

研究種目：若手研究（B）

研究期間：2007～2008

課題番号：19700061

研究課題名（和文） GPUグリッドによる京速計算のための計算資源管理

研究課題名（英文） Resource Management for 10-Petaflops Computing on the GPU Grid

研究代表者

伊野 文彦（INO FUMIHIKO）

大阪大学・大学院情報科学研究科・准教授

研究者番号：90346172

研究成果の概要：本研究の目的は、GPU（Graphics Processing Unit）を装備する多数のパソコンを一つの仮想的な高性能計算機 GPU グリッドとして動作させるための資源管理手法を開発することである。開発した手法を用いた結果、日常業務に用いる計算機に GPU を1つ追加すれば、単体 CPU を2台ほど専用し続ける場合と同等の性能を引き出せることが分かった。

交付額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	1,900,000	0	1,900,000
2008年度	1,400,000	420,000	1,820,000
年度			
年度			
年度			
総計	3,300,000	420,000	3,720,000

研究分野：総合領域

科研費の分科・細目：情報学・計算機システム・ネットワーク

キーワード：GPU, グリッドコンピューティング

## 1. 研究開始当初の背景

GPU（Graphics Processing Unit）はパソコンの標準的な付属部品であり、ディスプレイに画面を出力するものである（図1）。GPUは、映画やゲームなどの娯楽分野における高い需要を背景に、半年で2倍の速さで急速に性能を向上している。この速さは、CPUの向上速度（1年半で2倍）を超えている。結果、最近のGPUは、CPUを追い越すだけでなく、桁違いの演算性能を達成している。例えば、nVIDIA社のGPU（GeForce 7800 GTX）は1秒間当たりの浮動小数点演算数が200 GFLOPS（2千億回）を超える。一方、Intel社のCPU（Pentium 4 3-GHz）は12 GFLOPSである。



図1：GPU（Graphics Processing Unit）

この魅力的な演算性能に加え、最近のGPUはプログラム可能な演算器を備え、より柔軟な汎用演算器として進化しつつある。結

果、従来のグラフィクス用途のみならず、多様な分野への応用が盛んである。例えば、物理シミュレーション、数値計算、データベース処理などにおいて、CPUよりも1~3桁ほど高い性能を達成している。

## 2. 研究の目的

研究の全体構想は、この成長著しいGPUの将来性に着目し、10 PFLOPS (=10\*1000\*1000 GFLOPS) 級の京速計算を達成しうる次世代スーパーコンピュータ(京速計算機)を実現し、誰もが容易に利用できる高性能計算基盤を構築することである。具体的には、グリッド技術を駆使して、多数のGPUをインターネット上で統合し、それらを遠隔地から利用するための仕組みを提供する(図2)。ユーザは、自らの計算資源を提供する代わりに、グリッドの膨大な計算資源を共有でき、高速計算を容易に実現できる。

この構想の中での本研究の目的は、GPUを装備する多数のパソコンを1台の仮想的な高性能計算機GPUグリッドとして動作させるための資源管理手法を開発することである。さらに、京速計算を実現しうる計算基盤としてGPUグリッドの有用性を実用的な応用とともに示すことを目指した。

特に、ユーザ間でGPUを共有するGPUグリッド固有の技術的な課題を解決することに研究の焦点がある。具体的には、以下に挙げる2点を実現する。

(1) GPUを装備する多数のパソコンを一つの仮想的な高性能計算機GPUグリッドとして動作させるための資源管理手法を開発する。

(2) 京速計算を達成しうる計算基盤としてGPUグリッドの有用性を示す。具体的には、実用的な応用をGPUグリッド上に実装し、CPU版に対する高速化の度合いを検証する。

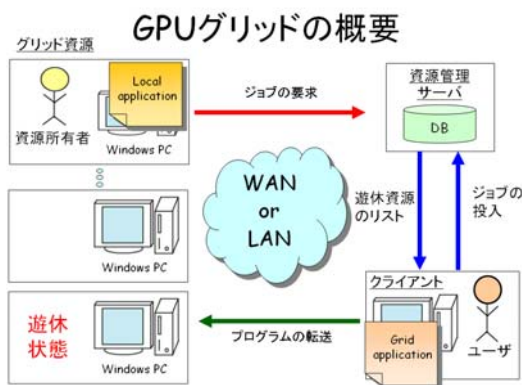


図 2 : GPU グリッドの概要

## 3. 研究の方法

GPUを装備するパソコンを用い、以下に挙げる方法で資源管理手法および応用の開発に取り組んだ。

(1) 計算資源の競合を低オーバーヘッドで解決できる資源管理手法を開発した。本来、GPUは画面描画処理に専有されることを前提にしており、複数のユーザが共有することを考慮しない。したがって、遊休か否かを監視する処理自体がアプリケーションの大幅な性能低下を引き起こしてしまう。

