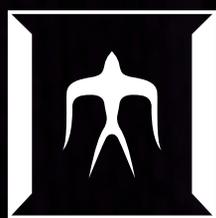
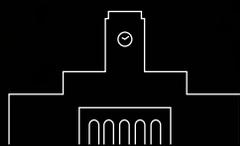


新世代に先駆ける「みんなのスパコン」 TSUBAMEの世界へ!

2013 年秋、多くの成果を上げた TSUBAME が、東京工業大学で大幅に性能向上。
ペタスケールかつ省電力を実現した日本を代表するスパコン “TSUBAME”。
その世界トップレベル性能の秘密はどこにあるのか?



Tokyo Tech
Supercomputer
UBiquitously
Accessible
Massstorage
Environment



東京工業大学
Tokyo Institute of Technology

圧倒的な計算性能を発揮する“ペタコン”

1442 個の計算ノードの計算性能合計は、5.7 ペタフロップス^{*1}=1秒間に 5700 兆回の浮動小数点演算が可能。

TSUBAME2.5 は TSUBAME 2.0 から GPU のアップグレードにより 2.4 倍の性能向上。

単精度での浮動小数点演算性能は 17 ペタフロップスと国内トップ！



スパコンはその国の科学研究や技術開発の基盤となる重要なインフラであり、国際的に激しい開発競争が繰り広げられています。2010 年 11 月に運用スタートした TSUBAME2.0 は理論性能が 2.4 ペタフロップス。日本初のペタフロップス・スパコンを実現し、スパコンの絶対性能を比較する世界ランキングである Top500^{*2}でも世界 4 位を獲得したのです。

その性能の秘密は、GPU を取り入れたハイブリッド型のアーキテクチャ。1400 個以上もある計算ノードは厳選された高性能な GPU、CPU、メモリなどの部品から構成されています。このような強力な計算ノードが大量に、しかも高速ネットワークでつながれているのです。2013 年 9 月には、更なるスパコンへの需要の高まりに対応するために GPU を最新の Kepler 世代 GPU にアップグレードを行い、TSUBAME2.0 の約 2.4 倍・5.7 ペタフロップスの性能となる TSUBAME2.5 へと進化しました。



■ GPU^{*3} 採用によるハイブリッド設計

業界標準 CPU によるスカラー演算と GPU によるベクトル演算を、大規模スパコン仕様で組み合わせた「混合ハイブリッドアーキテクチャ」が最大の特徴です。この発想の転換が、優れた計算性能を生み出したのです。



■ 高性能メモリ

計算ノードに組み込まれ CPU と GPU にデータを提供するメモリにも、高い性能が要求されます。合算メモリバンド幅は 720TB/s と、TSUBAME1.0 の 17TB/s に比べて 40 倍になっています。



■ 高速ネットワーク

各計算ノード間はハイエンド PC に比べて 80 倍速いネットワークで接続しています。さらに、世界最高速級のバイセクションバンド幅毎秒 200Tbps 以上を実現可能な光ネットワークを採用しており、すべてのノードが同時通信を行っても混雑が生じません。

※1：Flops(フロップス) / Mega(メガ)、Giga(ギガ)、Tera(テラ)、Peta(ペタ)、Exa(エクサ)

Flops は Floating point number Operations Per Second の略。1秒間に浮動小数点演算が何回できるかを表したもので、主に科学技術計算やシミュレーションなどを行うスパコンの計算性能を表すために使われます。ちなみにメガは 10 の 6 乗 (百万)、ギガは 10 の 9 乗 (十億)、テラは 10 の 12 乗 (兆)、ペタは 10 の 15 乗 (千兆)、エクサは 10 の 18 乗 (百京) ですから、1 ペタフロップスは 1 秒間に浮動小数点演算が 1 千兆回できることとなります。

