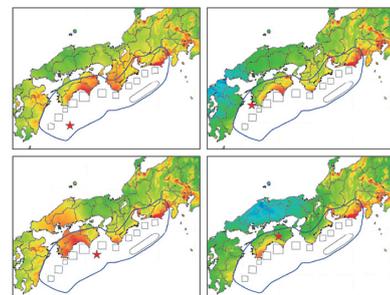


# TSUBAME ESJ.



## TSUBAME-KFC : 液浸冷却を用いた 世界一省エネなスーパーコンピュータ

TSUBAME-KFC : the Greenest Supercomputer in the World  
With Liquid Submersion Cooling

## 超並列計算機 TSUBAME の利用による 幾つかの有機化合物のシュレーディンガー解の計算

Solving the Schrödinger Equations of Some Organic Molecules  
with Superparallel Computer TSUBAME

## GPGPU による地震ハザード評価

Application of GPGPU to Seismic Hazard Assessment

# TSUBAME-KFC：液浸冷却を用いた 世界一省エネなスーパーコンピュータ

遠藤 敏夫 額田 彰 松岡 聡

東京工業大学学術国際情報センター

近年のスーパーコンピュータの性能の限界を決定する最大の要因は消費電力である。

次世代のTSUBAME3.0やエクサスケールスパコンに向けたプロトタイプとして、TSUBAME-KFCを構築した。

Green Revolution Cooling社などとの協働により、TSUBAME-KFCでは高密度高性能な計算ノード群の全体を冷媒油に液浸させ、最終的な排熱は蒸散冷却塔により外気は排出される。この方法により冷却に係る電力や半導体リーク電流を削減させることができる。その結果、TSUBAME-KFCは2013年11月のGreen500ランキングとGreen Graph500ランキングの双方で世界一を獲得した。

## はじめに

# 1

将来のエクサスケールスーパーコンピュータを実現するにあたって最大の課題のひとつは消費電力であり、スパコンの省電力性に関する研究の重要性はすでに広まっている。DARPA exascale report<sup>[1]</sup>ではエクサスケールを2018～20年ごろに20MWattで実現する（50GFlops/Wattに相当）という目標が提示されており、その実現のために解決の必要な課題についても低電力設計を中心に述べられている。

東工大学術国際情報センターにおいても、スパコンの省電力化の研究には早くから取り組んできた。その一例はメニーコアアーキテクチャに基づいたGPUなどのアクセラレータの、スパコンにおける大規模利用<sup>[2,3]</sup>である。その研究成果は2010年11月に稼働開始したTSUBAME2.0に集約された。TSUBAME2.0は日本初のベタスケールスパコンとなっただけでなく、スパコン省エネ性の世界ランキングであるGreen500（www.green500.org）において958MFlops/Wattを記録し、世界三位およびgreenest production supercomputer賞を獲得した。

しかしながら50GFlops/Wattという目標は依然遠く、我々はそれに向けてウルトラグリーンスパコンプロジェクトを文部科学省に提案し採択された。本稿で解説するTSUBAME-KFCシステムはこのプロジェクトの主要成果であり、2013年秋に完成した。

TSUBAME-KFCの最大の特徴は、計算ノード群の冷却のためにエネルギー効率の高い液浸冷却方式を用いていることである。これは、平均的スパコン・データセンターより高効率なTSUBAME2の冷却方式よりも、さらに省エネな方式である。そして計算ノードおよび冷却機器は無人のコンテナに格納されている。

TSUBAME-KFCは、2016年に導入を予定しているTSUBAME3.0スパコンに向けたプロトタイプの役割を持つ。高密度実装されたアクセラレータ利用技術、多数の電力・温度センサとそれを効率利用した自動制御技術などの研究プラットフォームとして活用していく予定である。TSUBAME-KFCにはたった一ラックに高性能な計算ノード群が格納されており、その合計理論演算性能は単精

度で645TFlops、倍精度で217TFlopsである。そして消費電力はTSUBAME2の一ラック（35kWatt）とほぼ同程度に抑えられている。

先進的な冷却・ハードウェア・ソフトウェア技術により、TSUBAME-KFCは省エネスパコンランキングであるGreen500およびGreen Graph500の両方において、2013年11月に世界一を獲得することができた。同ランキングで世界一となったのは、国内のシステムとして初めてである。

## TSUBAME2の消費電力の議論

# 2

TSUBAME2システムにおいて電力を大きく消費する要素として、冷却設備と半導体リーク電流が挙げられる。TSUBAME2ではHP社のModular cooling system (MCS) ラックにより、空冷・液冷のハイブリッド方式が用いられている。冷却水の温度は入口で8～9℃、出口で15～18℃となっている。冬季を除きこの温度は外気温より低いため、冷却水の温度を保つための圧縮機が大きく電力を消費していることが分かっている。

この方式は一般的なデータセンターなどの空冷より電力効率にすぐれているものの、上述した圧縮機やラックに搭載された冷却ファンのために冷却電力を押し上げる。計測によるとTSUBAME2の年間平均PUEは1.29であり（PUEの詳細な説明については4.2章参照）、計算機のシャーシファンなどを考慮すると、システム全体の電力のうち約30%も冷却に費やされている。

またGPUがフル稼働する際には、GPU温度は80～90℃にも達し、それがリーク電流を増大させるという問題も判明している。

**TSUBAME-KFC**

**3**

**3.1 概要**

冷却設備まで含めたスパコン全体の省エネ技術の実証のために、我々はTSUBAME-KFC (Kepler Fluid Cooling) システムを、NEC、NVIDIA、米国Green Revolution Cooling (GRC) 社らと協働で設計・開発した。このシステムでは液浸冷却技術、高密度高性能な計算ノード群、多数の電力・温度センサが統合されている。図1はTSUBAME-KFCのシステム全体の概要を、図2はシステムを格納し屋外に設置された20フィート型コンテナと、その隣の冷却塔の外観を示す。

冷媒油を用いた液浸冷却技術として、GRC社Carnot Jetシステム(図3)を用いている。TSUBAME-KFCでは体積1000リットル強で比較的高温の冷却油に、計算ノード群の全体が液浸されている。この方式により、標準的なラックマウントサーバに小さい改造(後述)を施すことによって、そのまま液浸することができる。図4に、計算サーバ群が冷媒油に浸されている様子を示す。

温度の上昇した油の冷却のために、熱交換器を介して熱を二次冷却水ループへ移動させる。冷却水は、TSUBAME-KFCコンテナに隣接する蒸散冷却塔によって冷却される。冷却水は冷却塔の上部からゆっくり流れ落ち、その際の蒸散熱によって冷却される。以上のような経路により、計算サーバ群の発熱は最終的に外気へ拡散されている。



図2 Tsubame-KFCの外観。  
屋外の20フィート型コンテナに格納されており、その隣に蒸散冷却塔が置かれている。

**TSUBAME-KFC : ウルトラグリーンスパコンテストベッド**



図1 Tsubame-KFCシステムの概要。  
計算ノードが発した熱は冷媒油、冷媒水へと移動し、最終的に外気へ拡散される。

## TSUBAME-KFC: 液浸冷却を用いた 世界一省エネなスーパーコンピュータ



図3 コンテナ内に設置されたGRC社液浸冷却ラック (Carnot Jetシステム)。

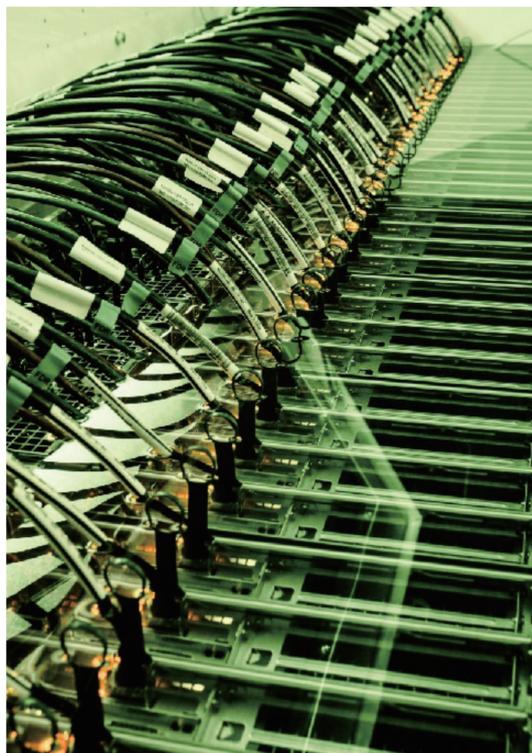


図4 Carnot Jetシステムに  
計算ノードが液浸されてる様子。

### 3.2 液浸冷却の詳細

TSUBAME-KFCの計算ノード群は液浸機能を持った42Uラックに格納されている。このラックは横倒しに設置され、風呂桶のように冷媒油が満たされている。現在格納されている計算ノード群の最大消費電力は40kW程度であるが、Carnot Jetシステム自身は最大100 kWの冷却性能を持つ。

我が国の法規上の問題をクリアするために、GRCのオリジナルのシステムに対して変更を加える必要があることが分かった。当初提案されていた冷媒油は、安全性が高く粘度が低いという良好な性質を備えていたものの、引火点が177℃であった。引火点250℃以下の液体は日本の消防法においては危険物に分類され、TSUBAME-KFCでの利用および、将来のさらなる大規模利用を想定すると、ふさわしくないと判断した。NECおよびGRCと検討した結果、より引火点の高いExxon Mobil社のSpectrasyn8 PAO (poly-alpha-olefin)を選定した。この冷媒油の引火点は260℃であり、危険物該当外であり、なおかつ十分な安全性を持つ。

