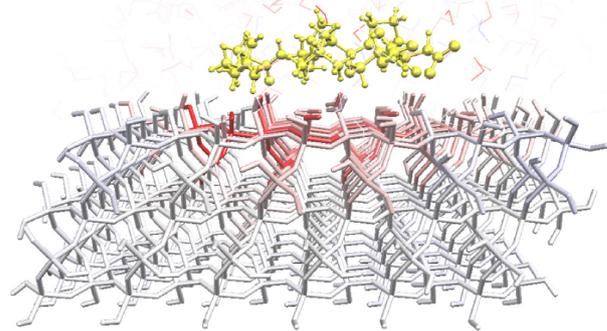
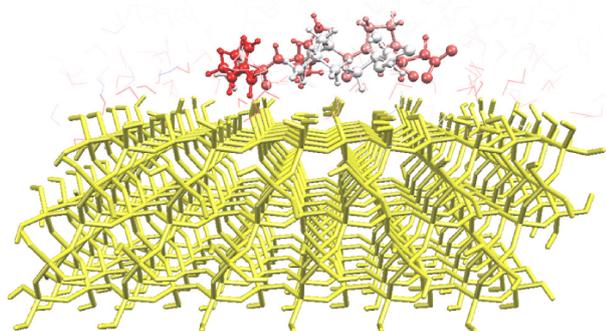
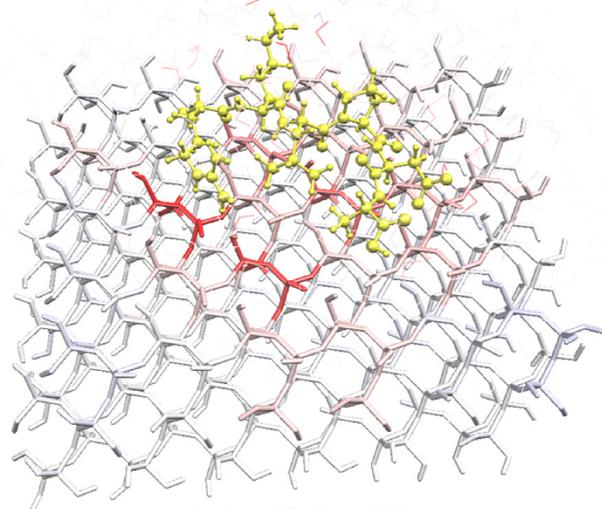
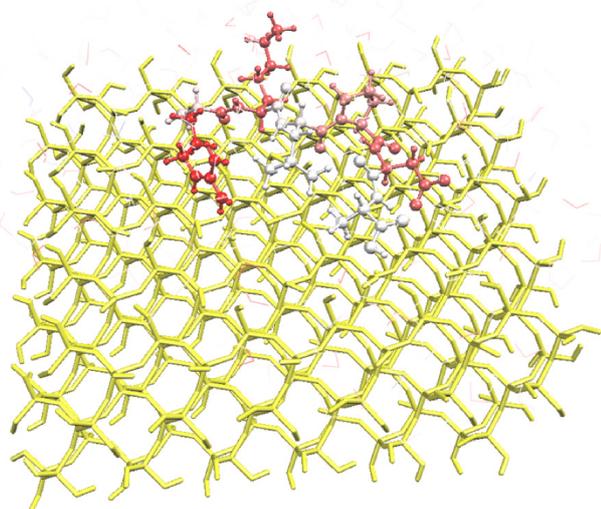


ものづくりにHPCを活用するためのツールとケーススタディー

計 算 工 学 ナ ビ

QUARTERLY NEWSLETTER / FALL 2014



Y. Okiyama et. al., Chemical Physics Letters 566 (2013) 25-31.
Reproduced from the reference with permission from Elsevier.

フラグメント分子軌道法プログラム ABINIT-MP

創薬からものづくりへ広がる FMO 法の可能性

立教大学 望月祐志 / 日本大学 福澤 薫

HPCI 戦略プログラム分野4『次世代ものづくり』

開発進む HPC プラットフォーム

ソフトウェアライブラリ

FrontFlow/violet

VOL.

5

創薬からものづくりへ広がる ABINIT-MPの可能性



立教大学 理学部化学科
反応解析グループ 教授
望月祐志

日本大学 松戸歯学部
化学教室 助教
福澤 薫

非経験的(ab initio)に分子の電子状態を計算する手法のひとつである「フラグメント分子軌道法」。それを使いやすく実装した国産プログラム"ABINIT-MP"が創薬やものづくりの分野で注目されています。開発の中心にいるふたりの研究者に、ABINIT-MPの特徴、応用分野、今後の展開について聞きました。

フラグメント分子軌道法（FMO法）はどのような意図で提案されたのですか？

タンパク質や水の分子が数百個から数千個ある凝集系を普通の分子軌道計算と同じ精度で実用的に計算したいというのが、1999年に北浦先生がこの手法を提案された意図です。大きな分子をフラグメントという部分部分に分解して計算することで、高い並列性を実現しながら、タンパク質のいろんな機能や描像が細かく精度よくわかるという特徴があります。