※2：Top500

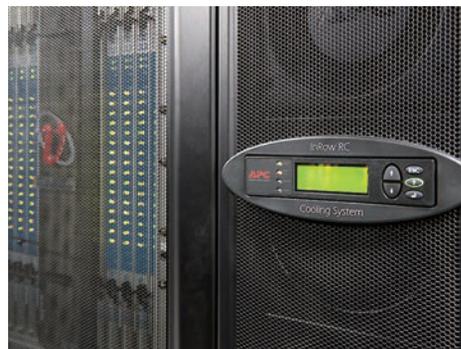
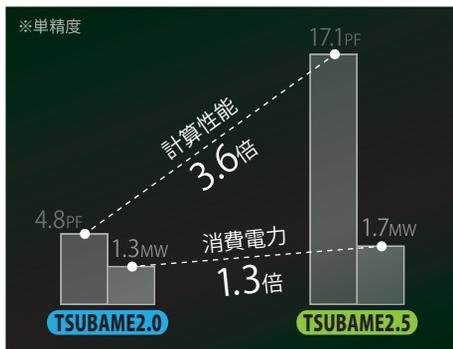
Top500 は世界で最も高速なコンピュータシステムの上位 500 位までをランキングするプロジェクトで、リストは毎年 6 月と 11 月に更新されます。その規準となるのが、コンピュータ上で線形代数の行列計算を行う LINPACK ベンチマークプログラム。コンピュータシステムの浮動小数点演算性能を評価します。もちろん Top500 のランキングは評価指標の一つです。ただ、車でいえばエンジン全開でテストコースを回り続けるような性能評価ですので、最も過酷な使用状況での評価といえるでしょう。

※3：GPU

Graphics Processing Unit (画像処理装置) の略。その名の通り、本来は画像処理に使われる装置です。画像処理では簡単な計算を数多くこなす必要があるため、GPU には CPU に比べて低い電力で動く単純な演算コアが多数搭載されています。そのため高い並列計算能力をもち、エネルギー効率がよいという特徴があります。TSUBAME は GPU を CPU と組み合わせて用いることで、計算を高速化し、同時に電力効率を向上させています。

世界一の“グリーン・スパコン”

TSUBAME の設計では消費電力の増加の問題に真正面から取り組みました。省電力に最も効果を発揮した GPU の本格活用や新冷却システムなどの技術的工夫で、大幅な省電力を実現。2010年11月の Green500 List^{※4}で、運用中スパコンとして世界一省エネという称号を獲得しました。TSUBAME2.5 では GPU のアップグレードによりさらなる電力効率の向上。



■ペタフロップスマシンなのにパソコンより省エネ
TSUBAME は、消費電力を抑えつつ計算能力を上げるさまざまな技術を導入。その結果、TSUBAME2.0 に比べて3.6倍の性能を持ちながら、消費電力の増加はわずか。ペタフロップスマシンなのに消費電力あたりの演算性能は通常のパソコンの約3倍という省エネ機となりました。

■性能向上に加え、省スペースも実現
CPU と比べチップあたりの演算性能の高い GPU を組み合わせることで高密度の実装が可能に。設置面積は TSUBAME1.0 の 3分の2に小さされました。スパコンの場合、設置に体育館並みのスペースを要するのが普通ですが、TSUBAME2.5 は教室2つ分とコンパクトに納まっています。

■スパコンは水で冷やす！
最先端の冷却設備として、ラック内に熱交換装置を内蔵した密閉型の水冷システム^{※5}を採用。電力使用効率 (PUE)^{※6}が大幅に低減できただけでなく、空冷ファンから発生するノイズも閉じ込めることができ、格段に静かになりました。扉にもそのための工夫がなされています。

巨大なデータを蓄え、 超高速でアクセス可能な“みんなのスパコン”

TSUBAME 1.0 の 10 倍超！ 11PB を超える階層的ストレージで、圧倒的な最大容量を誇ります。
この TSUBAME2.5 のストレージはネットワーク経由で日常的にも使えます。
東工大生は個人のパソコンからログインし、TSUBAME のストレージを利用できるのです。



