4011：*504*（GPU；CUDA；メソスケール・モデル）

GPU コンピューティング

1．GPU コンピューティングとは

パソコン等に装着するグラフィクス・ボードに搭載される画像処理プロセッサ GPU（Graphics Processing Unit）は，より高速に，より美しく，より精細な画像を表示することを求められたために描画機能が飛躍的に向上し，GPU を画像表示だけではなく汎用計算に使う試み「GPU コンピューティング」（GPGPU（General-Purpose computing on GPUs）と同義）が2000年頃から始まった．2006年に NVIDIA 社が自社の GPU に対して GPU コンピューティング用の統合開発環境として CUDA（Compute Unified Device Architecture）をリリースしたことが大きな節目となる（NVIDIA 2010；青木・額田 2009）．それまでは Cg（C for Graphics）言語や HLSL（High Level Shader Language）により画像処理の機能を汎用計算に置き換えてプログラミングする必要があったが，CUDA の登場により C 言語でプログラミングすることができるようになり，一気に GPU コンピューティングが広まり始めた．さらに2009年には CUDA の FORTRAN 版もリリースされ，GPU の性能向上に合わせた頻繁なバージョン・アップとともに機能もますます向上している．2009年末には特定の GPU だけでなく，AMD 社，Intel 社の CPU，GPU や Cell などでも同じプログラムが動作するような標準化として策定された Open CL も正式にリリースされ，より GPU コンピューティングが普及する環境が整ってきている．

GPU はポリゴンに対する幾何学的描画処理を高速化することに特化して開発されてきたため，単体で 1 TFLOPS を超えるようなピーク演算性能を持つ．また，GPU はパソコンのグラフィクス市場で展開する製品であるため，コストが極めて低く，手元のパソコンにも装着できる，という大きな特徴がある．さらに，GPU は消費電力当たりの演算性能が高いため，スパコンの高性能化と低消費電力化に向けたアクセラレータとして広く認識されるようになっている．2008年には東京工業大学学術国際情報センターが680個の GPU をスパコン TSUBAME 1.2に導入し，世界的に大きな注目を集めた．NVIDIA 社は2010年に倍精度浮動小数点演算性能の向上や ECC メモリに対応した Fermi コアを搭載した GPU（第 1 図）をリリースしており，GPU を HPC（High Performance Computing）の分野で利用する条件が整った．2010年11月のスパコン Top500のランキングでは，1 位と 3 位に中国の GPU スパコンが入り，4 位に総合演算性能2.4 PFLOPS の東京工業大学学術国際情報センターの TSUBAME 2.0（東京工業大学 2011；第 2 図）がランクインするなど，GPU マシンがスパコンの上位を独占している．

初期の頃の GPU コンピューティングは重力多体問題への適用で注目を集めた．GRAPE などの専用計算機と同様に演算負荷の高い部分に GPU をアクセラレータとして利用し，高い実行性能を引き出すことができた．GPU が ClearSpeed や GRAPE など従来のアクセラレータと大きく違う点は，広いメモリバンド幅を持つところである．そのため，GPU コンピューティングは限定された分野での計算だけではなく，物理，化学，金融，データ処理等の様々な分野のアプリケーションへの適用が進んでいる．

2．GPU のアーキテクチャとプログラミング・モデル

GPU のアーキテクチャは汎用 CPU と異なり，NVIDIA 社の2010年の最新 GPU では 1 チップ当たり500個以上の演算プロセッサ（CUDA コア）が搭載されている．さらに 8～32個の CUDA コアが一つのストリーミング・マルチプロセッサを構成していて，そこには L1キャッシュと共有メモリがある．GPU を搭載したボード上にビデオ・メモリがあり，GPU からビデオ・メモリへ100 GB/sec を超える高速なアクセスが可能なことやメモリが階層的構造になっている点にも特徴がある．