FMO法の実装系のひとつであるABINIT-

MPの特徴を教えてください。

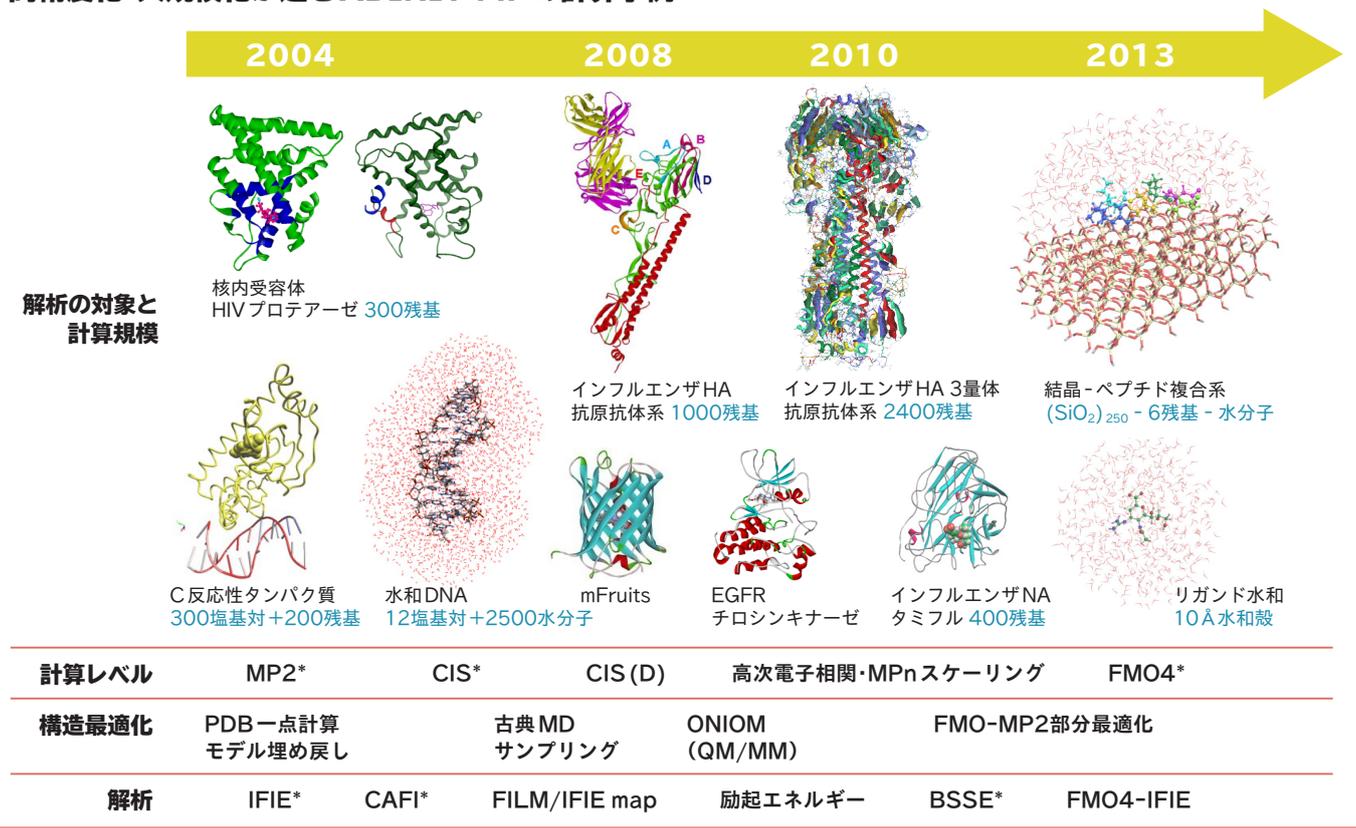
FMO法ではフラグメントを1量体（モノマー）と2量体（ダイマー）に分けて個別に計算を行うのですが、我々のABINIT-MPは3量体（トリマー）と4量体（テトラマー）にも対応しています（FMO4法）。この拡張によっていろんなことができるようになりました。たとえば、薬の分子がタンパク質に入ったとき、どのくらい安定に結合しているかを官能基ごとに細かく見ることができます。こういった解析的な使い方ができるところがFMO法のアドバンテージで、フラグメント化が計算を加速すると

同時に、そのまま解析ツールとしても使えるということを証明しています。

4量体まで扱えると計算量は膨大になるのではないのでしょうか？

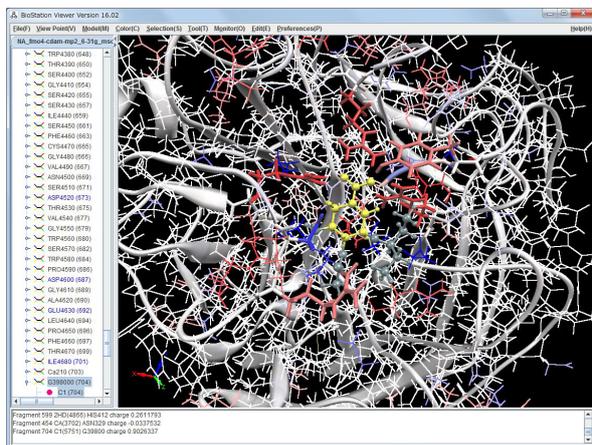
研究の結果、有効な4量体だけを計算すればいいということが分かって、8倍から10倍くらいの計算コストで実現することができました。高速化の過程で、京のような共有メモリ型の計算機に対応したことも重要です。フラグメント内の計算は共有メモリに適したOpenMP、フラグメント間は従来通りMPIで並列化しています。も