### 3.3 計算ノード群とその改造

TSUBAME-KFCは計算ノードとして、NEC/Supermicro 104Re-1Gを改造したものを40台用いている。この計算ノードは標準的な形状の1Uラックマウントサーバであり、1Uに2CPUと4GPUを備えるという高密度実装がなされている。ノード仕様は下記の通りである：

- Intel Xeon E5-2620 v2 (IvyBridge) 6 コア 2.1GHz x 2
- DDR3 メモリ 64GB
- NVIDIA Tesla K20X GPU x 4
- SATA3 SSD 120GB x 1+ 500GB x 2
- 4x FDR InfiniBand HCA x 1

ノードあたりの理論演算性能は単精度浮動小数で16.1TFlops、倍精度で5.4TFlopsである。40ノード全体では、単精度645TFlops、倍精度217TFlopsと、たった一ラックでペタスケールに近い性能を持っている。

標準的なサーバを基にしつつも、冷媒油中での動作を可能とするため、下記のような改造をNEC、GRCと協働で施した。

- 稼働パーツの除去：液体内での動作を可能とするため、サーバ内の冷却ファン(12個)を含む物理的な稼働パーツの除去を行った。これにより、ファン電力が必要なくなるという利点もある。なおハードディスクを用いる場合はそれへの対処も必要だが、我々は稼働部分を持たないSSDを用いているため、この点は問題ない。
- 熱伝導グリスの除去：通常、CPUやGPUとヒートシンクの間には熱伝導グリスが塗布されているが、冷媒油の中では溶け出す可能性がある。そのため、金属製熱伝導シートと置き換えた。

### 3.4 電力・温度センサ

TSUBAME-KFC は数多くの電力・温度センサを備えており、採取されたデータはネットワークを介して集計される。センサとその計測対象は下記の通りである：

- 電力センサ
  - ▶ 各計算ノード
  - ▶ ネットワークスイッチ
  - ▶ 冷却塔
  - ▶ 冷却油用ポンプ、冷却水用ポンプ
- 温度センサ
  - ▶ 各CPU、GPU
  - ▶ 冷却油、冷却水（それぞれ入口、出口）
  - ▶ 外気および、コンテナ内

このような細粒度なセンサデータの利用により、電力・温度の詳細制御を可能とし、さらなる省エネ化を実現する予定である。

これまでに述べた TSUBAME-KFC は 2013 年 10 月に完成し稼働開始した。2016 年春に導入が予定されている次世代 TSUBAME3.0 スパコンの省エネ設計に向けて、様々な実験を行う予定である。

## TSUBAME-KFC の電力性能評価

# 4

### 4.1 液浸冷却の計算ノードへの影響

液浸冷却方式がTSUBAME-KFCの計算ノードの消費電力や温度に与える影響を評価する。比較実験のために、上記計算ノードと全く同じ仕様(CPU/メモリ/GPUなど)を持ち、空冷されるノードを用いた。

図5で、空冷ノード (air-cooled) で吸気26℃の場合と、液浸ノード (submersion) の電力と温度を比較する。後者については、冷媒油の入口温度29℃の場合と19℃の場合を示す。計測中は、計算ノードに負荷をかけるために、全GPU上で倍精度密行列積演算を実行し続けた。

空冷ノードと液浸ノードの温度を比べると、冷媒油温度が高いときですら、後者のほうがCPUおよびGPUの温度が低いことが分かる。ノードの電力を比較すると、液浸ノードのほうが約7.8%低いことが分かった。これは冷却ファンの除去と、プロセッサ温度低下によるリーク電流低下の双方の影響と考えられる。

Cooling	Air-Cooled (26 °C)	Submersion (Oil 29 °C)	Submersion (Oil 19 °C)
Temp (°C)			
CPU1	46	42	33
CPU2	50	40	31
GPU1	52	47	42
GPU2	59	46	43
GPU3	57	40	33
GPU4	48	49	42
Node Power (W)	749	693	691

図5 TSUBAME-KFC液浸ノードと空冷ノードの温度・電力比較。

### 4.2 PUE評価

PUE (power usage effectiveness) はデータセンター等で広く電力使用効率を表す指標として用いられ、以下で求められる：

$$PUE = (P_E + P_I) / P_I$$

ここで、 $P_E$  はIT機器 (計算ノードやネットワーク機器など) の消費電力であり、 $P_I$  は付帯設備の消費電力であり、本稿では冷却に係る電力とする。PUE=1は冷却に電力が不要な理想的なケースを意味し、PUE=2は、冷却のためにIT機器と同等な電力を消費してしまうケースを意味する。

TSUBAME-KFCのPUEを評価するために、全計算ノードに前節と同様のGPU行列積演算の負荷をかけた。高負荷時の計算ノード、ネットワークスイッチ、冷却機器の電力をセンサによって計測し、その結果を図6に示す。比較対象として空冷時のケースを示すが、その値は下記のような仮定による：一ノードの空冷ノードに同じ負荷をかけ、その際の電力を外挿することにより計算ノード群の電力を見積もる。またTSUBAME2と同等の冷却方式を用いることを仮定し、PUE=1.29から冷却電力を見積もる。

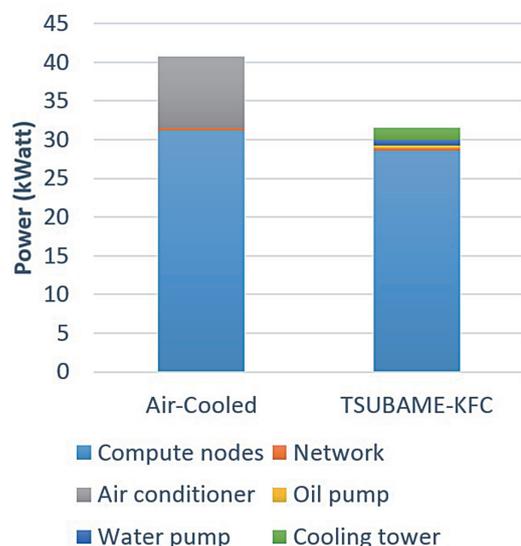


図6 TSUBAME-KFCと空冷システムのPUE評価結果。

## TSUBAME-KFC: 液浸冷却を用いた 世界一省エネなスーパーコンピュータ

計測の結果、TSUBAME-KFCのPUEは1.09と、TSUBAME2と比べても大きく1に近づけることができた。さらに計算ノードの電力も削減されていることを考慮し、空冷時のIT機器電力を基準とすると、修正後のPUEは1.001と、ほぼ理想値が実現されている。

### 4.3 Green500測定

本節ではTSUBAME-KFCのGreen500ランキングにおける結果について述べる。Green500ではスパコンの電力性能比のランキングを、Linpackベンチマーク実行時において比較する。つまり（Linpack速度（Flops）/電力（Watt））が高い方が高位となる。ここでの電力には冷却電力を含まなくて良いとされる。そのためTSUBAME-KFCにおいては、冷却電力自身の低下は本ランキングに直接関係しないものの、計算ノードの電力低下が有利に働く。

TSUBAME-KFC上で電力性能比を向上させるために、これまで述べた要素に加え、下記のようなアーキテクチャからソフトウェアレベルに渡る方策を取った：

- ノードアーキテクチャにおいては、GPUとCPUのプロセッサ比を2:1とし、一般的な1:1のケースよりもCPUなどの相対的な消費電力を削減した。
- 測定に用いたLinpackソフトウェアは、NVIDIA社による効率的なIn-coreアルゴリズムを採用したものである。
- Linpackの性能には、数多くの実行時パラメータが影響する。よく知られたブロックサイズ（NB）やプロセスグリッドサイズ（P&Q）などのパラメータ探索に加え、GPUクロック周波数を変化させながら性能測定を繰り返した。

数多くの性能測定を通して、我々は速度性能が良いことが必ずしも最良の電力性能比をもたらすわけではないことを観測した。図7に示すように、最速の実行のケースと比べて、電力性能比が24%も高いケースが存在することを確認した。

結果として、TSUBAME-KFCの電力性能比は4.503GFlops/Wattに達した。この記録は2013年11月のGreen 500ランキングにおいて世界一を獲得した。なお二位のシステムとは24%の大差をつけている。



図7 TSUBAME-KFC上の種々設定下でのLinpackベンチマークの速度性能と電力性能比。各点は各ベンチマーク実行に対応する。

### 4.4 Green Graph 500 測定

Green Graph 500 ランキング (green.graph500.org) は新しい省エネランキングであり、近年注目を浴びているビッグデータ処理における電力性能比に注目する。このランキングでは、大規模グラフ構造に対する幅優先探索 (BFS) ベンチマークを用いる。BFSの速度はTEPS (Traversed Edges per Second) 値で表され、Green Graph500ではTEPS/Watt値を競う。

Linpackベンチマークにおいては、GPUの利用が電力性能比向上に不可欠であったが、BSPベンチマークはより不規則な構造を持つため、その点が自明ではない。我々は最適化されたアルゴリズムに基づくCPU版とGPU版の実装を比較することによって<sup>[4]</sup>、前者を選択した。

その結果、TSUBAME-KFCの32ノードを用いて6.72 MTEPS/Watt (= 44.01GTEPS / 6.55kWatt)を記録した。この記録は2013年11月のGreen Graph 500ランキングのビッグデータカテゴリにおいて世界一を獲得した。前節のGreen500の世界一と合わせて、TSUBAME-KFCは二冠を獲得することができた。

