

4. アプリケーション調査研究グループ

4.1 概要

アプリケーション調査研究グループでは、計算科学・社会科学・データ科学のアプリ分野から構成するサブグループにおいて、次世代計算基盤により最先端の研究成果を実現するために必要な計算資源に関する調査を行う。本グループには、主にアプリ分野で活動する① 生命科学、② 新物質・エネルギー、③ 気象・気候、④ 地震・津波防災、⑤ ものづくり、⑥ 基礎科学、⑦ 社会科学、⑧ デジタルツイン・Society 5.0の8つのサブグループとこれらのサブグループの活動を支援し、他グループとの連携を図る計算機科学分野の⑨ 科学技術計算アルゴリズム、⑩ 機械学習アルゴリズム、⑪ ベンチマーク構築、⑫ 性能モデリング技術の4つのサブグループが置かれている。本グループの調査では、各分野のアプリケーションについて、総演算量、演算性能、メモリ使用量、ノード間通信性能といった一般的な指標に加え、Byte/Flops値やSIMD命令の活用率、生産性といったアプリケーションプログラムの性質を規定する様々な指標の調査を行う。また、アプリ分野における代表的なプログラムの特性を反映したベンチマークプログラムの整備を行う。本ベンチマークプログラムはアーキテクチャ調査研究グループ、システムソフトウェア・ライブラリ調査研究グループに提供し、これらの調査研究の推進に役立てる。2023年度は、アプリケーションの各サブグループに対してアーキテクチャ、システムソフトウェアグループのヒアリングを行い、ベンチマークやアーキテクチャ選定の現状について討議を行った。また、各サブグループにおいて、2030年頃に期待される成果に関して英語スライドを作成し、あわせて内容更新を行うとともに、ベンチマークの整備を適宜、実施した。

4.2 生命科学分野調査研究サブグループ

4.2.1 研究分野の概要

生命科学分野は、複雑で精緻な生命現象を解き明かし、精密医療やインシリコ創薬を通じて健康・長寿社会の実現に貢献する。本分野の特徴は、図 4.2.2.1に示すように原子レベルの生体分子から、細胞、臓器、個体のレベルまで異なるスケールの多層的で複雑な現象を扱う点にある。特に、実験では解き明かすことのできないダイナミクスを解析するために、分子動力学シミュレーションをはじめとする多種多様なシミュレーション技法が開発・活用されている。また、シミュレーション技術とAIによる新規医薬品の開発、医療・ゲノムデータ解析に基づく健康予測・個別化医療などが社会的に重要な応用例として期待される。近年、大規模言語モデルや拡散モデルを用いた生成AIを生命科学分野で活用する研究が大きく進展しており、生体分子の構造予測、新規機能分子の生成、マルチオミクスデータの統合、電子カルテや医療データの統合なども進んでいる。これら先進的なAI技術との連携も今後の生命科学分野では必須になる。

4.2.2 現行システムにおける解析の現況

本分野では多数のアプリケーションプログラム・計算対象が存在する。まず、代表的なコアアプリとそれらを組み合わせた手法の例について、特に無償利用可能なものを中心に図 4.2.2.2に示す。多岐に渡るため網羅することは困難であるが、計算対象と目的に応じて適切なアプリケーションを選択する必要がある。以下では富岳で実施されている課題を中心に、代表的な計算例に関して3種類示す。特に生体分子・生体内環境の分子動力学シミュレーションに関しては、4.2.4に記載したように、拡張アンサンブル型分子動力学シミュレーションの効率化が重要な課題となっており、この種のアプリケーションはシミュレーションだけではなくデータ処理や解析にも一定の計算負荷がかかるため、今後はgREST、McMD、ColDock/MSM等も代表例に加えて、解析を進める必要があると思われる。

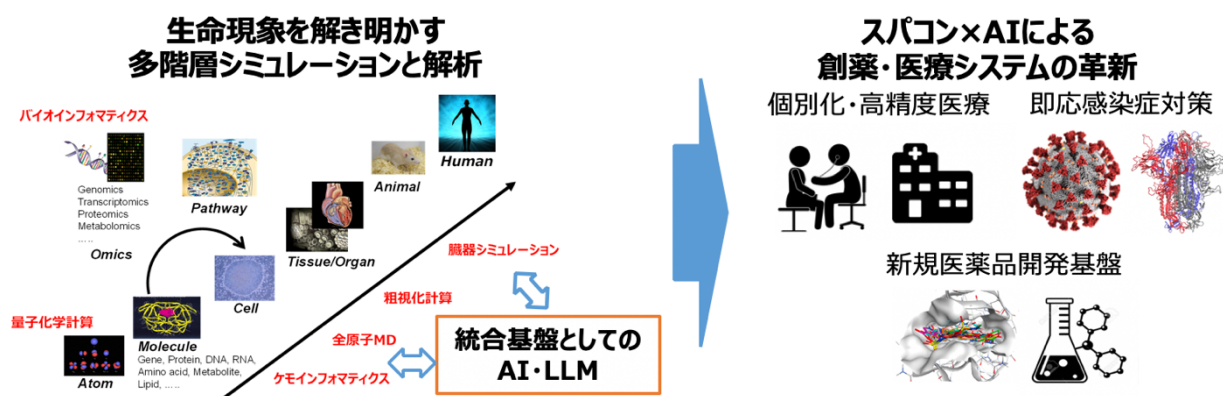


図 4.2.2.1 生命科学分野の概要

- 代表的なコアアプリ
 - 全原子分子動力学(MD)シミュレーション
 - GENESIS, Gromacs, AMBER, CHARMM, NAMD
 - ドッキングシミュレーション
 - Autodock Vina, Dock, ZDOCK
 - 量子化学シミュレーション
 - GAMESS, Gaussian, FMO, QM-MM
 - 疎視化分子動力学法
 - CafeMol, MARTINI
 - 細胞スケールシミュレーション
 - pSpatioocyte
 - ゲノム解析・遺伝子ネットワーク解析
 - INGOR, Genomon
 - 臓器・脳・全身
 - UT-Heart, 全脳循環代謝シミュレータ
- コアアプリを用いた複合的手法の例
 - シミュレーション結果解析手法・AI連携
 - MSM, TS-MD, ChemTS
 - 実験連携
 - MD-SAXS, テンプレートマッチング(Cryo-EM連携, XFEL連携)
- AI・最適化・統合フロー環境構築に必要な要素
 - Python (Miniconda環境, pytorch, TensorFlow等ライブラリ), Julia, GPU連携



図 4.2.2.2 代表的なアプリケーションの例

4.2.2.1 生体分子・生体内環境の分子動力学シミュレーション

タンパク質の機能を解析したり、薬剤分子との相互作用を推定するために用いられる分子動力学(MD)計算は本分野の代表的なシミュレーション手法の一つである。例えばコアアプリGENESIS2.0.0betaを用いて、癌治療の標的タンパク質であるEGFRを対象に(約10万原子)、標準的計算時間である1 μ 秒の計算を富岳24ノードを用いた場合8132NHを要する。これは比較的小さな系であり、変異の入った多数のタンパク質や多数の薬の候補(100~最大数十億)を検討する場合、より多くの計算を要する。またMDを用いてタンパク質の性質を解析する手法(MSM解析)を行う際には同様のEGFRの系でGENESIS2.0.0betaを用いて、総計で128 μ 秒の計算が必要となる。この場合768ノードを用いて880,440NHを要し、生成データ量は1.5TB程度になる。計算コストだけでなく、計算結果を手元のサーバに移動し解析する必要があり、データ転送もボトルネックの一つとなっている。従って、今後はMSM解析も富岳で実行することを視野に入れ、ハードウェアのコードザインを検討していく必要があると思われる。また、近年、gREST、McMD、Coldock、Hypersound-coldockなどの拡張アンサンブルMDが国内で開発されており、ポスト「富岳」では生命科学分野におけるこれらのアプリケーションの実用化が望まれる。そのためにはシミュレーション部分のみならずプレポスト処理・解析過程も含めて富岳で一気通貫にワークフローを実行できる計算環境が必要であり、この部分についても今後検討していく必要があると思われる。

4.2.2.2 遺伝子ネットワーク解析

疾患に関わる遺伝子の同定や遺伝子相互の発現制御の関係性を解析するために、多数の遺伝子の発現データから遺伝子の制御関係のネットワークを推定する必要がある。ベイジアンネットワークに基づく遺伝子ネットワーク推定手法INGORでは、100人(サンプル)分のデータからのネットワーク推定で、256ノードを用いて計200NH程度を要する。入出力のファイルはそれぞれ13MB、11MBと比較的小さい。しかし、今後遺伝子発現データは増大し続けると考えられ、それに対応した解析計算が必要となる。(2022年度と同様)

4.2.2.3 心臓シミュレーション

心臓シミュレーションUT-Heartにおける計算資源は、分子数に比例したノード数が必要となる。有限要素数は5,000~7,705要素、1要素内の分子数は再現精度に応じて16~75分子となる。「富岳」では、1ノード4プロセス(プロセス内12スレッド)で心臓の拍動1.0秒に約240時間を要する。最小規模5,000要素16分子で480万ノード時間、最大規模7,705要素75分子で約3,467万ノード時間が必要となる。ただし全ての研究ケ

ースにおいて拍動1.0秒が必要ではなく、0.1～0.5秒分のデータが標準的となる。(2022年度と同様)

4.2.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

2022年度の検討に引き続き検討を行なった。今年度の検討結果を以下に示す。

4.2.3.1 次世代計算機で期待されるブレイクスルー

分子から臓器・人体レベルまで、多階層でのシミュレーションによる解析技術が発展しており、これまで実験で捉えられなかった生体内のメカニズムが解析・予言可能になると期待される。例えば1細胞シミュレーション等により細胞全体の解析が可能になると期待される。また、生命現象の多階層のデータ及びシミュレーションをAI・LLMにより統合することで、計算による疾病メカニズムの解明、効果的な治療法の推薦などの高精度医療、小分子から抗体医薬品を含む様々な薬のインシリコ探索・設計などが可能になると期待される。

4.2.3.2 ブレイクスルーに必要な計算環境、解決すべき課題

1種類のシミュレーションによる解析・分析では不十分で、AI・多数のシミュレーション・データベースの高度な連携が必須である。個別アプリの高速化は必要条件に過ぎず、複数のアプリを容易かつ柔軟に組み合わせられる計算環境なしには解決すべき課題を解けない。このように複数のアプリやAIを大規模かつ自由に連携して動作せられる計算環境が必須になる。また近年、GPU等を用いた高速なアプリが世界中で開発されており、それらをうまく活用することで、新たな成果が生まれると期待される。すでに開発されている様々なアプリが特段の改変なく動作する環境が望まれる。また、中長期的な視点と予算に基づき安定的な人材の確保・育成が可能な状況を維持すること、産官学での好循環する連携体制の構築が望まれる。

4.2.4 ベンチマークの整備計画と状況

2022年度に引き続き、以下の表 4.2.4.1に示すようにベンチマークアプリケーションの整備を進めている。特に生体分子・生体内環境の分子動力学シミュレーションに関しては、現状ではベンチマークの対象からは外れているが、マルチカノニカルMD (McMD)、ColDock/MSM、Hypersound-ColDock/MSMなどの創薬DXプラットフォームに含まれるアプリケーションの実用化が求められている。これらのアプリケーションには、データ転送・プレポスト処理・解析といったシミュレーション以外に係る計算負荷も大きいいため、これらのパートの計算環境・処理速度も考慮したハード・ソフトウェアのコデザインを検討していく必要がでると思われる。

表 4.2.4.1 ベンチマークアプリケーションの整備状況

名称	目的・内容	整備状況
GENESIS	MD シミュレーション(特に大規模な生体分子系向け)	評価チームと共有済み。GPU 版も共有済み。
Gromacs	MD シミュレーション(特に小・中規模生体分子向け)	複数の系を評価チームに共有済み。GPU 版も共有済み。
INGOR	遺伝子ネットワーク推定	評価チームに提出済
ChemTSv2	生成モデル (RNN) と強化学習による分子デザイン	評価チームに共有済み。GPU 版も共有済み。
kMOL	グラフニューラルネットワークによる分子やタンパク質の性質予測	評価チームと共有済み。
VFVS-docking	大量ドッキングシミュレーションワークフロー	評価チームと共有済み。
テンプレートマッチング 計算フロー	テンプレートマッチング法を用いた XFEL 画像に対するタンパク質構造の推定	ベンチマークとして利用可能な箇所を準備中。
UT-Heart	心臓の拍動シミュレーション	評価チームに共有済み。GPU 版も共有済み。

4.3 新物質・エネルギー分野調査研究サブグループ

新物質・エネルギー分野では、多数の電子やイオンの量子力学的な振る舞い、あるいは量子力学的な電子状態から得られた原子間ポテンシャルに従う原子集団の古典的な振る舞いをシミュレーションすることで、原子・分子、結晶・非晶質固体の性質を予測することを目指してきた。

4.3.1 研究分野の概要

本研究分野では、結晶固体、アモルファス、分子に閉じ込められた多体電子に対するシュレディンガー方程式を解いて、それらの電子物性を予測してきた。半導体や金属、分子について第一原理電子状態計算、中でも密度汎関数理論に基づく近似手法が広く成功を収める一方、大規模な不規則系や分子系への応用には困難があり、電子の運動エネルギーと電子間相互作用の競合が物性に重要な影響を与える量子物質では、非自明な量子繰れの効果を取り入れる手法が必要となり、配置間相互作用法や量子モンテカルロ法などが用いられている。

様々なアルゴリズム、計算手法が用いられている中、第一原理電子状態計算に限っても、そのアルゴリズムは、1電子の波動関数をどのような基底関数系によって記述するか大きく依存している。実空間基底系を基礎とする第一原理計算コードRSDFTが「京」コンピュータで成功を収めたことから、RSDFTが「富岳」のコデザインに用いられた一方、様々なユーザが用いているVASPやQuantumEspressoなど平面波基底系に基づくアプリケーションが「富岳」のアーキテクチャに適合しているか否かについては議論がある。そこで本調査研究では、幅広い基底系のコードを取り上げることとした。本節では、NTChem (ガウス基底)、BigDFT (ウェーブレット基底)、AkaiKKR (グリーン関数)、QuantumEspresso (平面波基底)、およびOpenMX (局在・平面波の混合基底)について以下に調査結果を述べる(ただし、括弧内は基底名とする)。

また量子多体波動関数計算については、行列式の平方根にあたるパフィアン計算とモンテカルロ法を基礎とするmVMCとクリロフ部分空間法に基づくHΦを調査する。特に行列式やパフィアンは、電子数に対して指数関数的に増減するため、倍精度実数による演算が電子数の大きな大規模計算には不可欠となっている。

加えて、本調査研究では、AkaiKKR、mVMC、HΦのベンチマークセットを作成した。実空間基底を用いているSALMONについても、コード開発者である量子科学技術研究開発機構の山田俊介博士、乙部智仁博士、および筑波大学の矢花一浩教授からベンチマークの提供を受けている。さらに、ワークフロー型アプリケーションとして、HΦとベイズ最適化ライブラリPHYSBOを組み合わせたBEEMs、および第一原理計算による訓練を組み込んだAI加速分子動力学のベンチマークセットを作成した。