■これが階層的ストレージの実力だ

ストレージの第1階層では、各ノードに備えたスクラッチ用の合算 190TB の SSD で、毎秒 660GB 以上の超高速を実現。第2階層は 6PB を超える並列ファイル領域で、ペタバイト級のデータのやり取りを高速化。第3階層は 4PB からなるテープライブラリーシステムで、ファイルシステムとテープライブラリー間で透過的かつオンデマンドなデータアクセスが可能。学内はもとより学外の人による共同利用など、まさに「みんなのスパコン」としての利用が可能になっています。

■学内向けクラウド^{※7}ストレージ

学内クラウド用の 1.2PB のホーム領域は、さまざまなところからアクセス可能。東工大生は、学食や学内のコンビニからも無線 LAN を使って、自分のパソコンからアクセスし、まるで自分のパソコンの一部のような感覚で手軽に使っています。学生のストレージ容量は一人当たり保存用 25GB、計算用 25GB の計 50GB です。

※4：Green500

世界のスパコンの電力使用効率の良さを争うランキング。Top500 の計算性能を消費電力で割った「燃費」で順位を競うもの。ペタフロップスマシンで、2010年11月の Green500 で 10位以内にランクされたのは TSUBAME だけ。TSUBAME は、レーシングカーが軽自動車以上の燃費を実現したようなものなのです。

※5：密閉型の冷水システム

ラック内に熱交換システムを内蔵しています。サーバの吸入口に均質な冷却風を送り、94~97%の熱を水冷で取り除きます。完全自動温度制御で最適温度を保つとともに、冷却の消費電力も必要最小限に抑えることができます。スパコンが置かれた部屋全体を冷やすよりも、はるかに省電力となります。

※7：クラウド

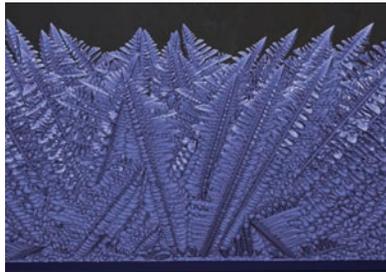
クラウドコンピューティングのこと。ユーザーがインターネット経由でさまざまなサービスまたはストレージなどコンピュータの提供する機能を、特にハードウェアやソフトウェアを意識せず利用できる形態をさします。ユーザーからは、インターネットとコンピュータが漠然とした雲（クラウド）のように捉えがたく感じられることからこのように呼ばれます。TSUBAME では OS を動的に変更する機構を使い、余剰の計算資源でクラウドサービスを行っています。たとえば学内の教育系・業務系システムが TSUBAME 上で動いているのです。

※6：電力使用効率 (PUE) Power Usage Effectiveness の略。計算用電力+冷却等電力を計算用電力で割ることで求められます。冷却に要する電力が減り、値が 1.0 に近づくほど効率がよいということになります。東工大では、TSUBAME3.0 に向けてさらなる省電力技術の研究開発を進めています。

TSUBAME2.5の切り拓く新たな研究の最前線

学術研究利用や産業利用でも高い評価を得た TSUBAME2.0 より、さらに高速な GPU 計算で、世界最先端の研究が可能になりました TSUBAME2.5 は、TSUBAME2.0 での経験がそのまま通じるため、短期間で多くの研究・産業上の成果が見込まれています。

ここでは東工大内外で現在進行中の最先端研究を紹介します。



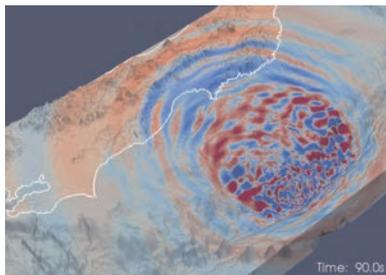
新しい合金材料開発のための凝固成長シミュレーション

自動車や鉄道等、私たちの生活の至る所で合金材料が使われています。これらの合金材料の強度は、マイクロな材料組織に強く依存し、そのマイクロ組織は凝固成長過程で形作られます。このため、新しい合金材料の開発では、凝固成長を予測する数値シミュレーションが必須です。TSUBAME では、4000GPU を用いた 2.0PFlops (単精度) という高性能計算に成功し、数 mm スケールという科学的に意味のある大規模計算に成功しました。この成果に対し、2011 年スパコン分野での最高栄誉と言われる ACM ゴードン・ベル賞が贈られました。



