

氏 名 小池 邦昭

学位（専攻分野） 博士（学術）

学位記番号 総研大甲第 1319 号

学位授与の日付 平成 22 年 3 月 24 日

学位授与の要件 物理科学研究科 天文科学専攻  
学位規則第 6 条第 1 項該当

学位論文題目 Development of Control Processor for Multipurpose  
Computer GRAPE-DR and its Performance Evaluation

論文審査委員 主 査 教授 近田 義広  
教授 吉田 春夫  
教授 富阪 幸治  
教授 坂上 仁志  
教授 梅村 雅之（筑波大学）

## 論文内容の要旨

本論文では超並列汎用計算機 GRAPE-DR の制御回路(GRAPE-DR 制御プロセッサ)の開発およびそれを用いて実行した重力多体問題シミュレーションと密行列の LU 分解の性能について述べる。論文申請者は GRAPE-DR 用の制御モデルを考案し、それを実回路として GRAPE-DR ボード上に実装した。この制御モデルでは制御プロセッサ側で GRAPE-DR プロセッサへの命令投入・データ転送・結果回収、さらにはホストへの転送およびホストからのデータ転送をすべて並行に行い、パイプライン的に行うことを可能にした。これは従来の付加的なプロセッサでは実現されていなかった。この並行動作を行うために、制御回路では単一のマイクロプロセッサではなく、各部に専用の制御回路を実装し、これらを並列に動作させる。これにより GRAPE-DR プロセッサとのインタフェース周波数が 200MHz となる低速の動作で、GRAPE-DR プロセッサとの通信において総計 6.775GB/s のバンド幅のデータフローを処理することに成功した。

重力相互作用は到達距離が無限大であるため、重力多体問題のシミュレーションでは単純な方法では計算量が粒子数の 2 乗に比例して増大するという問題がある。もちろん、ツリー法等で計算量のオーダーを  $O(N \log N)$  程度にできるが、それでも計算量が非常に多い。重力計算部分を高速に計算するハードウェアである GRAvity piPE(GRAPE)を汎用の計算機(ホスト計算機)に接続し、これを用いて重力の計算時間を短縮するという方法は非常に有効であり、広く用いられてきた。

しかし、GRAPE のような専用計算機を製作するためには計算を行うための専用 LSI を開発する必要がある。この開発コストが製造プロセスの微細化に従って増大したため、新規の開発は困難になっている。一方、重力相互作用には Single Instruction Multi Data (SIMD) 計算機も高い性能を実現できる。SIMD 型計算機は多数の演算器を同じ命令で実行する計算機であり、現在一般的になっている MIMD 型計算機に比べて同じトランジスタ数で多数の演算器を実現可能なため、1990 年頃まではいくつかの成功例があった。

しかしながらこのような SIMD 型計算機は 1990 年代以降新規開発がない。これは、SIMD 型計算機における通信の帯域幅の問題のためである。1990 年代以降、多数の浮動小数点演算器が 1 チップの LSI に集積できるようになったため、その多数の演算器にデータを供給するための外部メモリとの通信帯域が不足するようになった。また多数の演算器を 1 チップに集積して演算性能が向上すると、チップ間通信に必要な帯域幅も増大する。このために現状では 1990 年頃まで開発されていたようなアーキテクチャの SIMD 型の計算機は開発が困難になっている。

重力相互作用で相互作用している多体系の相互作用計算では多数の粒子から多数の粒子への力を計算するので、外部メモリへの必要な通信速度を計算速度に比べて小さくできる。また密行列演算も計算量は行列のサイズの 3 乗程に対してデータ量は 2 乗なので、計算量に比べて通信量を小さくできる。このような通信量に対して計算量が大きい応用に特化したアーキテクチャにすれば、必要な外部メモリへの通信帯域を削減できる。

このような形で SIMD 型の計算機を改良し、これを GRAPE と同じように重力計算等に使うのが GRAPE-DR の基本的なアイデアである。このアイデアに基づいて SING プロセッサチップが開発された。この SING プロセッサチップは演算に 512 個の簡略化した

