

10-NA12

GPGPU の地震ハザード予測シミュレーションへの適応性評価

青井 真 (独立行政法人 防災科学技術研究所)

概要 地震ハザードを予測するには高精度かつ大規模な地震波伝播シミュレーションを行う必要があり、実用的な計算においては格子数が数億から数十億に及ぶ極めて規模の大きなモデルとなることが多い。このように大きな計算機リソース (CPU パワー及びメモリ) を必要とする地震ハザード予測シミュレーションへの GPGPU (General Purpose Computation on Graphics Processing Unit) の適用性を評価した。

1. 研究の目的と意義

兵庫県南部地震を契機に、地震に関する調査研究の成果が国民や防災を担当する機関に十分に伝達され活用される体制になっていなかったという課題意識の下に、行政施策に直結すべき地震に関する調査研究の責任体制を明らかにし、これを政府として一元的に推進するため、地震防災対策特別措置法に基づき総理府に設置 (現・文部科学省に設置された) 機関として、地震調査研究推進本部が設置されている。地震調査研究推進本部の地震調査委員会において、地震に関する観測、測量、調査または研究を行う関係行政機関、大学等の調査結果を収集、整理、分析し、並びにこれに基づき総合的な評価を行っており、この中で地震動予測地図の高度化を進めている。日本周辺で発生する全ての大地震に関して、地震のリスク評価の基礎となり得る精度で地震ハザードを予測できるよう、手法・モデルの高度化を目指している。そのために、必要な精度、分解能を持つ地盤構造の開発を行うとともに、高精度かつ汎用性のある地震波伝播 (強震動) シミュレーション手法の開発に関する研究を行っている。

地震波伝播シミュレーションにおいて詳細な 3 次元地下構造を十分な精度で離散化し高精度な計算を行ったり、短周期の地震波まで計算するためには細かな格子が必要であるため、実用的な計算においては格子数が数億から数十億に及ぶ極めて規模の大きなモデルを扱う必要がある。近年の計算機環境の劇的な進歩の恩恵を受けたとはいえ、

大きな計算機リソース (CPU パワー及びメモリ) を必要とし、実務等で一般的に使用可能な計算機では数日以上計算となる事もしばしばである。このような状況を打開する手段として、多くの数値計算分野で広く使われるようになりつつある GPGPU (General Purpose Computation on Graphics Processing Unit) の利用が考えられる。

本研究では、GPU コンピューティングをスーパーコンピュータシステムに積極的に取り入れている TSUBAME 2.0 で地震波伝播シミュレーションを試行し、高度な地震ハザード・リスク評価への適応性を評価することを目的とする。

2. 当拠点公募型共同研究として実施した意義

地震の被害を軽減するためには、個々人の地震への防災意識を高め、地震に対する備えを促すことが不可欠である。このため、防災科学技術研究所では、日本全国で発生する地震を対象として、地震調査研究の成果の集大成である地震動予測地図を高度化し、地震ハザード・リスク評価に関する研究を行うとともに、WebGIS 等の技術を用いて、地震ハザード・リスク情報、地下構造データ等の関連情報を網羅的に提供可能な地震ハザード・リスク情報ステーションを構築している。

本研究では東京工業大学との超大規模数値計算応用分野での共同研究を行っており、これにより地震ハザード・リスク評価に用いられる地震波伝播シミュレーションに対する GPGPU に適用について、実務等に使用される汎用計算機環境から

TSUBAME 2.0 のような大規模なスーパーコンピュータシステムまでの様々な規模の計算機環境における評価が可能となる。

3. 研究成果の詳細と当初計画の達成状況

3.1. GMS による波動場の計算

本研究では、不連続な食い違い格子 (Aoi and Fujiwara, 1999) を用いた、空間四次・時間二次精度の差分演算子による実用コードである GMS (図 1: Ground Motion Simulator、青井・他、2004) をベースに GPGPU の適用性の評価を行なった。GMS は、防災科学技術研究所によってパッケージ化された、3次元有限差分法 (FDM) により地震波伝搬シミュレーションを行うためのツール群となっており、主に Fortran90 で書かれた差分計算ソルバはソースコードも公開している。