高精度化・大規模化が進むABINIT-MPの計算事例



MP2: Moller-Plessetの2次の摂動法
CIS: Single Excitation CI(配置間相互作用)法
FMO4: 4量体を含むFMO計算

IFIE: フラグメント間相互作用エネルギー
CAFI: 軌道相互作用解析
BSSE: 基底関数重ね合わせ誤差の補正



【図1】BioStation Viewer

分子を切断しフラグメント化する作業や、ABINIT-MPが生成した膨大なデータを可視化する段階では、自動処理と対話的な操作の両方に対応するユーザーインターフェイスが不可欠。BioStation ViewerはABINIT-MPと並行して開発が進められてきた専用GUIである。

うひとつ京で性能を出すために重要だったのが、コレスキー分解という行列計算の手法です。これによって64コア程度の一般的なPCクラスターでも実用的に使えるレベルまで高速化できました。

タンパク質を丸ごと精度よく解析できることで、どのような応用が可能になったのでしょうか？

早い段階から創薬に使えるという話がありました。女性ホルモン受容体を50残基にモデル化したタンパク質の電子状態を丸ごと計算したのが初期の例です。これを発表すると、製薬企業の方には衝撃だったらしく「ぜひやりたい」という声があがりました。創薬に繋げる上で重要なのは相互作用の解析です。薬の分子とまわりのアミノ酸残基との間の相互作用のエネルギーが分かれば、より良い医薬品を作る上での指標として使えます。

2010年には、インフルエンザ表面タンパク質とタミフルの相互作用を解析しました。なぜタミフルでやったかという、当時、薬剤耐性の問題が始まっていたからです。耐性をもつウイルスは化合物から認識されないようタンパク質が変化しています。それに対応するためには化合物側の構造を変える必要がある。つまり、タンパク質と化合物からなる系を丸ごと解いて、精密に相互作用の様相を捉える必要があります。こうした用途には4量体の計算までやって精度を担保しないとイケない。ABINIT-MPならそれが可能です。

創薬分野ですでに実用レベルと考えていてでしょうか？

はい。すでに十数社の国内企業が利用しています。2014年11月には『FMO創薬コンソーシアム』という新団体を設立しました。京やHPCIを用いて出てくるデータをまとめて、創薬に結びつける産学協同の場とすることが目的です。

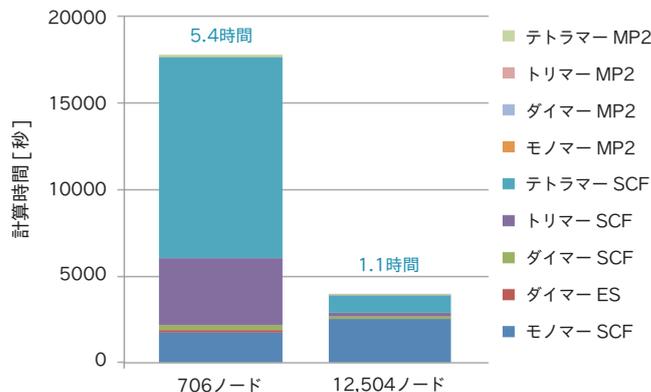
また、大規模・超高速FMO計算の結果をSPring-8などのX線構造解析施設に提供して、相補的に精度向上を図る試みも始まっています。計算と実験の連携により、構造決定がより難しいタンパク質の理解を進める意味があります。

ABINIT-MPをものづくり分野に適用する研究も進められていますね。

大規模な系であっても相互作用の細かいところが分かるABINIT-MPの特徴はものづくりにも有効です。すでにペプチドとシリカの相互作用を4量体のFMO計算で解くことに成功しています。これはナノバイオ系の界面の問題をシミュレーションで解明する試みと言っていいでしょう。具体的にはドラッグデリバリーやインプラントの技術開発に繋がります。

もうひとつ、ものづくり分野での新しい取り組みとして紹介したいのが、粗視化シミュレーションへの応用です。

膜やプラスチックを作るとき、最適な条件を求めるために使われてきた手法が粗視化動力学計算です。分子を4~5個ずつま



【図2】京コンピュータによる10万コア規模の計算にも対応

インフルエンザノイラミニダーゼとタミフルの複合系(386残基、706フラグメント)の4量体のFMO計算(FMO4)を京コンピュータ上でOpenMP/MPIの混成並列で実行し、各ステップの内訳を調べたグラフ。12,504ノード(100,032コア)ではモノマーSCF以外で高い並列処理能力が発揮されることが分かる。SCFは各フラグメントの分子軌道の組を最適化するSelf-Consistent-Field法、MP2は電子相関を記述する2次のMøller-Plesset法のこと。

