

jh190030-NAH

大規模津波浸水被害推計シミュレーションの マルチプラットフォーム向け最適化手法の研究

撫佐 昭裕（東北大学サイバーサイエンスセンター）

概要

本津波浸水被害推計シミュレーションは、内閣府の総合防災システムとして、南海トラフ地震発生時には SX-ACE を用いて 5 分以内に静岡県から鹿児島県までの津波被害推計を行うものである。本研究では、Intel Xeon, Xeon Phi, GPU などのスーパーコンピュータに本プログラムを移植・高速化し、将来、有事発生時にこれまで開発済みの SX-ACE だけでなく、JHPCN が提供している複数のスーパーコンピュータで被害推計が行えるようにするものである。今年度は、シミュレーションプログラムの分析をルーブリックモデルで実施し、その結果をもとに高速化を実施した。また、昨年度開発した並列処理のロード・インバランス改善モデルの性能評価を実施し、昨年度より 1/2 の実行時間でシミュレーションを行うことに成功した。さらに、GPU において OpenACC と CUDA Fortran によるモデルの開発も実施した。なお、昨年度の評価委員のアドバイスをもとに評価するシステムを増やし、商用クラウドの調査も実施した。

1. 共同研究に関する情報

(1) 共同研究を実施した拠点名

大阪大学, 東北大学

(2) 共同研究分野

超大規模数値計算系応用分野

(3) 参加研究者の役割分担

所属：東北大学サイバーサイエンスセンター		
1	氏名	撫佐 昭裕
	役割	全体統括, プログラム高速化検討
2	氏名	小松 一彦
	役割	Xeon Phi システムチューニング, GPU 高速化検討
3	氏名	下村 陽一
	役割	プログラム高速化検討・実装, システム構成検討
所属：東北大学災害科学国際研究所		
4	氏名	越村 俊一
	役割	副代表, モデル検討
5	氏名	阿部 孝志

	役割	モデル実装
所属：東北大学大学院情報科学研究科		
6	氏名	小林 広明
	役割	高速化の指針検討
7	氏名	佐藤 雅之
	役割	Xeon 高速化検討
8	氏名	岸谷 拓海
	役割	Xeon Phi システムチューニング, GPU 高速化検討・実装
所属：大阪大学サイバーメディアセンター		
9	氏名	伊達 進
	役割	システム構成検討, Octopus 高速化検討

2. 研究の目的と意義

2011 年の東日本大震災では、大規模な津波により北海道から千葉県の広範囲に津波浸水被害が及び、その全貌把握には数日を要した。そのため、救援、物資支援の遅れによる被害が拡大した。津

波災害では、被害の全貌をいち早く把握し、すばやく対応することが求められる。東日本大震災後、政府は災害発生直後に航空機等による被災地調査の体制を整備している。しかし、夜間や悪天の場合には対応できないことや写真画像などの処理に時間がかかり、被害把握には未だ時間を要している。この解決手段の一つとしてコンピュータシミュレーションの活用がある。短時間で津波の伝搬、遡上および建築物被害を推定することができれば、大規模地震発生直後に津波浸水範囲や建築物被害の全貌を把握することができる。そして、その推定結果をもとに救援、物資支援の対応計画をいち早く立案することが可能になる。

本研究グループは、大規模地震発生時に津波被害の全貌を即時に把握することを目標として、津波浸水被害推計を高精度に行う、リアルタイム津波浸水被害推計シミュレーションの開発を行っている。現在、当研究グループのシミュレーションプログラムは、内閣府総合防災システムの津波被害推計処理に採用され、南海トラフ地震発生時には、東北大学と大阪大学のスーパーコンピュータ SX-ACE を用いて、静岡県から鹿児島県までの沿岸約 6,000km を 30m の格子サイズで、5 分以内にシミュレーションを行うことになっている（地震データ受信からシミュレーション結果の送信までを 30 分以内）。