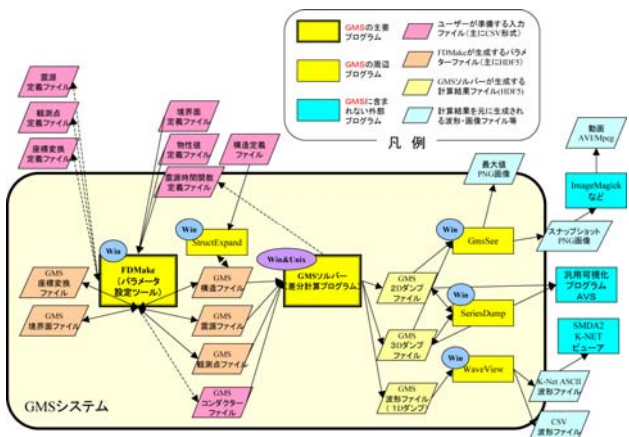


図 1 GMS (Ground Motion Simulator) の構成

差分法により地震波伝播シミュレーションを行う場合、均質な大きさの格子を用いる限り、格子サイズは計算すべき最低波長により決定される。低速度の媒質がモデルのごく一部にのみ存在する構造を扱う場合でも計算領域全体を小さな格子に分割せざるを得ず、大規模なモデル計算を行う際には大きな障害となる。実地震の再現計算や地震ハザード評価を行う際には、堆積平野構造を有するモデルなど地表に近いほど極端に地震波速度が遅い構造を取り扱うことが多いため、上記のような問題を回避するためには地震波速度構造に合わせた不均質な格子を用いることが有効である。

Aoi and Fujiwara (1998) や青井・他 (2004) は、大きさの異なる格子を組み合わせることにより効率的かつ高精度に計算を行うことの出来る不連続格子による差分法の定式化を提案した。図 2 に示すように、浅い部分 (領域 I) の格子点間隔は細かく、深い部分 (領域 II) の格子点間隔は領域 I の 3 倍の粗い格子点間隔を有する格子モデルを用いている。2つの領域はオーバーラップしており、波動場の連続性が保たれるよう内挿される。典型的な盆地構造モデルの計算において、均質な格子による場合と比較し数倍から十数倍程度効率がよいことが分かっている。

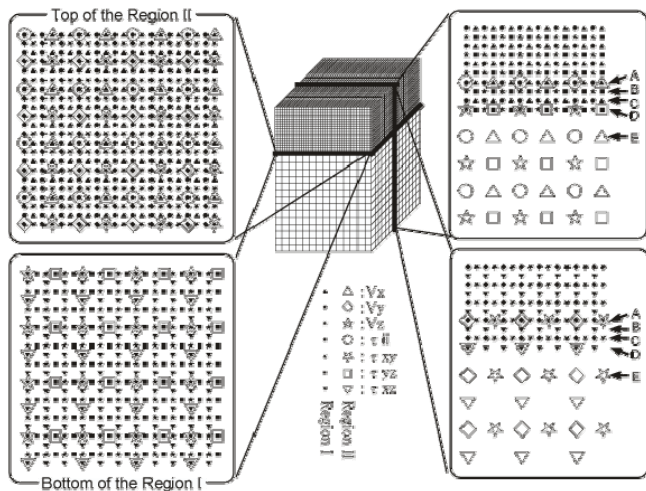


図 2 (中央) 計算に用いる不連続格子。(右) 不連続格子の垂直断面。領域 I と領域 II の接続部分で、内挿のために格子が重なっている。(左) 領域 II の最上面 (A 面) と領域 I の最下面 (D 面) における不連続格子の水平断面。

3.2. GPU での差分計算処理

GMS の差分計算ソルバを GPU での実行に対応させるため、NVIDIA 社から提供されている CUDA (Compute Unified Device Architecture) を使用し、差分法の計算処理についてはほぼ全て GPU で処理するように実装を行っている。CUDA では、演算処理を担当する GPU と、その演算処理に必要なデータを格納するビデオメモリ (グローバルメモリ) により構成されるデバイスをアクセラレータとして用いることを前提としている (図 3)。GPU

には単純な演算ユニットが多数搭載されており、これらを並列に稼働させることで極めて高い演算性能を実現している。差分法のように大量のメモリアクセスを伴う手法においては、GPU とグローバルメモリ間のメモリアクセスがボトルネックになることが多いため、GPU の高い演算性能を引き出すためには、効率的なメモリアクセスを行うことが重要となる。