とめて、一連のピースのように扱い計算するこの手法について、業界には長年の蓄積があるのですが、従来のパラメータ算定の方法には経験的な面があって、うまくいかない系もあります。ピース間の関係をうまく描けないことがあるんですね。たとえば電池系の膜にはそういうことが多い。FMO計算によって非経験的にパラメータを求めることでこの問題を解決できる可能性があります。予備的なテストとしてニトロベンゼン+ヘキサン系を対象に、数千サンプルを統計的に処理するパラメータ算定のプロトコルが有効であることを確認済みです。

最後に、セミナーの実施やプログラム公開など、ABINIT-MPの普及活動について教えてください。

2014年には高度情報科学技術研究機構(RIST)と共催で2度のセミナーを実施しました。専用GUIであるBioStation ViewerとFX-10上で動作するABINIT-MPを使うハンズオン形式のセミナーです。今後もこうしたセミナーを実施していく予定です。

プログラムの公開については、最新のIntel版バイナリーの提供を2014年7月に開始しました。現在はソースコードの公開に向けてマニュアル作成などを進めている最中です。詳細については『計算工学ナビ』のWebサイトでお知らせする予定です。

HPCI 戦略プログラム分野4における 成果公開・普及促進活動の状況



東京大学生産技術研究所
特任教授
畑田敏夫

HPCI戦略プログラム分野4『次世代ものづくり』は、ものづくりプロセスの質的・時間的ブレークスルーを実現するため、先端的なシミュレーション技術の開発に取り組んでいます。その成果が産業界で広く利用されることを重視するプロジェクトでもあり、成果公開・普及促進のために多くの施策が実施されています。

HPCI戦略プログラムは5つの分野からなる国家プロジェクトです^[図1]。そのなかで、分野4『次世代ものづくり』の特徴はどういったところにあるのでしょうか。プロジェクト内で成果の公開と普及に取り組んでいる東京大学生産技術研究所の畑田敏夫特任教授は次のように説明します。

「分野4には、ほかの分野とかなり違う特徴があると思います。ものづくりにおいて、シミュレーションは目的ではなく、あくまでも手段なんですね。スパコンを使って大規模かつ高精度な解析を行うのは、そこから得られる成果を実際の製品開発に適用するためです。結果をいかに利用するか、という所に踏み込まないといけない」

先端的で実用的なシミュレーション・ソフトウェアを開発することが、分野4の中心的課題ですが、それと同程度に重視されているのが、開発されたソフトウェアを産業界が実際に活用し、成果を享受できる環境作りです。

「分野4には2つの柱があります。ひとつは技術的な課題を設けてその研究開発をや

る。もうひとつは、その成果を活用、普及させることです。いかに成果を使いやすくするか、成果を活用できる人材をどうやって育てるか、いかにして成果を知ってもらうか……まず知ってもらいファンになってもらうことが重要で、企業の方に、これはうちでも使えそうだ、という認識を持ってもらえれば、その先はきっとうまくいくと考えています」

アウトリーチ活動とHPC/PF

分野4におけるアウトリーチ活動の主な目的は「HPC利用のメリットの明確化」。そのための施策としてまずあげられるのが、この『計算工学ナビ』です。webサイトを通じて、様々な立場のユーザーに向けてプロジェクトの成果を伝えています。さらに計算工学ナビは知識データベースを含むポータルサイトとしての側面も有します。「シンポ等で研究者の発表を聞いても、それだけでは自分のところで使えるかまでは分からないでしょう。知見をデータベース

化してほしいという要求が強くあります。京はいちばん大きな計算機ですけど、我々は京以外にもいろんなスパコンを使って解析している。その結果もDB化して順次公開していきます。そのDBは、スパコンと分野4のソフトを使うことで自分たちの製品がどう変わるのかを実感として認識するのに役立つはずですよ」

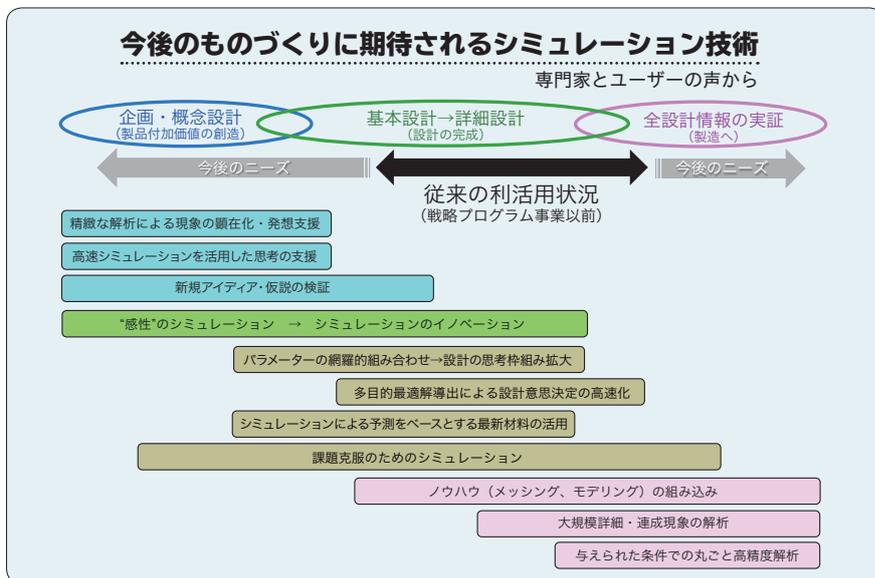
HPC利用のメリットが認識されたとしても、使うのがあまりに難しいようでは絵に描いた餅となってしまいます。同時に「HPC利用の敷居を低くする」取り組みが欠かせません。