そこで、開発した手法はGPUの稼働状況を直接観測するのではなく、間接的に推測することで性能低下を回避する(図3)。本手法は、インターネットに接続するパソコン上でスクリーンセーバとして動作する。パソコンに対して一定時間に渡り入力がない場合、スクリーンセーバがビデオメモリの使用量を調べることによりGPUの稼働状況を推測する。

この仕組みにより、共有されることを想定していない現在のGPUにおいても、パソコン所有者に対して外乱の小さい資源監視を実現できる。さらに、GPUだけでなくCPU使用率を監視することにより、資源全体で遊休なものを選択できる。

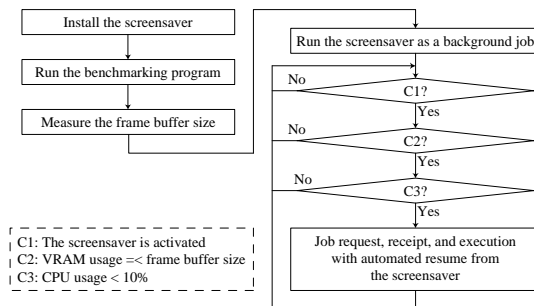


図 3 : 資源監視の手順

(2) 遊休GPUを持つパソコンを検出し、それらを用いてグリッド計算を行うための分散並列システムを開発した。開発したシステムは、オペレーティングシステムとしてWindows XPを前提とし、資源管理サーバおよびグリッド資源からなる(図2)。各々はインターネットで相互接続されている。サーバは、上記で述べた資源管理手法を基に、遊休状態にある資源の一覧を保持する。この際、各資源が装備するGPUの仕様やドライバのバージョンなども資源情報として収集する。

これらの静的な情報に加えて、動的な情報をもとに、ユーザは計算に必要とする資源の特徴を指定できる。例えば、GPUとしてnVIDIA GeForce 8800 GTXを持ち、CPUの駆動周波数が2 GHz以上であるもの、さらにドライバのバージョンが69.14以上のものなどという具

合に指定できる。

なお、スクリーンセーバのインストール時にベンチマークプログラムを実行し、実効性能情報を収集しているため、具体的なGPUを指定することなく、単にフィルレートなどの性能を指定することも可能である。

(3) 具体的な応用として、生物データベースに対して共通部分を探し出す配列アライメントの高速化に取り組んだ。データベースのサイズが加速度的に増大しつつあり、高速化の要望が強い応用の一つである。

標準的なアライメントアルゴリズムとしてSmith-Watermanアルゴリズムに着目し(図4)、OpenGLグラフィックスライブラリを用いてGPU上に実装した。このアルゴリズムは動的計画法によるものであり、並列化は容易でない。そこで、開発した手法では、アライメントが多数の配列間で実行されることに着目し、いくつかの組み合わせを同時に処理することで、高い並列化効率を得られるよう工夫している。配列データはテクスチャと呼ばれるグラフィックスのためのデータ構造に格納し、それらを参照しながら画面描画を繰り返すことで計算結果を得る。

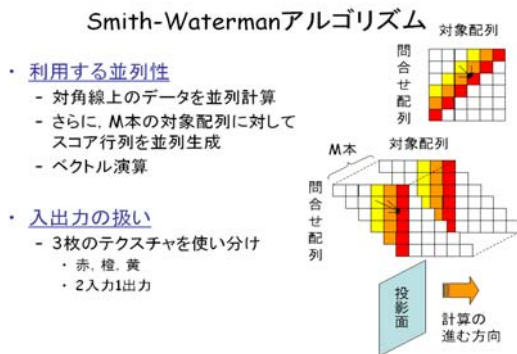


図 4 : Smith-Waterman アルゴリズム

#### 4. 研究成果

本研究における主な成果は、以下に挙げる3点に要約できる。

(1) 開発した手法により、パソコン所有者に対する外乱の小さい資源監視を実現した。資源監視のオーバーヘッドは260ミリ秒程度であり、そのうち200ミリ秒弱はスクリーンセーバの起動に要している。一方、ビデオメモリの使用量は2ミリ秒程度で確認できている。短時間で確認できている理由は、ドライバがその使用量を管理しているため、GPUを介する必要がないことにある。また、遊休GPUを持つパソコンを検出し、それらを用いてグリッド計算を行うための分散並列システムを開発した。

#### 実効性能の分析



図 5 : 配列アライメントの実効性能

(2) 単体のCPU上で5時間を要するアライメントに対し、単体のGPUを用いておよそ1時間に短縮できた。このとき、計算の核となる部分の実効性能はnVIDIA GeForce 8800 GTX上で47 GFLOPSであり、これは8台のCPU (Intel Core 2 Duo 2.4 GHz) で得られる実効性能に匹敵する(図5)。

また、実効性能がビデオメモリバンド幅により抑えられていることを示した。仕様上の理論性能が86.4 GB/sであるのに対し、開発した手法は、その値を上回る実測性能を達成できているため、効率が高い実装といえる。理論性能を上回った理由は、テクスチャキャッシュによる貢献があったためである。