第 1 図　NVIDIA 社製 GPU 搭載グラフィクス・ボード．

第 2 図　東京工業大学学術国際情報センターのスパコン TSUBAME 2.0．

GPU コンピューティングでは，多数の演算プロセッサを効率的に使うようにプログラミングすることが高い実行性能を達成するために必須である．このため，データ並列性の高い問題に対して GPU コンピューティングを適用することが重要である．マルチコアの CPU ではコア数とほぼ同数のスレッドを実行させることが多いが，GPU の場合は高速にスレッドを切り替えるハードウェア・コントローラーが多数搭載されているため，数百個の演算ユニットに対して数万以上の数のスレッドを実行しても切り替えのオーバーヘッドがなく，このような多数のスレッドで計算することにより始めて十分な性能を引き出すことができる．GPU のプログラミングはストリーミング・マルチプロセッサ単位では SPMD（Single Program Multiple Data）であり，その中では32個のスレッド毎に SIMD（Single Instruction Multiple Data）が実行される．このようなアーキテクチャを前提にプログラミングすることにより，高い実行性能を引き出すことができる．

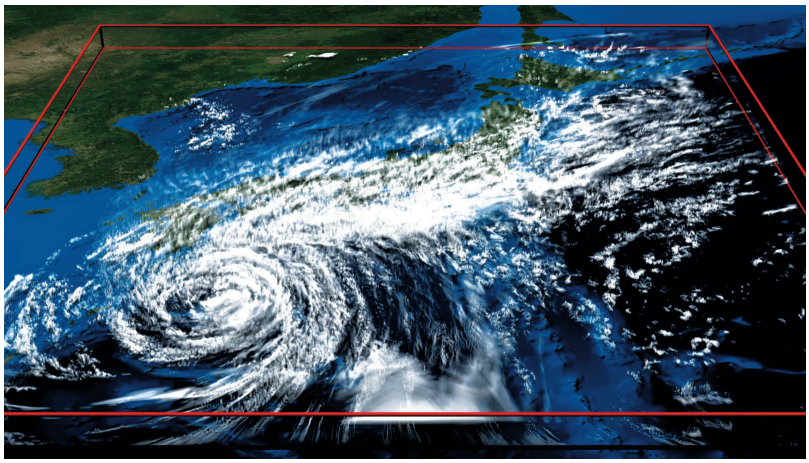
3．GPU コンピューティングによる気象計算

気象計算はスパコンを利用する代表的な大規模計算の一つである．特に非静力のメソスケール・モデルでは，雲を解像する格子間隔・格子点数が必要となり，より大規模な計算が必要となっている．できるだけ短時間で予報計算を終了させる目的があるため，WRF（Skamarock *et al*. 2008）の開発グループではいち早く GPU を利用する取り組みが行われている．WRF では物理過程における計算負荷の高いモジュールの一部を GPU に移植し高速化を図った（Michalakes and Vachharajani 2008）．計算全体は従来通り CPU 上で実行し，GPU 化したモジュールの部分を計算する際には，先に CPU のメモリから GPU のメモリに計算に必要なデータを転送しておく．GPU で計算した結果は GPU のメモリに出力されるため，CPU で計算を続行させるためには GPU から CPU のメモリへのデータ転送も必要になる．WRF の時間積分の毎ステップでこのような CPU-GPU 間の通信が発生し，この通信時間がボトルネックとなる．GPU に移植したモジュール単体では CPU の20倍の高速化に成功しているが，計算全体では30％程度の速度向上に留まり，GPU の持つ本来の性能を十分に発揮できない結果となった（Michalakes and Vachharajani 2008）．GPU コンピューティングにより高い実行性能を達成するには，可能な限り CPU-GPU 間の通信を排除する必要がある．そのためには，気象モデルの時間積分ループの中の全てのサブルーチン（関数）を GPU 化する必要がある．

東京工業大学学術国際情報センターの青木グループは気象庁数値予報課と協力し，次期気象予報モデルとして開発している ASUCA（Ishida *et al*. 2010）の物理過程と力学過程の全てのモジュールの GPU 化を行った．GPU 化するためには ASUCA のコード全体を一から CUDA に書き直す必要がある．物理過程の各モジュールは単体の GPU 化が可能であり，移植は比較的容易である．一方，力学過程の計算は隣接格子点へのアクセスを伴うため，予報変数は終始 GPU ボードのビデオ・メモリ上に確保しておき，一度に GPU 化する必要がある．力学過程の計算は圧縮性流体力学の方程式に基づいているため演算よりメモリアクセスが支配的である．GPU のビデオ・メモリへのアクセスは CPU のメインメモリへのアクセスと比較すると数倍以上高速であるため，GPU は力学過程においても CPU より十分高速な計算が期待できる．