「解析手順があまりに複雑だと誰も使ってくれないでしょう。ここを自動化するために知恵を絞らないといけない。我々のこの課題に対する提案がHPC/PFです（次ページ参照）。分野4で開発されたアプリをHPC/PFから使うことで、敷居が下がります。もうひとつの取り組みはハンズオンセミナーです。アプリの使い方を手取り足取りで何回も体験していただく。この2本立てによって、製品開発に使ってみよう

- 分野1 予測する生命科学・医療および創薬基盤
- 分野2 新物質・エネルギー創成
- 分野3 防災・減災に資する地球変動予測
- 分野4 次世代ものづくり
- 分野5 物質と宇宙の起源と構造

【図1】HPCI戦略プログラムの5分野

文部科学省「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(HPCI)の構築について」より



と決断する人が増えると考えています」

プロセス変革のツールとして

もともと熱工学分野の研究者であり、メーカーで空調機や人工衛星といった伝熱技術を応用した製品開発に携わった畑田氏は、自身の経験からシミュレーションを実験と相補的に使うことを重視しています。「シミュレーションの利用のしかたは3つ

考えられると思います。上流からいくと、まずアイデアを出す段階での利用。性能向上に繋がりそうなアイデアが出たときに、まずはシミュレーションで精度の高い推定ができることもすごく助かるはず。次が製品設計の段階。最適な設計解を見つけるために膨大な計算をして、その結果を設計者がジャッジする、いわゆるパラメータサーベイに使うことができます。3つ目は詳細設計が終わったあとの試験段階での活

用。できたものを丸ごとシミュレーションできれば、ムダな実験をせずに性能評価が可能です。コンセプトドリブン型ものづくりへ移行しつつある現在、とくに上流での発想を助けるような使い方ができる道具を提供したいというのが私の希望です」

実用段階に入った分野4のシミュレーション・ソフトウェアを、あなたの仕事でもぜひ活用してください。

基盤ソフトウェア HPC/PF について

HPC/PFとは、ウェブサイト『計算工学ナビ』からダウンロード可能な各種ものづくり系解析ソフトウェアを、パソコン上で簡単に使えるようにする基盤ソフトウェアです。システムの全体構成は図1に示すように、ネットワーク上の知識データベース、ユーザーPCで動作するHPC/PF-GUI、それからスパコン/PCクラスタで実行される解析ソフトウェアの3つで構成されています。

知識データベースは計算工学ナビの"Knowledge Base"として実装されています。各アプリの解析事例が豊富に掲載されており、国プロアプリに興味がある人にとって、各アプリの特徴や性能について知ることができる事例ショーケースとして利用できます。

一歩進んで国プロアプリを使ってみたい、と考える人にとっては、知識データベースは解析例題の提供場所となります。次に説明するHPC/PF-GUIと組み合わせて利用することで、簡単に例題の解析実行が可能です。

HPC/PF-GUIはwebブラウザ上で動作する、HPC/PFのフロントエンド・アプリケーションです。webサーバに相当するプログラムはユーザPCにインストールします。Webブラウザで知識データベースを閲覧し、興味のある解析事例を見つけたら、その事例に付属する再現実行用のデータセット「解析プロジェクト・アーカイブ」をダウンロードし、そのままシームレスにHPC/PF-GUIで解析実行が可能です。

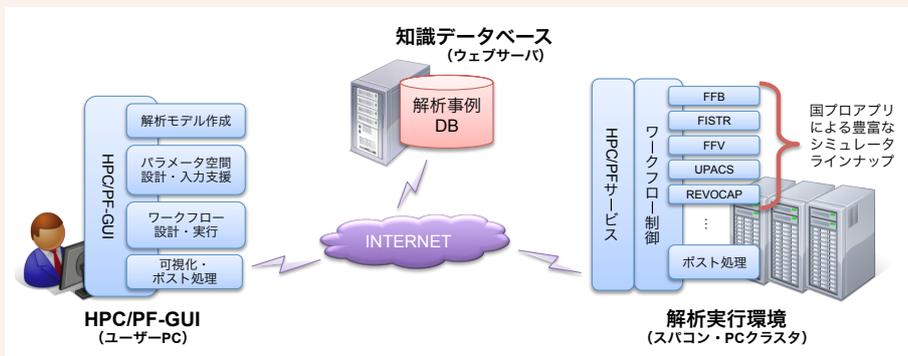
解析に必要なファイルのスパコンへの転送、スパコンシステムに依存したジョブスクリプトの生成・実行、結果ファイルの回収など、解析業務の本質と関係ない部分で手間の掛かる作業はすべてHPC/PFシステムが自動実行するので、ユーザーは解析アプリの学習に専念することができます。