この研究が進むと？

数値シミュレーションによって自由に材料組織の設計ができるようになれば、用途にあった軽量で高強度な新材料を開発することができます。軽量な自動車や鉄道等により物資を低燃費で輸送することができれば、低炭素社会の実現に大きく貢献します。



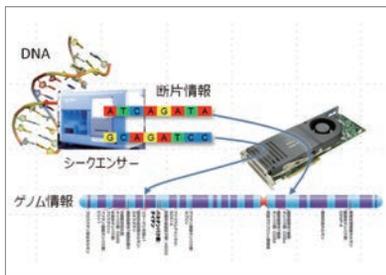
シミュレーションによって地震の揺れを予測する

地震の波は震源から放射され地球内部を伝わって地表に到達します。震源自体が複雑であることや地球内部の不均質性のために地震の揺れは非常に複雑なものになります。このような地震の揺れをあらかじめ予測することは地震科学の重要な課題の一つです。そこでコンピュータの中に「地球」を作り、その中を伝わる地震の波を理論的に計算する地震波シミュレーションの研究を TSUBAME で進めています。TSUBAME では多くの GPU により地震波シミュレーションを非常に高速に実行できるからです。



この研究が進むと？

地震の揺れの予測が進むことによって地震災害の想定が進展することに貢献できます。さらに観測データも用いると、震源自体の複雑な性質を高い精度で推定することも可能となります。そのことは地震の揺れの予測の高精度化にもフィードバックされます。



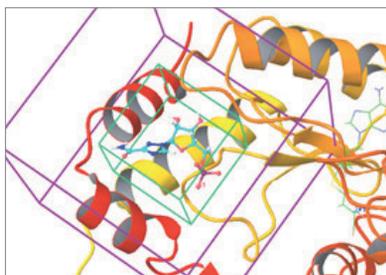
環境研究や医学研究に貢献する超高速メタゲノム解析

DNA 配列の読み取り技術の劇的な発展により、ヒトや多数の動植物について、生命の設計図であるゲノム情報が急速に集められています。近年ではさらに、海や川、田畑、工場、住宅等、さまざまな環境内で共存している膨大な種類の微生物のゲノムを環境ごとに丸ごと一気に解析するメタゲノム解析が盛んです。大量の DNA 断片配列を、これまでに知られている遺伝子と高精度に比較するため、GPU を活用した効率的なアルゴリズムを開発し、従来より数桁高速なメタゲノム解析の環境を開発しました。



この研究が進むと？

水源や農地や処理場などの環境のモニタリングに利用できるほか、ヒトの体内における微生物のバランス等の調査にも活用できます。たとえば、腸内や口腔内の細菌のバランスや遺伝子活動の変化は、特定の疾病の起きやすさとの関連も指摘されています。



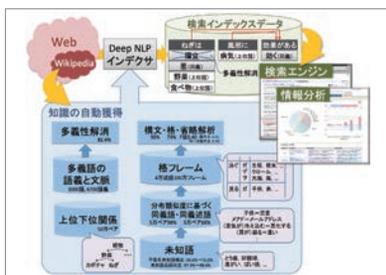
インシリコスクリーニングによる治療薬探索

私達の体の中で、タンパク質は他のタンパク質や化合物、ペプチドと結合することにより、さまざまな機能を発揮します。インフルエンザウイルス・HIV ウイルスなどもタンパク質等で構成されており、私達の体のタンパク質の働きを妨害しないで、ウイルスを構成するタンパク質の働きを妨害する化合物を見つけることが出来れば、これらの病気の治療薬候補となります。TSUBAME を用いることでタンパク質と化合物の相互作用を高速に見積もること出来、インシリコスクリーニングによる創薬支援が可能になります。