### 4.5 フェーズフィールド計算の評価

本章の最後に、実際のシミュレーションアプリケーションであるフェーズフィールド計算における電力性能比を示す。このアプリケーションはステンシル計算に基づいて金属結晶の凝固過程を計算するものであり、2011年にゴードンベル賞を受賞している<sup>[5]</sup>。我々は長年、このアプリケーションを用いて世代の異なる計算機システムの電力性能比向上を評価してきた。このため2006年に導入されたTSUBAME1.0の計算ノードからの性能比較を行うことができる。

図8に2006年以降の複数計算機システム上で本アプリケーションを実行した際の電力性能比の推移を示す（冷却電力を含まない）。TSUBAME-KFCでは、2006年の時点と比べて420倍もの電力性能比の向上がなされている。

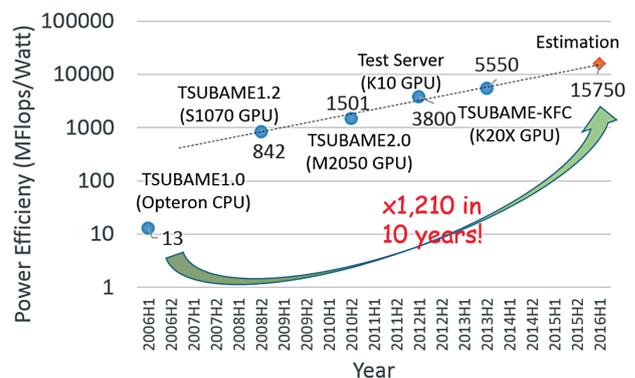


図8 異なる世代の計算機上におけるフェーズフィールド計算の性能電力費の推移。

グラフ中で 2006 年と 2008 年の間に大きな隔りがあるが、これは 2008 年以降のマシンでは GPU を用いているためである。2008 年以降の性能向上を外挿すると、2016 年ごろには 2006 年時点の 1210 倍の電力性能比である 15.5GFlops/Watt の実現が期待される。この成果は、エクサスケールにおける目標である 2020 年ごろの 50GFlops/Watt の実現へ向けて大きく前進するものである。

- [5] Takashi Shimokawabe, Takayuki Aoki, Tomohiro Takaki, Akinori Yamanaka, Akira Nukada, Toshio Endo, Naoya Maruyama, Satoshi Matsuoka. Peta-scale Phase-Field Simulation for Dendritic Solidification on the TSUBAME 2.0 Supercomputer. In Proceedings of the ACM/IEEE conference on Supercomputing (SC11), pp. 1--11, Seattle, November 2011.

## おわりに

# 5

本稿では次世代の超省エネスパコンに向けたプロトタイプである TSUBAME-KFC について解説した。先進的な液浸冷却技術を中心に、アーキテクチャレベル・ソフトウェアレベルの研究成果を統合することにより、Green 500 と Green Graph 500 の二つのランキングにおいて、電力性能比世界一を獲得した。特に液浸冷却技術の採用などにより、同じ種類の GPU を備えるシステム (TSUBAME2.5 を含む) よりも高い電力効率を実現している。

この TSUBAME-KFC を用いて更なる省エネ技術のための実証実験・研究開発を進める予定である。特に、外気温が 35℃ を超えるようなケースにおける本方式の冷却効率を測定するには、夏季を待って実験する必要がある。TSUBAME-KFC 上の研究成果を、次世代の超省エネ・ビッグデータ対応スパコンである TSUBAME3.0 の開発にフィードバックしていく予定である。

### 参考文献

- [1] P. Kogge, K. Bergman et al. ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems, DARPA Technical report 2008-13, 2008.
- [2] Akira Nukada and Satoshi Matsuoka. Auto-Tuning 3-D FFT Library for CUDA GPUs. In Proceedings of the ACM/IEEE conference on Supercomputing (SC09), 10 pages, ACM Press, Portland, November 2009.
- [3] Takashi Shimokawabe, Takayuki Aoki, Chiashi Muroi, Junichi Ishida, Kohei Kawano, Toshio Endo, Akira Nukada, Naoya Maruyama, Satoshi Matsuoka. An 80-Fold Speedup, 15.0 TFlops, Full GPU Acceleration of Non-Hydrostatic Weather Model ASUCA Production Code. In Proceedings of the ACM/IEEE conference on Supercomputing (SC10), pp.1-11, New Orleans, November 2010.
- [4] Koji Ueno and Toyotaro Suzumura, Parallel Distributed Breadth First Search on GPU, IEEE International Conference on High Performance Computing (HiPC 2013), India, Dec. 2013.

# 超並列計算機TSUBAMEの利用による 幾つかの有機化合物の シュレーディンガー解の計算

中辻 博 中嶋 浩之

量子化学研究協会・研究所

理論とシミュレーションの目的は現象の完全予測にあり、正確な予測こそがその使命である。化学の世界の支配原理はシュレーディンガー方程式であるので、これを正確に解くことができれば、化学現象を正確に予測することができる。2000年以降、中辻によりその正確な一般的解法がはじめて発見され、その理論体系は自由完員関数法 (FC法) としてまとめられた。さらにその変数の決定にあたって、分子上に採ったサンプリング点でシュレーディンガー方程式が成り立つ必要条件: Local Schrödinger Equation (LSE) 法を用いることで、任意の原子・分子系にも適用することができる。我々は、超並列計算機TSUBAMEのグランド・チャレンジ利用によって、幾つかの簡単な有機化合物のシュレーディンガー方程式の解を計算し、絶対エネルギーとして化学精度 (kcal/molの精度) を満足するシュレーディンガー解を得ることができた。計算の律速部分である波動関数の反対称化計算は、計算負荷が等価なサンプリング点の分配により各プロセスに均等に分散することができる。実際に、TSUBAMEの4600コアを用いたテスト計算では、111.5%の高い並列性が得られた。これらの成果は基礎科学として重要な意義を持ち、今後、化学現象の理論予測に応用する計算化学技術の確立に繋げたい。

## はじめに

# 1

現象の完全予測のためには、シミュレーションに用いる理論が原理的に正確であり、それを高速で解く手法が必要である。天体の運行予測や惑星探検を例にとると、ここではニュートン力学と僅かの相対論が基本原理であり、この原理を本来の精度で解く道筋は明快に確立されている。そのため、火星探検や「はやぶさ」の快挙など、素晴らしい実例が沢山ある。他方、よく知られているように、化学はシュレーディンガー方程式によって支配されている。従って、この方程式を正確かつ高速に解くことができれば、化学の世界でも正確な予測が可能となり、実験と並んで「シミュレーション理論化学」が正確な予測学として活躍している筈であった。しかし、量子力学の誕生以来80年の努力から、むしろシュレーディンガー方程式を正確に解くことは不可能とされ、量子化学は「理解と解釈のための学問」の地位に甘んじてこざるを得なかった。しかし、2000年以降、中辻によってシュレーディンガー方程式の正確な一般的解法が世界で初めて発見され、その理論体系は自由完員関数 (Free Complement (FC)) 法としてまとめられた。<sup>[1,2]</sup> 積分の困難を克服するために導入したLocal Schrödinger Equation (LSE) 法<sup>[3]</sup>を用いて、どのような原子・分子系にも適用することのできる汎用プログラム「プリンシピア」を開発している。

本研究では、水素、炭素、窒素、酸素などの軽い元素から成る簡単な有機化合物のシュレーディンガー方程式の精密解を、2013年10月8-15日、超並列計算機TSUBAMEのグランド・チャレンジ利用によって計算した。これらの分子の正確なシュレーディンガー解は、未だどの量子化学理論でも計算されておらず、その成果は基礎科学としても重要な意義がある。また、我々の方法は、理論の性格上高い並列性が得られることが予測され、TSUBAMEを用いた高並列計算でそれを確認することも意義深い。

本研究で得られる基礎的計算データの蓄積と、理論・計算技術の開発・改良により、今後の有機化学分野での「正確なシミュレーション理論化学」の構築に向けて努力したい。

## 自由完員関数法の流れ

# 2

まず、我々の理論的基礎である自由完員関数 (FC) 法を説明する。詳細は、レビュー<sup>[4]</sup>等を参照して欲しい。図1に自由完員関数法の流れを示した。

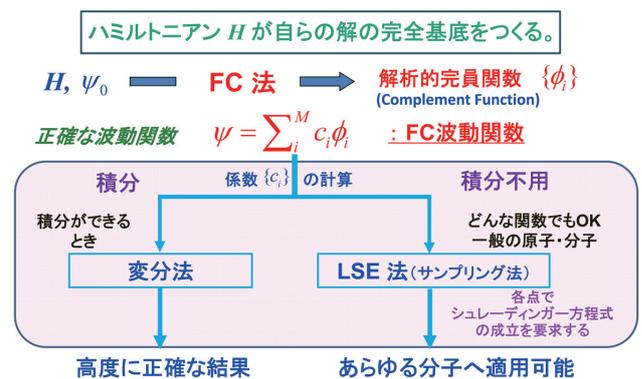


図1 自由完員関数法の流れ

### 2.1 FC法: 完員関数の生成

自由完員関数法は、正確な波動関数の構造論に基づき、「系のハミルトニアンが自らの解の完全な関数空間を作る」という理論である。この理論により、(1)式に示されるように、ハミルトニアン  $H$  を適当な初期関数  $\psi_0$  に何度か作用させることにより、正確な波