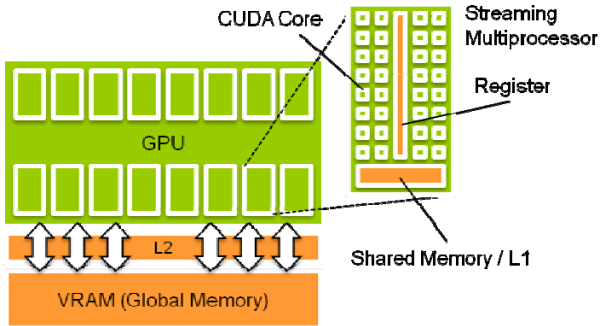


図 3 GPU のアーキテクチャ

本実装では、スレッド (GPU で計算処理を行う際の実行の最小単位) の集合であるブロックを 2 次元で構成し、計算の対象となる x-y 平面にブロックを敷き詰めるように配置している。そのため、生成されるスレッドの数は x-y 方向の格子数に対応することになる。ブロック内の各スレッドは z 方向の始点から終点まで 1 格子点ずつ計算処理を進め、3 次元領域全体の計算処理を行う (図 4)。差分法の計算処理では計算の対象となる格子のデータの他に隣り合う格子のデータが必要となるが、各スレッドが隣り合う格子のデータをグローバルメモリから読み込んだ場合、スレッド間では重複して同じデータが読み込まれることになる。グローバルメモリへのアクセスは GPU 内部に搭載されているレジスタやシェアードメモリと比べレイテンシが高く低速であるため、このような重複したデータの読み込みは性能低下の要因となる。そこで、各スレッドが担当する格子のデータをブロック内で共有される高速なシェアードメモリに読み込み、隣接する格子のデータを参照する際にはシェアードメモリに読み込んだデータを利用することによりグローバルメモリへのアクセスを軽減している。

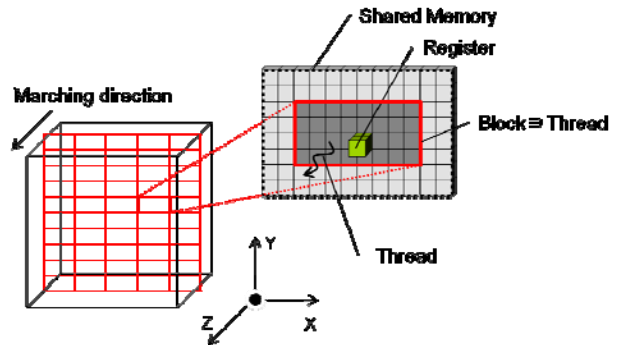


図 4 GPU での差分法計算処理

3.3. 複数 GPU を用いた並列化

複数 GPU を用いた並列計算を行う際は、モデルを水平 (x-y) 方向に分割し、各 GPU に分割された部分領域を割り付けて計算処理を行う (図 5)。各部分領域の境界面では隣接する部分領域のデータが必要となるため、袖領域 (重複領域) を設け、通信が必要なデータを MPI (Message-Passing Interface) 等を用いて交換する。

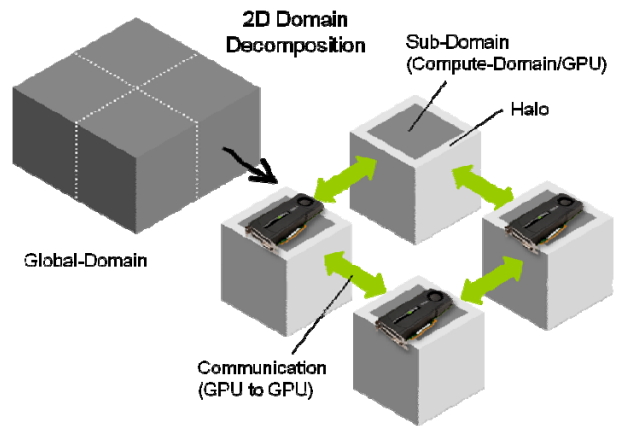


図 5 複数 GPU での並列計算 (水平方向の領域分割に並列化)

GPU の演算速度は極めて高速であるために相対的に通信に要する時間の割合が大きくなる。さらに、GPU 同士の直接の通信は出来ないために CPU を介した間接的な通信を行うことになるため、通信のオーバーヘッドは極めて大きい。従って複数 GPU を用いた並列計算では、計算と同時に通信を行う「隠蔽」が重要となる。通信の隠蔽を行う際には、通信の対象となる袖領域の計算を事前に行

い、内部領域を計算している間に並行して通信を行う手法がとられることが多いが、このような方法では、GPU が苦手とする不連続なメモリアクセスが生じるために効率が上がりにくい (図 6)。本研究では、GMS が格子最適化手法として採用している不連続格子が 2 つの異なる格子サイズを持つ領域 (領域 I、領域 II) からなることに注目し、一方の領域を計算する間に他方の領域の通信を行うことで袖領域のみの計算を別途行う必要性を回避した (図 7)。