4.3.2 現行システムにおける解析の現況

2022年度に得られた現況に加え、演算加速部への対応状況の調査結果を以下に示す。

4.3.2.1 NTChem (ガウス基底)

標準的な解析

4.3.2.1.1 NTChem [hybrid O(N) DFT] の標準的な解析に関する情報:

- (1) 計算条件: 8000原子、ダブルゼータ基底
- (2) メモリ要求: 16TB程度
- (3) B/F: 不明
- (4) プロセス間通信: MPI_IAllGatherv
- (5) 6144ノード時間 (1ジョブ6時間@1024ノード)

4.3.2.1.2 NTChem (RI-MP2) の標準的な解析に関する情報:

- (1) 計算条件: 360原子、9840軌道(最大: 720原子、19680軌道)
- (2) メモリ要求: 16GB/node (最大: 25GB/node)

- (3) B/F: 不明
- (4) プロセス間通信: MPI_Isend/Irecv, MPI_Allreduce
- (5) 170.7ノード時間@1280ノード(最大: 5021ノード時間積@17720ノード)

大規模な解析

hybrid O(N) DFT [(i)-1.]については100万原子系、RI-MP2[(i)-2.]については2000原子系が大規模な解析となる。

4.3.2.2 BigDFT (ウェーブレット基底)

標準的な解析

- (1) 5000原子 (最大: 50,000原子)
- (2) メモリ要求: 24GB程度/node
- (3) B/F: 不明
- (4) プロセス間通信: MPI one sided communication
- (5) 1ジョブ1時間程度[64ノード時間@「富岳」] (最大: 768ノード時間@AMD Irene ROME, TGCC Joliot Curie)]

大規模な解析

50,000原子系のO(1000)回計算によるhigh throughput mutation screening。

4.3.2.3 AkaiKKR (グリーン関数)

標準的な解析

- (1) AkaiKKR自体はグリーン関数のk点・エネルギーメッシュ点のグループに対してOpenMP並列を適用
- (2) 材料パラメータ並列には、MPIもしくはArray Jobを利用
- (3) AkaiKKRのホットスポットはダイソン方程式を解く際の逆行列計算
- (4) 演算量: 不規則系磁石材料 (2-17-3系) を対象とした5万件の計算で4万ノード時間 (「富岳」)
- (5) データ量: 不規則系磁石材料1件当たり12GBメモリを使用 (「富岳」)
- (6) データ転送量: 不明
- (7) 通信量: OpenMPのReduction
- (8) I/O量: 不規則系磁石材料1件当たり: インプット (数KB)、アウトプット (数百KB)、ポテンシャル (数MB)
- (9) 1解析あたりの計算時間: 不規則系磁石材料1件当たり1時間程度@「富岳」

大規模な解析

約15万件の不規則系磁性材料を対象とした、「富岳」における自動網羅電子状態計算による物性データベースの構築。

4.3.2.4 QuantumEspresso (平面波基底)

標準的な解析

- (1) ハイブリッド並列になっている (MPIとOpenMP)、バンド、k点メッシュ、スピン、基底、FFT
- (2) ホットスポットはFFT、直交化など
- (3) 演算量: 不明
- (4) データ量: 113.48 GB (下記リンクのベンチマークの「PSIWAT」)
- (5) データ転送量: 不明
- (6) 通信量: MPIの1体1通信、集団通信
- (7) I/O量(PSIWAT): 擬ポテンシャル(入力)=1.5MB、電荷密度(出力)=59MB、波動関数(出

力)=9GB

- (8) 1解析あたりの計算時間：基本的な系のSCF計算、もしくは大規模系のSCF計算について、以下に報告されている。

https://www.hpci-office.jp/documents/appli_software/Fugaku_QE_performance.pdf

大規模な解析

100原子程度の系のハイブリッド汎関数計算、フォノン計算。

4.3.2.5 OpenMX (局在・平面波の混合基底)

標準的な解析

- (1) ハイブリッド並列。

通常のDFT: MPIはスピン、k点、LAPACKの対角化ルーチンの並列化を使用し、OpenMPも併用
オーダーNのDFT: MPIはスピン、原子数、LAPACKの対角化ルーチンの並列化を使用しOpenMPも併用

例えば1000原子系であれば並列計算において約40000 CPUコアまで加速

- (2) ホットスポットはLAPACKによる対角化
(3) 演算量: 不明
(4) データ量: メモリ使用量はCPUコア当たり1~2GB
(5) データ転送量: 不明
(6) 通信量: MPI、1体1通信、集団通信
(7) I/O量: Fe1000原子、Si1280原子の周期系計算をオーダーNで行うと、電子密度などで1GB。
(8) 1解析あたりの計算時間: Siの大規模系で 1 SCFステップ1.5秒、5000MDステップ(5ps)の計算で20時間

大規模な解析

10万原子を超える大規模系のオーダーN計算、機械学習ポテンシャルを用いた高速化、および2次元材料の物性データベース（通常のDFT計算×約12000種類の化合物候補）作成。

4.3.2.6 mVMC

標準的な解析

- (1) 計算条件: 900電子系(2次元ハバード模型)、1000MCS、1000最適化ステップ
(2) 136GFLOPS/node (ピーク性能比6%、ピーク性能3.3792TFLOPS)
(3) メモリ要求: 2BG/node
(4) B/F値: 1.6
(5) 実測メモリスループット215.7GB/s (ピーク性能比21.1%)
(6) ノード間通信: Reduction
(7) 通信量: $1000 \times 900 \times 42 \times 16 = 0.56\text{GB}$
(8) I/O量: 0.2GB
(9) 220ノード時間@東京大学情報基盤センターWisteria/BDEC-01、A64FX1ノード (8スレッド、6プロセス)

大規模な解析

計算条件: 10,000電子系の計算: 変分パラメータ数160,000超すなわち 1.6×10^5 次元のCG法計算、1000倍の演算量。最大のI/O量が1TBに達する1,000電子系の高温度超伝体のシミュレーション。

4.3.2.7 HΦ

標準的な解析

- (1) 計算条件: 24量子ビット(2^{24} 次元)、1000回の疎行列-ベクトル積
- (2) 21.7GFLOPS /node (ピーク性能比9.1%)
- (3) 要求メモリ: N量子ビットについて($16 \times 2N$ /プロセス)B/core
- (4) B/F 1.9
- (5) メモリスループット 40.4GB/s
- (6) ノード間通信: 複数プロセスの場合:ステップごとに16Byteの1対1通信を各ノード[(2^{24})/プロセス]数回
- (7) I/O量: ほぼ無視できる
- (8) 0.4ノード時間@FX10 1 ノード (16スレッド)
- (9) ベイズ最適化の1ステップに使える計算規模

大規模な解析

50量子ビットを超える計算 (メモリ要求量1024PB)

4.3.2.8 新物質・エネルギー分野におけるソフトウェアの演算加速部への対応

ソフトウェアごとに異なる構造の演算が律速過程となるため、演算加速部とCPUとのデータ転送頻度や、加速部メモリ上のデータの再利用可能性に差異があり、演算加速部 (主にGPGPU) への対応にはばらつきがあるのが現状である。

4.3.2.8.1 量子化学・分子化学

量子化学・分子化学では、Hartree-Fock(HF)近似や多体摂動論(MP2など)が多用される。これらの計算では四つ足のクーロン行列の計算が多くの計算時間を占める場合があり、国内外で演算加速部の利用が進んでいる。

特にGAMMESS ECPなど、米国DOEが主導する国際プロジェクトにおいて、 $O(10^4)$ のGPUを用いた大規模なHFおよびMP2計算のためのプログラム整備が行われている。同時に、アーキテクチャに依存しない加速部への対応のためにkokkosを始めとする演算加速部の利用を前提としたC++ラブリ整備が進んでいる。それと同時にプログラミング言語におけるC/C++のシェアが増えている。

NTChemについても、4,000 GPU上でのMP2計算が行われてきた。一方、BigDFTでは、SYCLを用いてhybrid functional計算部分の演算加速部への対応が図られている。

4.3.2.8.2 物質科学のための第一原理

密度汎関数理論に基づくコードの多くで高速フーリエ変換FFTがよく用いられているため、量子化学分野とは異なるアプローチが必要となる。国内外で広く用いられているオーストリアの第一原理計算ソフトウェアVASPについては、NVIDIAが同社のGPGPU向けのチューニングを行なっている。国内では、Fortranを用いたコードが多く現存するが、GPUの利用に向けたチューニングが進められている。筑波大学/QSTに開発拠点を置くSALMONでは、OpenACCおよびOpenACCとCUDAを組み合わせたGPGPUへの対応が始まっている。AkaiKKRでも対応が進んでいる。

4.3.2.8.3 物質科学のための量子多体計算

量子多体系のための計算では、演算密度に比べデータ移動が多い。そのため、HΦやmVMCでは演算加速部への対応に遅れが見られる。HΦは疎行列-ベクトル積によるクリロフ部分空間法が核となるコードであるが、密行列対角化部分のみGPGPU向けライブラリMAGMAに接続されている。一方で、mVMCは未対応となっている。mVMCでは多体波動関数としてパフィアンに基づく表現が用いられているが、国外では、深層ニューラルネットワークが用いられ始めており、コード自体がAI向けライブラリで書かれていたため、そもそも並列

GPU環境に対応できている。今後、国内のコードもOpenAI TritonなどのAI向けライブラリによる実装を検討していく必要がある。

4.3.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

- (1) 富岳NEXTは新物質・エネルギーの分野に以下のような様々なインパクトをもたらすことが予測される:
 - (2) 不規則系磁性材料の大規模物性データベース構築による新永久磁石材料およびスピントロニクス材料開発期間の短縮
 - (3) AIで加速された第一原理シミュレーションによる全固体電池材料の性能劣化および太陽電池材料の量子収率予測
 - (4) トポロジカル電子状態の超高速光応答の第一原理シミュレーションによる新規量子デバイス動作原理の開拓
 - (5) 極短パルスレーザーと固体との相互作用の第一原理シミュレーションによる高精度な微細加工の開拓
 - (6) 高温超伝導体やトポロジカル量子計算基盤(量子スピン液体)における量子繚れ構造の観測と制御
- これらの成果が、高性能・高信頼性蓄電池、高性能モータ、省電力スピndeバイスの開発、量子デバイスにおける量子繚れ制御技術に波及することが期待される。

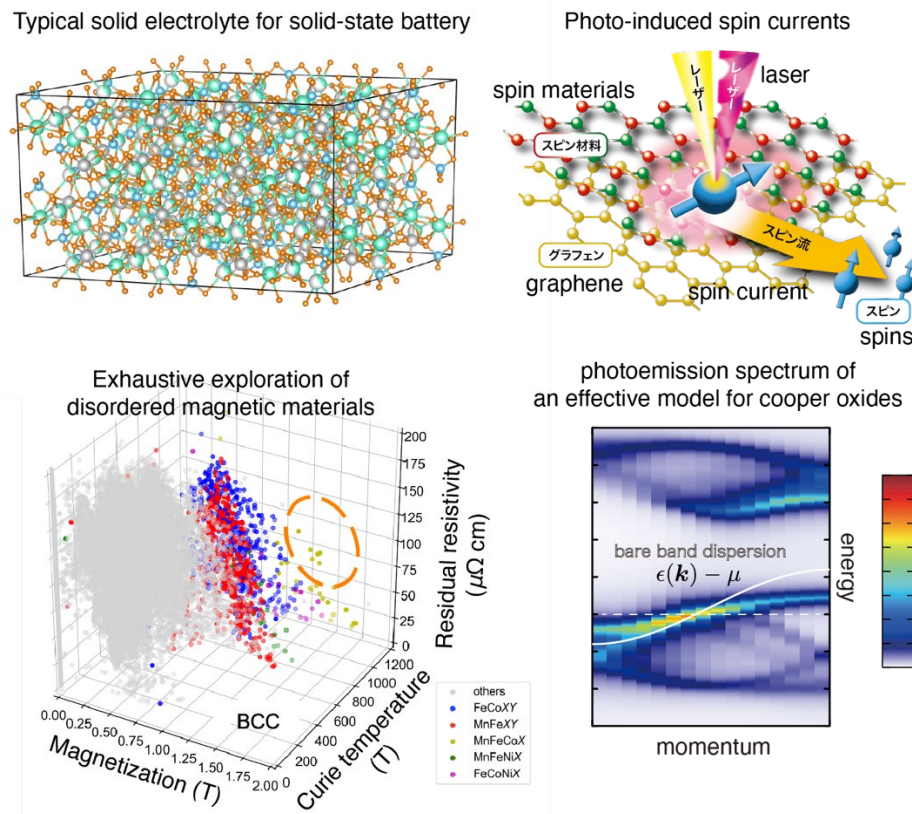


図 4.3.3.1 高性能・高信頼性蓄電池のための電解質(左上)、磁性材料データベース(左下)、省電力スピndeバイス原理の提案(右上)、量子もつれ状態の分光スペクトル計算(右下)

上記のブレークスルーをもたらすためには、「富岳」数千ノードを擁する計算を日常的に数十個並列実行可能な計算能力が、富岳NEXTに期待される。加えて、演算加速部の有効利用を可能とするコンパイラが必要となるとともに、アプリケーション開発者は演算加速部を有効に活用するため、チューニング並びにアルゴリズム改変による演算密度の向上が必要となる。

以下では、2022年度に調査を行った、ソフトウェアごとに期待される成果と必要資源を列挙する。

4.3.3.1 NTChem

期待される成果

- 100 万原子系の密度汎関数理論(DFT)計算による生体分子系の電子物性解析

必要資源

- 10^{10} ノード時間積

4.3.3.2 BigDFT

期待される成果

- 5 万原子からなる現実的なデバイスサイズのモデルを用いた物質へのドーピング効果の解析

必要資源

- 3×10^7 ノード時間積

4.3.3.3 AkaiKKR

期待される成果

- マテリアルズインフォマティクスに資する高品質の物性データ創出
富岳 NEXT では、大規模ハイスループット計算によって無機材料分野の「スモールデータ問題」を解決し物性ビッグデータを実現することが可能になり、広範囲の材料空間を効率的に探索できる。
- 有限温度における磁性・スピントロニクス材料のデザイン
富岳 NEXT では、有限温度における素励起を考慮した物性（電子、磁気、伝導特性）が計算可能になる。バルクのみならず MTJ 素子構造等をデザインするための、ハイスループット計算が期待できる。

必要資源

- 計算資源：不規則系バルク 200 万個、不規則系超格子 5 万個を対象としたハイスループット計算を行うため、
 - 約 4×10^7 ノード時間積
 - メモリ要求：1 プロセスあたり O(10~20)GB
- プログラム・スクリプト言語：Fortran、Python
- 並列化：1 つの材料に対しては OpenMP、材料パラメータ並列には MPI か Array job

4.3.3.4 mVMC

期待される成果

- 強相関量子物質の分光光学スペクトル
量子物質に創発する情報キャリアの素性を解明する上で重要な役割を果たしてきた中性子/X線散乱実験や光電子分光測定とシミュレーションの定量比較が可能となる。それによって、高温超伝導体の発現機構、および量子スピン液体候補物質における(マヨラナ粒子をはじめとする)分数粒子発現が検証可能となる。

必要資源

- 計算資源：144 電子の共鳴非弾性 X 線散乱スペクトル 1 計算条件あたり
 - 5×10^7 ノード時間
 - メモリ要求：1 プロセスあたり O(100)GB
- プログラム・環境：C 言語、機械学習最適化(python/julia)、及び外部の古典/量子アニーラとの接続