動関数に収束をすることが保証される。<sup>[1,2]</sup>

$$\psi_{n+1} = [1 + C_n g(H - E_n)] \psi_n \quad (1)$$

ここで、 $C_n$ 、 $E_n$ は、繰り返し(またはオーダー) $n$ における変分係数とエネルギーである。 $g$ は原子・分子のハミルトニアンを持つクーロン発散項を抑える関数である。<sup>[2]</sup>(1)式の右辺を独立関数に展開することで、系の正確な波動関数を張る解析関数: 完員関数(Complement Function (cf))  $\{\phi_i\}$ が生成される。即ち、(2)式のように、正確な波動関数はこの完員関数の線形結合で書かれることが証明される。

$$\psi = \sum_i^M c_i \phi_i \quad (2)$$

ここで、 $M$ は完員関数の数(次元)、 $c_i$ はそれぞれの完員関数の未知の線形係数であり、変分法あるいはサンプリング法に基づくLSE法により決定する。

## 2.2 LSE法: 積分不要なサンプリング法に

基づく解法変分法は、完員関数のハミルトニアン行列と重なり行列を積分により計算できる時可能であり、高度に正確な結果を得ることができる。しかし、積分可能な系は小さな原子・分子に限られる。一方、LSE法は、局所的なサンプリング点でシュレーディンガー方程式が成立するという必要条件、

$$H\psi(\mathbf{r}_\mu) = E\psi(\mathbf{r}_\mu) \quad (3)$$

を用い解を決定する方法である。<sup>[3]</sup>ここで、 $\mathbf{r}_\mu$ はサンプリング点である。そのため、解析的な積分計算は不用で、原理上いかなる原子・分子、いかなる関数形にも適用可能である。

計算の最も律速部分である波動関数の反対称化計算(2.3で示す)は、計算負荷が等価なサンプリング点を、各プロセスに分配することで、容易に高い並列性が得られる。

## 2.3 パウリ原理と反対称化法

電子のシュレーディンガー方程式を解く際には、同時に波動関数がPauli原理を満たすことが要求され、これを保証することは、一般に多電子系の正確な波動関数の計算において大きな負担となる。

我々は、行列式をベースに、露わな相関関数を含む完員関数においても、分子軌道計算とほぼ同じ速度(電子数の約3-5乗に比例)で反対称化演算を行うことのできる理論を考案し、これをNk法と呼んだ。<sup>[5]</sup>この計算手法はほぼ確立されており、これまで多くの原子・分子系に適用し、実績を積んでいる。本研究では、主にこのNk法を用い、有機化合物の計算を行った。

他方、最近、分子系の反対称化の負担を飛躍的に軽減することのできるinter Exchange (*i*Exg)理論を発見し<sup>[6]</sup>、計算負荷が系の大きさに対しオーダー  $N$ にしか増えないことを理論的に示した。この理論は、今後、大きな分子のシュレーディンガー解を計算していく上で、明るい見透しを与えてくれる。この手法については、今後、論文・学会等で発表していく。

## FC-LSE法の並列計算アルゴリズム

FC-LSE法の並列計算アルゴリズムを図2にまとめた。以下、その手順を簡単に説明する。

### ステップ1,2:

系のハミルトニアンに基づき、完員関数を生成する。このステップは解析的な演算部位であるため、計算量は僅かである。そのため、あらかじめ低並列計算機で実行する。

### ステップ3.1:

LSE法に基づき、サンプリング点  $\mathbf{r}_\mu$  においてシュレーディンガー方程式が成立する条件((3)式)を計算する。(3)式は行列固有値方程式、

$$\mathbf{A}\mathbf{C} = \mathbf{B}\mathbf{C}\mathbf{E} \quad (4)$$

に置き換えることができる。ここで、行列  $\mathbf{A}$ 、 $\mathbf{B}$ の要素は

$\mathbf{A}_{\mu\nu} = H\phi_\nu(\mathbf{r}_\mu)$ 、 $\mathbf{B}_{\mu\nu} = \phi_\nu(\mathbf{r}_\mu)$ であり、 $\mathbf{C}$ 、 $\mathbf{E}$ はそれぞれ係数ベクトルと対応するエネルギー(対角行列)である。

各サンプリング点に対する  $\phi_i(\mathbf{r}_\mu)$  と  $H\phi_i(\mathbf{r}_\mu)$  の計算は、電子の反対称化計算を伴うため、FC-LSE法全体で最も計算律速な部分である。しかし、計算負荷が等価なサンプリング点を各プロセス(コア)に均等に分配することができるため、極めて高い並列性能が期待できる。さらに、このステップではプロセス間の通信も発生しない。サンプリング点は通常  $10^{6-7}$ 個の点を用いるため、Peta-Flops超並列マシンでは各プロセスは1-10点のみを担当することになる。

ステップ3.1の計算コストは、 $M \cdot N_s \cdot \alpha(N_e^{3-5})$ となる。ここで、 $N_s$ 、 $N_e$ はサンプリング数及び電子数である。

### ステップ3.2:

完員関数の次元  $M$ よりも多くのサンプリング点を用いる場合、(4)式の $\mathbf{A}$ 、 $\mathbf{B}$ 行列は長方形列となる。そこで、 $\mathbf{B}^\dagger$ を(4)式の左からかけることで、

$$\mathbf{H}\mathbf{C} = \mathbf{S}\mathbf{C}\mathbf{E} \quad (5)$$

を得る。ここで、 $\mathbf{H} = \mathbf{B}^\dagger \mathbf{A}$ 、 $\mathbf{S} = \mathbf{B}^\dagger \mathbf{B}$ であり、それぞれハミルトニアン行列と重なり行列に対応する。 $\mathbf{S}$ 行列は常に正定値対称行列であり、 $\mathbf{H}$ 行列も $\mathbf{A}$ 行列に比べ対称行列に近づくため、固有値問題(ステップ4)を比較的安定に解くことができる。

このステップでは、 $\mathbf{H}$ 、 $\mathbf{S}$ 行列を計算するために行列積の演算が必要であるが、計算機環境に最適化されたBLAS3ライブラリを使用する。ここでの計算コストは  $M^2$ 、 $N_s$ となり、プロセス間の通信は発生しない。

# 超並列計算機TSUBAMEの利用による 幾つかの有機化合物のシュレーディンガー解の計算

## ステップ3.3:

このステップでは、各プロセスで計算した、**H**、**S**行列を集約する。ノード間にまたがる場合は、MPI\_Reduceにより和をとりながら集約する。データ転送量は  $M^2$  である。

## ステップ4:

完成した行列方程式 (5 式) の固有値問題を解く (行列対角化)。ここでは、一般的な並列化固有値問題ライブラリを利用し、中程度の並列化性能を期待することができる。しかし、並列計算機資源を有効に利用するために、ステップ3終了後、いったん計算を打ち切り、低並列度での計算を行う。

## ステップ5:

ステップ4で得られた固有ベクトル (波動関数) を用い、エネルギーや各物理量を計算する。波動関数の正確さを示す指標である H-square Error も計算する。このステップの計算量は僅かで、メモリやディスク資源もほとんど必要としない。

## 簡単な有機化合物の シュレーディンガー解の計算

# 4

我々は、FC-LSE法を図3に示す幾つかの簡単な有機化合物のシュレーディンガー解の計算に適用した。

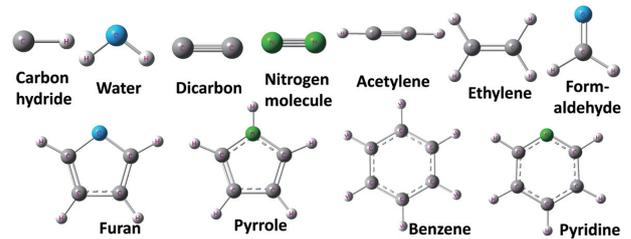


図3 FC-LSE法を適用した有機化合物

### ステップ1. ハミルトニアンに基づく完員関数の生成

$$\psi_{n+1} = [1 + C_n g(H - E_n)] \psi_n$$



### ステップ2. FC波動関数

$$\psi_{n+1} = \sum_i^M c_i \phi_i$$

解析的な演算  
(計算量はわずか)

### ステップ3. LSE方程式

最も律速な部分

$$3.1. \quad \mathbf{AC} = \mathbf{BCE} \quad \mathbf{A}_{\mu i} = H\phi_i(\mathbf{r}_\mu), \quad \mathbf{B}_{\mu i} = \phi_i(\mathbf{r}_\mu)$$

- サンプルング点の分配 `call MPI_Send(rμ)` `call MPI_Recv(rμ)`
- 各サンプルング点での  $H\phi$  と  $\phi$  の値を評価 ( $10^6$ - $10^8$  点)
 
$$\phi(1, 2, \dots, N_e) = A[f_{1,2}(1, 2) \cdots f_{r-1,r}(r-1, r) \cdot o_1(1) \cdots o_{N_e}(N_e)]$$

