GRAPE-5 および GRAPE-6 計画の現状

牧野淳一郎 (東京大学大学院総合文化研究科)

1　はじめに

ここでは、現在我々が開発を進めている次世代重力多体問題専用計算機 GRAPE-5 および GRAPE-6 について、それぞれ概要と開発の進行状況を述べる。

まず GRAPE-5 について、次に GRAPE-6 について述べる。

2　GRAPE-5

GRAPE-5 は低精度型の GRAPE-3[OME$^+$93] の後継機である。現在、プロセッサチップの設計がほぼ終了した段階であり、システム全体の詳細についてはまだ検討中である。

2.1　パイプラインプロセッサの概要

表 1 に GRAPE-3 との比較の形で GRAPE-5 パイプラインプロセッサの概要をまとめる。

表 1: GRAPE-3 と GRAPE-5

	GRAPE-3	GRAPE5
動作速度	20–25 MHz	80–100 MHz
チップ内パイプライン数	1	2
デザインルール	1μm	0.5μm
動作電圧	5V	3.3V
計算速度	0.6–0.75 Gflops	4.8–6 Gflops
単一粒子からの力の相対精度 (RMS)	2%	0.3%
位置座標精度	19 ビット	32 ビット
カットオフ	なし	可能
外部バス分周比	1:1	1:6

データの内部表現や演算方式などは GRAPE-3 のものをほぼ踏襲している。改良点の重要なものは以下の 4 つである :

1. 動作速度を 4 倍 (予定) に向上させた
2. チップ内に 2 本の演算パイプラインを集積し、さらに 2 倍の性能向上を実現した。
3. 位置座標の分解能、ペアワイズの相互作用の相対精度を向上させた。
4. P^{3}M 法、 Ewald 法などの周期境界を扱う方法で必要になる、有限距離での相互作用のカットオフを実装した。

動作速度と集積度の向上の結果、チップ当たりの性能はほぼ一桁向上する。すなわち、現在の GRAPE-3A ボード一枚と同等な性能がチップ単体で得られる。実際のシステムでは、 GRAPE-3 と同様に 8 個程度のチップを載せるとして、ボードあたり 50 Gflops 程度の性能がえられることになる。直接計算を行なったとして、 10 万体が 1 ステップ 6 秒で計算できる。

計算精度に関しては、今回は必要精度ぎりぎりにするというよりはある程度高い精度にして応用範囲を増やすという方針をとった。GRAPE-3 では、データの内部表現は底が 2 の対数フォーマットをとり、その小数点以下（ほぼ、普通の浮動小数点表現の仮数部に相当）に 5 ビットを使った。これにたいし、GRAPE-5 ではここを 8 ビットとし、相対精度をほぼ一桁改善した。これにより、GRAPE-3 では若干精度に不安があった（もっとも、数値実験の結果は大きな問題はないことを示してはいる [APB97]）円盤銀河などの場合でも問題なく計算可能にする。また、位置座標についても同様にレンジをひろげる。

GRAPE-3 に比べた新しい機能として、有限距離での相互作用のカットオフを実装した。これは単純な RAM によるテーブルで実現されていて、力、ポテンシャルに対して、距離のある区間で定義された任意関数を掛けることができる。これにより、P^3M 法、Ewald 法などの周期境界を扱う方法が低精度の GRAPE を使っても高速に計算できるようになる。

Brieu ら [BSO95] は、GRAPE-3 を使って P^3M 法を実現した。彼らは、ソフトニングを変えて複数回重力を計算し、その線形結合をつくることで近似的にカットオフを実現した。しかし、この方法では実効性能が低下するだけでなく、精度的にも問題が多い。ハードウェアでカットオフを実現することで、大きく計算精度を上げるとともに計算速度も向上できる。

2.2 GRAPE-5 システムの概要

すでに述べたように、GRAPE-5 システム全体の設計はまだ確定していない。以下に、現時点での案を示す。

設計において、最も重要なのはホストインターフェースの選定である。GRAPE-3 では VME バスを使ったが、これは GRAPE-3 の段階でも性能のボトルネックになる場合があった。特に GRAPE 用ツリーコード [Mak91] では、ホスト計算機の性能が向上するに従ってインターフェース速度が大きな問題になってきていた。