4.3.4 ベンチマークの整備計画と状況

AkaiKKR、mVMC、HΦについて、2022年度末にベンチマークセットを提出した。また、SALMONについて、コ

ード開発者である量子科学技術研究開発機構の山田俊介博士、乙部智仁博士、および筑波大学の矢花一浩教授からベンチマークの提供を受けた。2023年度は、ベイズ最適化アプリケーションPHYSBOと量子多体問題ソルバーアプリHΦとを統合するpythonコードBEEMs、および第一原理計算によって訓練を行うAI加速分子動力学法に基づくワークフロー型ベンチマークを作成した。

4.3.4.1 BEEMs

BEEMs (<https://github.com/k-yoshimi/BEEMs>)は、図 4.3.4.1のようにHΦによるクリロフ部分空間法を実行するinner loopとPHYSBOによるベイズ最適化を行うouter loopからなる。BEEMsでは、inner loopの物理量計算結果と所与の実験データとの二乗誤差を目的関数として、多体電子系の有効ハミルトニアンを最適化する。outer loopはinner loopの結果に基づく目的関数を最大化する有効ハミルトニアンのパラメータをベイズ最適化によって探索することになる。BEEMsでは、inner loopとouter loopのインターフェイスでは目的変数のみを受け渡すため、inner loopの計算はほぼ独立に実行される。

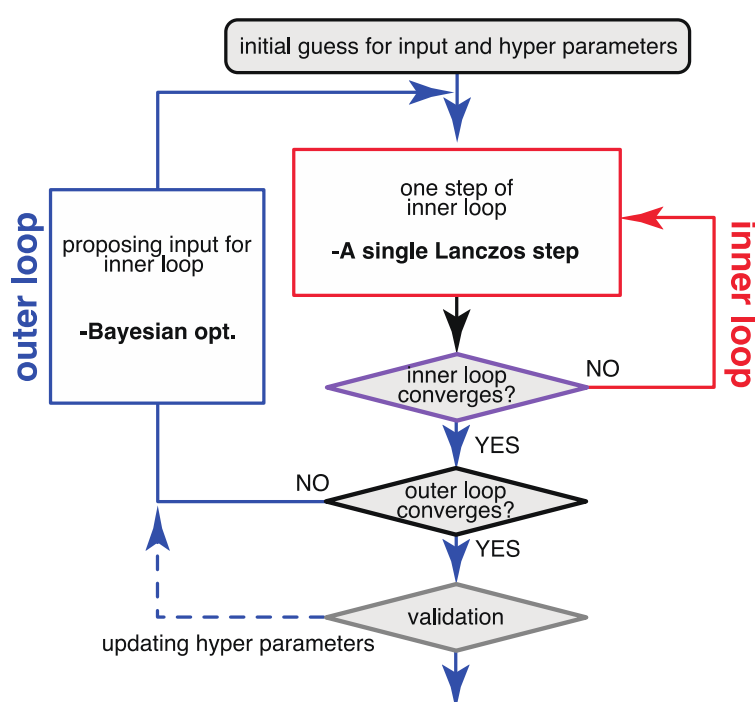


図 4.3.4.1 BEEMs における inner loop と outer loop の概要

4.3.4.2 AI 加速分子動力学

分子動力学を用いて結晶構造最適化を行うために、人工ニューラルネットワーク原子間ポテンシャルを学習するワークフロー型アプリケーションであるPIMD

(<https://ccse.jaea.go.jp/software/PIMD/index.jp.html>) を用いた自己学習ハイブリッドモンテカル (SLHMC) 法のベンチマーク作成を行った。PIMDから第一原理計算を呼び出す際、Quantum Espresso は内部でリンクを行い、VASPなどは外部リンクとなる。後者ではディスクIOを用いることになる。

本ベンチマークセットでは、DFT計算部分にQuantum ESPRESSO (<https://www.quantum-espresso.org/>) を用い、人工ニューラルネットワーク (ANN) ポテンシャル作成及びそれに基づく機械学習分子動力学 (MLMD) 計算については、aenet (<http://ann.atomistic.net/>) を使用している。

4.4 気象・気候分野調査研究サブグループ

4.4.1 研究分野の概要

気象現象や気候・地球システムの理解・予測のため、対象とする時空間スケールや主な用途（研究・現業）に応じて多種多様な気象・気候モデルが開発されている。基礎的な研究とともに、現業予測での利活用や気候変動適応策といった社会実装も活発に進められていることが分野の特徴である。スーパーコンピュータの黎明期からのヘビーユーザであり、モデルの高解像度化、大アンサンブル化、長期積分化、プロセス精緻化などの有望な施策を今後も着実に推進するため、HPCへの期待は今後も高まり続けるだろう。

国内におけるモデルの開発・利用現況を俯瞰し、今後のco-designに向けたターゲットアプリの選定や分野の振興につなげるため、昨年度までに以下のモデルについてアプリ情報を収集済みである。

- 気象庁の現業数値予報モデル
 - **asuca**：主に領域を対象とする非静力学大気モデル
 - **asuca-Var**：asuca ベースの変分法データ同化システム
 - **GSM**：全球を対象とする大気モデル
- 領域モデル
 - **SCALE-LETKF**：非静力学大気モデル（SCALE）および局所アンサンブル変換カルマンフィルタ（LETKF）に基づくデータ同化システム
 - **SCALE-LT**：SCALE に雷の頻度・分布の計算部分を追加（気象雷モデル）
 - **SCALE-SDM**：SCALE に粒子法に基づく雲微物理スキーム（超水滴法）を追加
 - **kinaco**：非静力学海洋モデル
- 全球モデル・マルチスケールモデル
 - **NICAM, NICOCO**：非静力学大気モデル（NICAM）、および NICAM と海洋モデル COCO が結合した大気海洋結合モデル（NICOCO）
 - **NICAM-LETKF**：NICAM および LETKF に基づくデータ同化システム
 - **MRI-ESM**：大気・海洋・化学・エアロゾルモデルが結合した地球システムモデル
 - **GAIA**：大気大循環、電離圏、電気力学モデルが結合した大気圏電離圏結合モデル
 - **MSSG**：全球・メソから都市街区までをカバーするマルチスケール大気海洋モデル
- 力学コア
 - **ISPACK3/SPML**：スペクトル変換ライブラリ（ISPACK3）およびその Fortran90 ラッパー（SPML）
 - **SCALE-DG**：不連続ガラーキン法に基づく大気力学コア

以上に加えて、**AFES**（全球モデル）の開発者が協力者として議論に参加している。

4.4.2 現行システムにおける解析の現況

気象・気候分野のモデルの大きな傾向として、Time-to-solutionが重要な現業モデルでは1解析あたりのノード数やノード時間積が小さい。研究モデルでは1解析あたり数Nhから最大1600万Nh程度（富岳一般課題の年間上限近く）まで幅広い。一般に気象・気候モデルは要求B/F比が高く実行性能は数%止まりという特徴があるが、SCALE-LT/-DGなど精緻なスキームを用いたモデルではその限りではない。ノードあたりのメモリ使用量は1GBを切るモデルがある一方で数十GBを必要とするモデルもあり多様である。通信パターンについては格子法ベースのモデルでは隣接1対1通信、スペクトル法ベースのモデルでは全通信が基本であり、特に後者は一般に通信負荷が大きい。I/Oノードを用いる場合は、計算ノードとの大容量通信も生じうる。入出力サイズについては多くのモデルでPBオーダーの入出力を要するためI/O負荷は高い。なお、一部のアプリはOpenACCによるGPU化に対応している。

4.4.2.1 個別モデルの現況（標準解析・大規模解析）

（2022年度報告書より再掲；個別モデルの詳細については2022年度報告書の表4.4.1を参照）

asuca：（標準）東アジア域5 kmメッシュ、78時間積分（メソモデル）

（大規模）1 km局地モデル（格子数1,400×1,400×76）リアルタイムシミュレーション

asuca-Var：（標準）東アジア域5 kmメッシュ（メソ解析）

（大規模）ハイブリッド同化のための30メンバのメソ解析

GSM：（標準）水平解像度20 km、264時間積分（2024年3月時点の最新現業モデルでは水平解像度13kmに高解像度化している）。（大規模）水平解像度5 km

SCALE-LETKF：（標準）解像度500 m、50メンバ、30秒サイクルを1時間分実行

（大規模）解像度500 m、1,000メンバ、30秒サイクルを合計30日分実行

SCALE-LT：（標準）36時間積分。（大規模）格子数1200×600×60、24時間積分

SCALE-SDM：（標準）領域サイズ60 km×60 km×20 km（解像度125 m、3,700万格子）、超粒子密度128/格子（47億個）、2時間積分（72,000時間ステップ）

（大規模）領域サイズ60 km×60 km×20 km（解像度62.5 m、3億格子）、超粒子密度128/格子（377億個）、2時間積分（72,000時間ステップ）

kinaco：（標準）解像度100 m以下で数100 kmスケール領域、もしくは解像度1 m以下で数kmスケール領域の積分。（大規模）標準的な解析の100倍程度

NICAM：（標準）3.5 kmメッシュ（格子数 $10 \times 4^{11} \times 78$ ）を用いた1月積分

（大規模）① 220 mメッシュ（格子数 $10 \times 4^{15} \times 78$ ）実験、2時間積分。② 3.5 kmメッシュ（格子数 $10 \times 4^{11} \times 78$ ）、10年実験（4200万時間ステップ）

NICAM-LETKF：（標準）格子数267,632,640、256メンバ

（大規模）：格子数4,282,122,240、256メンバ

MRI-ESM：（標準）大気120 kmメッシュ、海洋1×0.5度メッシュを用いた1ヶ月積分

（大規模）大気10 kmメッシュ、海洋10 kmメッシュを用いた11日積分

GAIA：（大規模）T42、12年データ同化計算

MSSG：（標準）領域サイズ2 km×2 km×400 m、2 m解像度（格子数1,000×1,000×200）

（大規模）①領域サイズ40 km×40 km×20 km、2 m解像度（格子数20,000×20,000×340）、6時間積分。②領域サイズ35 km×25 km×2.5 km、2 m解像度（格子数17,500×12,500×270）、24時間積分

ISPACK3/SPML：（標準）T682の解像度で2次元乱流の数値計算、100万ステップ

（大規模）TL524287（格子数1,048,576×524,288）、球面調和関数①逆変換、②正変換

SCALE-DG：（標準）領域サイズ10 km×10 km×3 km、10 m解像度（節点数 2.5×10^8 ）、4時間積分。（大規模）標準的な解析をベースに2.5 m解像度、30分積分

4.4.3 2030年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

これまで述べてきた多種多様な気象・気候モデルは、それぞれ異なる目的・ブレイクスルーを目指して開発されている（4.4.3.1節）。これらのモデルから次世代計算機で期待されるブレイクスルーとしては、

- 毎年災害をもたらす集中豪雨・台風の予測精度の大幅向上に向けたりアルできめ細やかな気象予測システムの開発。竜巻・雷頻度予測の実用化、シナリオを網羅する確率予測、3次元高密度観測とデータ科学による高頻度リアルタイム予測の実現。

- 気候変動の理解と予測を質的に高める精緻な地球システムモデルの実現。特に、気候変動適応策で実利用可能な台風・極端現象・都市気候予測、パリ協定へ資する温室効果ガス吸排出量推定、気候系における海洋中規模渦の役割の解明。
- これまで不可能であった大規模数値実験の実現による、微小から惑星スケールに至るまでの現象の包括的理解の獲得。さらに精緻な物理的表現が可能な次世代型の力学、雲微物理、放射スキームの開発と実証。低精度、AI 代理モデルによる高速化。

と包括的にまとめることができる。ブレイクスルーの具体例を図 4.4.3.1に示す。

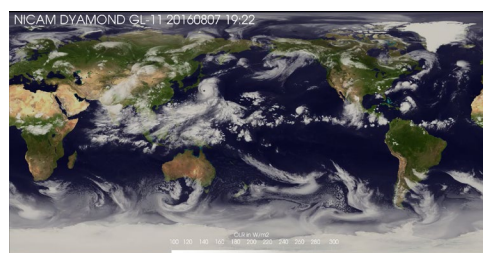
ブレイクスルーを達成するために必要な計算環境、解決すべき課題としては、

- ワークフロー全体で富岳比数十倍以上の実効性能向上。
- 加速演算器を利用するための高いソースコード書き換えコストおよび効率的利用に向けた大幅なアルゴリズム改変が課題。
- リアルタイム予測や全系実験を見据えたシステム高可用性。
- 高速・大容量な global/local ストレージ。
- 多様な力学・物理スキームやターゲット問題に対応可能な、メモリバンド幅、メモリ量、キャッシュ、IO、通信等の設計。

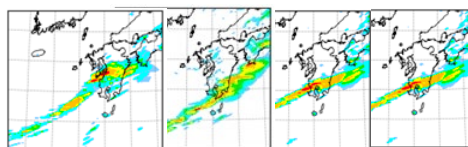
といった点が挙げられる。



データ同化



全球雲解像モデル



大アンサンブル

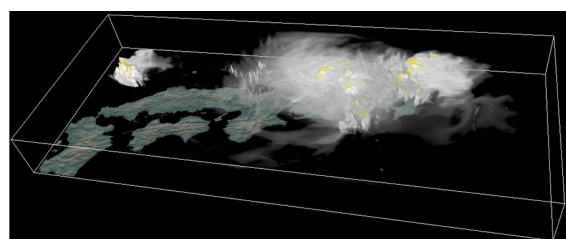


図1: 気象雷モデルの計算例。日本を横断した台風周辺の(白)雲と(黄色)雷の分布

気象雷モデル



動画

超水滴法による雄大積雲のシミュレーション結果。松嶋(神戸大)作成。

粒子法雲微物理スキーム

図 4.4.3.1 期待されるブレイクスルーの例

4.4.3.1 個別モデルから期待される成果の概要

個別モデルから期待される成果の概要を以下に、必要な計算資源の見積もりを表 4.4.3.1に示す。

asuca : より精度の高い予測のためのリファレンスとなる高精度モデル構築 (**asuca-Var**, **GSM**共通)。

①水平解像度1 km 1,000メンバ18時間予測。②水平解像度200 m 18時間予測

asuca-Var : インナーモデルの水平解像度15 km（アウターモデルの解像度5 km）、40万程度の観測データを同化するアンサンブルメンバ数1,000のハイブリッド同化モデル

GSM : 水平解像度5 km 1,000メンバ 132時間予測

SCALE-LETKF : 30秒更新の竜巻予報システム。解像度10 m（アダプティブメッシュ的なネスティングを想定）、1,000メンバ、30秒サイクル、30日分実行

SCALE-LT : 雷の頻度・分布および関連現象の予測と社会実装。格子数2,400×2,400×60、72時間積分、20アンサンブルメンバ

SCALE-SDM : 積乱雲集団のライフサイクル再現と形成メカニズム理解。格子数147億、超粒子数1.89兆、時間ステップ数720万、200時間積分、領域サイズ1,200 km×1,200 km×20 km、格子サイズ125 m、超粒子数密度128 /grid

kinaco : ① LES スケールでの沿岸域シミュレーション。1 mメッシュ、領域サイズ30 km×30 km（30,000×30,000×100格子）、dt=0.1秒、10日積分程度。② 非静力学スケールでの広域シミュレーション。30 mメッシュ、1,000 kmスケール、dt=5秒、1年積分程度