プロセッサコアを集積し、500MHz 駆動で単精度で 512Gflops、倍精度で 256Gflops の性能値を持つ。各演算器はスループット 1 の倍精度浮動小数点加算器とスループット 1/2 の倍精度浮動小数点乗算器(単精度乗算時のスループットは 1 となる)、32 ワードのレジスタファイル、256 ワードのローカルメモリ、整数 ALU から構成されている。SING プロセッサチップは命令実行サイクルごとに外部バスから命令を供給されて動作する。また、データの入出力ポートも単純な同期インタフェースであり、メモリや I/O バスを直結できない。

GRAPE-DR プロセッサボードは SING プロセッサ、FPGA、オンボードメモリ(DDR2-SDRAM)から構成され、SING プロセッサの制御回路は FPGA 上に実装される。制御プロセッサは SING プロセッサへのデータ入出力インタフェース、オンボードメモリインタフェース、ホストインタフェースを搭載して SING プロセッサと外部メモリ及びホスト計算機とのデータのやり取りを制御する。そのために、制御プロセッサは命令コントローラ、外部メモリインタフェース、ホスト側と SING プロセッサ側とのデータ変換回路および SING プロセッサ間の結果縮約回路から構成されている。命令コントローラはホスト計算機からの計算起動指令で SING プロセッサへ送出する命令列を生成する。外部メモリインタフェースはオンボードメモリへの書き込みおよび読み込みを制御する。データ変換回路は、ホスト側の 64 ビットのデータを SING プロセッサ側の 72 ビットデータに変換し、同様に逆の変換も行う。プロセッサ間結果縮約回路はプロセッサボード上の 4 つの SING プロセッサの出力した結果を縮約して、結果回収に必要なデータ量を削減するとともに PCI Express リンクの帯域幅を有効に利用できるようにする。ホスト側とのインタフェース回路と高速シリアルリンクについては既存の回路を使用した。

重力相互作用ライブラリは Hermite 積分法に対応した重力相互作用計算を SING プロセッサ上で計算するためのホスト計算機側のライブラリである。SING プロセッサの動作周波数が 333MHz、粒子数が  $N=262144$  の場合に重力相互作用での演算性能は 377Gflops となった。これは通信を無視した場合の漸近性能の 88%となる。

密行列のブロック化 LU 分解は行列積が計算時間の大部分を占めるため、行列積の部分を SING プロセッサ上で実行することにより、密行列の LU 分解をホスト計算機のみで実行する場合に比べて大幅に加速することができる。SING プロセッサの動作周波数が 400MHz の場合で、密行列の乗算性能は行列サイズが 34816 行 2048 列の行列と 2048 行 34816 行の行列積の場合に、635Gflops と理論ピーク性能の 79.3%の性能値となった。行列データの転送と計算実行を並行して行えるようにしたため、行列乗算では比較的小さな行列でも高い実効性能を出すことに成功した。SING プロセッサ用に最適化した密行列乗算ライブラリを用いて LINPACK ベンチマークの性能評価を行った。その結果、LU 分解全体では 433Gflops と理論ピークの 54%の性能値となった。

このように重力多体問題および密行列 LU 分解といった実際のアプリケーションで高い理論ピーク性能に対する実効性能を達成し、準汎用型のアクセラレータの実用性を実証した。

大規模数値計算向けの計算機の速度は、過去 60 年にわたってほぼ 10 年で 100 倍という指数関数的な発展をしてきた。しかし、必ずしも半導体技術の進歩を計算機の速度向上に有効に利用できていたわけではない。最近 20 年ほどの主力のアーキテクチャはマイクロプロセッサを多数並べて MIMD (Multi Instruction Multi Data)型、すなわちプロセッサ毎に別々のプログラムを実行する方式の分散メモリ並列計算機を構築するものであったが、このようなアプローチではプロセッサチップ上のトランジスタのうち演算器に利用されている部分の割合は極めてわずかであり、また年々低下してきている。

SIMD (Single Instruction Multi Data)型の超並列計算機では、多数の演算ユニットが単一の制御部を共有するためにトランジスタ利用効率を大きく向上させることができる。このため、1980 年代まではいくつかの成功例があったが、90 年代に入って、複数の演算器を単一の LSI チップに集積できるようになると、MIMD 型に対する相対的な利点であった高い集積度を有効に利用することが困難になった。これは、当時の SIMD 型超並列計算機的设计で前提になっていた、高いバンド幅でのローカルメモリや隣接プロセッサとの通信が実現困難になったためである。