GRAPE-5 では、ここには GRAPE-4 用に開発された PCI インターフェースボード [KFTM97] を使う。これにより、GRAPE-3 では実効 2–3 MB/s であった転送速度が 60 MB/s 程度と 20 倍程度にできる。また、PCI バスを使うことで、安価な Intel マシンから超並列計算機、あるいはベクトル計算機まで多種多様なホスト計算機を使えることになる。

さらに、GRAPE-3 では必要であったホスト計算機の I/O バスと VME バスの間をつなぐアダプタや VME バスのための匡体、バックプレーンなどが不要になり、安価にシステムを構成出来る。

GRAPE-4 では PCI インターフェースボードにまずコントロールボードがつながり、その先にプロセッサボードがつながるという構成をとった (図 1) が、GRAPE-5 ではそうしないで PCI インターフェースボードに直接プロセッサボードをつなぐ。こうすることで、最小構成のシステムを安価に構成できる。この構成の問題点は、複数ボードを使う時にボード数に比例した枚数の PCI インターフェースボードと PCI スロットが必要になることである。これに対しては、ボード 1 枚に載せる GRAPE-5 チップの数をなるべく増やすことで対応したい。

ツリーコードに対して予想される性能は、川井 [KFTM97] の GRAPE-4 のデータから推測するとホスト計算機が 500 MHz の Alpha、GRAPE-5 が 8 チップであった場合に 100 万粒子で 1 ステップ 20 秒程度、その内訳は GRAPE の計算が 3 秒、通信が 4 秒、残りがホストでの計算である。ホストはこれからも速くなるであろうし、GRAPE の計算速度はチップ数を増やせば改善できるので、1 ステップ 10 秒程度までは大きな困難はないであろう。Dubinski[Dub96] は 256 プロセッサの Cray T3D で 100 万体 1 ステップ 20 秒程度の速度を得ている。これと比べると、GRAPE-5 一枚でかなり大きな超並列計算機と同等の速度が得られるということがわかる。

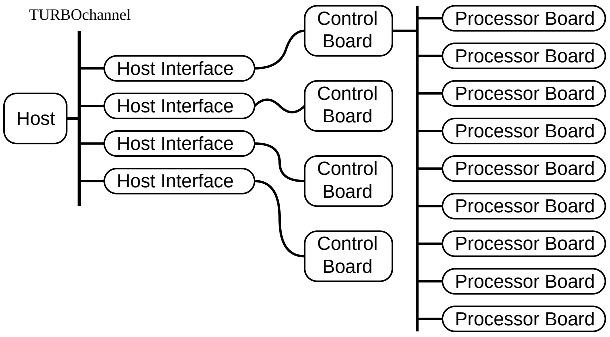

図 1: GRAPE-4 の構成

2.3 開発状況

1998 年 1 月現在において、チップの設計は完了し、現在メーカー側が作業を行なっている。一年以内にシステムとして完成させられる見込みである。

3 GRAPE-6

GRAPE-6 は高精度型の GRAPE-4[MTES97] の後継機である。日本学術振興会未来開拓学術研究推進事業の研究分野「計算科学」のプロジェクトの一つとして、1997 年度から開発を始めた。現在、システム全体とプロセッサチップの概念設計がほぼ終了した段階であり、プロセッサチップの詳細設計を進めているところである。

3.1 システム全体の構成

GRAPE-6 は、図 2 に示すように

ピーク性能 200 テラフロップス程度の重力専用プロセッサと、粒子系に専用化したアーキテクチャを持つが相互作用を計算するパイプラインを再構成可能論理（FPGA、Field Programmable Gate Array）で実現した多用途粒子系プロセッサを汎用のホスト計算機につないだものになる。FPGA を使ったプロセッサは、集積度において専用プロセッサに劣るが、そのかわり多様な問題に適用可能である。自己重力系への応用としては、

- SPH による流体計算
- Ewald 法による周期境界条件での重力計算
- ツリーコード [BH86]、FMM（例えば [BHER95]）などの高速計算法における多重極展開の評価
- SCF 法 [HO92] の実現