また、GUIパラメータエディタによるパラメータサーベイ空間の設定、サーベイ数に応じて複数ジョブが自動実行され

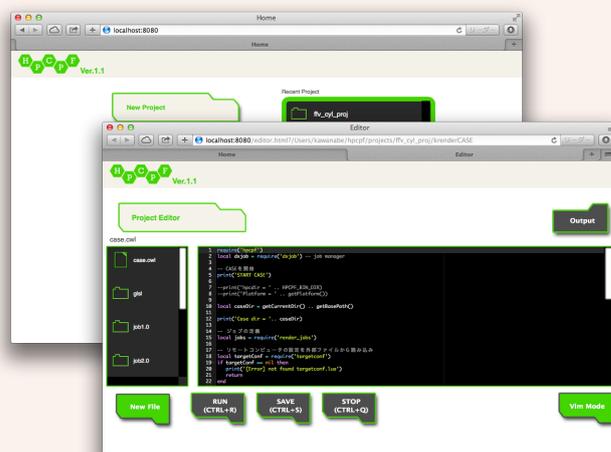
東京大学生産技術研究所
革新的シミュレーション研究センター 特任研究員
川鍋友宏

る機能が利用できます。ユーザスクリプトを記述してプリ/ポストとの連携ワークフローの自動実行と組み合わせれば、実設計業務にも応用可能です。

HPC/PFは、ほかにもスパコンを利用した解析処理に必要なさまざまな機能を提供しています。詳しくは計算工学ナビのサイトに解説記事を掲載していますので、そちらをご覧ください。また、2014年12月1日に第1回HPC/PFハンズオンセミナーを開催します。今後も定期的にセミナーを開催予定です。



【図1】HPC/PF構成図



【図2】HPC/PF-GUI

三次元非定常熱流体解析プログラム

FrontFlow/violet [フロントフロー バイオレット]

理化学研究所 計算科学研究機構
可視化技術研究チーム
小野 謙二

FrontFlow/violet (FFV) は、短時間でシミュレーションを実行できる三次元非定常の熱流体解析プログラムです。

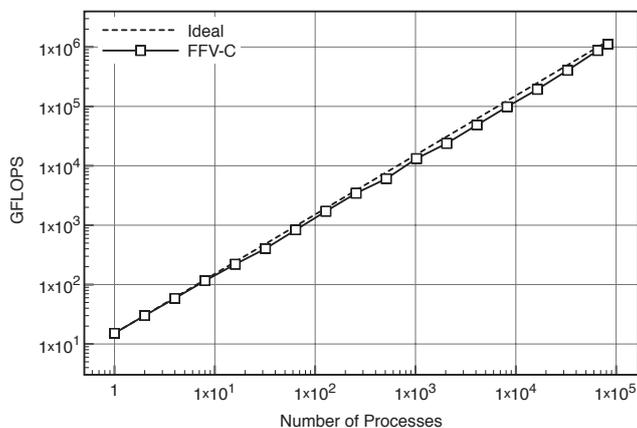
検討に必要な計算結果を短時間で作成し、さらにそこから検討に必要な情報を抽出することにより設計作業を支援することを目的としています。

特に設計上流での活用を想定し、短時間で解析プロセスを実行するために様々なアプローチを採用しています。

例えば、流体計算に必要な計算格子の作成は、解析プロセスの中で最も作業者の労力を必要とする処理ですが、この格子生成を自動化するアプローチを採用しています。また、短時間化のために有効な並列計算についても、最近の計算機アーキテクチャの性能をうまく引き出せる実装方法をアルゴリズムレベルから検討することによって、「京」コンピュータの全ノードの性能を引き出すことにも成功しています。^[図1]

開発ヒストリー

FFVのルーツは、開発者が某自動車会社に勤めていた1990年半ば頃に開発した直交格子系の解析システムに端を発してい



【図1】京コンピュータの82944ノードを使ったハイブリッド並列(OpenMP+MPI)の弱スケール性能。全ノード利用時に1.2PFLOPS、約90%の並列実行効率