NICAM, NICOCO : ① 全球LES。110mメッシュ、398層、2日積分、dt=0.25秒。② 全球雲解像渦解像気候実験。大気3.5 kmメッシュ・78層、海洋0.1度メッシュ、100年積分

MRI-ESM : 日本周辺の黒潮や親潮の再現性を高めた高解像度地球システムモデルによる蓋然性の高い日本周辺域気候予測情報の生成。大気・海洋の水平解像度10 km（目標は大気・海洋の水平解像度kmスケール）。50年スピンアップ後に100年間のシナリオ実験。10メンバ並列実行

GAIA : 1太陽サイクル分のデータ同化・予測実験による超高層大気・宇宙天気予測の改善。水平T106、鉛直150層、12年、20アンサンブル計算

MSSG : ① 都市型豪雨を対象とした建物・雲解像豪雨シミュレーション。領域サイズ40 km×40 km×20 km、解像度1 m（格子数40,000×40,000×370）、6時間積分。② 自治体スケール領域を対象とした建物・樹冠解像の微気候シミュレーション。領域サイズ35 km×25 km×2.5km、解像度1 m（格子数35,000×25,000×300）、24時間積分

ISPACK3/SPML : 超高解像度の球面調和関数変換の高速化実現による、スペクトル法モデルやスペクトル解析の超高解像度化・高速化。ガス惑星大気モデル（非弾性系 3 次元回転球殻熱対流モデル）を経度・緯度・動径方向解像度 4096×2048×1536 格子点で1000惑星日計算。

SCALE-DG : 高精度化した力学コアによる広範囲での大気 LES 実験。領域サイズ800 km×800 km×15 km、解像度50 m（節点数640億）、20時間積分

表 4.4.3.1 2030 年頃に期待される成果を達成するために必要な計算資源の見積もり

アプリ名	問題	ノード時間積 (total)	入出力ファイル サイズ (total)	解析数	計算時間 /解析	総メモリ量 /解析
asuca	①	4.6 億 Nh	25.2 PB	1,000	10.3 日	75 TiB
	②	9,920 万 Nh	1.8 PB	1	10.3 日	7.5 PiB
asuca-Var		6,900 万 Nh	81 TB	1	12 日	4.8 PiB
GSM		8.2 億 Nh	56 PB	1,000	8.9 日	124 TiB
SCALE-LETKF		5,000 万 Nh	24 PB	1	20-30 日	検討中
SCALE-LT		2,100 万 Nh	100 TB	20	< 1 日	360 GiB
SCALE-SDM		7,900 万 Nh	6.4 PB	1	9.2 日	41 TiB

kinaco		3,900 万 Nh	検討中	1	1 日	91 TiB
NICAM/NICOCO	①	5.8 億 Nh	35 PB	1	10 日	42 PiB
	②	1.7 億 Nh	19 PB	1	2,250 日	1.3 TiB
MRI-ESM		1,550 万 Nh	83 PB	1	511 日	28 TiB
GAIA		2,000 万 Nh	100 TB	1	10 日	1 TiB
MSSG	①	2,900 万 Nh	5.9 PB	1	7.5 日	8.9 PiB
	②	5,100 万 Nh	2.6 PB	1	14 日	3.8 PiB
ISPACK3/SPML		2,000 万 Nh	1 PB	1	30 日	930 GiB
SCALE-DG		300 万 Nh	2 PB	1	40 日	36 PiB

※ノード時間積および計算時間は富岳タイプのアーキテクチャを想定して推定

4.4.4 ベンチマークの整備計画と状況

ベンダによる性能評価に資するため、本年度は代表アプリとしてSCALE-LETKFを選定し、ベンチマークを整備・提供するとともにGPU対応を実施した（4.4.4.1節）。富岳においてSCALEおよびNICAMを用いたベンチマークを実施し、フルモデルにおけるI/O性能の重要性を確認した（4.4.4.2節）。高I/O負荷のタスクを取り出してアルゴリズム改善を実施するとともに、I/O性能評価用のベンチマークセットを整備・提供した（4.4.4.3節）。

なお、気象庁の現業数値予報モデルについては、気象庁モデル評価・解析サブグループの活動として、asuca、asuca-Var、GSMをベンチマークプログラムとして気象庁気象研究所から理化学研究所に提供し、富岳を用いたベンチマークが実施された。

4.4.4.1 SCALE-LETKF の GPU 化

本年度は、SCALE-LETKFの気象シミュレーションパートであるSCALE-RMを対象にOpenACCを用いてGPU対応を行い、現実大気のシミュレーションをGPU上で実行可能にした。計算性能向上のため、時間積分計算中はデータをすべてGPU上のメモリに載せたままで計算を行うことで、CPUとGPU間のデータ転送を極力排除するとともに、その他のいくつかの高速化のための検証および実装を行った。性能検証として、東京大学情報センターのWisteria/BDEC-01を用いてCPUとGPUの計算速度の比較を行った。CPUとしてFX1000を4ノード、GPUとしてNVIDIA A100 4基を用いて測定を行った結果、GPUを用いた場合の方がCPUを用いた場合に比べて約1.5倍高速に計算できた。この比率は、概ね両者のメモリ転送速度の理論ピーク比に相当しており、GPUにおける計算性能がおおむね妥当であると考えられる。

ベンチマークとして、SCALEを用いたシミュレーション結果を入力としてLETKFによるデータ同化を実行するアプリケーションを整備し、ベンチマーク構築サブグループに提供した。また、より包括的な性能評価を可能にするため、SCALEのシミュレーションを含めたアプリケーション全体の実行を行うベンチマークの整備も行った。ただしこのベンチマークに含まれるSCALEのバージョンには、上記のGPU対応が含まれていない。今後は、2023年度末にリリース予定のSCALE ver.5.5を採用し、GPUを用いた場合のベンチマーク評価にも対応する予定である。

4.4.4.2 SCALE と NICAM のベンチマーク性能比較

SCALEモデルの中で一般的に用いられているSCALE-RM 5.2.6とベンチマーク構築時のNICAMの最新安定版であるNICAM.21を対象に、モデル全体および各スキームの性能を比較するため富岳において計算時間の計測を行った。表 4.4.4.1に実験設定を、図 4.4.4.1に計測結果を示す。それぞれのモデルで用いている力学・物理スキームやタイムステップは異なるため単純な比較は難しいが、両モデルにおいてweak-scalingおよびstrong-scaling性能ともに、力学過程に比べて物理過程で良く、力学過程の中ではスモールステップ（HEVI

ソルバ) において良い、という結果が得られた。両モデルは配列インデックスの順序やループ構造などが異なるものの、総じて両モデルの測定結果に大差はなく、主たる計算部分においては全体として良いスケーリング性能が得られていることが確認できた。

一方で、モデルの高解像度化に伴いプロセス数を増やすことで、データ書き出しの過程が全体の実行時間に占める割合が増加することが分かった(図 4.4.4.1 b)。そこで、表 4.4.4.1のNICAM実験③について、富岳におけるI/O高速化手段の一つであるストライピング設定を変更する感度実験を行ったが、実行時間のばらつきが大きく、現時点で有望な結果は得られていない(図省略)。多くの変数出力を必要とする気象・気候モデルでは、ハード面でのI/O性能向上、およびI/O性能を最大限に引き出すシステムソフトウェアへの期待は非常に大きい。

表 4.4.4.1 各モデルにおける実験設定

	SCALE-RM					NICAM				
	①	②	③	④	⑤	①	②	③	④	⑤
プロセス数	36	144	576	36	144	40	160	640	40	160
水平格子間隔[km]	1.0	0.5	0.25	0.5	0.25	56	28	14	28	14
格子数/プロセス数	4096		16384		4096		16384		16384	

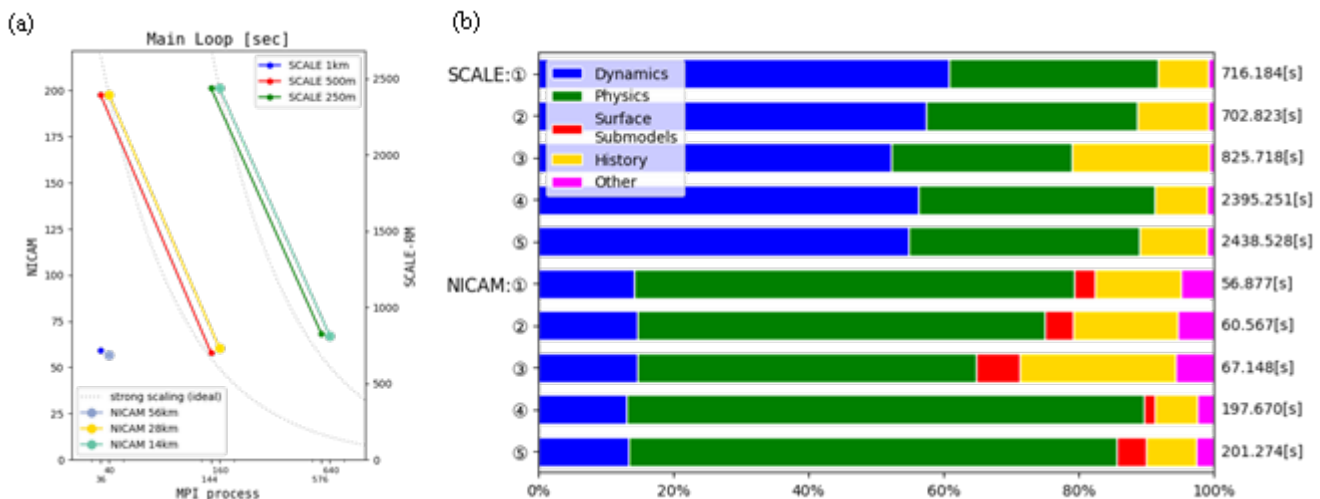


図 4.4.4.1 (a) モデル全体の計算時間のMPIプロセス数および水平解像度依存性。モデルによって縦軸が異なることに注意。(b) 力学、物理、陸面、出力 (History) およびそれ以外の過程の全体の計算時間への寄与。丸囲みの数字は表 4.4.4.1を参照。

4.4.4.3 I/O 評価用ベンチマークの整備

気象・気候シミュレーションでは大量のデータが入出力される(表 4.4.4.1)。一般に計算性能の向上に対してI/O性能の向上は緩やかであることから、実用的なワークフローに含まれるI/Oプロセスの性能評価は今後ますます重要になっていくと考えられる。そこで本年度は、典型的な高I/O負荷タスクである座標変換

(remapping) を対象としてI/O評価用ベンチマークセットを作成した。このベンチマークセットでは、NICAMで計算された全球14~0.87km格子の正二十面体格子データ[1]を入力とし、解析に適した緯度経度格子データへ変換してファイル出力する処理を想定する。独自フォーマットである正二十面体格子データはシミュレーション時にMPIプロセス毎に出力されるため、表 4.4.4.2に示す通り水平解像度に応じて640~20,480ファイルに

分割されている。座標変換後の全球の緯度経度格子データはNetCDF形式で保存され、as-isのアルゴリズムではmaster rankに集約して変数毎にファイルへ出力する（Algorithm-1）。ファイル出力プロセスを並列化するため、変数・鉛直層毎のファイル細分割による並列出力（Algorithm-2、およびMPI通信量を削減したAlgorithm-3）およびMPI-IOライブラリによる変数毎のファイル出力（Algorithm-4）という複数のアルゴリズムを実装して測定を実施した[2]。

図 4.4.4.2は4タイムステップのデータを座標変換した場合の実行時間およびメモリ使用量である。as-is版（Algorithm-1；灰色）に比べて改良版（Algorithm-2, 3, 4；緑, 青, 赤）では実行時間およびメモリ使用量が大幅に改善した。実行時間の観点からはファイル細分割を行ったAlgorithm-3が最も優れているが、高解像度化によるメモリ使用量の急増という難点がある。一方、MPI-IOを用いたAlgorithm-4はメモリ使用量の観点で最も優れており、実行時間およびファイル数の観点からもバランスが取れている。しかし、Algorithm-4ではMPI-IOライブラリに起因したstrong-scaling性能の低下が見られ、高解像度データの高速変換に課題がある。ここで示した多数のファイルを単一ファイルに集約するような処理は気象・気候シミュレーションにおける典型的な処理パターンであり、特に大並列時におけるMPI-IOライブラリの性能向上は気象・気候分野にとって恩恵が大きいと考えられる。なお、本I/O評価用ベンチマークセットはIO・ストレージ・ファイルシステムサブGへ提供済みである。

表 4.4.4.2 正二十面体格子データの仕様

水平解像度	全球格子数	ファイル数	データサイズ/ステップ ^o
14 km	• 2,621,442	• 640	• 11.5 GiB
7 km	• 10,485,762	• 2,560	• 46.1 GiB
3.5 km	• 41,943,042	• 2,560	• 184 GiB
1.7 km	• 167,772,162	• 10,240	• 738 GiB
0.87 km	• 671,088,642	• 20,480	• 2.95 TiB

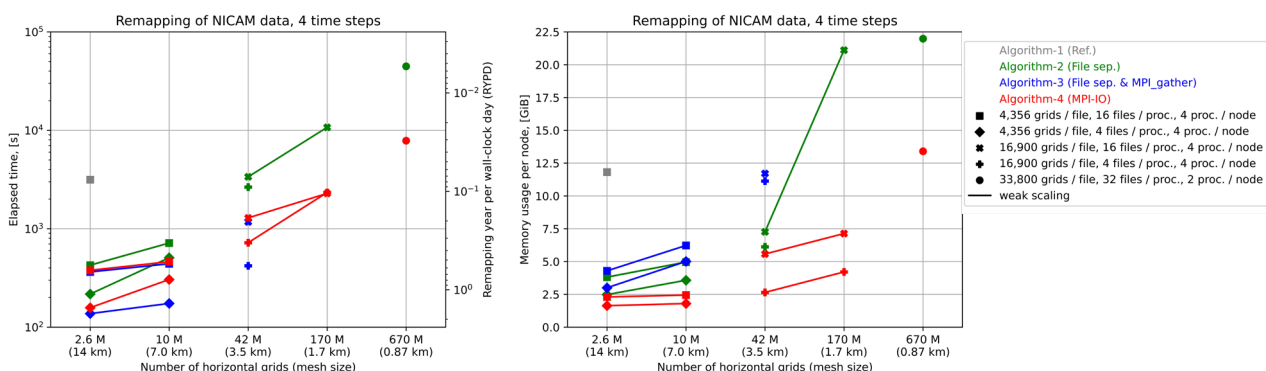


図 4.4.4.2 富岳における NICAM I/O ベンチマーク測定結果[2]。(左) 実行時間、(右) ノードあたりのメモリ使用量。

参考文献：

- [1] Y. Miyamoto, Y. Kajikawa, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita (2013), Deep moist atmospheric convection in a subkilometer global simulation. Geophys. Res. Lett. 40, 18, 4922–4926. doi:10.1002/grl.50944.
- [2] C. Kodama, H. Yashiro, T. Arakawa, D. Takasuka, S. Matsugishi, and H. Tomita (2024), Parallelized remapping algorithms for km-scale global weather and climate

simulations with icosahedral grid system, HPCAsia '24: Proceedings of the International Conference on High Performance Computing in Asia-Pacific Region,
doi:10.1145/3635035.3635040.