反対称化: 時間がかかる
- サンプルング点を各プロセスに均等に分配することができる (データ転送は必要ない)
- Peta-Flops超並列マシンでは 各プロセスは1-10点を担当
- 計算コスト:  $M_n \times N_s \times O(N_e^{3.5})$ 
  - $M_n$ : 完員関数の数
  - $N_s$ : サンプルング数
  - $N_e$ : 電子数

$$3.2. \quad \mathbf{HC} = \mathbf{SCE} \quad \mathbf{H} = \mathbf{B}^\dagger \mathbf{A}, \quad \mathbf{S} = \mathbf{B}^\dagger \mathbf{B}$$

$$H_{ij}^{(Local)} = \sum_{\mu}^{(Local)} \phi_i(\mathbf{r}_\mu) H \phi_j(\mathbf{r}_\mu)$$

$$S_{ij}^{(Local)} = \sum_{\mu}^{(Local)} \phi_i(\mathbf{r}_\mu) \phi_j(\mathbf{r}_\mu)$$

BLAS3 library

- ローカルな  $H, S$  行列を作成
- 計算コスト:  $M_n \times M_n \times N_s$

### 3.3. 各プロセスからデータを集約し行列積を計算

$$H_{ij}^{(All)} = \sum_{n_{proc}} H_{ij}^{(Local)}$$

$$S_{ij}^{(All)} = \sum_{n_{proc}} S_{ij}^{(Local)}$$

`call MPI_Reduce(H, MPI_SUM)`  
`call MPI_Reduce(S, MPI_SUM)`

- データ転送:  $M_n \times M_n$

### ステップ4. 対角化 (固有値問題)

並列固有値ライブラリの利用: 中程度の並列性能を期待

### ステップ5. 物理量の計算

エネルギー:  $\langle E_L \rangle = \frac{\sum_{i,j} C_i C_j H_{ij}}{\sum_{i,j} C_i C_j S_{ij}}$     H-square error:  $\sigma^2 = \langle E_L^2 \rangle - \langle E_L \rangle^2$

図2 FC-LSE法の並列計算アルゴリズム

FC法((1)式)に従い生成される完員関数は、一般に(6)式のように表される。

$$\phi_i(1,2,\dots,N_e) = \prod_i^{N_e} \left( \chi(\mathbf{r}_{iA_i}) \cdot x_{iA_i}^{k_x,iA_i} y_{iA_i}^{k_y,iA_i} z_{iA_i}^{k_z,iA_i} \left| \vec{r}_{iA_i} \right|^{k_{r,iA_i}} \right) \times \prod_{A(\neq A_i)}^{N_A} \left| \vec{r}_{iA} \right|^{k_{iA}} \times \prod_{j(>i)}^{N_e} \left| \vec{r}_{ij} \right|^{m_{ij}} \cdot \sigma_i \quad (6)$$

ここで、 $N_e$ 、 $N_A$ は、それぞれ電子数と原子核の数である。 $i$ 、 $j$ は電子、 $A$ は原子核のインデックスを示す。 $A_i$ は電子 $i$ の属する原子を示す。 $\chi(\mathbf{r}_{iA_i})$ は電子 $i$ の属する原子軌道を示し、ここでは原子のHartree-Fock軌道を用いた。<sup>[7]</sup>  $\sigma_i$ はスピン関数を示し、分子の化学結合は原子価結合法に基づき構成した。FC法は、オーダー $n$ を上げることにより、どのような出発関数からも正確な解に収束することが保証されているため、原子価結合法に見られるような多くの配置を採用する必要はない。

分子のサンプリング点は、原子のサンプリング点をLocal Sampling法により合成し、各サンプリング点の局所的分散値を基準にサンプリング点を最適化するH-square Error Stationary(HSES) Algorithmも導入した。<sup>[8]</sup> また、計算を簡単化するため、電子相関

の局所性に基づき原子間にまたがる露わな相関関数(Inter- $r_{ij}$ )は採用しなかった。Inter- $r_{ij}$ の効果については、現在より詳細に検討中である。

計算は、TSUBAMEグランドチャレンジ課題(カテゴリーB)で行った。表1に、CHからH<sub>2</sub>CO(16電子系)まで、FC法のオーダー $n=2$ で計算した結果を示した。実験値等から見積もられた正確なエネルギーの推測値と計算値のエネルギー差は、表1のすべての分子でkcal/mol程度であり、絶対エネルギーとして、ほぼ化学精度を満足している。また、表2に、36電子系のフラン、ピロール、42電子系のベンゼン、ピリジンの $n=1$ でのテスト計算の結果を示す。 $n=1$ であるため、正確なエネルギーの推測値とはまだ少し差があるが、例えばベンゼンにおいて、従来の量子化学に基づく進んだ電子相関理論(MP2-F12法)による計算値:-231.8356 a.u.と比べれば、<sup>[9]</sup> FC-LSE法の計算結果は絶対エネルギーでかなり正確な解と言える。本計算は、まだテスト段階であり、今後高いオーダーの本格計算を実施し、より精密なシュレーディンガー解を算出したい。

TSUBAMEの超並列計算機パワーにより、これらの分子の計算すべてが1週間で完了し、得られた計算結果の成果と共に、結果をすぐに方法論・アルゴリズムの改良に反映することができ、有意義な研究を遂行できた。

表1 簡単な有機化合物のシュレーディンガー解の計算(FC法のオーダー: $n=2$ )

分子	電子数	次元	エネルギー (a.u.)		$\Delta E = E_{FC-LSE} - E_{exact}$ (kcal/mol)
			FC-LSE	Estimated exact (experiment)	
Carbon hydride (CH)	7	1503	-38.480 41	-38.479 0	-0.88
Water (H <sub>2</sub> O)	10	2075	-76.456 78	-76.457 8	0.67
Dicarbon (C <sub>2</sub> )	12	1976	-75.923 69	-75.926 5	1.76
Nitrogen molecule (N <sub>2</sub> )	14	1121	-109.542 07	-109.542 7	0.39
Acetylene (C <sub>2</sub> H <sub>2</sub> )	14	1709	-77.333 31	-77.335 7	1.49
Ethylene (C <sub>2</sub> H <sub>4</sub> )	16	2628	-78.577 95	-78.587 4	5.93
Formaldehyde (H <sub>2</sub> CO)	16	4083	-114.505 35	-114.508 0	1.66

表2 中程度の有機化合物のテスト計算(FC法のオーダー: $n=1$ )

分子	電子数	次元	エネルギー (a.u.)		$\Delta E = E_{FC-LSE} - E_{exact}$ (a.u.)
			FC-LSE	Estimated exact (experiment)	
Furan (C <sub>4</sub> H <sub>4</sub> O)	36	161	-229.860 1	-230.027	0.167
Pyrrrole (C <sub>4</sub> H <sub>5</sub> N)	36	174	-209.974 3	-210.173	0.199
Benzene (C <sub>6</sub> H <sub>6</sub> )	42	398	-232.409 3	-232.248	-0.161
		5092 <sup>a</sup>	-232.195 8 <sup>a</sup>		
Pyridine (C <sub>5</sub> H <sub>5</sub> N)	42	386	-247.704 1	(-248.290)	0.586

<sup>a</sup> 炭素原子の完員関数がより高次

# 超並列計算機 TSUBAMEの利用による 幾つかの有機化合物のシュレーディンガー解の計算

## 並列化効率

# 5

FC-LSE法の並列化効率をテストするために、ベンゼンのテスト計算での並列化加速を測定した。図4に、460並列の計算時間を基準に、2300並列、4600並列の計算でのスピードアップを示した。その結果、4600並列で、111.5%の超並列の加速を得た。大きな原子・分子系では、最も計算律速な反対称化計算(図2のステップ3.1)の演算が増すため、より並列化効率は良くなると考えられる。並列化効率が100%を超えている理由は、単体プログラムのチューニング不足が考えられるが、並列度を上げることによるメモリの分散効果によるものが大きいと考えている。

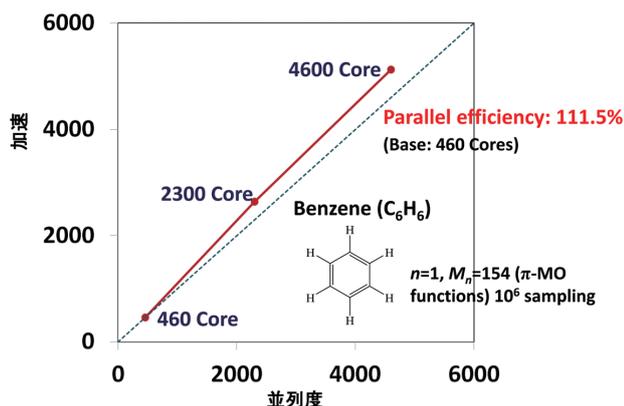


図4 ベンゼンのテスト計算での並列化効率

## おわりに

# 6

原子・分子のシュレーディンガー方程式の正確な解を求めることは、化学の夢であり、本研究により簡単な有機化合物の計算を行うことができた。計算結果について、さらに検証を重ねる必要があるが、従来の量子化学では到達できない精密な解を得ることができた。方法論と計算アルゴリズムは、まだまだ改良の余地があり、TSUBAMEグランドチャレンジは、理論とアルゴリズム・プログラムの検証という立場で大変有用であった。既存の量子化学理論は、精度の高い理論ほど並列化が難しい。一方、FC-LSE法は、普通の計算機は勿論、超並列計算機の本来の力をそのまま反映することのできる理論体系を持っており、非常にフレキシブルである。最近、分子の波動関数の合成法や、新しい反対称化理論を提案