など、さまざまなものが考え得る。以下、まず重力専用プロセッサについて、次に多用途粒子系プロセッサについてその概要を簡単に述べる。

3.2 重力専用プロセッサの概要

表 2 に GRAPE-4 との比較の形で GRAPE-6 重力専用プロセッサの概要をまとめる。

データの内部表現や演算方式などは GRAPE-4 のものをほぼ踏襲している。改良点の重要なものは以下の 4 つである：

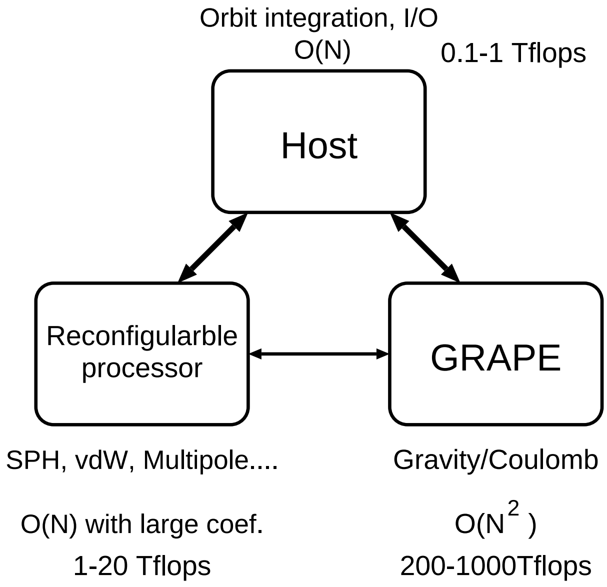

図 2: システム全体の構成

1. 動作速度を 4 倍程度に向上させた
2. チップ内に 6 本の演算パイプラインを集積し、3 サイクル必要とする GRAPE-4 に比べてさらに 18 倍の性能向上を実現した。
3. 予測子計算パイプラインも内蔵し、システム設計を単純化した。
4. GRAPE-5 と同様な有限距離での相互作用のカットオフを実装した。

動作速度と集積度の向上の結果、チップ当たりの性能はほぼ 50 倍に向上する。すなわち、現在の GRAPE-4 ボード一枚と同等な性能がチップ単体で得られる。実際のシステムでは、16 個程度のチップを載せるとして、ボードあたり 500Gflops 以上、すなわち現在の GRAPE-4 に匹敵する性能がえられることになる。直接計算を行なったとして、100 万体が 1 ステップ 1 分で計算できる。システム全体ではさらにこの 200 倍以上の性能になり、100 万体が 1 ステップ 1 秒以下となる。

以下に、重力プロセッサ全体のアーキテクチャを簡単にまとめる。

基板 1 枚に 16 プロセッサをのせてプロセッシングボードとし、プロセッシングボード 16 枚とコミュニケーションボード 1 枚で 1 クラスタ、16 クラスタでシステム全体を構成する。これらの 16 のクラスタは、インターフェースを集中制御するコントロールユニットを介してホスト計算機に接続される。

なお、プロセッサチップはそれぞれ自分のメモリを持つ構成とし、2-4 チップをまとめてモジュールとする。従って、そのモジュールが載る基板には通信ネットワークだけがあることになる。

ホスト計算機とコントロールユニットの通信速度は最低 2 GB/s 程度が必要になる。なるべく多様な応用を可能にするためには、それ以上の通信速度があることが望ましいので、必要に応じて拡張可能な構成とする。

プロセッサボード内およびボード間のネットワークのトポロジーと物理的な実装方式についてはこれから検討する必要がある。最近になって、Gigabit Ethernet や IEEE-1394 などの非常に高速なシリアル接続技術が安価に利用可能になってきたので、これらを利用して高速でしかも柔軟なネットワークを構成する。

表 2: GRAPE-4 と GRAPE-6 重力プロセッサ

	GRAPE-4	GRAPE6
動作速度	32 MHz	100–125 MHz
チップ内パイプライン数	1/3	6
デザインルール	1μm	0.25μm
動作電圧	5V	2.5V
計算速度	0.6 Gflops	36–45 Gflops
単一粒子からの力の相対精度	24 ビット	24 ビット
位置座標精度	64 ビット浮動小数点	64 ビット固定小数点
カットオフ	なし	可能
外部バス分周比	1:2	1:4
総チップ数	1692	4096
ピーク性能	1.08 Tflops	150–180 Tflops