4.5 地震・津波防災分野調査研究サブグループ

4.5.1 研究分野の概要（2022 年度報告書より再掲）

それぞれ1,000点前後の観測点がある地震波形観測のためのHi-netや地表面変形観測のためのGEONETなどをはじめとして、日本列島には多くの観測網が設置されている。本分野では、これらをさらに有効活用するために必要とされている数値解析手法の準備を、「京」・「富岳」プロジェクトを通じて進めてきている。観測データと大規模シミュレーションの融合は、地震分野における様々なサイエンスへの貢献が期待されており、例えば、日本列島島弧の挙動予測、地震活動度推定、活動帯の高解像地殻変動・地震動を介したメカニズム解明、津波評価に必要な地殻変動推定、さらには、重要かつ困難な課題である地震発生予測をも視野に入れるインパクトのある成果が期待される。これらの解析を行う上で、不確定性を考慮するため、地殻構造や初期応力の異なるアンサンブル計算が必要となる。さらに、連続リアルタイムデータに対して適用し続けるために超大規模計算を高速に行うことが求められるため、次世代計算基盤への強い期待がある。

上記のような3次元複雑形状・非均質構造を求解対象とする地震分野においては、陰解法の低次非構造格子有限要素法が主要な計算手法の一つとなる。低次構造格子有限要素法に基づく地殻変動解析(静弾性解析・粘弾性解析)・地震動解析(動的な線形・非線形波動場解析)において基礎アルゴリズム・カーネルは共通であり、反復法ベースのソルバ内での疎行列ベクトル積が主要カーネルとなる。このカーネルにおいてはランダムアクセス・データレカレンスが卓越し、メモリバンド幅ネックのカーネルとなる。また、疎行列ベクトル積における1対1通信と反復法ソルバにおける1要素のAllreduceが主要な通信パターンとなる。本サブグループにおいては上記アプリケーションの調査・検討を実施した。

4.5.2 現行システムにおける解析の現況（2022 年度報告書より再掲）

- 当該計算手法で現在行われている標準的な解析例
 - 地殻変動解析・地震動解析ともに富岳数百～数千ノード・GPU 数十～数百枚を使った解析規模となる。計算特性は後述の大規模解析、グランドチャレンジ解析例と同様である。
- 当該計算手法で現在行われている大規模解析、グランドチャレンジ解析例
 - 地殻変動解析：問題サイズは 0.54 兆自由度、20 年間の解析で、現状、富岳 73728 ノードにて 3 時間程度の解析規模となる。メモリは各計算ノードあたり 16GB 程度使用している。通信時間のほとんどは疎行列ベクトル積における 1 対 1 通信に要しており、全計算時間の 10%程度を占めている。I/O 量は数十 TB オーダーとなる。
 - 地震動解析：問題サイズは 0.32 兆自由度、48 秒間の解析で、現状富岳 55296 ノードにて 4.5 時間程度の解析規模となる。メモリは各計算ノードあたり 16GB 程度使用している。通信時間のほとんどは疎行列ベクトル積における 1 対 1 通信に要しており、全計算時間の 10%程度を占めている。I/O 量は数十 TB オーダーとなる。

4.5.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

- 次世代計算機で期待されるブレイクスルー
 - 全球スケールの大規模マクロ地殻変動シミュレータと特定地域に絞った高解像度ミクロ地殻変動・地震動のシミュレータを組み合わせたマルチスケールの 3 次元地殻変動・地震動シミュレータを開発し、日本列島における地震波形データと地殻変動データの双方に整合するマルチスケールモデルを構築
 - これにより、一定規模の地震が発生した後、周囲での地震発生の推移を予測できるか検証できるようになると期待
- ブレイクスルーに必要な計算環境、解決すべき課題

- ① 上記のブレークスルーに必要な計算機について：富岳の5～10倍のメモリ容量・5～10倍程度のメモリ帯域、15～30倍程度の演算性能を想定(B/F 0.1程度)。ただし、ランダムアクセス・データレカレンスにより性能が著しく劣化しない計算機構が必要(hardware accelerated atomic add機構など)
- ② 将来アプリに向けての課題、ロードマップ等：ブレークスルーの実現に向け、観測データの高効率な取り込み方、また、B/Fの小さい将来システムにおいては従来の計算アルゴリズムでは性能を出し切れなくなることが課題として予想されるため、これらの課題の解決に向け手法開発を進める計画である。

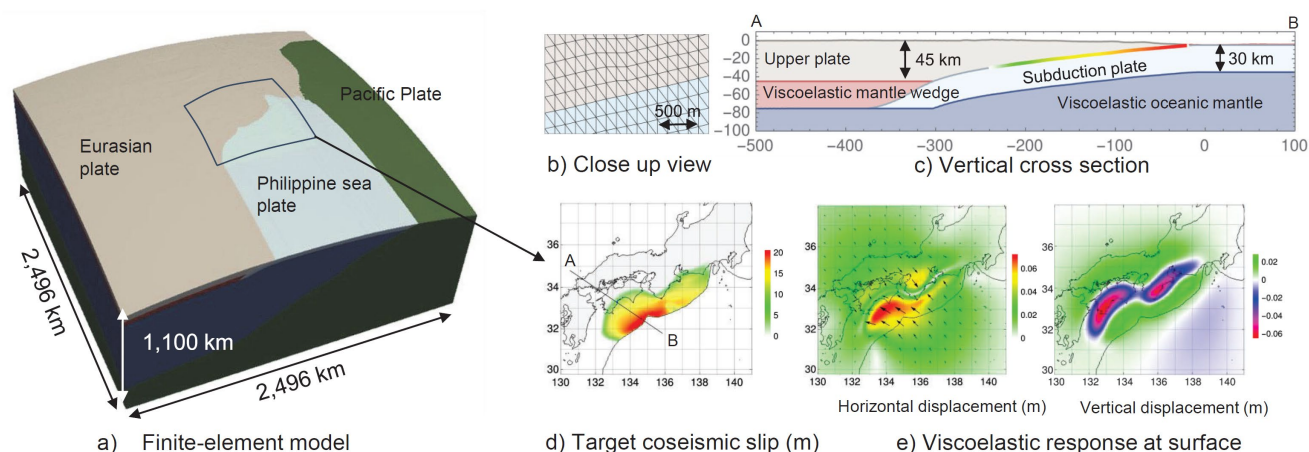


図 4.5.3.1 非構格子有限要素法を用いた地殻変動解析の例

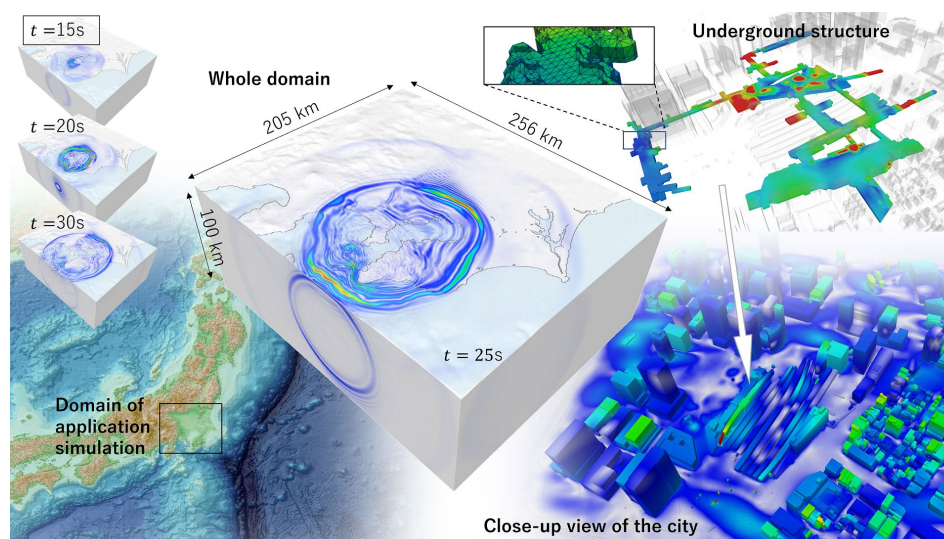


図 4.5.3.2 非構格子有限要素法を用いた地震動解析の例。図 4.5.1 の解析と組み合わせることで、地殻変動解析と地震動解析を組み合わせたマルチスケールシミュレータを開発し、地殻変動と地震動の双方の観測に整合するモデルを構築

● 必要な計算資源

- 全球スケールの大規模マクロ地殻変動解析：マンツルの底までを対象とした重力下での粘弾性応答計算・粘弾性完全緩和のため長期間の応答計算における必要計算資源量は先述の富岳全系の Grand Challenge Capability 問題の 30 倍程度(演算数 30 倍、必要メモリ容量：富岳全系の約 15 倍)と見積もっている。

- 特定領域のマイクロ地殻変動・地震動解析：南海トラフなど、1,000 km×1,000 km×500 km オーダーの解析における 1 ケースあたりの必要計算資源量は富岳全系の Grand Challenge Capability 問題と同様だが、多数ケースを実施することとなる。
- 計算資源に関する要望：本サブグループにおいては、計算資源量の見積もりと共に、次世代計算基盤において上記の計算を実施するにあたり重要となると考えられる性能項目について検討した。その結果、計算加速器(accelerator)において演算性能だけでなくメモリ容量・メモリ帯域も確保される構成が望ましいことが分かった(演算性能あたりメモリ容量・帯域が富岳と同等から 1/2 程度までに保たれることが理想的)。また、疎行列ベクトル積の計算においては足しこみ競合が発生するため、高速な hardware accelerated atomic add が実装されていることが性能上重要となることがわかった。また、1 対 1 通信においてはインターコネクトのバンド幅だけでなく計算加速器と CPU 間のバンド幅が律速になることが多いことから、計算加速器から CPU を介さずに通信するなど計算加速器からの通信性能が最大化される構成が望ましい。また、上記の計算を実施するにあたり、高速計算用の native なプログラミング環境・コンパイラに重点を置いた開発をすることが性能向上において重要であると考え（計算加速器/CPU の双方で実行可能なプログラミング環境(OpenACC や OpenMP)もあると良いが native なプログラミング環境で性能が出ることが重要）。

4.5.4 ベンチマークの整備計画と状況

2022年度においては、並列非構造格子有限要素法プログラムにおけるCPU用、および、GPU用に主要計算部である疎行列ベクトル積のカーネルプログラムの整備を進めたが、2023年度においては新たにフルアプリケーションの性能評価のため主要機能のみを残したアプリケーションプログラムをCPU用及びGPU用に整備した（以下、ベンチマークアプリケーションプログラムと呼ぶ）。ベンチマークアプリケーションプログラムが対象とする方程式、計算アルゴリズムおよび計算科学的特性について説明する。

- 対象方程式
 - 非線形波動方程式 ($\rho \ddot{\mathbf{u}} - (\nabla \cdot \mathbf{c} \cdot \nabla) \cdot \mathbf{u} = \mathbf{f}$, ここで、 ρ は密度、 $\mathbf{u}$ は変位ベクトル、 $\mathbf{c}$ は履歴に応じて変化する剛性テンソル) を空間方向を四面体二次要素で離散化し、Newmark- β 法による陰的時間積分を適用した結果得られる正定値対称な大規模疎行列に関する連立方程式 $\mathbf{Ax}^i = \mathbf{b}^i$ を各時間ステップ i について求解する。
- 計算アルゴリズム、および、各計算コンポーネントの性質
 - 3x3 ブロックヤコビ前処理を用いた共役勾配法を用いて $\mathbf{Ax}^i = \mathbf{b}^i$ を解く。主要計算は①Element-by-Element (EBE)法による疎行列ベクトル積、②共役勾配法におけるベクトル計算、③ベクトルの内積計算となる。すべての計算が FP64 で実施される。
 - ①疎行列ベクトル積においては Element-by-Element (EBE)法を用いる。この方法は、メモリに直接全体行列を格納し疎行列ベクトル積の度にメモリから読み出し右辺ベクトルとかけ合わせる方法(e.g., CRS 形式)と比べ、要素行列を計算するために演算量は多くなるもののメモリ使用量・メモリ読み出し量を削減することができ、演算性能に対してメモリ量・メモリ帯域が限られると予想される次世代計算基盤において適した方法となる（図 4.5.4.1 図 4.5.4.1 参照）。この方法においては、要素毎の結果ベクトル $\mathbf{Ape}$ を全体左辺ベクトル $\mathbf{Ap}$ に足しこむ際に足しこみ競合が起きるため、この部分の扱いが性能向上に重要となる。CPU 用にはスレッド数に応じた要素ループ i のカラーリング、GPU 用では $\mathbf{Ap}$ の足しこみ時に hardware accelerated atomic add を活用する。MPI プロセス間では袖領域の同期が必要となるため、MPI_ISEND、MPI_IRECV を用いた point-to-point 通信を実施する。全体領域分割においては METIS などのグラフ分割ソフトウェアを用いるため、非一様なランダム性の高い通信パターンとなる。なお、通信は疎行列ベクトル積における内部領域計算とオーバーラップ可能である。

- ②共役勾配法におけるベクトル計算：メモリバンド幅律速のカーネルとなる。MPI 通信は含まれない。
- ③ベクトル内積計算：プロセス内ではメモリバンド幅律速のカーネルとなり、通信パターンとしては 1 要素の MPI_Allreduce となる。

- ベンチマークアプリケーションプログラムの実装

- CPU 用プログラムは OpenMP+MPI で実装され、1 NUMA 領域あたり 1 MPI プロセスで実行することを想定している（例：富岳においては 1 CMG あたり 1 MPI プロセス 12 OpenMP スレッドで実行）。GPU プログラムは OpenACC+MPI で実装されており、1 GPU あたり 1 MPI プロセスで実行することを想定している。

今後、新規アーキテクチャ・演算加速器などに対応できるようアルゴリズムの検討・プログラム整備を継続して進める計画である。なお、上記ベンチマークアプリケーションプログラムは各種アーキテクチャ上での実装のしやすさ、及び、各アーキテクチャの性能比較がしやすいよう、マルチグリッドやデータ駆動型アルゴリズムを用いたソルバ前処理や、精度混合演算などの工夫がされていないスタンダードなアルゴリズムを対象とした。今後、ターゲット問題およびターゲットアーキテクチャにあわせて適切な前処理等を組み込んだアプリケーションプログラムを開発・検討することで、より信頼性の高い性能推定が可能になると期待される。

```

Ap(:) = 0
do ie = 1, ne
  Ae(1:m,1:m) = element_wise_matrix(ie)
  do j1 = 1, m
    pe(j1) = p(local_to_global(j1,ie))
  enddo
  Ape(1:m) = matrix_vector_product(Ae,pe)
  do j1 = 1, m
    Ap(local_to_global(j1,ie)) += Ape(j1)
  enddo
enddo

```

local_to_global: mapping from element-wise index to global index
 ne: number of elements Ae: element-wise matrix
 m: number of DOF per element Ap: results of the matrix vector product
 p: a vector to be multiplied by A

図 4.5.4.1 非構造格子有限要素法における疎行列ベクトル積カーネル(EBE 法)の概要

4.6 ものづくり分野調査研究サブグループ

4.6.1 研究分野の概要

ものづくり分野においては、現象解明を目的とした第一原理的な高精度解析や、様々な拘束条件を考慮した多目的最適化に対して、HPCの適用が進んでいる。特に近年では、多種・多様な市場ニーズを満たすために製品開発期間の短縮が求められていることから、HPCを活用した設計プロセスの革新への期待が高まっている。本サブグループでは、ものづくり分野において用いられるアプリケーションを対象として最新の動向やニーズを調査し、その結果を踏まえ、次世代計算基盤におけるターゲット問題やそのために必要な計算資源をまとめる。また、対象アプリケーションを用いてベンチマーク問題を設定し、アーキテクチャ調査研究グループによるベンチマークテストの結果を踏まえ、アルゴリズムの検討、必要計算資源の見直しを行う。