青木グループは FORTRAN プログラミング言語で記述された ASUCA のコードを一旦 C/C＋＋言語に書き直し，その後，CUDA に書き換えている．さら

にストリーミング・マルチプロセッサ内の共有メモリをキャッシュ的に使うアルゴリズムやレジスタを有効利用するなどの多数の技法を導入し，最終的に Intel CPU Xeon X5670の１ソケット（６コア）に対して，NVIDIA GPU Tesla M2050の１ソケット（448 CUDA コア）が約12倍高速に計算できることを示している（Shimokawabe *et al.* 2010）．



第３図　水平解像度500 m 格子を用いて次世代メソスケール気象モデル ASUCA で計算した雲分布．

一つの問題点は GPU ボード上のビデオ・メモリはせいぜい数 GB であるため，実際の気象モデルを動作させるには複数 GPU を用いた大規模計算を行う必要がある．CPU で計算するときと同じように単体 GPU カードのメモリで計算可能なサイズにまで計算領域を分割し，それぞれを各 GPU に割り当てる．現時点の CUDA ではノードが異なる GPU のメモリ間で直接データ通信を行うことができず，CPU 側のメモリを介して通信する必要がある．大規模計算では，この GPU 間のデータ通信が大きなオーバーヘッドになるため，通信と計算のオーバーラップの技術も開発されている．これらにより，ASUCA の GPU 版は TSUBAME 2.0 の3990個の GPU を使って145 TFLOPS という非常に高い実行性能を達成している（Shimokawabe *et al.* 2011）．

第３図は気象庁メソモデル（MSM）のデータを初期条件・境界条件とし，TSUBAME 2.0 の437個の GPU を用いて GPU 版の ASUCA により4792×4696×48（水平500 m 格子）の格子で，初期時刻2009年10月６日15 UTC から計算した９時間後の雲の分布を可視化している．GPU コンピューティングにより，実運用を目指している気象予報モデルが CPU で計算するよりも10倍以上高速に計算でき，さらに10倍以上少ない消費電力で計算できることが示された．

参 考 文 献

青木尊之，額田　彰，2009：はじめての CUDA プログラミング．工学社，249pp.

Ishida, J., C. Muroi, K. Kawano and Y. Kitamura, 2010：Development of a new nonhydrostatic model ASUCA at JMA. CAS/JSC WGNE Research Activities in Atmospheric and Oceanic Modeling, (40), 5.11–5.12.

Michalakes, J. and M. Vachharajani, 2008：GPU acceleration of numerical weather prediction. Proceedings of the IEEE International Parallel and Distributed Processing Symposium, 1–8.

NVIDIA, 2010：CUDA Programming Guide 3.2. http://developer.download.nvidia.com/compute/cuda/3_2/toolkit/docs/CUDA_C_Programming_Guide.pdf (2011.4.30閲覧).

Shimokawabe, T., T. Aoki, C. Muroi, J. Ishida, K. Kawano, T. Endo, A. Nukada, N. Maruyama and S. Matsuoka, 2010：An 80-fold speedup, 15.0 TFlops full GPU acceleration of non-hydrostatic weather model ASUCA production code. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, New York, USA.

Shimokawabe, T., T. Aoki, J. Ishida, K. Kawano and C. Muroi, 2011：145 TFlops performance on 3990 GPUs of TSUBAME 2.0 supercomputer for an operational weather prediction. First International Workshop on Advances in High-Performance Computational Earth Sciences：Applications and Frameworks (IHPCES), Singapore.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. D. Duda, X.-Y. Huang, W. Wang and J. G. Powers, 2008：A Description of the Advanced

Research WRF Version 3. National Center for Atmospheric Research, 113pp.
東京工業大学，2011：TSUBAME 2.0 の仕様．http://www.gsic.titech.ac.jp/node/392（2011.4.30閲覧）．

（東京工業大学学術国際情報センター　青木尊之）