{F}_i(\text{Clb})$ を実空間からの力 $\vec{F}_i(\text{re})$ と波数空間からの力 $\vec{F}_i(\text{wn})$ に分けて計算する。

$$\vec{F}_i(\text{Clb}) = \vec{F}_i(\text{re}) + \vec{F}_i(\text{wn}) \quad (1)$$

ここで、 $\vec{F}_i(\text{re})$ と $\vec{F}_i(\text{wn})$ は以下のように表される。

$$\begin{aligned} \vec{F}_i(\text{re}) &= \frac{q_i}{4\pi\epsilon_0} \sum_j^{r_{ij} < r_{\text{cut}}} q_j \left[\frac{\text{erfc}(\alpha r_{ij}/L)}{r_{ij}} \right. \\ &\quad \left. + \frac{2\alpha}{\pi^{1/2}L} \exp(-\alpha^2 r_{ij}^2/L^2) \right] \frac{\vec{r}_{ij}}{r_{ij}^2} \quad (2) \\ \vec{F}_i(\text{wn}) &= \frac{q_i}{2\pi\epsilon_0 L^3} \sum_{\vec{k}}^{k < k_{\text{cut}}} \frac{\vec{k}}{k^2} \exp(-\pi^2 L^2 k^2/\alpha^2) \\ &\quad \times \left[\sin(2\pi\vec{k} \cdot \vec{r}_i) \sum_j q_j \cos(2\pi\vec{k} \cdot \vec{r}_j) \right] \end{aligned}$$

$$- \cos(2\pi\vec{k} \cdot \vec{r}_i) \sum_j q_j \sin(2\pi\vec{k} \cdot \vec{r}_j) \Big]$$

ここで \vec{r}_i は粒子 i の位置、 \vec{r}_{ij} は粒子 j から粒子 i への相対ベクトル、 r_{ij} は粒子 i と粒子 j との間の距離、 q_i は粒子 i の電荷、 \vec{k} は波数ベクトル、 k は \vec{k} の大きさ、 L はシミュレーション空間の大きさ、 r_{cut} と k_{cut} は相対ベクトルと波数ベクトルの切断距離、 ϵ_0 は真空の誘電率、 α は実空間と波数空間の計算量を調整するパラメータ、そして $\text{erfc}(x)$ は補誤差関数である。

以下では MDM において実空間および波数空間がどのように計算されているかを説明する。

2.1 実空間の計算

MDGRAPE-2 では、粒子 i へ働く力 \vec{f}_i を以下のようにして計算する。

$$\vec{f}_i = \sum_{j=1}^{N_{\text{int}}} b_{ij} g(a_{ij} r_{ij}^2) \vec{r}_{ij} \quad (4)$$

ここで $g(x)$ は任意の中心力、 a_{ij} および b_{ij} は粒子 i と粒子 j の原子種によって決まる係数、 N_{int} は粒子 i に力を及ぼす粒子の数であり、切断距離内の粒子の数で

$$N_{\text{int}} \simeq \frac{4}{3} \pi r_{\text{cut}}^3 \left(\frac{N}{L^3} \right)$$

で表される。

例えばクーロン力の実空間部分の計算は、

$$\begin{aligned} g(x) &= \frac{2 \exp(-x)}{\pi^{1/2} x} + \frac{\text{erfc}(x^{1/2})}{x^{3/2}} \\ a_{ij} &= \alpha^2 L^{-2} \\ b_{ij} &= q_j \\ \vec{F}_i(\text{re}) &= \frac{q_i}{4\pi\epsilon_0} \vec{f}_i \end{aligned}$$