SIMD 方式とは違うアプローチでトランジスタ利用効率の向上を実現してきたのは、銀河のように非常に多数の質点からなる系の重力的自己発展を計算するために開発された GRAPE に代表される専用計算機である。GRAPE の場合、2 粒子間の相互作用を専用パイプラインで計算することで、SIMD 型並列計算機に比べてもトランジスタ利用効率を大きく向上させている。さらに、演算結果をバケツリレーするパイプライン方式にすることでメモリバンド幅への要求が下がり、また複数のパイプラインが入力データを共有することでさらにメモリバンド幅への要求を下げ、非常に多数の演算器を 1 チップに集積し、しかも高い実行効率を実現することに成功してきた。

このような背景のもと、専用計算機の高い効率を生かしつつ、プログラム可能なアーキテクチャにすることで応用範囲を広げることを目指して開発されたのが GRAPE-DR プロセッサである。出願者は、この GRAPE-DR プロセッサと、外部メモリ、ホスト計算機とのインターフェースとなり、また命令を供給する制御プロセッサの開発をおこなった。GRAPE-DR プロセッサは 512 個の SIMD 並列動作するプロセッサを内蔵し、400MHz 動作時に単精度 800 Gflops、倍精度 400 Gflops のピーク性能を持つ。

しかし、プロセッサ内部にプログラムカウンタや命令シーケンサをもち、命令を実行サイクル毎に供給される必要がある。また、データ通信のためのインターフェースは単純なプロトコルのものであり、外部メモリ等に直結できるようになっていない。このため、外付けの DRAM メモリへのインターフェース、命令供給のためのシーケンサ、ホスト計算機への PCI-Express インターフェース等をまとめた制御プロセッサを FPGA (Field Programmable Gate Array: ユーザの手許で内部論理回路を定義・変更できる集積回路)

によって実現してシステムを構築している。出願者は、この FPGA の内部回路の開発と、それを使った実行ライブラリを開発を行ない、これらの性能評価を重力計算と密行列計算についておこなった。

出願者が開発した制御プロセッサの特徴は、ホストからのデータ転送、DRAM から GRAPE-DR プロセッサへのデータ転送、実際の命令実行、GRAPE-DR プロセッサからのデータ出力とホストへの転送といった複数のオペレーションが並列実行できるように設計されていることである。密行列の乗算ではこの並列実行の効果は非常に大きく、理論的には乗算する行列の両方のサイズが大きい極限で 2 倍、片方のみが大きい場合には最大 3 倍の性能向上が可能である。このような並列動作を可能にするため、内蔵プロセッサによるソフトウェア制御方式をとらず、上の機能毎にシーケンサを実装し、それらを協調動作させることで並行動作を実現している。このような、ハードウェア化した制御プロセッサにより、PCI-Express を通したデータ転送ではパケットヘッダやペイロード長から決まる理論限界に近い転送速度を実現し、また一度の命令起動のオーバーヘッドも 3 マイクロ秒程度と極めて小さくできている。GRAPE-DR と類似の SIMD 型加速ボードである GPGPU(General-Purpose computation on Graphics Processing Units) では典型的には命令起動(GPGPU 側での関数呼び出し) のオーバーヘッドは数十ミリ秒と極めて大きく、これが様々なアプリケーションの実効性能を制限しているが、出願者が開発した制御プロセッサによって GRAPE-DR ではアプリケーションの実行効率を極めて高いものにできている。例えば行列乗算の場合、理論ピークの 80% 以上と高い効率を実現した。これにより、連立一次方程式の求解のベンチマークである LINPACK で使われる LU 分解で、430 Gflops 以上と理論ピークの 50% 以上、またホスト計算機の速度の 10 倍以上と大きな加速に成功した。GPGPU や他の加速ボードで、LU 分解の速度をボードなしの場合の 3 倍以上に上げることに成功したのは本研究が初めてであり、大きな成果である。また、重力計算においても、ツリー法で 100 万粒子/秒と、加速ボード 1 枚当りの性能としては世界最高水準を実現した。

出願者は博士論文に以上の結果を示すと共に、他の方式との比較を行い、SIMD 方式の加速プロセッサが、ホスト計算機との通信と計算の並行実行によって高い効率を実現でき、アプリケーション全体を加速することができることを示した。これは、この分野の研究にとって十分意義のある成果と判定された。これにより審査委員会は、全員一致で本論文が博士論文として十分な価値を有し、合格であると判定した。