ます^[1]。

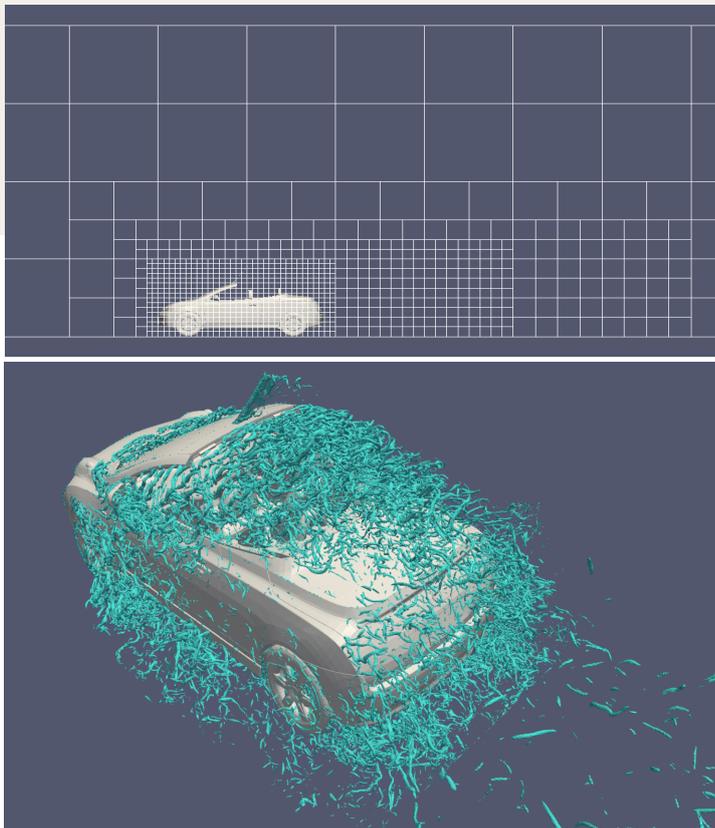
当時、エンジンルーム内の熱流体解析のための格子を作成するのに2ヶ月程度かかっていましたが、これではまったく使えないシステムになるという危惧と、何よりヒトの労力を格子生成に使うことはもったいないので、それまでのBFC格子^[註1]を用いた方法から格子生成を自動化できる直交格子法に変更しました。直交格子法は格子生成の労力を軽減できるポテンシャルはあったのですが、同時に増加する計算格子数に対処するための高速計算技術や任意の形状を扱う高精度化技術などブレークスルーが必要なアプローチでした。しかし、発展していく計算機の能力でカバー出来るだろうという展望から、研究開発の軸足を移したのです。

その後、システムを改善しながら、現在のFFVという形をとっています。

目指す方向

FFVはCFDシステムの家電化を目指しています。つまり、わかりやすいインターフェイスで誰でも簡単に使えるということ。背後に複雑なメカニズムや高性能な仕組みを持っていたとしても、利用者は出来るだけ簡単に使うことができることを目指しています。

設計者にとってはCFDを実施することが目的ではなく、ものをつくるための情報を得るために解析を行うのですから、解析よりもむしろ解析結果を設計に活かすという観点を支援できるシステムの構築を目指しています。そのため、マルチプラットフォーム、高速化、並列化、自動化、ユーザーインターフェイスなどのキーワードを考慮した開発を行っています。利用展開の点から、ライセンスはできるだけ活用に制限のない



【図2】階層的な直交格子による車の周りの流れの計算

[註1] BFC格子、物体適合格子、一般曲線座標系格子といわれ、物体の形状に沿った比較的滑らかな構造型の格子で、計算精度と計算効率がよい計算が可能なのが特徴ですが、非常に複雑な形状には対応できません。

修正BSDライセンスとしています。

オープンな実験プラットフォーム

FFVが従来の他の流体解析システムと大きく異なるところは、その開発スタイルや構築方法だと思います。

ルーツまで遡ると、フルスクラッチでの再構築、その後4回の大きなリファクタリングを繰り返してきました。その度に、記述言語の変更、オブジェクト指向開発、ライブラリ化の推進、プログラムの構造や開発スタイルを変更し、機能・性能や可搬性の拡張を果たしてきました。長期にわたり開発を続けてきた中で、ソーシャルコーディングツールであるGitHubやautotoolsパッケージ、Jenkinsテストツールなどを活用することによって効率的で完成度の高いソフトウェア開発が少人数でも可能になってきました。

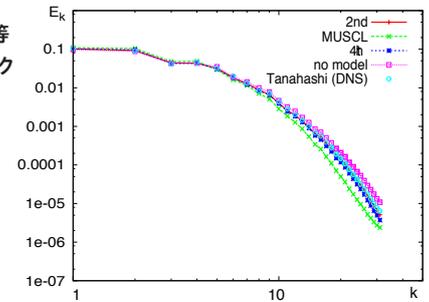
最新版のパッケージでは、領域分割型の並列計算を記述するライブラリを構築し、そのフレームワークを使って流体シミュレータを開発するという計算工学的なアプローチ^[2]を採っており、プログラム開発の方法自体も実験しながら、高性能化や可搬性・メンテナンス性などを検討しています。

特に、オープンソースでソーシャルコーディングというスタイルを採用することにより、新しいアイデアを持っている方ならば誰でも開発に参加できるようになった点は大きなメリットと感じています^{[3][5]}。

派生バージョン

開発過程の中で、FFVシミュレータに

[図3] スキームによる一様等
方形乱流のエネルギースペク
トルの違い (t=3.792 [s])