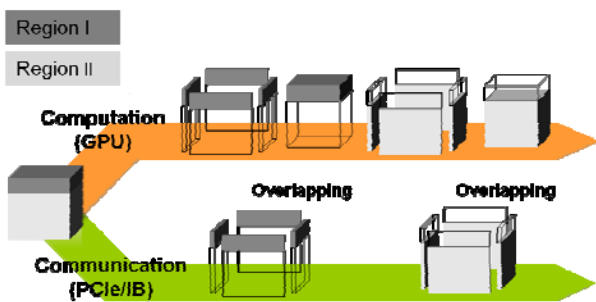


図 6 従来から用いられている通信の隠蔽方法 (袖領域の事前計算が必要)

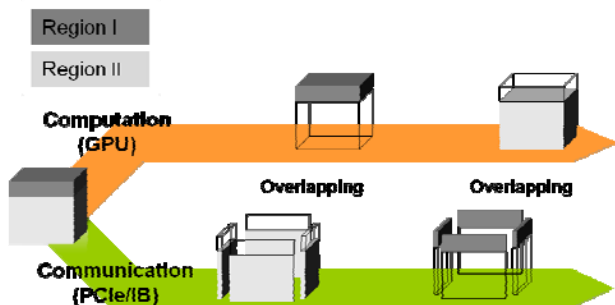


図 7 本研究で開発した通信の隠蔽方法 (袖領域の事前計算が必要ない)

3.4. 性能試験

GPU での実行に対応した差分計算ソルバについて TSUBAME 2.0 を用いて性能試験を行った。使用したノードは GPU に NVIDIA Tesla M2050、CPU に Intel Xeon X5670 2.93GHz が搭載され、ノード間は InfiniBand 4x QDR に接続されている。

まず、単体 GPU での性能を評価するため、GPU 版プログラム、および比較用として CPU 版プログラムを 1 並列で実行した際の性能の比較を行った。

なお、1 並列での実行であるため GPU 版プログラムでは 1GPU、CPU 版プログラムでは 1CPU コアを使用して実行している。計算規模は領域 I が $420 \times 420 \times 100$ 、領域 II が $140 \times 140 \times 200$ の約 2.2 千万格子からなるモデルを使用した。ステップ数は 1000 とし、時間計測においてはファイル入出力等を除いた差分法の計算処理部分のみを対象とした。また、浮動小数点演算は、GPU 版プログラム、CPU 版プログラム共に単精度で行っている。

単体 GPU での性能試験の結果から、CPU 版プログラムに比べて GPU 版プログラムは約 20.4 倍の性能となり、相当な高速化が図られたことが確認された (図 8)。また、性能解析ツールを用いて GPU とグローバルメモリ間のメモリ転送速度を計測したところ、対理論性能に換算すると最大で 70% を超えることから律速の要因はメモリ転送であることが分かった。計算の主要部分においてメモリ転送が理論性能に比して十分に高速に行われており、ハードウェアの性能を生かしていることから単体 GPU での性能については更なる高速化の余地はそれほど大きくはないと考えられる。

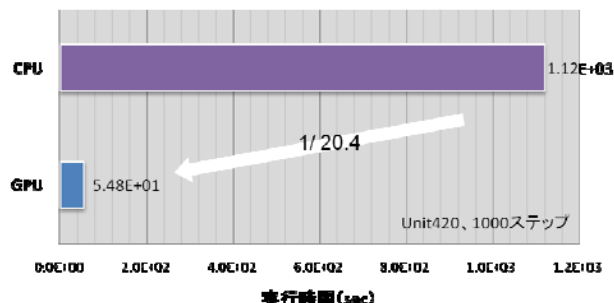


図 8 単体 GPU での性能試験 (GPU 版プログラムと CPU 版プログラムとの比較)

続いて、複数 GPU を使用した際の並列性能について評価を行った。データは、単体 GPU での性能試験で用いた約 2.2 千万格子からなるモデルを単位データとし、これを水平方向に 2x2、3x3、4x4 個並べて、全体の格子数が 4 倍、9 倍、16 倍となるようなモデルを使用した (図 9)。

複数 GPU での性能試験の結果から、モデルのサイズ (格子数) を用いる GPU の数 (並列数) に比例して大きくする weak scaling については、16GPU