調査対象アプリケーションは、ものづくり分野における主要課題である乱流解析から選定した。この理由は、乱流解析ではメモリサイズやメモリバンド幅、ノード間通信性能がアプリケーションの適用範囲を決めることが多く、データ移動を重視する次世代計算基盤のアーキテクチャにより大幅な性能向上が期待できるためである。このことから、メモリアクセス特性の異なる2つアプリケーション、具体的には、FrontFlow/blue (以下、FFB)およびFFVHC-ACEを調査対象とした。

FFBは非圧縮性・圧縮性流体の非定常流動を高精度に予測可能なLarge Eddy Simulation (LES)に基づいた汎用流体解析コードである。形状適合性に優れた有限要素法を採用し、ファン／ポンプ等の流体機械や複雑形状周りの非定常乱流解析および流れから発生する騒音の予測が可能である。また、FFBはベクトル計算機およびスカラ型超並列計算機上で高速動作するように最適化されており、自動化された最適領域分割・統合処理を実装した領域分割法によって、大規模超並列計算に対応する。実際、「京」、「富岳」をはじめ多くの大規模並列計算機での実績があり、International Conference for High Performance Computing, Networking, Storage, and Analysis (SC20) では、ゴードン・ベル賞の最終候補に選出されている。一方、FFVHC-ACEは航空機前期複雑形状周りの高忠実な圧縮性流体LES解析を実現する流体解析コードであり、3つのキー技術、(1)複雑形状に対して完全自動格子生成を可能とする階層型等間隔直交格子法、(2)高レイノルズ数流れのLES解析を可能とする壁面モデルLES、(3)高忠実な圧縮性流体LES解析を実現する数値計算スキーム

(KEEPスキーム)を持つことを特徴とする。FFBとFFVHC-ACEの大きな違いは、流体場情報を格納するための内部データ構造である。FFBは非構造データを持つ。そのため、主要な演算部ではリストベクトルが多用されている。一方、FFVHC-ACEは構造データを持ち、ステンシル計算が主となる。

4.6.2 現行システムにおける解析の現況 (2022 年度報告書より再掲)

① FFB

● 概要

- 非圧縮性流体、圧縮性流体を計算対象とする汎用流体解析コード。乱流中の渦のダイナミクスを直接計算することにより乱流現象を高精度に予測。

● 計算手法

- 空間離散化：有限要素法
- 計算要素タイプ：六面体、四面体、三角柱、ピラミッド
- 乱流解析：ダイナミクスマゴリンスキーモデルによる Large Eddy Simulation。
- マトリックスソルバー：残差切除法を内部ソルバとする BCGSTAB 法

● 標準的な解析

- 解析規模：200 億グリッド、200,000 ステップ
- 演算量：80 EFLOP

- データ量：120 TB
- データ転送量：320 EB
- 通信量：24 EB (種類：1対1)
- IO量：80 TB
- 1解析あたりの計算時間：24 hour
- ノード数：10,000
- 大規模解析、グランドチャレンジ解析
 - ループ試験の代替計算
 - ◇ ポンプ・圧縮機の実機等レイノルズ数 10^7 以上の流れに対し、最大 20 兆グリッド規模の準直接計算を実施する。これによりループ試験の代替を実現することができ計測に匹敵する性能評価が可能となる。
 - ターボ機械のデジタル試作
 - ◇ 圧縮機、ポンプ、船舶用プロペラ等、レイノルズ数 $10^6 - 10^7$ オーダーのターボ機械内部流れに対し、数 100 億グリッド規模の計算格子を用いた準直接計算を実施し、一般性能・吸込み性能・騒音等を目的関数とした多目的最適計算により、数百個体のターボ機械をデジタル試作する。
- ② FFVHC-ACE
 - 概要
 - 階層型等間隔直交構造格子法を用いた圧縮性流体解析プログラム。流体现象の本質をとらえた準第一原理的手法を用いて航空機実機フライト環境を忠実に再現し、かつ高速に解析することが可能。
 - 計算手法
 - 階層型等間隔直交構造格子法
 - 高精度圧縮性解法
 - 壁面形状モデル
 - LES (Large Eddy Simulation)壁面モデル
 - 標準的な解析
 - 解析規模：650 億グリッド
 - 演算量：56 ZFLOP
 - データ量：0.26 PB
 - データ転送量：364 ZB
 - 通信量：36 EB (種類：1対1)
 - IO量：1 PB
 - 1解析あたりの計算時間：340 hour
 - ノード数：15,360
 - 大規模解析、グランドチャレンジ解析
 - 高速飛行性能を決める高速飛行条件における遷音速バフェット境界ならびにフラッター境界の予測評価、離着陸時の機体騒音の予測評価
 - ◇ 遷音速バフェット、騒音解析：100-1000 億格子規模の壁面モデル LES 解析
 - ◇ 飛行領域全体解析（含フラッター、運動解析）：1-10 兆格子規模の壁面モデル LES 解析
 - 飛行条件全体における性能や安全性に関する要求を統合的に評価する多目的・多設計変数の機体設計
 - ◇ 100 億格子規模の壁面モデル LES、1000-10000 個体規模の統合的多目的・多設計変数最適化

4.6.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

(1) 次世代計算機で期待されるブレイクスルー

- HPC を活用した現象解明とそれに基づいた最適設計が様々な産業分野の設計・製造プロセスで実用化されることで、製品の更なる高性能化や、多種・多様なニーズに対応した製品設計の迅速化等の実現が期待される
- ポンプ・圧縮機の実機等におけるレイノルズ数 10^7 以上の流れに対し準直接計算を実施することで、計測に匹敵する性能評価が可能となり、ループ試験の代替が実現
- 圧縮機、ポンプ、船舶用プロペラ等ターボ機械内部におけるレイノルズ数 $10^6 - 10^7$ オーダーの流れに対し、準直接計算に基づいた多目的最適化計算を実施することで、数百個体のターボ機械のデジタル試作が実現
- 航空機の高速飛行条件に関して、遷音速バフェット境界ならびにフラッター境界の予測評価、離着陸時の機体騒音の予測評価が実現。
- 飛行条件全体における性能や安全性に関する要求を統合的に評価する多目的・多設計変数の機体設計が実現

(2) ブレイクスルーに必要な計算環境、解決すべき課題

(2-1) 必要計算資源

③ FFB

- 成果
 - 30 倍程度のウィークスケーリング
- 必要な計算資源
 - 演算量：2.4 ZFLOP
 - データ量：3.6 PB
 - データ転送量：1.0 ZB
 - 通信量：720 EB
 - IO 量：2.4 PB
 - 1 解析あたりの計算時間：24 hour

④ FFVHC-ACE

- 必要な計算資源
 - 演算量：3.7 YFLOP
 - データ量：6.0 PB
 - データ転送量：24 YB
 - 通信量：283 EB (1 対 1)
 - IO 量：23PB
 - 1 解析あたりの計算時間：969 hour

(2-2) 将来アプリに向けての課題

- ウィークスケーリング向上のためのプロセス間通信の効率化（ランクマップの最適化等）
- 前後処理（格子生成、可視化）の高速化、および、各処理間のデータフローの効率化（ファイル I/O の削減）
- データ科学との連携による設計空間探索の効率化（設計空間の次元縮約、サロゲートモデルの構築、不確実性解析等）

4.6.4 ベンチマークの整備計画と状況

昨年度の研究で、ベンチマーク調査対象アプリケーションを、ものづくり分野における主要課題である乱流解析から選定した。この理由は、乱流解析ではメモリサイズやメモリバンド幅、ノード間通信性能がアプリケーションの適用範囲を決めることが多く、データ移動を重視する次世代計算基盤のアーキテクチャにより大幅な性能向上が期待できるためである。このことから、メモリアクセス特性の異なる2つアプリケーション、具体的には、FrontFlow/blue（以下、FFB）およびFFVHC-ACEを調査対象とした。

FFBは非圧縮性・圧縮性流体の非定常流動を高精度に予測可能なLarge Eddy Simulation (LES)に基づいた汎用流体解析コードである。形状適合性に優れた有限要素法を採用し、ファン／ポンプ等の流体機械や複雑形状周りの非定常乱流解析および流れから発生する騒音の予測が可能である。また、FFBはベクトル計算機およびスカラ型超並列計算機上で高速動作するように最適化されており、自動化された最適領域分割・統合処理を実装した領域分割法によって、大規模超並列計算に対応する。実際、「京」、「富岳」をはじめ多くの大規模並列計算機での実績があり、International Conference for High Performance Computing, Networking, Storage, and Analysis (SC20) では、ゴードン・ベル賞の最終候補に選出されている。一方、FFVHC-ACEは航空機前期複雑形状周りの高忠実な圧縮性流体LES解析を実現する流体解析コードであり、3つのキー技術、(1)複雑形状に対して完全自動格子生成を可能とする階層型等間隔直交格子法、(2)高レイノルズ数流れのLES解析を可能とする壁面モデルLES、(3)高忠実な圧縮性流体LES解析を実現する数値計算スキーム

(KEEPスキーム)を持つことを特徴とする。FFBとFFVHC-ACEの大きな違いは、流体場情報を格納するための内部データ構造である。FFBは非構造データを持つ。そのため、主要な演算部ではリストベクトルが多用されている。一方、FFVHC-ACEは構造データを持ち、ステンシル計算が主となる。

今年度の研究では、アーキテクチャ調査研究グループおよびベンチマーク構築サブグループによるヒアリングを受け、優先ベンチマークとしてFFBを選定した。選定した理由は以下の通りである。

- FFB の主要演算部ではリストベクトルが使用されており、メモリ性能を評価するのに好適
- FFB は富岳全資源の 10%以上を使用するほどユーザが多い
- FFB はすでに GPU 対応が済んでいる

また、全系システムの性能推定に関するヒアリングにおいて、以下の点を踏まえ、FFBのウィークスケール特性を評価することとした。

- ものづくり分野ではなるべく解像度の高い計算がしたいという需要がある（キャパシティコンピューティング）
- その実現には、ノード内のメモリ容量を増やすことが望ましい

次年度以降は、アーキテクチャ調査研究グループおよびベンチマーク構築サブグループの調査結果を踏まえ、次世代計算基盤のアーキテクチャに適したアルゴリズムとデータ構造について検討する。

4.7 基礎科学分野調査研究サブグループ

4.7.1 研究分野の概要

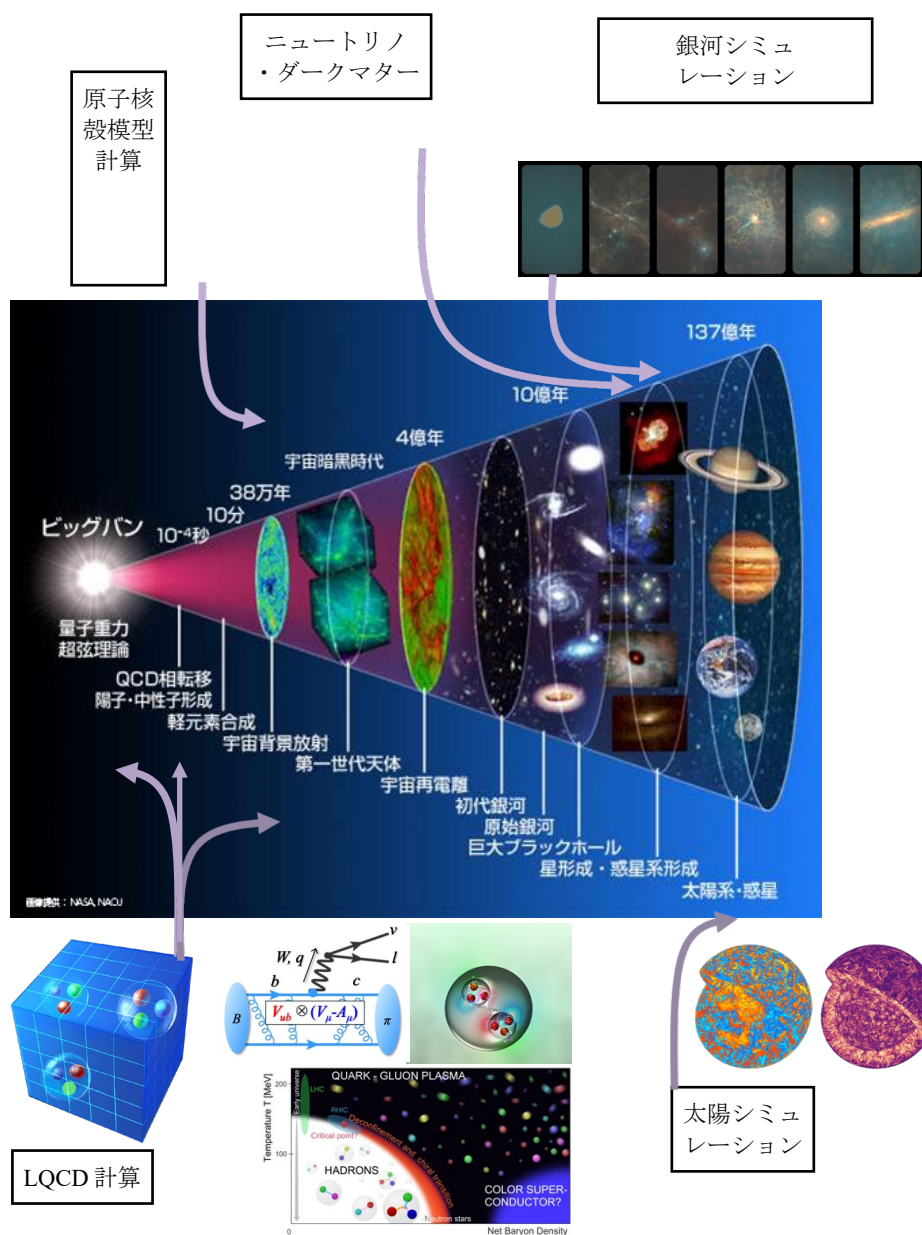


図 4.7.1.1 基礎科学ランドスケープ

基礎科学分野は素粒子のミクロな基本法則の検証に始まり、基本法則の(準)第一原理計算に基づく素粒子・原子核のミクロと惑星・宇宙のマクロな物理現象をカバーし、物質と宇宙のなりたち、またそこに登場する様々なスケールの物理現象の理解を目指す。

素粒子原子核における大規模数値計算としては、格子量子色力学(LQCD)に基づく3つの典型的な計算と原子核殻模型に基づく原子核計算の合計4つを代表アプリとして選定した。

量子色力学(QCD)はミクロな世界の物理法則である素粒子標準模型中、クォーク、グルオンの動力学を司る。力学を特徴付ける「カイラル対称性」が自発的に破れることにより原子核に質量が与えられ、ひいては目に見える宇宙の99%の質量を説明することになる。この動力学は遠距離では強大な引力と複雑な量子効果が渾然一体となり、遠距離相関から求められる核子質量や崩壊、散乱の定量的理解には大規模数値計算が唯一の武器となる。