のように行う。

2.2 波数空間の計算

WINE-2 では、粒子 i への波数空間部分の力 $\vec{F}_i(\text{wn})$ を以下の二つのステップで計算する。まず、最初のステップで以下の離散フーリエ変換 (DFT) を行う。

$$S_n = \sum_{j=1}^N q_j \sin(2\pi\vec{k}_n \cdot \vec{r}_j), \quad (5)$$

$$C_n = \sum_{j=1}^N q_j \cos(2\pi\vec{k}_n \cdot \vec{r}_j), \quad (6)$$

Table 1: MDM に用いたハードウェア

Component	Product	Manufacturer
Node computer	Enterprise 4500	Sun
CPU	Ultra SPARC-II 400 MHz	Microsystems
Network	Myrinet	Myricom
Switch	16-port LAN switch	
Network card	LAN PCI card (LANai 4.3)	
Link	Bus bridge	SBS
Interface card	PCI host card/CompactPCI backplane controller card	Technologies
Bus	CompactPCI (WINE-2) / PCI (MDGRAPE-2)	PCI local bus spec. rev. 2.1

ここで N は粒子数である。次のステップでは、 S_n と C_n を使って逆離散フーリエ変換 (IDFT) を以下のようにして行う。

$$\vec{F}_i(\mathbf{w}_n) = \frac{q_i}{\pi \epsilon_0 L^3} \sum_n^{N_{wv}} a_n \left[C_n \sin(2\pi \vec{k}_n \cdot \vec{r}_i) - S_n \cos(2\pi \vec{k}_n \cdot \vec{r}_i) \right] \cdot \vec{k}_n \quad (7)$$

ここで a_n は

$$a_n = \frac{\exp(-\pi^2 L^2 k_n^2 / \alpha^2)}{k_n^2} \quad (8)$$

であり、 N_{wv} は長さが k_{cut} 以下の波数ベクトルの数の半分である。

$$N_{wv} \simeq \frac{1}{2} \cdot \frac{4}{3} \pi L^3 k_{cut}^3 \quad (9)$$

で表される。

3 MDM System

MDM は WINE-2、MDGRAPE-2、およびホストコンピュータで構成される (図 3)。

WINE-2 (図 2b) は 24 個の WINE-2 クラスタで構成されており、それぞれ 6 枚の WINE-2 ボードが接続されている。WINE-2 ボード (図 4b) は 16 個の WINE-2 チップが搭載されている。WINE-2 チップ (図 4c) は新たに開発した LSI であり、8 つの WINE-2 パイプラインを持っている。WINE-2 パイプラインは、DFT モードでは離散フーリエ変換 (式 5 および 6) を行い、IDFT モードでは逆離散フーリエ変換 (式 7) を行い、クーロン力の波数空間部分の力を計算する。図 6a は DFT モードでのパイ

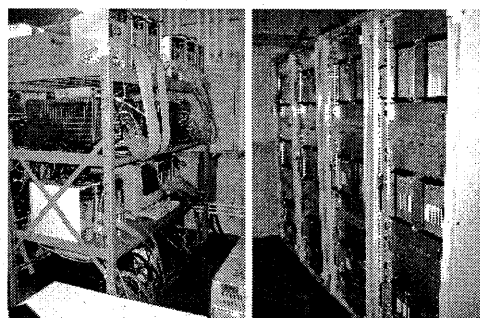


Figure 2: (a) ホストコンピュータは 4 つのノードコンピュータで構成される。36 枚の MDGRAPE-2 ボードが 12 個の PCI イクスパンションボックスの中に装着されている (左)。(b) 144 枚の WINE-2 ボードが 4 つのキャビネットに装着されている (右)。

プラインの構造を示す。IDFT モードもほぼ同じである。WINE-2 チップのピーク速度は、66.6MHz のクロック周波数の時約 20 Gflops である。WINE-2 は全部で 2,304 個の WINE-2 チップで構成され、合計 46 Tflops の計算速度を持つ。

MDGRAPE-2 (図 2a) は 12 個の MDGRAPE-2 クラスタで構成される。それぞれのクラスタには 3 つの MDGRAPE-2 ボードが接続されている。MDGRAPE-2 ボード (図 5b) には 2 個の MDGRAPE-2 チップが搭載されている。MDGRAPE-2 チップも我々が開発した LSI である。MDGRAPE-2 チップ (図 5c) は 4 つの MDGRAPE-2 パイプライン (図 6b) を持ち、式 4 の計算を行う。MDGRAPE-2 チップのピーク速度は、