し、理論の大きな進展があった。<sup>[6]</sup> 本研究での成果を基礎に、新しい理論を発展させ、それを実現できる方法論とプログラムを開発していく。

## 謝辞

本研究は、TSUBAMEグランドチャレンジ課題(カテゴリーB)の採用により、東京工業大学学術国際情報センターの大型計算機(TSUBAME2)を主に利用した。また、自然科学研究機構計算科学研究センターの計算機もその事後の計算処理に利用した。ここに記して深く感謝したい。さらに、川上量生氏にはQCRIの研究に対するサポートを深く感謝したい。

## 参考文献

- [1] H. Nakatsuji: Structure of the Exact Wave Function, J. Chem. Phys., Vol. 113, pp.2949-2956 (2000). H. Nakatsuji and E. R. Davidson: Structure of the Exact Wave Function. II. Iterative Configuration Interaction Method, J. Chem. Phys., Vol. 115, pp.2000-2006 (2001)
- [2] H. Nakatsuji: Scaled Schrödinger Equation and the Exact Wave Function, Phys. Rev. Lett., Vol.93, pp.030403-1-4 (2004). H. Nakatsuji: General Method of Solving the Schrödinger Equation of Atoms and Molecules, Phys. Rev. A, Vol. 72, pp.062110-1-12 (2005)
- [3] H. Nakatsuji, H. Nakashima, Y. Kurokawa, and A. Ishikawa: Solving the Schrödinger Equation of Atoms and Molecules without Analytical Integration Based on the Free Iterative-Complement-Interaction Wave Function, Phys. Rev. Lett., Vol.99, pp.240402-1-4 (2007)
- [4] H. Nakatsuji: Discovery of a General Method of Solving the Schrödinger and Dirac Equations That Opens a Way to Accurately Predictive Quantum Chemistry, Acc. Chem. Res., Vol. 45, pp.1480-1490 (2012)
- [5] H. Nakashima and H. Nakatsuji: Efficient antisymmetrization algorithm for the partially correlated wave functions in the free complement - local Schrödinger equation method, J. Chem. Phys., Vol. 139, pp. 044112-1-16 (2013).
- [6] H. Nakatsuji: ACS meeting, Indianapolis, U.S.A. (2013)
- [7] E. Clementi and C. Roetti: Roothaan-Hartree-Fock Atomic Wavefunctions, Atomic Data and Nuclear Data Tables, Vol. 14, pp.177-478 (1974)
- [8] H. Nakatsuji: Local sampling method and H-square error stationary algorithm, to be submitted.
- [9] D. Yamaki, H. Koch, and S. Ten-no: Basis set limits of the second order Møller-Plesset perturbation energies of water, methane, acetylene, ethylene and benzene, J. Chem. Phys., Vol. 127, pp.144104-1-5 (2007)

# GPGPUによる地震ハザード評価

青井 真\* 前田 宜浩\* 青木 尊之\*\*

\*防災科学技術研究所 \*\*東京工業大学 学術国際情報センター

将来発生する地震によりどのような揺れが生じるかを予測する地震ハザード評価では高精度かつ大規模な地震波伝播シミュレーションが必要となる。防災科学技術研究所が開発してきた地震動シミュレータGMSをTSUBAME上に実装し、格子数が数百億規模の極めて大規模かつ現実的なモデルまで高いスケーラビリティ(弱スケーリング)が得られることが分かった。また、本研究で開発したシミュレーションシステムを応用し、南海トラフで発生すると考えられる海溝型巨大地震を対象とした長周期地震動のハザード計算について紹介する。

## はじめに

# 1

3次元差分法は、長周期地震動をはじめとする地震ハザードを評価する際に行われる複雑な地下構造を考慮した地震波伝播シミュレーションにおいて、現在最も標準的に用いられている手法である。しかしながら、詳細な3次元地下構造を十分な精度で離散化し、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規模のモデルを扱うことになる。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パワーおよびメモリ)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以上の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GPGPU (General Purpose Computation on Graphics Processing Unit) の利用が考えられる<sup>[1][2][3]</sup>。本稿では、東京工業大学が運用するTSUBAMEのGPUを用いた3次元差分法の大規模並列シミュレーションについて紹介する。

## GPUによる 3次元差分シミュレーション

# 2

### 2.1 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<sup>[4]</sup>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 (Ground Motion Simulator)<sup>[5]</sup>を用いる。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次元差分法により地震波伝播シミュレーションを行うためのツール群であり、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されている。

差分法による地震波伝播シミュレーションを行う場合、格子サ

イズは計算すべき最短波長により決定されるため、均一の格子サイズで離散化する場合には、モデルのごく一部のみが低速度の媒質である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の大きな障害となっていた。そこで、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が提案されている<sup>[4][5]</sup>。

GMSでは、軟弱な(=地震波速度が低い)表層が存在する浅い部分(領域I)は細かい格子点間隔とし、深い部分(領域II)は領域Iの3倍の粗い格子点間隔とした格子モデルを用いている(図1)。領域Iの下端部と領域IIの上端部は格子が重なっており、波動場の連続性が保たれるよう内挿される。この不連続格子を用いることで、典型的な盆地構造モデルの計算において、均一な格子による場合と比較し数倍から十数倍程度効率が向上することが分かっている。

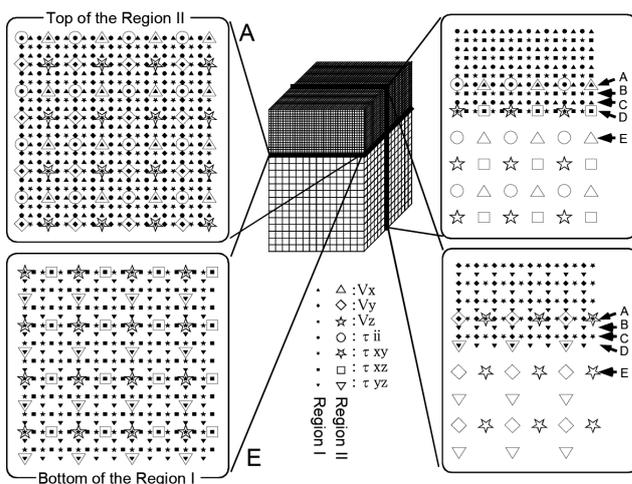


図1 (中央) 計算に用いる不連続格子。(右) 不連続格子の垂直断面。領域Iと領域IIの接続部分で、内挿のために格子が重なっている。(左) 領域IIの最上面(A面)と領域Iの最下面(D面)における不連続格子の水平断面。

2.2 GPUでの差分計算処理

GMSでは、NVIDIA社から提供されているCUDA (Compute Unified Device Architecture) を使用して、差分計算処理のほぼ全てをGPU上で行うように実装されている。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2左)。

差分法の計算を行う際は、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。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2右)。差分法のように大量のメモリアクセスを伴う計算ではメモリバンド幅がボトルネックになることが多いため、GPU上に搭載されている高速なレジスタやシェアードメモリをキャッシュのように使用して、低速なグローバルメモリへのアクセスを軽減する工夫を行っている。

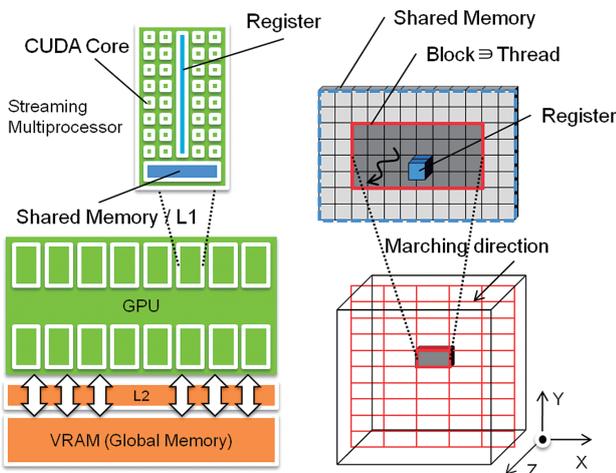


図2 (左) GPUのアーキテクチャ、(右) GPUでの計算処理。x-y面方向の隣接参照にはshared memory、z方向の隣接参照にはregisterを使用。

2.3 複数GPUを用いた並列化

複数GPUを用いた並列化においては、水平二方向 (x方向、y方向) の領域分割法を用いている。分割した各部分領域には、それぞれ1つのGPUを割り付けて計算処理を行う。各部分領域には、隣接する部分領域のデータを取る袖領域を設け、データが更新された際に、各部分領域間でMPI (Message-Passing Interface) ライブラリを用いてデータを交換する。

GPUの演算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り、CPUのみを用いて並列化した場合に比べて、並列性能が低くなる傾向がある。GMSでは、これを改善する

ためGPUでの計算処理と通信とを同時に行うことで、通信に要する時間の隠蔽を行っている。一般に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袖領域の通信を行う手法がとられることが多い。しかしこのような方法では、袖領域を計算する際にGPUが苦手とする不連続なメモリアクセスが生じるため、袖領域の事前計算そのものが足かせとなってしまいう可能性がある (図3)。そこで本研究では、GMSが2つの領域 (領域I、領域II) からなる格子モデルを採用していることに注目し、一方の領域を計算する間に他方の袖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4)。なお、この方法は、不連続格子を用いていないアルゴリズムにも拡張が可能で、事前計算などの余分なオーバーヘッドが無いことから、通信を隠蔽するには有効な手段と言える。