(3) 4台構成のGPUグリッドを用いてアライメントを並列実行した。各々のパソコンは100 Mb/sのイーサネットに接続されている。また、パソコンの所有者はGPUに関する研究開発を遂行している大学院生である。この結果、日常業務に用いるパソコンにGPUを1つ追加すれば、単体CPUを2台ほど専用し続ける場合と同等の性能を引き出せることが分かった。

以上の研究成果により、複数のユーザがGPUを共有できる実行環境を整えることができた。特に、GPUの遊休状態をサーバ側で積極的に監視し、適切な資源を選択できるシステムは、我々が知る限り初めてである。この成果は、平成20年度情報処理学会山下記念研究賞として選出されている。

今後の展望としては、より大規模かつ広域に分散した環境における評価が挙げられる。現在のシステムは、会社内あるいは学内などのネットワーク遅延の短い環境で有用性を示している。しかし、上記のような遅延の長い環境においては、ネットワーク遅延の長さがGPUによる高速化の効果を低減する可能性がある。解決策としてグリッドジョブの大きさを調節する仕組みなど必要である。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計10件)

- ① Fumihiko Ino、Yuki Kotani、Kenichi Hagihara、Harnessing the Power of Idle GPUs for Acceleration of Biological Sequence Alignment、Proceedings of the 2nd Workshop on Large-Scale Parallel Processing、8 pages (CD-ROM)、(2009)、有
- ② Fumihiko Ino、Seiji Yoshida、Kenichi Hagihara、RGBA Packing for Fast Cone Beam Reconstruction on the GPU、Proceedings of the SPIE Medical Imaging、8 pages (CD-ROM)、(2009)、有
- ③ Yusuke Okitsu、Fumihiko Ino、Kenichi Hagihara、Fast Cone Beam Reconstruction Using the CUDA-enabled GPU、Proceedings of 15th International Conference on High Performance Computing、pp.108-119、(2008)、有
- ④ Tomohiro Okuyama、Fumihiko Ino、Kenichi Hagihara、A Task Parallel Algorithm for Computing the Costs of All-Pairs Shortest Paths on the CUDA-compatible GPU、Proceedings of the 6th International Symposium on Parallel and Distributed Processing and Applications、pp.284-291、(2008)、有
- ⑤ Yuki Kotani、Fumihiko Ino、Kenichi Hagihara、A Resource Selection System for Cycle Stealing in GPU Grids、Journal of Grid Computing、6、pp.399-416、(2008)、有
- ⑥ Yuma Munekawa、Fumihiko Ino、Kenichi Hagihara、Design and Implementation of the Smith-Waterman Algorithm on the CUDA-Compatible GPU、Proceedings of the 8th IEEE International Conference on Bioinformatics and Bioengineering、6 pages (CD-ROM)、(2008)、有
- ⑦ 吉田征司、伊野文彦、西野和義、萩原兼一、GPUによる高速なコーンビーム再構成:円軌道装置のためのRGBAデータへの詰め込み、情報処理学会論文誌:コンピューティングシステム、1、pp.41-53、(2008)、有
- ⑧ Daisuke Nagayasu、Fumihiko Ino、Kenichi Hagihara、A Decompression Pipeline for Accelerating Out-of-Core Volume Rendering of Time-Varying Data、Computers and Graphics、32、pp.350-362、(2008)、有
- ⑨ 川崎康博、伊野文彦、萩原兼一、スケール不変な格子を生成する適応格子細分化

法アプリケーションのための計算量予測手法、電子情報通信学会論文誌、J90-D、pp.2691-2703、(2007)、有

- ⑩ 伊藤信悟、伊野文彦、萩原兼一、GPGPUアプリケーションの開発を支援するための性能モデル、情報処理学会論文誌:コンピューティングシステム、48、pp.235-246、(2007)、有

[学会発表] (計4件)

- ① 荻田章博、伊野文彦、萩原兼一、GPUグリッドにおいて描画および科学計算を並行処理するための制御手法、情報処理学会第110回システムソフトウェアとオペレーティング・システム研究発表会、2009年1月28日、蒲郡情報ネットワークセンター
- ② Fumihiko Ino、GPU-accelerated volume reconstruction、Joint JSPS-SNSF Seminar on Computer Aided Surgery、2007年9月24日、大阪大学
- ③ 伊野文彦、汎用アクセラレータとしてGPUを駆使する試み GPGPU について、サイエントیفック・システム研究会 HPC フォーラム 2007、2007年8月28日、汐留シティセンター
- ④ 伊野文彦、小谷裕基、萩原兼一、GPUグリッドによる高速な塩基配列アライメント、情報処理学会第111回ハイパフォーマンスコンピューティング研究発表会、2007年8月1日、旭川国際会議場

[その他]

ホームページ等

<http://www-hagi.ist.osaka-u.ac.jp/~ino/>

## 6. 研究組織

### (1) 研究代表者

伊野 文彦 (INO FUMIHIKO)

大阪大学・大学院情報科学研究科・准教授  
研究者番号: 90346172

### (2) 研究分担者

### (3) 連携研究者