現在、本シミュレーションを防災システムとして利用するためには以下の 3 つの課題を有している。①大規模地震では広い範囲で建屋の倒壊・損壊や停電などの地震被害をもたらす、東北大学と大阪大学の冗長化だけでなく、他サイトのスーパーコンピュータシステムとの冗長化や分散配置が必要であること。②30m の格子サイズでは、地方自治体が整備している津波ハザードマップの 10m 格子サイズより精度が低いこと。③太平洋沖での巨大地震では、北海道から鹿児島まで津波が襲来する可能性があり、延べ 10,000km 以上に及ぶ海岸部の津波浸水被害推計を行う必要があること。

これら課題を解決するためには、有事発生時に被災を免れたスーパーコンピュータを用いてシミュレーション

を実行するマシン環境の整備（①の対策）と、10m 格子サイズによる全国規模のシミュレーションを短時間で実行するためのプログラム高速化（②、③の対策）が必要となる。

本研究の目的は、複数のスーパーコンピュータにおいて、津波浸水被害推計シミュレーションプログラムの高速化を行い、全国規模のシミュレーションを実現するための要件を明らかにし、今後の有事発生時のシミュレーション環境の指針を策定するものである。今年度は昨年度の評価委員からの助言を踏まえ、JHPCN が提供するスーパーコンピュータだけでなく、商用クラウドシステムの活用を念頭に置き、ターゲットとするスーパーコンピュータを、Intel 製 Xeon, Xeon Phi, NVIDIA 製 GPU, NEC 製 SX-ACE, SX-Aurora TSUBASA とした。

3. 当拠点公募型研究として実施した意義

当研究グループは、津波工学の専門家によるシミュレーションモデルの研究と計算機科学の専門家による高速化研究を融合し、スーパーコンピュータ SX-ACE において効率良く動作する津波浸水被害推計シミュレーションを開発してきた。

本研究は、この実績をもとに複数の異なるアーキテクチャのスーパーコンピュータにおいて、津波浸水被害推計シミュレーションを効率良く動作させるための高度化と大規模並列化を行うものである。これまで最適化を実施してきた SX-ACE の知見を活用して、Xeon, Xeon Phi, GPU, SX-Aurora TSUBASA を対象とした高速化の検討、開発を行う。

このため、それぞれのシステムにおける高速化実績を有する研究者との共同研究体制が必要になってくる。さらに、本研究では、全国規模の津波浸水被害推計シミュレーションを実現するため、大規模なシミュレーションを実行できるスーパーコンピュータの計算環境も必要である。

以上より、本公募型共同研究を通して共同研究体制の構築を行い、大規模スーパーコンピュータを活用した研究を実施するものである。

4. 前年度までに得られた研究成果の概要

(1) 性能評価

性能評価を SX-ACE, Xeon Gold, Xeon Phi (KNL) で実施した. 表 1 に各システムの諸元を, 表 2 に使用したコンパイルオプションを, そして表 3 に津波浸水被害推計シミュレーションの条件を示す.

表 1 システム諸元

	SX	Gold	Phi
CPU 型番	-	6126	7210
周波数 (GHz)	1	2.6	1.3
コア数	4	12	64
コア性能 (単精度) (Gflops)	64	166.4	83.2
CPU 性能 (単精度) (Gflops)	256	1996.8	5324.8
LLC 容量 (MB)	1	19.25	16
メモリ容量 (GB/s)	64	96	192
メモリバンド幅 (CPU 当たり) (GB/s)	256	128	115.2
CPU 当たりの B/F	1	0.064	0.021
MCDRAM 容量 (GB/s)	-	-	16
MCDRAM バンド幅 (GB/s)	-	-	490 FLAT mode

表 2 コンパイルオプション

CPU	項目名	コンパイル環境
SX-ACE	-	-Chopt -pi (指示行なし)
Xeon Gold	Def	-O3 -ipo (指示行なし)
	Direc.	-O3 -ipo (指示行あり)
	AVX512	-O3 -ipo -xCORE-AVX512 (指示行あり)
Xeon Phi	Def	-O3 -ipo (指示行なし)
	Direc.	-O3 -ipo (指示行あり)
	AVX512	-O3 -ipo -xMIC-AVX512 (指示行あり)