この数値計算を実現する理論が連続時空を4次元格子グリッドに分割して定義した格子量子色力学(LQCD)である。LQCDのアプリとして、カイラル対称性が重要な問題にはカイラルフェルミオンの一つであるドメインウォールフェルミオン(DWF)が用いられる。一方、格子上のカイラル対称性を犠牲にしても、体積、統計量、小さい格子間隔などを優先すべき問題には、DWFより低コストなWilsonフェルミオン(WF)が用いられる。DWFでは標準模型の検証とそれを超える物理の探索、また、初期宇宙で起こったであろう高温でのQCD相のカイラル転移を扱う。WFでは、核力やハイペロン力、また重クォークハドロン間相互作用など、様々なハドロン間力を決定する。これにより、QCDに直接基づく原子核物理学の構築を行うと共に、中性子星の構造など、宇宙天文物理の謎を解明する。

QCDの動力学の産物である原子核の多体系を扱う殻模型計算とは、与えられた少数個の軌道からなる模型空間の多体問題を正確に解く手法である。十分に多くの軌道をとることで第一原理的な核構造計算にも応用できる。厳密対角化計算の限界を超えるため、少数の精選された多体基底によって原子核の波動関数を構成する、モンテカルロ殻模型および準粒子真空殻模型が提唱されている

天文学、宇宙物理学はギリシア時代から続く歴史ある学問である。古くはニュートンが微積分学を発達させた経緯となるなど、基礎科学を進める原動力となっている。主要な研究対象は太陽などの星やその集合体である銀河などである。他にも銀河ができる前のガス、星が重力崩壊後成長したブラックホールなども対象になる。研究手法としては物質の塊やガスの時間進化を追うメッシュ法の流体計算や N 体計算が主となる。一般にメッシュ法の計算ではメモリバンド幅が不足しがちであり、N 体計算では演算力が不足しがちになるが、研究対象によっては違うボトルネックが現れる。宇宙分野では全てのスケールを第一原理的に解けないことが多く、一段階小さいスケールの物理を導入するためサブグリッドモデルが用いられることが多い。そのようなサブグリッドモデルによって、計算上のボトルネックの箇所が変わり得ることに注意されたい。

天文学、宇宙物理学のコミュニティでは研究対象によって独自にコード開発することも多く、典型的なアプリを選ぶことは難しい。このベンチマークでは富岳で高い性能を実現したアプリ 2 つ、音速抑制法を用いた太陽シミュレーションと Vlasov 法を用いた ニュートリノダークマターシミュレーションと、より複雑なメモリアクセスパターンをもつ二つのアプリ、Athena++コードを用いた銀河シミュレーションと ASURA-FDPS コードを用いた銀河シミュレーションを紹介する。前者はメッシュ法が使われており、後者は N 体法の SPHを用いており、メモリアクセスパターンが異なる。

4.7.2 現行システムにおける解析の現況 (2022 年度報告書より再掲)

素粒子原子核分野

● LQCD: DWF-HMC

DWFは4次元時空に仮想5次元を加えカイラル対称性がほぼ厳密な格子フェルミオン法で、統計アンサンブル生成に分子動力学を確率棄却と組み合わせたハイブリッドモンテカルロ法(HMC)を用いる。律速箇所はまずは5次元フェルミオン大規模疎行列線形ソルバ(ステンシル計算)と、次いでゲージ場スメアリングの逆解法が上げられる。現時点での大規模計算の典型的な計算量は、ゼロ温度の1パラメータあたり15,000EFLOP、有限温度では $10\text{--}100 \text{ EFLOP} \times O(100)$ パラメータほどである。演算あたりのメモリデータ転送量は1.8Byte/Flop、通信量は0.03Byte/Flop(隣接通信)。問題サイズはゼロ温度では最大 $72^3 \times 144 \times 12$ 、有限温度では $48^3 \times 12 \times 12$ となる。コードセットは Grid、Bridge++ (どちらもC++)を用いている。

● LQCD: Wilson-HMC

Wilsonフェルミオンを用いたハイブリッドモンテカルロ計算により様々な物理量計算で用いられる配位を生成する。4次元フェルミオン大規模疎行列線形ソルバ(ステンシル計算)とそれに次いで、ゲージ場スメアリングの逆解法が上げられる。演算量は60,000 EFLOP。演算あたりのメモリデータ転送量は0.84Byte/Flop、通信量は0.011Byte/Flop(隣接通信)。問題サイズは 96^4 。逐次的にジョブを実行する必要のある capability

computing。コードはLDDHMC (QWS 機能取り込み)を用いる。

- LQCD:配位解析

現在行っている大規模計算においては、Wilson-HMCで生成した配位を用い、1ジョブあたり富岳の600ノード程度の並列計算を行っている。こちらは並列に多数のジョブを実行できる capacity computing であり、律速箇所がHMCと異なり、HMCと同等以上の演算量が必要な典型的な例として、ハドロン相互作用を選定した。より小さな規模の計算は、上記の数分の1から数十分の1の規模となる。コードはLDDHMC + Hadron-Force code。

- 原子核殻模型: モンテカルロ殻模型計算コード「rmcsm」

現在富岳にて行われている大規模な解析として、核力に基づいた第一原理的な計算を想定する。模型空間を調和振動子基底 7 主殻とし、核子数 10-20 程度、モンテカルロ殻模型の基底数 200 とする。1 解析あたりの演算量は 70EFLOP となる。メモリ量は 1 ノードあたり 20GB、B/F 値は最低でも 0.1 必要となる。通信の種類は Reduction、通信量は 1 ノードあたり 100GB 程度。I/O 量は無視できるほどに小さい。

宇宙惑星分野

- Athena++ - 銀河シミュレーション

Athena++は国際共同研究で開発している公開コードで、宇宙物理学の様々な問題で利用されている。ここでは銀河スケールの星形成シミュレーションを例にとり記述する。計算の目的は半径 $\sim 30\text{kpc}$ 、厚み数百 pc 程度の銀河円盤を 4pc 以下の高分解能で分解した磁気流体計算により、銀河全体の構造と分子雲スケールの星形成過程を整合的に理解する。以下のインプット物理が考慮されている。磁気流体力学、銀河重力、加熱・冷却、超新星爆発、（自己重力）。流体のソルバは以下である。磁気流体部-HLLD+CT 法（2次精度）、自己重力-Multigrid 法（2次精度）。適合格子法で効率的な計算を行うため、タスクリストによる動的スケジューリングをするのが Athena++の大きな特徴となっている。 10^9 yr 程度計算し、準定常的な構造が実現するまでシミュレーションを行う。

- ASURA- FDPS - 銀河シミュレーション

宇宙項があり、質量としては冷たい暗黒物質が支配的な宇宙において、宇宙初期からの暗黒物質およびバリオン(通常物質)の重力流体力学的進化を解き、銀河形成過程を明らかにする。シミュレーションのコア部分は、遠距離力である重力と近距離力である流体相互作用からなる。重力計算はツリー法を用いた N 体計算を、流体力学計算には Smoothed ParticleHydrodynamics 法を用いる。重力計算には全ノード通信が、流体力学計算には近接ノードへの 1 対 1 通信が必要となる。現在の銀河形成シミュレーションのグランドチャレンジになる銀河の恒星一つ一つを分解できる質量分解能を用いると、暗黒物質およびバリオンの計算に1ステップあたりの演算数は 63 Peta flops となる。宇宙年齢のシミュレーションにはおよそ 10^7 ステップ必要になるため、総演算数は **6.3 Yotta flops** となる。かなりの計算時間が必要になるため、時間刻み幅の短い領域と長い領域をハミルトニアン分割テクニックを使って分割し、独立に計算することで時間刻み幅を長く取る方法を実装している。

- 音速抑制法-太陽恒星

太陽表面にしばしば現れる強磁場領域である黒点は、11 年の周期を持って変動するが、この周期維持のための物理機構は未だ明らかになっておらず問題となっている。この周期は乱流と磁場、回転の相互作用の結

果として発現すると考えられている。しかし、観測を定量的に説明できるような首尾一貫したモデルは提示されていない。未だ太陽内部の理解には大きく改善の余地がある。この周期に関係するフレア・コロナ質量放出と言った爆発現象は、地球を含んだ太陽系環境に直接影響する。11 年周期の解明は、これらの現象の根源的理解、長期予測にも関連し、宇宙環境学の推進への貢献も大きい。この問題を音速抑制法を用いた3 次元高解像度磁気流体計算を用いて解明する。

- Vlasov – ニュートリノダークマター

従来の宇宙大規模構造形成の数値シミュレーションでは、物質分布を座標空間・速度空間からなる 6 次元の位相空間上で超粒子近似によって統計的にサンプリングする N 体シミュレーションが行われてきたが、計算結果に数値的ノイズが内在するのが最大の欠点である。その結果、自己重力系における無衝突減衰や 2 流体不安定性を精度良く取り扱うことができない。特に、宇宙空間に多量に存在すると考えられるニュートリノの宇宙大規模構造形成での無衝突減衰は、N 体シミュレーションでは精度よく計算することが困難な物理課程である。宇宙大規模構造形成におけるニュートリノの力学的影響を調べるシミュレーションでは、ダークマターやバリオンの運動は従来の N 体シミュレーション、ニュートリノの運動を Vlasov シミュレーションで解くハイブリッドシミュレーションを行う。

4.7.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

次世代計算機で期待されるブレイクスルー

素粒子・原子核・宇宙・惑星：ミクロから超マクロへいたる物理現象を扱う

- LQCD から標準模型検証とそれを超える物理の探索、宇宙や物質のなりたちを解明
 - カイラルフェルミオンを用いた理想的なシミュレーション
 - u, d, s, c, b クォーク物理点直上計算→素粒子標準模型の多角的超精密検証
 - QCD 有限温度相構造の解明、ダークマターの制限
 - ハドロン間相互作用の精密決定のための大体積シミュレーション
 - 2-3 体相互作用→ストレンジネス核物理実験、連星中性子合体での重元素合成解明
- 原子核殻模型：
 - 質量数 40 程度までの軽い核の第一原理的計算
 - 中重核領域での原子核の存在限界やエキゾチックな核構造の理論的解明
- 銀河シミュレーション/銀河形成
 - Athena++: 銀河全体から分子雲の内部構造まで一貫したシミュレーションを行う
 - ASURA-FDPS: 宇宙に存在する多様な銀河の形成進化史を論じられるようになる
- 太陽恒星
 - 太陽内部での磁場生成から黒点形成までを一貫して追う
- ニュートリノダークマター
 - 銀河団・銀河群スケールのダークマターハローを分解できる数値シミュレーション

ブレイクスルーに必要な計算環境、解決すべき課題

- アプリ x ハード 最終性能として「富岳」x 100 倍程度必要 (計算科学ロードマップ)
- ハードへの要望：
 - 広いメモリバンド幅とその有効利用のため十分なレジスタや OoO リソース
 - 複素数演算やステンシル計算へのハードウェアサポート

- 低遅延な隣接通信
- コードの書き換えは厭わないが、それを将来的な資産として活かせること
- ソフトウェア/アプリケーション
 - ワークフロー全体を通した性能向上
 - LQCD: 半精度演算利用、通信回避、様々な加速技法(Mutigrid, AI などの)R&D

4.7.4 ベンチマークの整備計画と状況

2022年度に全てのアプリケーションのカーネル部分の提出を完了しており、2023年度は素粒子原子核分野の一部のアプリの機能追加、さらにはワークフローの追加を行った。

機能追加

- LQCD:DWF-HMC(カーネル) – 既提出「富岳」最適化バージョンに加え、x86 最適化、OpenACC

ワークフロー

- LQCD:DWF-HMC: 上記カーネルを用いるハイブリッドモンテカルロ ワークフロー

なお、2022年度までに整備済みのベンチマークは以下のとおりである

素粒子原子核分野

- LQCD:DWF-HMC
- LQCD:Wilson-HMC
- LQCD:配位解析
- 原子核殻模型

宇宙惑星分野

- Athena++ (銀河シミュレーション)
- ASURA-FDPS
- 音速抑制法-太陽恒星
- Vlasov – ニュートリノダークマター

4.8 社会科学分野調査研究サブグループ

4.8.1 研究分野の概要（2022 年度報告書より再掲）

数理的に社会を理解し、取り扱うための方策として、直接社会を操作する代わりにシミュレーションを実装することで、多様な事例・事象を発生させ、社会を再現し、ある事象の原因や結果を分析することができる。これら多数の事象と予め設定したパラメータの間には、何らかのシンプルな法則性が成り立っている可能性がある。この分野ではエージェント・ベース・モデリング(以降、ABM)がよく用いられる。このモデルの特徴は、人間や車ひとつひとつを個別にコンピュータ内で再現するもので、極めてパラメータ数が多くなる傾向がある。したがってABMの系が全体として生じる法則性は数学的には自明であるかもしれないが、簡単に相図を書くこともできない複雑さを有する。加えて、各事象があるパラメータセットに対してユニークなものとなりがちであるため、これらの中から統計則を導き出すには各パラメータに乱数で幅を持たせたうえできわめて多数回実行することが必要となる。これらの複雑性に対処し、有意な法則性を抽出するためには、大きな計算能力が必要となる。

上記の社会シミュレーションの試みの中で、現在、検証込みでもっともうまくモデル化できているのが交通流の分野である。1本道の直線道路における渋滞は「基本図」等として挙げられる数理的なモデル化によりかなりよく理解されていて、ABMを用いて再現できることが確かめられ、校正もなされている。こうしてモデル化された多数の自動車・道を相互につなげたネットワークを構築することで、仮定した交通需要が正しければ、都市スケールのデジタルツインが構築できる。一方、こうした都市スケールの道路ネットワークの中で何が起きるか、パラメータに対してどのように応答するかについてはあまり分かっておらず、予測に耐えうる精度を持つ基礎方程式が存在しない段階にある。

上記の状況において、無数にあるパラメータセットから生じた大量のシミュレーション結果をたとえば機械学習による次元削減等の手法で処理し、社会の事象をいろいろな「相」に分離することが考えられる。これにより、パラメータセットを設定した際に、社会がどの相に位置しているのか、そして別の相へと転移するにはどのパラメータを変化させればよいのか推定できるようになることが望ましい。さらに、こうした法則性を見出すことの利点は、現実には起こりうることは十分考えられるものの、実は一度も生じていないという、現実の外側を含む空間をサーチできることである。たとえば、多数のシミュレーション結果から極端事象を抽出し、その発生条件を特定することで、災害時や、自発的に生じる混乱の兆候を前もって察知することが可能となるため、そこに期待される社会的価値は大きい。

計算機資源の観点から述べると、本分野においては基礎方程式が明らかでないために、我々の実行したい計算は「精度が保証されている複雑な連立(微分)方程式を多重並列させて解くための、HPCに特化した標準的なアプリケーションが与えられていて、その実行により高精度な予測が可能」という性質のものではなく、むしろ比較的単純な系をきわめて多数個用意し、多様なパラメータセットから結果への写像を獲得し、法則性を見出す手法が高い価値を持つ性質のものである。このため単純並列計算能力、および、各パラメータセット(シナリオ)を計算時間制限内で取り扱えるだけの十分な計算能力があることが望ましい。