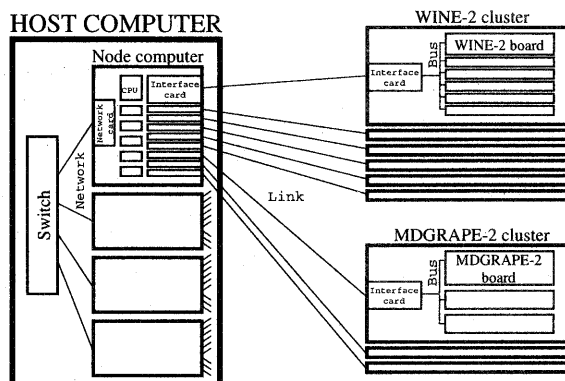


Figure 3: Molecular Dynamics Machine のブロック図。MDM は WINE-2、MDGRAPE-2 およびホストコンピュータで構成される。ホストコンピュータは 4 つのノードコンピュータで成り、ネットワークで接続されている。それぞれのノードコンピュータは 6 つの WINE-2 クラスターと 3 つの MDGRAPE-2 クラスターがリンクを通じて接続されている。1 つの WINE-2 クラスターには 6 つの WINE-2 ボードが、1 つの MDGRAPE-2 クラスターには 3 つの MDGRAPE-2 ボードがバスを介して接続されている。

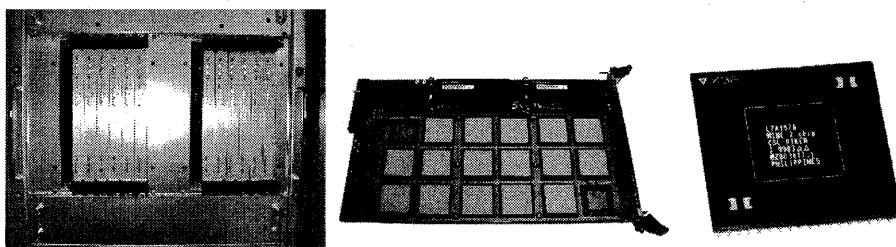


Figure 4: (a) 一つのサブラックに二つの WINE-2 クラスターが入っている (左)。 (b) WINE-2 ボードには 16 個の WINE-2 チップが chips 搭載されている (中)。 (c) WINE-2 チップ (右)。

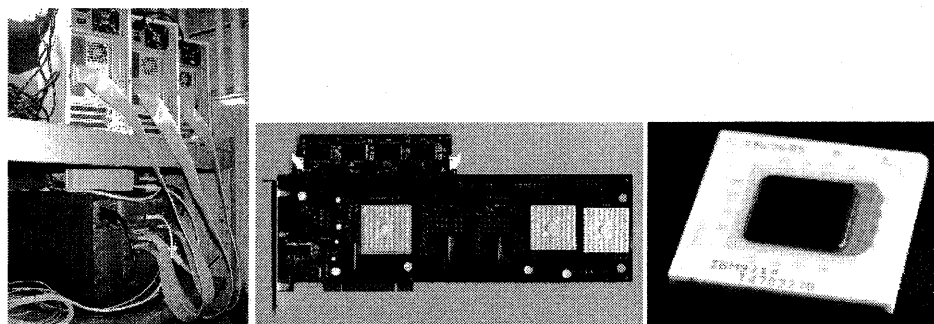


Figure 5: (a) 一つのノードコンピュータには三つの MDGRAPE-2 クラスターが接続されている (左)。 (b) MDGRAPE-2 ボードには二つの MDGRAPE-2 チップが搭載されている (中)。 (c) MDGRAPE-2 チップ (右)