表 3 シミュレーションの条件

対象地域	高知県
範囲	1288 km x 1025 km

階層格子	多角形格子
格子サイズ	810 m, 270 m, 90m, 30m
格子数	4.8×10^6
Δt	0.5 秒
再現時間	6 時間

図 1 が SX-ACE, Xeon Gold, Xeon Phi の 1 コアにおける実行時間で, 表 2 の各項目について測定を行った. Xeon Gold, Xeon Phi とともにベクトル化指示行の挿入と AVX512 命令の使用によって高速化ができることを明らかにした.

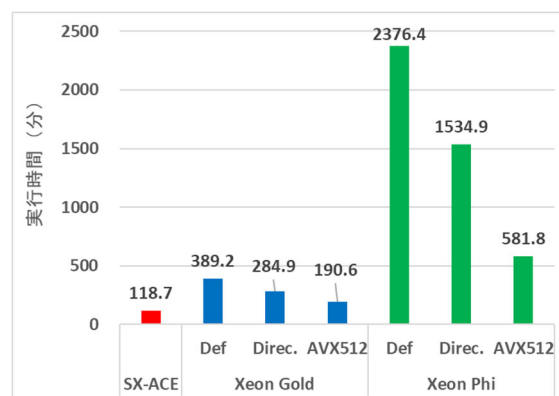


図 1 1 コアの実行時間

マルチコアにおける実行時間を図 2 に示す. 各システムとも数分でシミュレーションを完了することができ, リアルタイムシミュレーションのポテンシャルを有することを明らかにした.

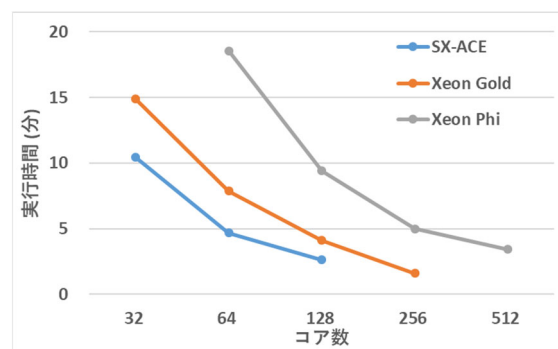


図 2 マルチコアにおける実行時間

(2) プログラムの GPU への移植

本プログラムを NVIDIA Tesla P100 へ移植を行った. 移植にあたっては OpenACC を利用した. 本プログラムのコストは二つのサブルーチンでプロ

グラム全体の 96%を占めていることから、この二つのサブルーチンを GPU で実行するように移植した。図 3 が 30 分間の津波現象を Xeon Gold 1 コアと GPU でシミュレーションした時の実行時間である。GPU の演算時間は Xeon Gold の約 1/10 になっているが、ホストとデバイス間の転送時間が 14 分もかかっている。また、GPU の性能を十分に引き出すまで至っていなかった。

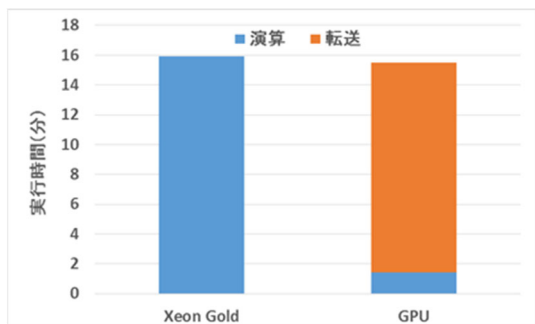


図 3 GPU での実行時間

(3) Xeon Phi システムチューニング手法

Xeon Phi はタイルの分割や MCDRAM の利用モードなどプログラムに合わせた最適構成が取れる仕組みになっている。その組み合わせは 300 通りになる。本研究では STREAM ベンチマーク等の著名なベンチマークプログラムの Xeon Phi における性能情報とシミュレーションプログラムの特性から 300 通りの性能評価を行わなくてもパラメータを絞り込みできる手法を開発した。これによりチューニング時間の大幅な短縮が可能となり、本プログラムでは表 4 で示したパラメータが最適であることを明らかにした。