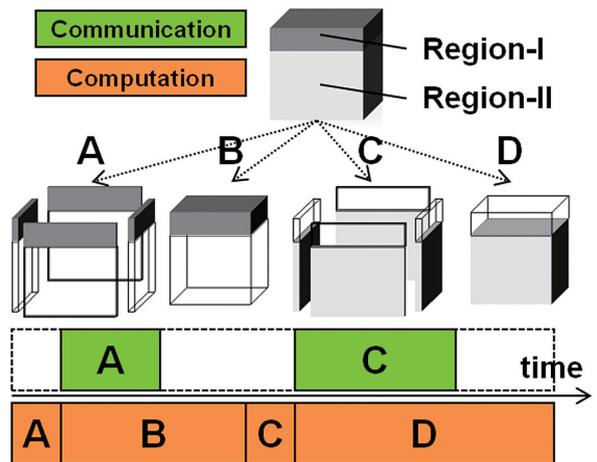


図3 通常行われる通信時間の隠蔽方法 (袖領域の事前計算が必要)。

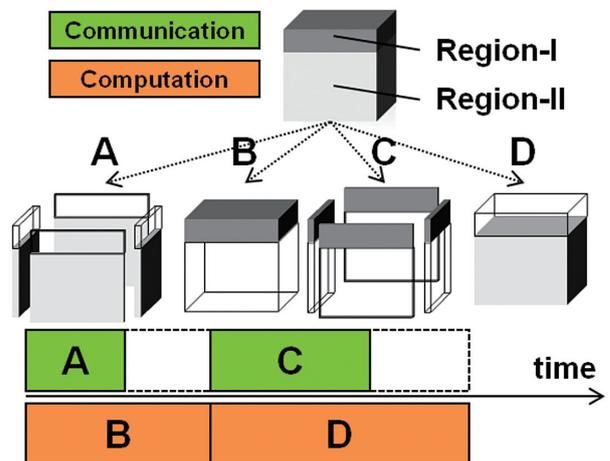


図4 本研究で開発した通信時間の隠蔽方法 (袖領域の事前計算が必要ない)。

## 2.4 並列計算の性能試験

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。使用したノードにはGPUにNVIDIA Tesla M2050が3基、CPUにIntel Xeon X5670 2.93GHzが2基搭載されており、ノード間はInfiniBand 4x QDRにて接続されている。

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版プログラムの1並列での性能試験を実施した。計算モデルには領域Iが $420 \times 420 \times 100$ 、領域IIが $140 \times 140 \times 200$ の21,560,000格子からなるモデル(以下、Unit420と呼ぶ)を使用し、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となる。比較用のCPU版プログラムの実行にはIntel Xeon X5670の1CPUコアを使用し、コンパイラにはIntel Compiler 11.1を用いている。単体GPUでの性能試験の結果(表1)から、CPU版プログラムに比べてGPU版プログラムは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。また、NVIDIA CUDAに付属する性能解析ツールを用いてGPU-VRAM(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、最大で理論性能の70%以上となることから、律速の要因が演算処理ではなくメモリ転送速度であることが確認された。計算の主要部分において、メモリ転送が理論性能に対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、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。

### Unit420 1000ステップ

プログラム	実行時間 (sec)	FLOPS (FP32)	高速化率
GPU版プログラム	5.48E+01	7.95E+10	20.42
CPU版プログラム	1.12E+03	3.89E+09	1.0

※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。

表1 単体GPUでの性能評価

続いて、複数GPUを用いた並列性能の評価を実施した。計算モデルには、Unit420を単位モデルとし、これを水平方向に $2 \times 2$ 、 $3 \times 3$ 、 $4 \times 4$ 、 $8 \times 8$ 、 $10 \times 10$ 、 $16 \times 16$ 、 $32 \times 32$ 個並べて、全体の格子数がそれぞれ4倍、9倍、16倍、64倍、100倍、256倍、1024倍となるようなモデル(それぞれ、Unit840、Unit1260、Unit1680、Unit3360、Unit6720、Unit13440と呼ぶ)を使用し、最大で1024GPU(342ノード)を用いた評価を行った。複数GPUでの性能試験の結果(図5)から、1GPU当りのモデル規模を固定しGPU数を増加させる弱スケーリングに関しては、ほぼ線形(実際には線形を超える)の性能が得られていることが分かる。一方、モデルを単位モデルに固定し使用するGPU数を増加させる強スケーリングに関しては、4GPUでは3.2倍の演算性能が得られるものの、16GPUでは7.3倍の性能となり、高並列時に効率が低下することが分かった。

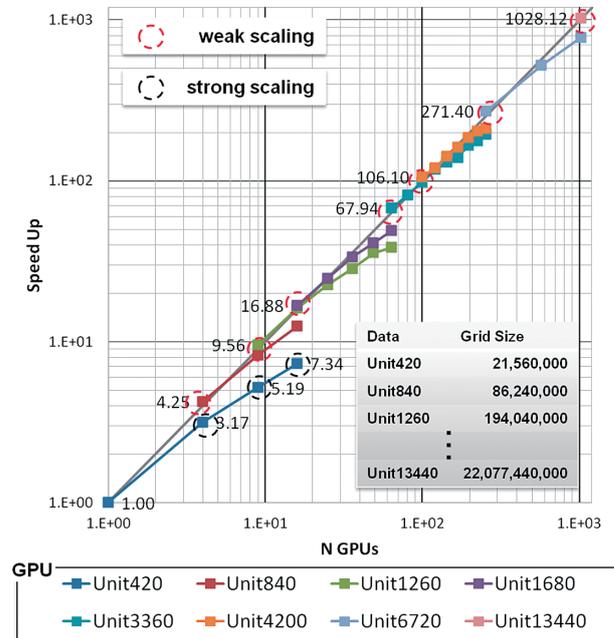


図5 並列数(N GPUs)と-高速化率(Speed Up)の関係。

これは、並列数の増加に伴い各GPUが担当する格子数が減少するため、相対的に通信の時間が大きくなること、およびGPUで生成されるスレッド数が減少することが要因として考えられる。通常行われるシミュレーションのタイム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。

また、実行時間と浮動小数点演算数を用いてFLOPS (Floating-point Operations Per Second)を算出した(図6)。GPU版プログラムでは、Unit6720(約55億格子)の256GPUによる性能が21.1TFLOPS、Unit13440(約220億格子)の1024GPUによる性能が79.7TFLOPSになることが確認された。CPU版プログラムではUnit13440の1024並列での性能が2.3TFLOPSであることから、約34倍の性能が得られたことになる。

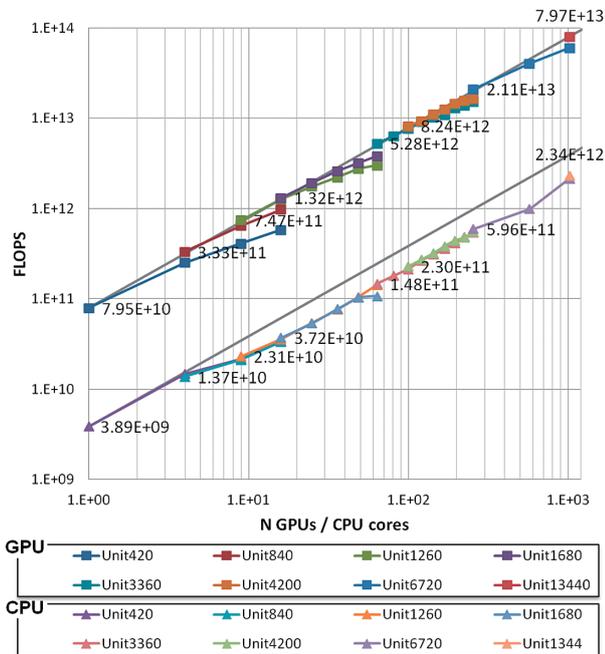


図6 並列数とFLOPSの関係 (GPU：GPU版プログラム、CPU：CPU版プログラム)。

### 海溝型巨大地震の長周期地震動の評価

# 3

本研究で開発したシミュレーションシステムを用いて、南海トラフで発生すると考えられる海溝型巨大地震を対象とした長周期地震動の試計算を行った。南海トラフの地震の震源域は、最大で駿河湾から日向灘まで及ぶと考えられており<sup>[6]</sup>、計算対象領域は九州から関東までを含む広大なものとなる。地震波伝播シミュレーションに用いる地下構造モデルと震源モデルは、次のように設定した。まず、地下構造モデルは、全国1次地下構造モデル<sup>[7]</sup>から、南北950km、東西1150km、深さ100kmまでの領域を切り出して、水平方向200m、深さ方向100m (8km以深ではその3倍の大きさ)の格子点間隔で離散化し、約32億点からなる格子モデルを構築した。また、震源モデルは、「南海トラフの地震活動の長期評価(第二版)について」<sup>[6]</sup>を参考に設定した。震源(破壊開始点)で始まった断層運動が、地震波を強く励起する領域(アスペリティ)を破壊しながら同心円状に拡がる様子を、約43万個の点震源によってモデル化し、差分格子上に配置した。

今回の計算では、このモデルを水平方向に9×9の部分領域に分割して、それぞれ1つのGPUを割り付けて計算処理を行った。計算結果は、GMSの1D Dump機能を利用し、地動速度の1成分を1ファイルとして東西、南北、上下の3成分について出力した。1ファイルには、陸域に約2km間隔で設定した約8万観測点の毎タイムステップ(全部で72000ステップ)の地動速度値が記録されており、1ファイルで約90GBである。計算された波動場を可視化した例として最大地動速度値の空間分布を図7に示す。これらは同じ震源域とアスペリティを持っているものの、破壊開始点の違いにより地震動の空間分布が大きく異なることがわかる。また、震源域から離れた堆積平野において長周期地震動が増幅される特徴が再現されている。