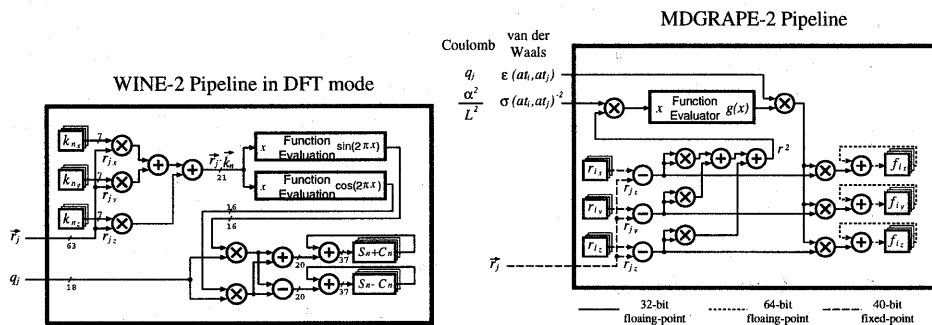


Figure 6: (a) DFT モードでの WINE-2 チップのパイプラインのブロック図 (左). (b) MDGRAPE-2 チップのパイプラインのブロック図 (右).

100MHz のクロック周波数の時約 16 Gflops である。MDGRAPE-2 は全部で 72 個の MDGRAPE-2 チップで構成され、合計 1 Tflops の計算速度を持つ。

表 1 は我々の使ったハードウェアを示す。ホストコンピュータとしては 6 つの Ultra SPARC II 400 MHz CPU をもつ Sun Enterprise 4500 を使った。Myrinet [7] ネットワークはノードコンピュータ同士を接続するために使用した。WINE-2 ボードや MDGRAPE-2 ボードとホストコンピュータを接続する部分には PCI バスを用いた。

4 パフォーマンス

ここでは MDM のパフォーマンスについて述べる。まず、WINE-2 だけによるパフォーマンスを述べ、次に現在の構成で行った MD シミュレーションのパフォーマンスについて述べる。

4.1 波数空間部分のみ

図 7(上) は、WINE-2 でクーロン力の波数空間部分のみを計算した場合の計算速度 (Tflops) を表している。2 千 4 百万粒子で、波数ベクトルが 50 万のとき、ピークの 63% である 29 Tflops を達成した。この実効速度は、世界最速のスーパーコンピュータのピーク速度の 7 倍以上である。1 ステップの計算は、ノードコンピュータの 1 個の CPU で行えば 7 ヶ月以上かかるのに対し、WINE-2 では 20.8 秒で行える。しかし、粒子数や波数が少ない場合にはそれほど実効速度が良くない。図 7(下) は、系のサイズを固定した時、ボード数を増やすにつれて実効速度がどのように上がるかを示している。理想的には

ボード数に比例して計算速度が向上するはずだが、大きくずれている。これは、ノードコンピュータと WINE-2 ボードとの通信、およびノードコンピュータ同士の通信がボトルネックになっているからである。詳しい議論は別の論文 [3] に述べられている。

4.2 MD シミュレーション

現在の MDM システムで NaCl の MD シミュレーションを行い、実効 1.02 Tflops を得た。5,926,176 ペアの NaCl イオン (11,852,352 原子) を一辺 728 Å の立方体に入れ 1200K で NVT アンサンブルのシミュレーションを行った。イオン間には以下の Tosi-Fumi potential [8] を用いた。

$$\phi(r) = \frac{q_i q_j}{r} + A_{ij} b \exp\left(\frac{\sigma_i + \sigma_j - r}{\rho}\right) - \frac{c_{ij}}{r^6} - \frac{d_{ij}}{r^8}$$

ここで、 A_{ij} , b , σ_i , σ_j , ρ , c_{ij} , d_{ij} はパラメータである。クーロン力の実空間部分および分子間力の切断距離 r_{cut} は 22.6 Å、クーロン力の波数空間部分の切断距離は $Lk_{cut} = 63.9$ 、このとき $\alpha = 72.8$ とした。時間刻は 2 fs で、800 ステップの計算に約 6 時間半 (23,155 秒) かかり、14.4 Tflops の計算速度を得た。ただし、現在の MDM システムでは波数空間部分の計算速度が実空間部分に対して約 50 倍速いため、波数空間部分の計算量が圧倒的に多い。仮に汎用計算機で行ったならば、実空間と波数空間バランスを取るために、 $r_{cut} = 68.9$, $Lk_{cut} = 20.6$, $\alpha = 23.5$ が最適と考えられる。この計算量の無駄まで考慮に入れると、実効 1.02 Tflops になることが分かった。詳しい議論は別の論文 [4] に述べられている。