表 4 最適なパラメータ候補

項目	パラメータ
クラスタモード	Quadrant, SNC-4
メモリモード	Flat MCDRAM
スレッドアフィニティ	Scatter, Balanced
スレッド数	64, 128

(4) 並列効率の改善

本プログラムは多角形領域接続のネスティング手法を用いている。本プログラムの並列化において、それぞれの矩形ごとに領域分割法を用いて並

列化を行っていた。図 4 の左図がそのイメージで、太い線で囲まれた領域が一つの矩形ある。P0 から P5 の 5 プロセスで分割したイメージである。矩形ごとに演算領域を分割した場合、矩形の大きさがそれぞれ異なるため、分割された領域の演算量にばらつきがあった。昨年度は、図 4 の右に示したように矩形を跨いだ分割を行うプログラムを開発した。

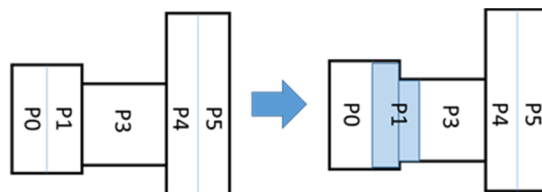


図 4 領域分割のイメージ

(5) システムの検討

(4) で高速化したプログラムを使用して、日本の太平洋全沿岸を 10m 格子でリアルタイムシミュレーション（実行時間 5 分以内）を行うための資源量を SX-ACE で算出した。表 5 が SX-ACE 必要 GPU 数である。太平洋全沿岸のシミュレーションを 5 分以内に行うためには約 6,000 個の GPU が必要であり、東北大学、大阪大学、地球シミュレータの 3 カ所のシステムへの分散配置が必要であることが判明した。

表 5 10m 格子による日本太平洋全沿岸をシミュレーションするために必要な GPU

	高知県	西日本	太平洋全沿岸
SX-ACE	74	1,998	5,994

5. 今年度の研究成果の詳細

本年度は、新しくリリースされたスーパーコンピュータ SX-Aurora TSUBASA を含めて研究を行った。表 6 に SX-Aurora TSUBASA の諸元を示す。

表 6 SX-Aurora TSUBASA の諸元

	数値
CPU 型番	10AE
周波数 (GHz)	1.584
コア数	8

コア性能(単精度) (Gflops)	608.25
CPU 性能(単精度) (Gflops)	4866.0
メモリバンド幅/CPU(GB/s)	1350
CPU 当たりの B/F	0.277

(1) プログラムの分析

4 節(1)で評価したプログラムについてルーフラインモデルを用いて分析を行った。図5はSX-ACE, SX-Aurora TSUBASA, Xeon Goldの1コアにおけるルーフラインモデルである。ここでは, SX-Aurora TSUBASA はソケットのメモリバンド幅でなく, キャッシュのメモリバンド幅に律速されるのでキャッシュのメモリバンド幅を使用した。また, Xeon Glod の場合は1コア当たりのメモリバンド幅を公開していないので, ソケットのメモリバンド幅を使用した。

本プログラムのソースコード上の B/F 値は 1.8 であり, 図5に破線でプロットした。B/F 値が 1.8 の場合, 本プログラムは SX-ACE においては演算器律速のプログラムになり, SX-Aurora TSUBASA と Xeon Gold ではメモリバンド幅律速のプログラムとなる。しかし, 各システムで実測した B/F 値は, SX-ACE で 0.62, SX-Aurora TSUBASA で 0.26, Xeon Gold で 0.73 であった。これらの結果を図5の■で示した。これは, コンパイラによるコードの最適化とキャッシュによるメモリアクセス数の削減により B/F 値 1.8 から減少したものである。この測定結果より, 本プログラムは各システムにおいて演算器律速のプログラムであることが明らかになった。