以上の計算の実行時間は、TSUBAME2.5の81GPU (27ノード)を使用して2.5時間程度であった。一方、同じ震源モデル、地下構造モデルに対して、防災科学技術研究所で使用している計算機(Altix4700)で252core (Intel Itanium 1.66GHz)を用いた場合の実行時間は約43時間であった。南海トラフなどで発生する海溝型巨大地震については、その発生パターンの多様性が指摘されており、多数の震源モデルに基づいた地震ハザード評価が必要となる。TSUBAMEを使用することで100ケースを超える多数の計算を現実的な時間で実行できる環境が得られたといえる。

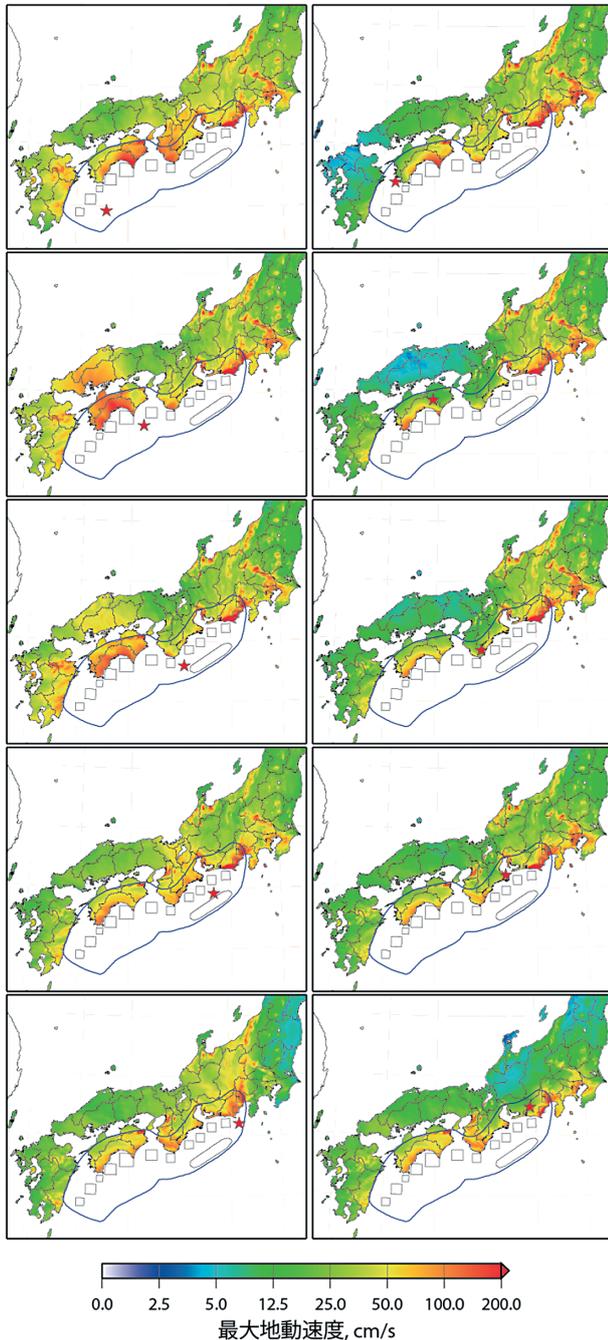


図7 南海トラフの最大クラス地震を想定した地震波伝播シミュレーションによる最大地動速度値の空間分布。青線で囲んだ領域は震源断層面、震源断層面内の黒線で囲んだ領域はアスペリティ、星印は破壊開始点を示している。

3次元差分法は、複雑な地下構造を考慮した地震波伝播シミュレーションにおいて最も標準的に用いられている手法である。本研究では、防災科学技術研究所が開発している3次元地震波伝播シミュレーションシステムGMSの差分計算ソルバのGPU化を行い、大幅な計算性能向上を達成した。東北地方太平洋沖地震の経験から、ハザード評価において巨大地震の発生パターンの多様性を考慮することが重要な課題となっている。そのためには可能性のある多数の発生パターンを対象とした高精度かつ大規模な地震波伝播シミュレーションが必要となる。これまでは計算機リソースの制約により困難であったこれらの計算がGPGPUを用いることで現実的な時間で行うことが可能となりつつあり、この環境を活かし、今後の地震ハザード評価を行っていく予定である。

#### 謝辞

本研究では、学際大規模情報基盤共同利用・共同研究拠点、および、革新的ハイパフォーマンス・コンピューティング・インフラの支援によりTSUBAMEを利用しました。また、本研究の一部は、文部科学省の「長周期地震動ハザードマップ作成等支援事業」の支援によるものです。記して謝意を表します。

#### 参考文献

- [1] 青井真・西沢直樹・青木尊之, “GPGPUを用いた三次元波動伝播シミュレーション”, 日本地震学会2009年秋季大会講演予稿集, A12-09, (2009)
- [2] D. Michéa, and D. Komatitsch, “Accelerating a three-dimensional finite-difference wave propagation code using GPU graphics cards,” *Geophysical Journal International*. vol. 182, pp. 389-402, (2010)
- [3] T. Okamoto, H. Takenaka, T. Nakamura, and T. Aoki, “Accelerating large-scale simulation of seismic wave propagation by multi-GPUs and three-dimensional domain decomposition,” *Earth, Planets and Space*. vol. 62, pp. 939-942, (2010).
- [4] S. Aoi, and H. Fujiwara, “3-D finite difference method using discontinuous grids,” *Bulletin of the Seismological Society of America*. vol. 89, pp. 918-930, (1999)
- [5] 青井真・早川俊彦・藤原広行, “地震動シミュレータ: GMS,” *物理探査*, vol. 57, pp. 651-666, (2004)
- [6] 地震調査委員会, “南海トラフの地震活動の長期評価(第二版)について”, 94pp. (2013)
- [7] 地震調査委員会, “「長周期地震動予測地図」2012年試作版—南海地震(昭和型)の検討—”, 59pp. (2012)

● **TSUBAME e-Science Journal vol.11**

2014年6月20日 東京工業大学 学術国際情報センター発行 ©  
ISSN 2185-6028

デザイン・レイアウト：キックアンドパンチ

編集：TSUBAME e-Science Journal 編集室

青木尊之 ピパットポンサー・ティラポン

渡邊寿雄 佐々木淳 仲川愛理

住所：〒152-8550 東京都目黒区大岡山 2-12-1-E2-6

電話：03-5734-2085 FAX：03-5734-3198

E-mail：tsubame\_j@sim.gsic.titech.ac.jp

URL：http://www.gsic.titech.ac.jp/

# TSUBAME

## TSUBAME 共同利用サービス

『みんなのスパコン』TSUBAME共同利用サービスは、  
ピーク性能 5.7PFlops、18000CPUコア、4300GPU搭載  
世界トップクラスの東工大のスパコンTSUBAME2.5を  
東工大以外の皆さまにご利用いただくための枠組みです。

### 課題公募する利用区分とカテゴリ

共同利用サービスには、「学術利用」、「産業利用」、「社会貢献利用」の3つの利用区分があり、さらに「成果公開」と「成果非公開」のカテゴリがあります。  
ご利用をご検討の際には、下記までお問い合わせください。

#### TSUBAME 共同利用とは…

他大学や公的研究機関の研究者の **学術利用** [有償利用]

民間企業の方の **産業利用** [有償・無償利用]

その他の組織による社会的貢献のための **社会貢献利用** [有償利用]

### 共同利用にて提供する計算資源

共同利用サービスの利用区分・カテゴリ別の利用課金表を下記に示します。TSUBAMEにおける計算機資源の割振りは口数を単位としており、1口は標準1ノード(12 CPUコア、3GPU、55.82GBメモリ搭載)の3000時間分(≒約4ヵ月)相当の計算機資源です。  
1000 CPUコアを1.5日利用する使い方や、100 GPUを3.75日利用する使い方も可能です。

利用区分	利用者	制度や利用規定等	カテゴリ	利用課金(税抜)※
学術利用	他大学または研究機関等	共同利用の利用規定に基づく	成果公開	1口: 120,000円
産業利用	民間企業を中心としたグループ	「先端研究基盤共用・プラットフォーム形成事業」に基づく	成果公開	トライアルユース(無償利用) 1口: 120,000円
			成果非公開	1口: 480,000円
社会貢献利用	非営利団体、公共団体等	共同利用の利用規定に基づく	成果公開	1口: 120,000円
			成果非公開	1口: 480,000円

※ 平成26年度の利用課金です。最新の利用課金については、下記 URL をご参照ください。  
<http://www.gsic.titech.ac.jp/kyodou/kakin>

### 産業利用トライアルユース制度 (先端研究基盤共用・プラットフォーム形成事業)

東工大のスパコンTSUBAMEを、より多くの企業の皆さまにご利用いただくため、初めてTSUBAMEをご利用いただく際に、無償にてご試用いただける制度です。

(文部科学省 先端研究基盤共用・プラットフォーム形成事業による助成)

詳しくは、下記までお問い合わせください。

## お問い合わせ

- 東京工業大学 学術国際情報センター 共同利用推進室
  - e-mail [kyoyo@gsic.titech.ac.jp](mailto:kyoyo@gsic.titech.ac.jp) Tel. 03-5734-2085 Fax. 03-5734-3198
- 詳しくは <http://www.gsic.titech.ac.jp/tsubame/> をご覧ください。