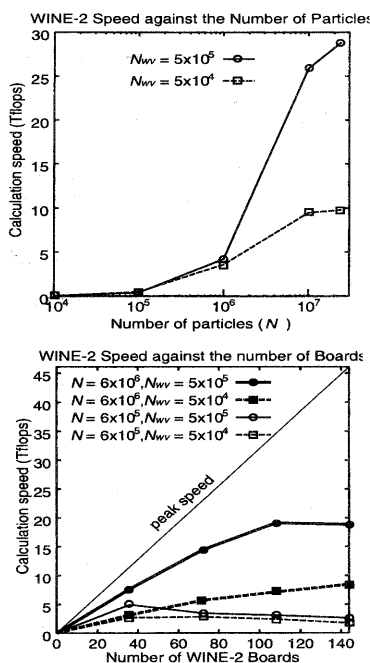


Figure 7: (上) WINE-2 の実効計算速度。 $N = 2.4 \times 10^7$, $N_{wv} = 5 \times 10^6$ の時 29 Tflops を得た。(下) システムサイズを固定してボード数を増やした時の計算速度の増加。システムサイズが小さい時はボード数を増やしても実効計算速度はあまり増加しない。

5 おわりに

現在の MDM システムは、ピーク速度の 47 Tflops に対して MD シミュレーションの実効速度はそれほど高くない。これは以下の三つの理由による。

まず、WINE-2 と MDGRAPE-2 のパフォーマンスのバランスが悪い。これによって実効速度は 10 倍損している。2000 年末までに MDGRAPE-2 チップを 2000 個程度に増加させる予定で、これによって解決できると思われる。次に、ボードの計算速度に対しボードとの通信速度が遅い。現在 32-bit の PCI バスを使っているのを 64-bit の PCI バスに変更する予定であり、これによって 2 倍程度通信速度が向上すると思われる。最後に、ノードコンピュータ同士の通信速度がボードの計算速度に対して遅い。今後 Myrinet カードを新しいバージョンの物に入れ換える予定であり、これによって 3 倍程度通信速度が向上すると思われる。

MDM の完成は 2000 年末を予定している。完成時には、百万粒子の Ewald 法を使った MD シミュ

レーションの一ステップが、0.19 秒で計算できる予定である。このとき、1.6 ns のシミュレーションに一週間しかかからない。これは仮にそのとき最速のスーパーコンピュータで同じ計算を行ったのに比べて 10 倍以上速いと予想される。このような大規模な MD シミュレーションは今後タンパク質の機能の解明などに役立つことと信じている。

References

- [1] Ewald, P.P., *Ann. Phys.*, **64**, 253 (1921).
- [2] Narumi, T., Susukita, R., Ebisuzaki, T., McNiven, G., and Elmegreen, B., *Molecular Dynamics Machine: Special-purpose Computer for Molecular Dynamics Simulations*, *Mol. Sim.*, **21**, pp. 401-415 (1999).
- [3] Narumi, T., Susukita, R., Furusawa, H. and Ebisuzaki, T., 46 Tflops Special-purpose Computer for Molecular Dynamics Simulations: WINE-2, in *Proceedings of the 5th International Conference on Signal Processing*, Beijing, (2000), in press.
- [4] Narumi, T. et al., Susukita, R., Furusawa, H., Kawai, A., Koishi, T., Yasuoka, K. and Ebisuzaki, T., 1.02 Tflops Molecular Dynamics Simulation for NaCl with a Special-Purpose Computer: MDM, SC2000, Dallas, (2000), in press.
- [5] Sugimoto, D., Chikada, Y., Makino, J., Ito, T., Ebisuzaki, T., and Umemura, M., "A Special Purpose Computer for Gravitational Many-body Problems," *Nature*, **345**, pp.33-35 (1990).
- [6] GRAPE Project, <http://grape.c.u-tokyo.ac.jp/grape/>
- [7] Myricom, Inc., <http://www.myri.com/>
- [8] Tosi, M. P. and Fumi, F. G., *J. Phys. Chem. Solids*, **25**, p. 45, (1964).