ここで, B/F 値の減少をみると SX-Aurora TSUBASA が大きく減少していることがわかる。SX-Aurora TSUBASA のキャッシュ容量は Xeon Gold より少ないことから, キャッシュ容量以外の要素が関係している可能性がある。その候補として Xeon Gold と SX-Aurora TSUBASA のベクトル処理の違いや, ベクトル長の長さが考えられる。今後, この減少の違いについて調査を行っていく予定である。

また, 各システムにおける演算性能はルーフラインモデルが示す性能の上限とは隔たりがある。

そこで, 本プログラムの高コストルーチンの演算命令を SX-ACE を用いて調査した。図6が各命令の実行数から求めた演算パイプラインの実行時間の割合である。最も命令数が多かった LOGIC を 100% として表示している。この図より, 本プログラム

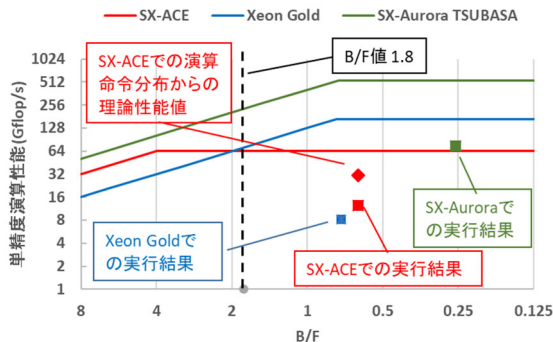


図5 ルーフラインモデル

の実行時間は LOGIC(論理演算パイプライン)の処理時間で決まることがわかる。そして, 浮動小数点演算は, MULT(乗算パイプライン)が LOGIC の 29%, ADD(加算パイプライン)が 63%, DIV(除算パイプライン)が 39%であることから, 理想的に演算器が動作した場合でも浮動小数点演算性能は 31 Gflop/s となり(図5の◆), ルーフラインモデルにおける演算性能(64 Gflop/s)の約半分になる。また, 演算器は演算相互の依存関係やデータのロード・ストアなどの待ち時間が発生し理想的に動作はしないので, 実際の演算性能が 12 Gflop/s になっていると考えられる。以上より, 演算性能がルーフラインモデルが示す上限と隔たりがあるのは, 論理演算が多いことと, 浮動小数点演算器がフルに稼働していないためである。

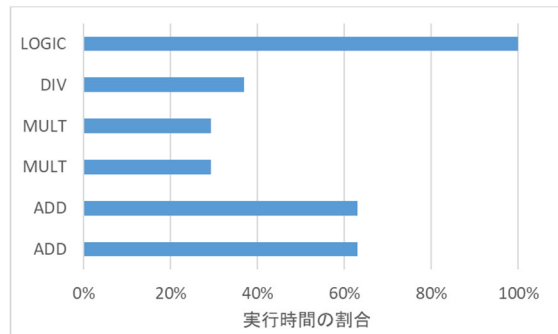


図6 SX-ACE の演算パイプラインの稼働率

(2) プログラムに高速化

(1)の結果より論理演算を削減するため、高コストルーチンにおける IF 文の削減について検討を行った。その結果として、流量フラックスを求めるサブルーチンにおいて以下の二つの処理を削減した。

- ① 津波の浸水を判断するため、各格子点で“水の有無”を判定
- ② 計算領域の境界条件を処理するため、各格子点が境界であるかを判定

削減方法として、①は深さが深いところでは必ず“水有り”になるので、潮汐現象や津波の振幅を考慮して、10m 以上の深水の格子点から①の処理を削減した。また、②においては、境界条件となる格子点は決まっているので、その他の格子で②の処理を削減した。その結果、論理演算の処理を17%削減した。図7が1コアでの実行時間である。青色が本高速化前の実行時間で、赤色が本高速化の実行時間である。本高速化により、それぞれのシステムでの実行時間を17%から30%削減を実現した。