この研究が進むと？

インシリコスクリーニングを実現することで、抗ウイルス薬や抗原虫薬の探索が可能になります。特に今まで生化学実験でコストがかかっていた部分等を効率化することで、「顧みられない熱帯病」のように創薬が行われなかった病気の治療薬開発が可能になると考えられます。



超大規模ウェブテキスト集合からの知識獲得とその利用(京都大学 河原大輔准教授)

SF 小説、映画では、コンピュータやロボットが人間と自然にコミュニケーションをしています。このような世界を実現するには、それらがことばを理解する必要があり、このためには人間がもっているような常識的な知識をもつことが不可欠です。近年、ウェブを中心に膨大な量のテキストが集積されるようになり、ここから知識を自動的に獲得することが可能になりました。TSUBAME の並列計算環境を利用することによって、超大規模のウェブテキストからの知識の自動獲得とその利用について研究を進めています。



この研究が進むと？

このような知識を生かしたコンピュータプログラムによって、私達が日々使っている検索エンジンならびに自動翻訳などが飛躍的に賢くなります。さらに、人間と自然にコミュニケーションできるコンピュータやロボットも実現できる日が来るでしょう。



スパコンが拓く未来 一少ない水で汚れを落とす TOTO 株式会社の挑戦

ものづくりの現場においても、スパコンは大いに活用されています。少ない水で汚れを落とす環境に優しい節水型サニタリー製品開発のため、TOTO 株式会社は TSUBAME を利用し、自社で開発した空気と水の複雑な界面を精緻に計算する気液二相流体解析の、精度と計算速度を飛躍的に向上させました。東京工業大学 学術国際情報センターは『みんなのスパコン』TSUBAME 産業利用制度により、先進的企業の研究開発のため、多くの計算資源を提供しております。



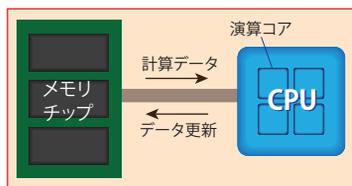
この研究が進むと？

企業の生産活動においてシミュレーションは不可欠であり、シミュレーションに対する要求はより高度になっています。皆さんの想像のおよばない、ものづくりの先端でスパコンが活用されています。皆さんの安全のため、便利のため、安心のために。

.....スパコンの謎にせまる!



そもそもスパコンって何? パソコンと何が違うの?



パソコンとスパコンの基本的な仕組み。スパコンは、この部分を大量に用意してつなぎ、並列に使用します。パソコンの場合は、大抵1つ。

みなさんの使うパソコン(PC)もスパコンも、基本的な仕組みは同じです。CPU(演算装置)と呼ばれるチップ(電子部品)内の演算コアが、メモリ・チップにあるデータを処理・更新する計算を繰り返し、その結果や途中経過が数値や画像として表示されたり通信されたりします。

一昔前までのスパコンは、特別仕様の高機能なCPUやメモリを使い、普通のパソコンにはできない、きわめて高速な計算を行うように設計されました。しかし、電子部品の高速化がそろそろ限界に近づく一方、普通のパソコンに使われるチップも、以前のスパコンに匹敵するほどになりました。

現在のスパコンでは、単に高機能なチップを使うだけでなく、それを大量にしかも並列に使うことで計算能力を上げています。皆で手分けをして一斉に仕事をすれば処理能力が上がるのと同じです。TSUBAMEには約17,000個の演算コア(CPUコア)が使われています。しかも各CPUには、非常に大容量のメモリがつながっています。

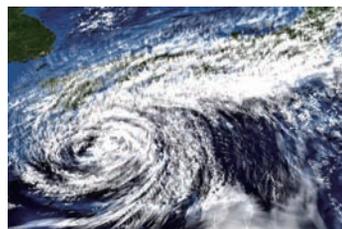
そうするとCPU同士が非常に高速に通信できる仕組みが必須です。重要なのは演算・メモリ・通信をバランス良く設計すること。そこではじめて超高速・超高性能計算が可能となるのです。

スパコンはまだ速くなる必要があるの?