(モデルサイズは約 3.4 億格子) までほぼ並列数に比例した性能が得られた (図 10)。一方、モデルサイズを 2.2 千万格子に固定した strong scaling については、4GPU では 3.2 倍程度の演算性能が出るものの、16GPU では 7.3 倍程度の性能しか出ず、急速に並列化効率が悪くなることが分かった。これは、各 GPU に割り当てられる計算の粒度が小さくなるため、相対的に通信の割合が大きくなり通信の隠蔽が破綻することに加え、各 GPU で実行されるスレッドの数が少なくなることにより効率的な計算が行えないことに起因する。通常行われるシミュレーションの時間ステップ数は数万ステップであり、モデルサイズに応じた計算機資源が確保されている場合には計算に要する時間 (turn around time) は数分から数十分程度となるため、過剰な並列度による更なる高速化の必要性は実用的にはほとんど無いと考えられる。

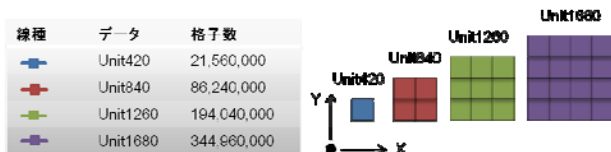


図 9 性能試験に用いたデータ

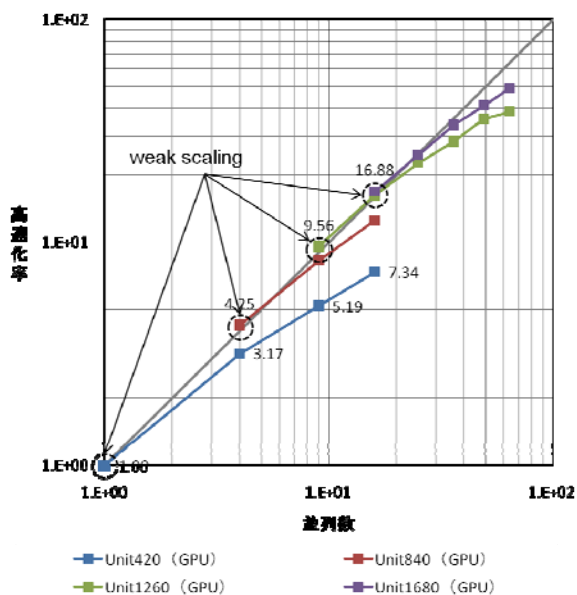


図 10 複数 GPU での性能試験 (並列数-高速化率の関係)

3.5. まとめ

計画していた複数 GPU での並列処理への対応および TSUBAME 2.0 での性能評価について当初計画通りに研究が完了し、GPU あたりの計算量を適切に設定することにより複数 GPU を用いた計算で高い並列性能が得られることが実証された。大規模な地震波伝播シミュレーションを伴う地震ハザード予測を GPU により実用的な計算コストで実行するための基本的な技術開発が確立できたことは大きな意義がある。

4. 今後の展望

2011 年 3 月に発生した東北地方太平洋沖地震や、今後発生が懸念されている南海トラフの海溝型巨大地震は、断層が巨大であるためモデルサイズが極めて大規模なものとなる。高い精度で地震ハザード予測を行うためには、格子数が数億から数十億に及ぶ規模のモデルを取り扱う必要があることから、今後 1000 台程度の GPU による並列計算の実施を計画している。その上で、現実的な震源モデルや地下構造モデルを用いた大規模シミュレーションにおいても高い並列性能を得るための技術開発を行っていく計画である。

ハザードマップを作成するためにはシミュレーション結果の面的な出力が不可欠である。また、データ同化などの必要性から三次元出力が必要となる場合もある。モデル規模が大きくなると出力のすべきデータも膨大になるが、一般的に出力の並列性能を高めることは困難である。計算と出力を同時に行う隠蔽アルゴリズムの開発を行い、出力の性能評価を行う必要がある。

5. 研究成果リスト

(1) 学術論文

青井真・西沢直樹・青木尊之, 2010, GPU を用いた GMS による強震動シミュレーション, 第 13 回日本地震工学シンポジウム論文集, 1915-1920.

(2) 国際会議プロシーディングス

(3) 国際会議発表

(4) 国内会議発表

青井真・西沢直樹・青木尊之, 2009, GPGPU を用いた三次元波動伝播シミュレーション, 2009 年日本地震学会秋季大会

青井真・西沢直樹・青木尊之, 2010, 強震動シミュレータ GMS の GPU 化, 第 122 回物理探査学会学術講演会

青井真・藤井公輔・青木尊之, 2010, マルチ GPU による三次元波動伝播シミュレーション, 日本地球惑星科学連合 2010 年大会