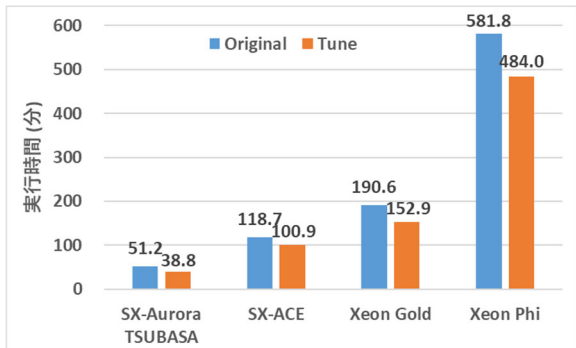


図7 1コアでの実行時間

図8に本最適化によるルーファインを示した。●が本高速化のものである。本最適化によってSX-ACEとSX-Aurora TSUBASAはB/F値が増加しながら演算性能が向上している。一方、Xeon GoldはB/Fはほぼ変わらず演算性能が向上している。しかし、ルーファインモデルとしては、最適化前と最適化後で大きく変化はしていない。これは、論理演算が未だ多いことと浮動小数点演算器がフル稼働していないことによるものである。

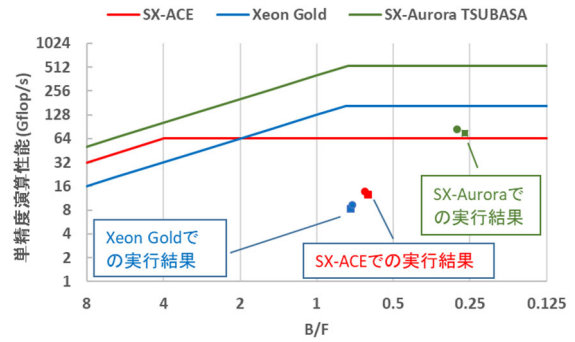


図8 高速化後のルーファイン

(3) マルチコアの性能評価

昨年度、本プログラムのロード・インバランスを改善したプログラムを開発した。そのプログラムの評価を実施した。図9に実行時間を示す。破線がロード・インバランスの改善前の実行時間で、実線が改善後の実行時間である。各コア数での実行時間はそれぞれのシステムで2/3から1/2になっている。また、実行時間5分を見てみると、各システムのコア数が改善前のコア数の半分になっている。このロード・インバランスの改善により、より少ないGPUでシミュレーションを行うことが可能となり、限られたコンピュータ資源の環境において、より多くの地域をシミュレーション対象とすることができる。

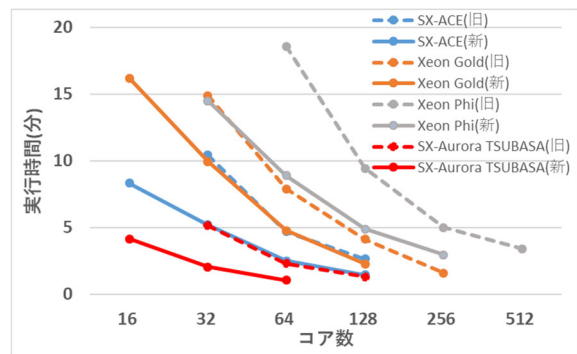


図9 マルチコアの実行時間

(4) GPU への最適化

昨年度はOpenACCを用いて2つの高コストサブルーチンをGPU(NVIDIA Tesla P100)へ実装した。しかし、CPUとGPU間のデータ転送が多く、Xeon Gold 1コアと実行時間が変わらなかった。本年度は本プログラムの時間発展部のすべてのサブルー

チンを GPU へ実装を行い、時間発展部における CPU と GPU 間のデータ転送を削減した。また、GPU への実装には OpenACC と CUDA Fortran の二つを用いた。図 10 がその実行時間である。OpenACC と CUDA とほぼ同じ実行時間で Xeon Gold の約 1/10 の実行時間となった。図 11 が GPU のルーフラインであり、メモリバンド幅律速になっていることがわかる。ここでは、OpenACC の方が CUDA Fortran より B/F 値が低くて性能が高くなっている。これは、CUDA Fortran においてシェアードメモリを利用した実装ができていないことが考えられるので、今後、その実装を行ってさらなる改善を行っていく予定である。