3.3 多用途粒子系プロセッサ

多用途粒子系プロセッサは、要するにこれまでの GRAPE では市販の LSI を配線したりカスタム LSI を作ったりして実現していた演算パイプラインを、 FPGA (Field Programmable Gate Array) というソフトウェアで実現論理が変更可能な LSI に置き換えようというものである。これにより、パイプラインの中身を変えていろいろな計算に使おうというわけである。

FPGA を使って計算機を作る試みはいろいろあるが、主に画像処理などのデータ長が短く、しかも入出力が多い処理に使われている。そのために、これまでに開発された計算機は強力な入出力機能を持つものがおおい。

しかし、粒子系への応用を考えると、強力な入出力機能は不要である。特に、演算結果を高速に取り出したり、ボード上で複数の FPGA が通信したりする必要がない。このために、比較的簡単なハードウェアで高い性能を実現できる。

昨年 11 月に、プロトタイプシステムである PROGRAPE-1 を完成させた。現在応用プログラムを実装する作業中である。プロジェクトの目標としては、以下のようなものを想定している。

- 単体チップとしては、1999 年頃に利用可能になると思われる公称ゲート数で 40 万程度のものを用い、75MHz 程度のクロックで動作させる。
- システム全体としては、このチップを 1024 個並列に動作させる

主要部の演算精度に単精度を要求した場合で、チップ単体の性能は 1Gflops、システム全体の性能は 1Tflops 程度となる。これはもっとも悪い場合で、例えば実効的な精度が 12 ビット程度で良ければ 4Tflops、さらに 8 ビット程度なら 10Tflops 以上の速度が期待できる。

システム構成としては、重力／クーロン用計算機と同様に各プロセッサチップがメモリユニットを持つ構成とし、プロセッサボードやホストとの接続インターフェースのハード・ソフトについては重力／クーロン用計算機と完全に同一のものを利用することで開発の手間を押える。

3.4 開発計画

予定としては、これから 1 年程度のうちに重力プロセッサチップを完成させる。全体システムとして完成するのは 2000 年の予定である。

参考文献

[APB97] Athanassoula E., Puerari I., and Bosma A. (1997) Formation of rings in galactic disks by infalling small companions. *Monthly Notices of Royal Astronomical Society* 286: 284–302.

[BH86] Barnes J. and Hut P. (1986) A hiearchical o(nlogn) force calculation algorithm. *Nature* 324: 446–449.

[BHER95] Board J. A. J., Hakura Z. S., Elliott W. D., and Rankin W. T. (1995) Scalable variants of multipole-based algorithms for molecular dynamics applications. In Bailey D. H., Bjorstad P. E., Gilbert J. R., Mascagni M. V., Schreiber R. S., Simon H. D., Torczon V. J., and Watson L. T. (eds) *Proceedings of the Seventh SIAM Conference on Parallel Processing for Scientific Computing*, pages 295–300. SIAM, Philadelphia.

[BSO95] Brieu P. P., Summers F. J., and Ostriker J. P. (1995) Cosmological simulations using special purpose computers: Implementing p 3m on grape. *The Astrophysical Journal* 453: 566–573.

[Dub96] Dubinski J. (October 1996) A parallel tree code. *New Astronomy, vol. 1, no. 2, p. 133-147.* 1: 133–147. Provided by the NASA Astrophysics Data System.

[HO92] Hernquist L. and Ostriker J. P. (1992) A self-consistent field method for galactic dynamics. *The Astrophysical Journal* 386: 375–397.

[KFTM97] Kawai A., Fukishige T., Taiji M., and Makino J. (1997) The pci interface for grape systems: Pci-hib. *Publications of the Astronomical Society of Japan* 49: 607–618.

[Mak91] Makino J. (1991) Treecode with a special-purpose processor. *Publications of the Astronomical Society of Japan* 43: 621–638.

[MTES97] Makino J., Taiji M., Ebisuzaki T., and Sugimoto D. (1997) Grape-4: A massively parallel special-purpose computer for collisional n-body simulations. *The Astrophysical Journal* 480: 432+.

[OME$^+$93] Okumura S. K., Makino J., Ebisuzaki T., Fukushige T., Ito T., Sugimoto D., Hashimoto E., Tomida K., and Miyakawa N. (1993) Highly parallelized special-purpose computer, grape-3. *Publications of the Astronomical Society of Japan* 45: 329–338.