スパコンの最も重要な利用法はシミュレーションです。台風や津波など、私たちの身の回りの自然現象は、方程式を用いて表わすことができます。しかし、その方程式を数学だけで解くことができないため、台風がどれだけ成長しどんな進路を取るのかを予測することができません。そこで、たとえば東京を5km×5kmのマスキに区切り、各マスで雲の発生や風の方向を数値で計算し、眼に見える形で表すシミュレーションが重要になってくるのです。

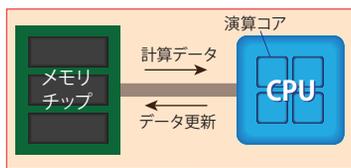
こうしたシミュレーションはスピードが命。被害を防ぐために台風の動きを知りたいのですから、通過後に計算結果がわかっていても意味がありません。そこでスパコンが必要になるというわけです。さらに計算が高速になるほど同じ時間内に多くのデータを扱えます。先ほどの台風ですが、細かいマスキで区切って計算するほど、現実に近いシミュレーションができるようになります。右図下は500m×500mのマスキで分けたものですが、上と比較していっそうリアルに台風を表現できるのがわかります。TSUBAMEならばそれが可能なのです。

また最近では、とにかく多量のデータを処理しなければ研究にならない分野も登場しています。遺伝子解析などに代表されるバイオ情報分野もそのひとつで、多量のデータや莫大な組合せを分析するためにスパコンが活躍しています。ここでも急速に増えるデータ処理のため、いっそうのスピードアップが望まれています。

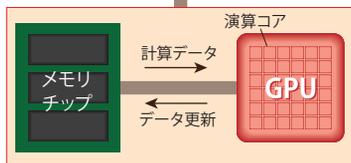


細かなマスキで分けて計算した方が(右図)、雲が詳細に表現されています。

TSUBAMEの特徴 GPUって何?



データ更新 ↑ ↓ 計算データ



GPUは、CPUから計算データを受け取り、より多くの演算コアで高速に処理を行います。

高速化のために大量のCPUを使おうとすると、スパコンはどんどん大きくなります。いまや最先端のスパコンには、体育館くらいの大きさのスペースが必要ですし、消費電力も莫大です。

もっと小さな演算コアはないのでしょうか? ありました!

パソコンなどで画面を制御するために使われているビデオカードに使われるGPUです。GPUは非常に多くの演算コアを用いて並列に画像データを高速処理するように設計されています。CPUほど自由はききませんが、数百の演算コアがとても小さくまとまっているではありませんか。これは使えるかも!?

それをスパコンで最初に試みたのが、先代のTSUBAMEで、結果は大成功でした。そこでTSUBAME2.0は、本格的にGPUを取り入れたスパコンとして設計されました。アップグレードされたTSUBAME2.5では、全体で4224枚のGPUカードが使われており、1個のGPUには2688個の演算コア(CUDAコア)が入っているので、その総数は何と1135万個にもなります。このお蔭で、TSUBAMEは教室2つ分の大きさで世界有数の計算能力を実現することができたのです。



未来を創り出すTSUBAMEのチャレンジ

教育分野での活用

TSUBAME は大学教育への活用はもちろん、広く国内の情報科学教育に利用されています。



Super Con (スーパーコン) 夏の電腦甲子園

スパコン TSUBAME を活用し、高校生・高専生を対象とした 2~3 人のチームによるプログラミングコンテストを毎夏開催しています。予選を通過した 20 チームが数日間かけて課題の問題を解くプログラムを作成。その正確さと速さを競います。1995 年から始まった歴史あるコンテストです。スパコンも原理はパソコンと一緒にですが、パソコンが乗用車ならスパコンはレーシングカーに相当します。そのスーパーな性能を引き出すプログラミングのアイデアと技術が問われるのです。本選の課題には、科学技術のさまざまな分野から最先端の問題が選ばれます。高校生向けにわかりやすくしてはありますが、なかなかの難問です。