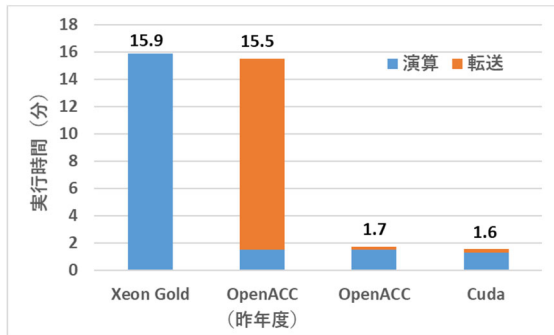


図 10 Xeon と GPU の実行時間

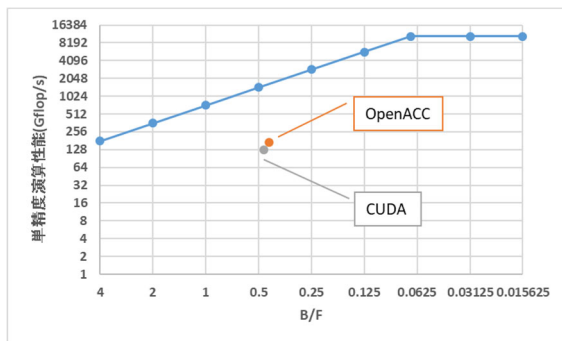


図 11 GPU のルーフラインモデル

(5) 全国規模に向けたシステム検討

本研究の目的は、10m の格子サイズで全国規模の津波シミュレーションをリアルタイムに行うことである。ここで、(3)のプログラムにおける高知県 10m 格子の実行時間を図 12 に示す。SX-Aurora TSUBASA, SX-ACE, Xeon Gold において 104 コア (13CPU), 296 コア (74CPU), 492 コア (41CPU) で 5 分以内にシミュレーションを完了している。

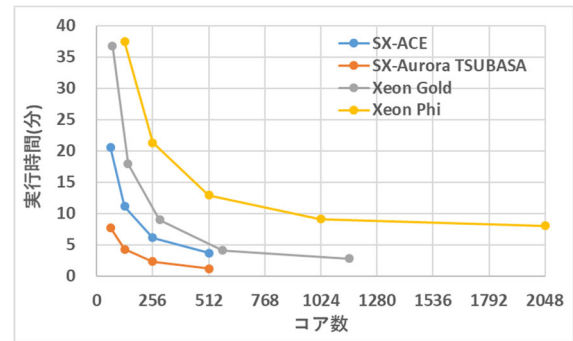


図 12 高知県 10m 格子での実行時間

西日本の太平洋沿岸の演算量は高知県の 27 倍、日本全国の太平洋沿岸は西日本の 3.5 倍になる。なお、昨年度の研究では、日本全国の太平洋沿岸は西日本の 3 倍になるとしていたが、本年度の研究で、千島海溝において巨大地震が発生した場合、津波が西日本領域にも襲来することが判明したため、その津波の伝搬をシミュレーションするために演算領域を拡大している。表 7 に SX-Aurora TSUBASA, SX-ACE, Xeon Gold での CPU 数を記載した。Xeon Gold では 4,000 CPU に近い大規模なシステムが必要になる。しかし、この規模を有する計算センターは多くないため、県や地域ごとに複数のシステムに分散してシミュレーションを行うことを検討する必要がある。

	高知県 10m	西日本 10m	太平洋 全沿岸
SX-ACE	74	1,998	6,993
SX-Aurora TSUBASA	13	351	1,229
Xeon Gold	41	1,107	3,875

表 7 10m 格子での必要 CPU 数

そこで有力なシステム候補は、世界中にあるクラウドシステムである。本年度はクラウドシステムの調査を実施した。津波発生時には短時間でシステムを起動する必要がある。アマゾンの AWS を用いて、必要なインスタンスが利用できるまでの時間を測定した。その結果、1 分から 2 分でシミュレーションを実行できる状態になり、有事発生時に商用クラウドが利用できる可能性があること