はデータ構造によっていくつかの派生版の開発が並行して進んでいます。

ひとつはFFV-C (Cartesian) で、直交等間隔格子を用いますが、現在不等間隔格子にも対応中です。

また、FFV-CにはASD (Active Sub Domain) という機能もあります。ASDは並列計算で領域分割を行ったあと、計算しないサブドメインにはMPIプロセスを割り当てずに効率的に計算する手法です。この方法は通信コストは最小化を維持したまま、不規則な計算領域の形に対応できるため、弱スケール性能を維持できるメリットがあります。

次は階層的な直交格子であるFFV-HC (Hierarchical Cartesian) です^[図2]。FFV-HCは空間解像度を局所的に密にする制御ができるため、FFV-Cよりも少ない格子点数で計算が可能になる利点があります。

もうひとつはFFV-OCT (Octree、八分木版) ですが、八分木版はまだ実験段階です。

HPC/PFとの連携

FFVシステムの方向性は設計開発に役立つ流体解析システムであることは先にお伝えしましたが、流体計算を実行するだけではその目的を達成することができません。迅速な計算を行うためには、大規模並列計算が必須になってきます。

現時点では、企業においては1000並列を超えるような計算が日常的に実行されていることはあまりありませんが、今後の計算機の性能価格比の向上に伴い、大規模並列計算が日常化することは明らかです。

大規模並列計算では、計算規模の増大に伴い、これまでは許容できていたファイル管理やポスト処理の問題が顕在化し、解析作業を困難にすることになります。

解析システムとしては、CFDに関わる前処理、計算処理、ポスト処理のすべてが大規模並列環境に対応していく必要があります。

このような問題点を緩和し、大規模並列計算の実行支援を行う実行環境としてHPC/PFシステムの開発を進めています^[4]。

FFVはこのHPC/PFの機能を十分活用できるように連携して開発を行っています。一例として、KDBを用いた検証事例のデータベースを紹介します。

このデータベースはシミュレータの機能や性能、使い方のガイドラインを示し、ユーザの活用を支援するものです。基本的な流れの計算により、どの程度の信頼性があるのかなどを知るために役立つでしょう^[図3]。

設計の革新にむけて

ものづくりにおいては、設計上流での十分な設計検討が、コスト低減・期間短縮・高い品質、魅力ある製品づくりに大きく影響します。

ここでは、性能や機能、コスト、生産性など異なる指標を俯瞰した、全体的な見地からの意思決定が最も重要なタスクとなります。

この意思決定を効果的に支援できるようなCAEシステムを研究開発していくことが、今後のFFVの進むべき方向です。

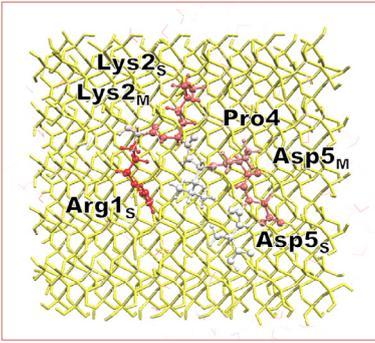
[1] Ono, et al, "Applications of CFD Using Voxel Modeling to Vehicle Development," Proc. of the 3rd ASME/JSME Joint Fluids Engineering Conference, (1999).

[2] Ono, et. al., "Data Centric Framework for Large-scale High-performance Parallel Computation," Procedia Computer Science, Vol. 29, pp.2336-2350, (2014).

[3] http://avr-aics-riken.github.io/ffvc_package/

[4] http://www.cenav.org/20141104_hpcpf_intro/

[5] <http://www.cenav.org/kdb/>



今号の表紙

ABINIT-MPによる人造ペプチド-シリカ基板の相互作用の解析

ナノバイオ系ものづくりシミュレーションの一例として、人造ペプチドとシリカ基板との相互作用エネルギーを4量体のFMO法で計算して可視化した図（赤が安定化、青が不安定化を表す）。ペプチド側のアミノ酸は主鎖と側鎖に分割して細かく評価している。左側の2枚はシリカ側を総和（黄色）、右側の2枚はペプチド側を総和して表示。

出典論文：Okiyama et al., Chemical Physics Letters 566 (2013) 25-31.

(立教大学 望月祐志)

編集後記

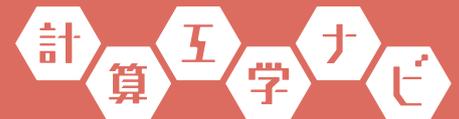
本誌を印刷物として受け取った方も、ぜひ一度下記のサイトを開いてみてください。本プロジェクトで開発されたソフトウェアや日々の活動についてお伝えしています。記事で紹介したとおり、解析事例のデータベースも拡充を図っています。皆さんの研究開発の一助となれば幸いです。(F)



計算工学ナビ オフィシャルサイト

本誌のPDF版やソフトウェアライブラリ、ニュースなどのコンテンツを提供しているWebサイトは下記のURLからアクセスできます

<http://www.cenav.org/>



計算工学ナビ Vol.5

発行日：2014年11月30日

発行：東京大学生産技術研究所

革新的シミュレーション研究センター

〒153-8505 東京都目黒区駒場4-6-1

office@ciiss.iis.u-tokyo.ac.jp