Supercomputing Contest <http://www.gsic.titech.ac.jp/supercon/main/attwiki/>

過去5回の本選課題と受賞校

	[1位]	[2位]	[3位]
2013年 第19回 宇宙の起源探索問題	imishinn (久留米工業高等専門学校)	YOSUBATO (麻布高等学校)	Answer (早稲田高等学校)
2012年 第18回 非出会い系最長乗車パターン計算問題	ima1000 (灘高等学校)	PANAI (開成高等学校)	UNAGEEL (栄光学園高等学校)
2011年 第17回 落ち物ゲーム問題「なくろん」	PANAI (開成高等学校)	YAMERO (甲陽学院高等学校)	H1TOHA (早稲田高等学校)
2010年 第16回 レンガ敷き詰め問題	zatoriku (筑波大学附属駒場高等学校)	warosu (早稲田高等学校)	nemui (八千代松蔭高等学校)
2009年 第15回 恒星写真探索問題	zatoriku (筑波大学附属駒場高等学校)	potassio (筑波大学附属駒場高等学校)	H5N1 (八千代松蔭高等学校)

学部生からスパコンを使える教育用システム

東工大では学生証に IC カードを用いており、学内の情報機器へはこの IC カードでアクセスできるようになっています。もちろん TSUBAME も学内情報機器のひとつ。東工大生は 1 年生から TSUBAME にログインできるのです。また学部教育で用いられるコンピュータシステムも TSUBAME が背後で支える構造で、演習室の PC と TSUBAME が境目なく利用できる仕組みになっています。

この環境を活かし、東工大では学部・学科を問わず、スパコンを使ってさまざまな研究に挑戦していくテクニックを学べる授業を提供しています。

新しい計算手法を学ぶ公開講座

TSUBAME に用いられている GPU はさまざまな可能性を持っていますが、その性能をうまく引き出すには技術が必要です。GSIC が運営している GPU コンピューティング研究会では、GPU コンピューティング (CUDA) 講習会を開き、その技術を基礎から学ぶ機会を提供しています。高校生・高専生も受講可能です。



GPU コンピューティング研究会 <http://gpu-computing.gsic.titech.ac.jp/>

スパコンのさらなる進化をめざして

国内の大学や研究所と協力しつつ、TSUBAME を活用した次世代スパコンへの挑戦がすでに始まっています。



次世代スパコンとの連携

2012 年には 10 ペタフロップス級のスパコン「京」が、神戸に誕生しました。いま多くの研究者が一丸となって、この「京」の能力を最大限に活かすための仕組みを作ろうとしています。全国のスパコンを持つ国立大学のコンピュータセンターや研究機関も「京」を中核とするスパコン連携である HPCI に参加し、TSUBAME もその一環となって活躍しています。



全国の国立大学系スパコンセンター



TSUBAME3.0 へ向けた更なる最先端研究

これからのスパコンでは消費電力の更なる削減が重要な鍵となります。TSUBAME2.5 の後継機である TSUBAME3.0 では、スパコン本体の省電力化も推進するのに加え、冷却に必要な電力を限りなく 0 に近づけることを目指します。そこで GSIC ではスパコンの超省電力化技術の研究開発プロジェクトを 2011 年度より開始。その一環として、液浸冷却と大気冷却の融合という新しい手法の実験を、コンテナデータセンター型の実験設備である TSUBAME-KFC により行っています。



学際大規模情報基盤共同利用・共同研究拠点

東京工業大学 学術国際情報センター

Global Scientific Information and Computing Center, Tokyo Institute of Technology

〒152-8550 東京都目黒区大岡山 2-12-1

TEL : 03-5734-2087 FAX : 03-5734-3198 MAIL : office@gsic.titech.ac.jp

URL : <http://www.gsic.titech.ac.jp/>