がわかった。しかし、アマゾン、マイクロソフト、オラクルなどの商用クラウドでは、InfiniBandによるクラスタシステムを未だサポートしていない、現在、米国の Penguin 社のシステムが有力な候補になっている。

6. 今年度の進捗状況と今後の展望

今年度の計画は、①プログラムの分析、②プログラムの高速化、③マルチコアの性能評価、④GPUへの最適化、⑤全国規模に向けたシステム検討であった。

①については、ルーフラインモデルと命令レベルの分析によってプログラムのボトルネックを明らかにした。②については、論理演算を削減するシミュレーションモデルを開発し、4つのシステムで17%から30%の実行時間を削減した。③についても、4つのシステムで評価を行い、実行時間を2/3から1/2に削減を行った。④については、OpenACCとCUDA FortranでCPUに実装を行った。⑤については、10m格子における太平洋沿岸の全国規模シミュレーションに必要なシステム規模を明らかにした。以上より、本年度予定していた研究はほぼ達成した。

今までは、ベクトル化や並列化の一般的な手法4)を用いて高速化を行ってきた。来年度はそれぞれのシステムアーキテクチャに適した高速化を検討していく。また、本プログラムのMPI通信の分析を行い、通信の隠蔽等について検討を行っていく。

7. 研究業績一覧（発表予定も含む）

(1) 学術論文

なし

(2) 国際会議プロシーディングス

・ A. Musa, T. Abe, T. Kishitani, T. Inoue, M. Sato, K. Komatsu, Y. Murashima, S. Koshimura, H. Kobayashi, “Performance Evaluation of Tsunami Inundation Simulation On SX-Aurora TSUBASA,” In: Rodrigues J. et al. (eds) Computational Science - ICCS 2019. ICCS 2019.

Lecture Notes in Computer Science, vol 11537, pp.363-376, Springer, June 2019. (査読あり)
 ・ T. Kishitani, K. Komatsu, M. Sato, A. Musa, H. Kobayashi, “Importance of Selecting Data Layouts in the Tsunami Simulation Code,” 14th International Workshop on Automatic Performance Tuning, May 2020 (Accepted). (査読あり)

国際会議発表

・ T. Kishitani, K. Komatsu, A. Musa, M. Sato, H. Kobayashi, “An Application Parameter Search Method Based on the Binary Tree Algorithm for Performance Tuning,” ISC High Performance 2019, Frankfurt, Germany, June 16-20, 2019. (査読あり)

・ S. Koshimura, T. Inoue, Y. Ohta, R. Hino, A. Musa, Y. Murashima, M. Kachi, Y. Sato, H. Kobayashi, Y. Murashima. “Real-time Tsunami Inundation and Damage Forecasting in Japan – Present and Future,” 27th International Union of Geodesy and Geophysics General Assembly 2019, Montreal, Canada, July, 2019. (招待講演)

国内会議発表

・ 撫佐昭裕, “SINETの防災利用 リアルタイム津波浸水被害推計システムでの利用”, 学術情報基盤オープンフォーラム2019, 学術総合センター, 5月29日, 2019. (査読なし)

・ 瀧川 陽平, 渡場 康弘, 伊達 進, 撫佐 昭裕, 佐藤 佳彦, “ジョブ管理システム Slurm の緊急ジョブスイッチング機能の検証”, 第17回 ディペンドブルシステムワークショップ, 日本ソフトウェア科学会, 東京, Dec. 2019. (査読なし)

その他（特許, プレスリリース, 著書等）

・ 越村俊一, 阿部孝志, 井上拓也, 撫佐昭裕, 村嶋陽一, 鈴木崇之, 太田雄策, 佐藤佳彦, 加地正明, 小林広明, “リアルタイム津波浸水被害予測の全国展開に向けた検討”, SENAC Vol. 52, No. 2 (2019. 4). (査読なし)