上記の性質はアプリケーション開発の姿勢にも影響を与えてきた。これまで、色々なオープンソース・アプリケーションが開発され、これらを継いで研究が進展してきた背景がある。例えば次章で紹介するSUMOやCrowdWalk等の交通シミュレータ(いずれもABM)はオープンソースソフトウェアであり、広く使われていて無料で公開されているため追試も容易であり、バグ出し等・Fix等のリソースもコミュニティの活動に期待することができる。これらの整備状況・開発能力は我々の社会における重要な資産であり、研究の要である。これらは一般的なPC、ワークステーション、クラスター等で数個程度のシナリオについて計算させることが想定されていて、HPCに特化しているわけではないため、これらをHPCにおいて利用しようとする際には困難が生ずる。使用する側の立場から見ると、公式で提供されている標準的なインストール方法を忠実になぞる作業によりインストールできることが望ましいが、京・富岳等へのインストールはそうした例に該当しない。パッケージに同梱されたmakefile・シェルスクリプトなどを、エラー出力を頼りに丹念に書き換えコンパイルする作業コストが発生し、動作させるまでに数ヶ月程度を費やす必要が生じるか、あるいは使用を断念する等の問題が生じてきた。

一方で、富岳にはSingularity仮想技術が投入されており、これによってオープンソース・アプリケーションの移植が飛躍的に容易となった。市販のノートPC上で、標準的なインストール手順により構築したDockerイメージをアップロードし、Singularityイメージに変換することで、実行したい計算が実行できる環境が1週間程度で整備できることが明らかとなった。一方で、次章で述べる通り、これらの実行は一般的なPCで実行したものに対して、数倍程度の計算時間を要する傾向にあるため、CPUの計算能力の向上、あるいはこれらの仮想化のさらなる最適化が望ましい。

今後10年程度、あるいはそれ以上にわたり「社会の基礎方程式」は不明であり続ける一方、研究は進展するため、部分的に法則が発見される可能性がある。その場合、新たなモデル・アルゴリズムを導入した新たなアプリケーションが開発され、そちらへの乗り換えが必要となる可能性も考えられる。それらのオープンソース・コミュニティによって開発されたパッケージを少ないリソースで導入できれば、社会シミュレーション・デジタルツインの研究開発において計算資源の利用価値が飛躍的に高まると考えられる。

4.8.2 現行システムにおける解析の現況

解析プログラムの現状としては、ABMの並列化は高性能な実装が困難であるため、シナリオ別に自明的な並列処理を用いることが多い。シナリオは多数あり、大規模並列処理が可能である。現在は倍精度浮動小数点数を用いているが、単精度演算の利用は考えられる。それ以下の精度の演算では問題が発生する可能性がある。シミュレーション全体としては、高速なソーティングアルゴリズムがあるとよい。上位の要素だけを知りたいというケースもある。ファイルIOへの要求も比較的高いと考えられる。20万エージェントのシミュレーションで、30時間で1TBを書き出す。

人流シミュレータと自動車交通シミュレータをとりあげる。手元のPCで標準的な解析、大規模なものとしては単純並列でパラメータを変えながら極めて多数のシナリオを投入する点は同様である。

4.8.2.1 SUMO(自動車交通シミュレータ) (2022 年度報告書より再掲)

オープンソースABMシミュレータである。自動車交通の研究業界でよく使われている。OSM、GISデータをインポートする機能が実装済みであり、任意の都市・地図における自動車交通を再現できる。

標準的な解析は、区ないし市程度の広がりをもつ領域を、数シナリオについて再現する規模である。例として、神戸市中心部のみを再現したシミュレーションを、現実に近い平常時の交通(約1万台)を投入し、シミュレーション内時間を6時間、学習に基づく渋滞回避を繰り返し回数を6回に設定して実行すると、2007年当時において一般的だったPC・クラスタにおいて約4時間を要した。出力データは約2GB/コアである。京では約8時間を要した。記憶容量の問題から、1ノードにつき2ないし4並列でシナリオを投入していた。また、最大で約1200個の試行(パラメータセットを乱数で振ったもの、計算が軽いものも含まれる)を単純並列で投入し、12時間以内に全て完了した実績がある。

これと同一のシミュレーションを、富岳と同一のCPUをもつFX700上でSingularityイメージにインストールして実行すると、約14時間を要した。富岳が京に比べて若干長い実行時間を要するのは、仮想化によるオーバーヘッドのほか、ABMであるためエージェントの行動を決定するif文が大量に含まれ、複雑な処理を要する箇所が、Arm CPUアーキテクチャに対して最適化されていないことに起因する可能性がある。それぞれについて約5倍程度、計算時間が延長すると仮定すれば、富岳は実質的に京の5～10倍程度の能力で計算できていると推定される。

2022年度にはより広いシミュレーション内の領域を実装すべく、神戸市中心部から近畿全域に拡大した新たなシミュレーションを構築した。道路は緊急輸送道路(主要国道など)のみに限定し、近畿地方の自動車登録台数の10%(約10万台)を12時間にわたって走行する設定で、現行のPC・クラスタを用いてこのシミュレーションを実行すると約22時間を要した。出力データは約1.4 GB/コアである。

大規模な解析は、上記のシナリオ数を標準解析における数個程度から、数千～数万程度に単純並列で増加させるものである。近未来において近畿全域を含むシミュレーションを100万シナリオ実行し、統計分析・機械学習によって処理し、相図を抽出する計画である。

4.8.2.2 CrowdWalk(人流シミュレータ) (2022 年度報告書より再掲)

オープンソースABMシミュレータである。OSM、GISデータをインポートする機能が実装済みであり、任意の都市・地図における人流を再現できる。

標準的な解析として、ポートアイランドを除く神戸市中央区をほぼ全域再現し、平常時・イベント開催時・災害発生時について、数シナリオ程度を再現するものである。実行時間はおおむねエージェント数に比例する特徴がある。2300人が避難し、完了するまで打ち切らないシナリオを設定すると、手元のArm アーキテクチャのPC (apple M1 max Macbook Pro)においては30・15分程度(それぞれGUI・CLIモード)を要し、富岳上で約50分(CLIモードのみ実行可能)を要した。富岳における実行は手元のPC(GUIモード)の約2倍の時間を要すると考えられる。さらに、上記を拡張した解析として、地図は変更せず、9万人が帰宅困難者となって運行を停止した電車が来るのを待ち、駅周辺で大混雑が発生するシナリオでは、GUIモードで約20時間を要した。富岳で実行していないが、先ほど述べたように倍の時間を要すると考えると、約40時間と予想できる。

大規模な解析は、上記のシナリオ数を標準解析における数個程度から、数千～数万程度に単純並列で増加させるものである。近未来において、現行の1.5倍程度のエージェントを追加したシミュレーションを数万シナリオ実行し、統計分析・機械学習によって処理する計画である。ただし、上記計算時間ではsmallのWall Timeの制限から通常実行が難しくなるため、可能であれば地図を複数の場所に分割し、それぞれについてシミュレーションを実行する計画である。その場合、1つのシナリオを複数のノード(またはコア・スレッド)に分割して実行することになるため、確保が必要となる資源数はさらにファクター倍になると予想される。

4.8.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題 (2022 年度報告書より再掲)

自動車交通について、主要国道等のみを抽出して粗視化すれば、市区にとどまらず地方レベルまでは計算可能であるが、通常実行の限界に近いことを上記で示した。2030年頃には、CPU能力の向上により、粗視化しない全ての道路を含んだ地方レベルのシミュレーション、もしくは全国を対象とする粗視化されたABMモデルが構築・計算可能となるのではないかと考えられる。上記の例で見た近畿地方は5県を含み、全国の軒数と比較すると10倍程度の計算規模が必要で、粗視化しない道路の数もその程度に収まる(収める)と考えられるからである。また、広がったシミュレーション領域に応じてパラメータも増えると考えられるため、並列数も10～100倍程度に増やす必要があると考えられる。6時間あれば東京・大阪間を車で移動することが可能だが、そのような広域のシミュレーションに市街地まで含まれていた例はこれまでに存在しない。これが可能となれば、災害によって発生した広域の避難や物資輸送などにおいて、きちんと繋がった(省略によって途切れていない)ネットワークを前提に、どこにどのような渋滞が生じるか、その渋滞の原因は何であるか(ネットワークのどこに設計の良くない点があるか)等が推定可能となる。

人流シミュレーションについても、神戸市中央区のみを再現するという計算量を下げる方向に作用する制限が存在するにも関わらず、現時点の9万人でもかなり通常実行計算時間の限界に近いが、倍以上の(現実的な見積りに基づく)23万人を投入したい、という希望が行政から寄せられている。2030年頃にはこれらの制限が撤廃でき、県レベルでの人の移動をシームレスに再現できる可能性が高い。これは、交通インフラの計画・需要予測や、災害が生じた際の大規模な人の避難計画等に資すると考えられる。

4.8.4 ベンチマークの整備計画と状況
2022年度に整備した環境を引き続き用いている。

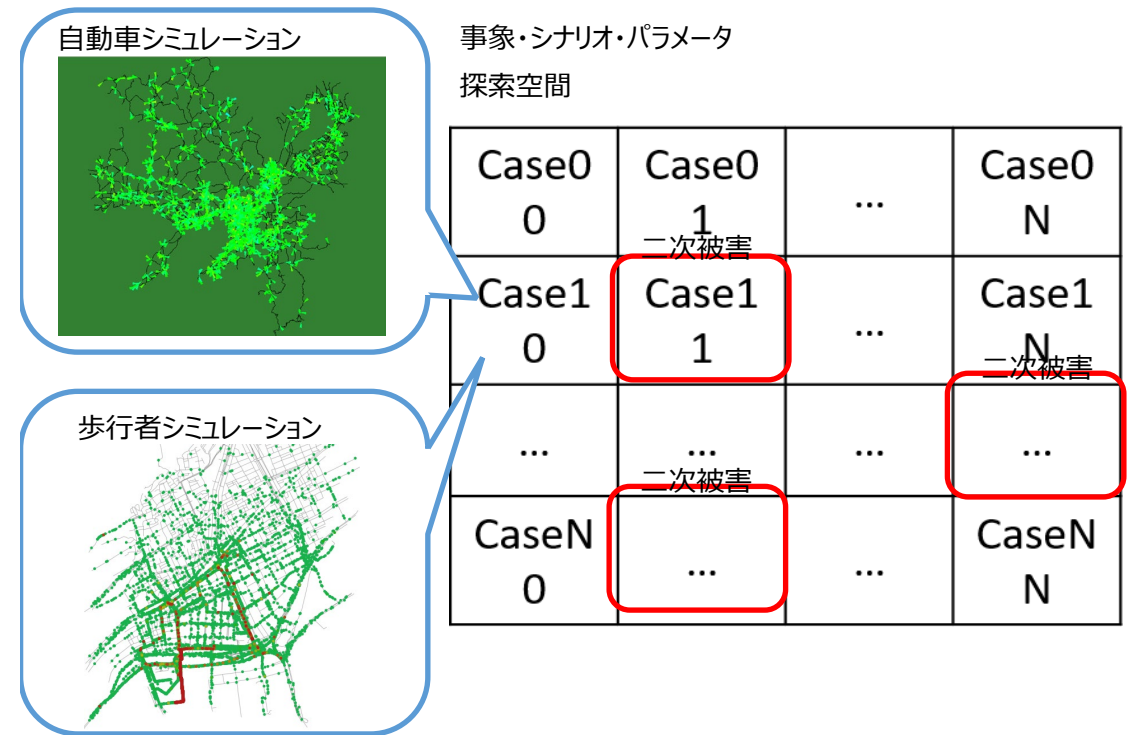


図 4.88.1 自動車シミュレーションと歩行者シミュレーションの連携による事象シナリオ探索空間

4.9 デジタルツイン・Society 5.0 分野調査研究サブグループ

4.9.1 研究分野の概要 (2022 年度報告書より再掲)

近年、センサから取得したデータを用いシミュレーションや機械学習などによってサイバー空間に現実世界を再現するデジタルツインや、そのサイバー空間とフィジカル空間を高度に融合したシステムを形成して実現する Society 5.0 の実現に向けた取り組みが始まってきている。デジタルツイン・ Society 5.0は、気象・気候、マテリアル、ものづくり、創薬、都市計画、社会基盤などの様々な分野で活用されることが期待される。この分野では、アプリケーションを実行し、現状分析や将来変化の予測を行う。将来的には、これらのアプリケーションを高度化し、大規模に実行することが予想される。

本分野調査研究サブグループでは、デジタルツインがどのように構築され、どのように活用されているか、Society 5.0を実現するシステムがどのように構築されているか、を調査し、それに必要な計算資源や高度な融合を実現するために必要な周辺技術の調査・分析を行う。特に、デジタルツイン・ Society 5.0では、データ取得、シミュレーション、分析、機械学習、データ処理、可視化など、計算や通信特性が異なる複数のアプリケーションを組み合わせ実現されることが多く、そのワークフローの分析が重要となる。

4.9.2 現行システムにおける解析の現況 (2022 年度報告書より再掲)

本調査研究で調査したアプリケーションについて説明する。

(a) NICAM-AI

全球非静力学大気モデルNICAMの計算プロセスの一部（現在は雲物理計算を対象）をAIによる計算で代替する。計算は(1) 対象プロセス（雲物理計算）の入力変数から出力変数を再現するようにAIに学習させる学習フェーズと、(2) 学習結果を用いて入力変数からAIによって計算された出力変数を返す計算フェーズに分けられる。なお、現在の調査は学習フェーズに限定したうえで更に予備実験段階の結果を対象としている。

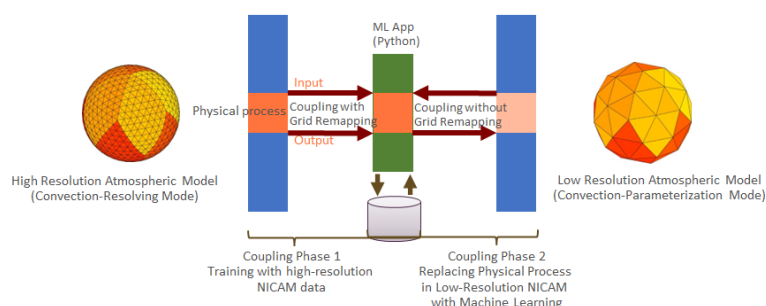


図 4.9.2.1 STYLEREF 2 ￥s 4.9. SEQ 図 ￥* ARABIC ￥s 2 1 NICAM-AI の
学習フェーズと計算フェーズ

図 4.9.2.1左半分の学習フェーズでは高解像度NICAMの雲物理計算の入力変数と出力変数が図中央緑色で表されるAIに送られる。この際、カブラの機能によって高解像度格子のデータは低解像度格子に変換（補間）される。変換されたデータはAIに渡され入力変数から出力変数を再現するように学習が行われる。計算フェーズ（図 4.9.2.1の右半分）では低解像度NICAMの雲物理計算の入力変数をAIに送り、AIが再現した出力変数がNICAMに返される。

現在は予備実験段階であり、次の条件および資源量で計算されている。NICAM側は、水平格子数163840、鉛直層数78の計算領域に対して東京大学情報基盤センターのWisteria-Odyssey 40ノードを用い160プロセスで実行される。プログラム内の通信は袖領域の1対1通信である。一方で、AI側はPyTorchで実装されており、水平格子数10240、鉛直層数40の計算領域となっている。東京大学情報基盤センターの

Wisteria-Aquarius 1ノードを用いて20プロセスで実行される。なおPyTorchによる学習計算は1プロセスで行っている。そのためデータ集合通信を行っている。

現状の学習フェーズの計算では、6時間積分の計算に対して、NICAM側で時間ステップ120秒で実行し、AI側は時間ステップ3600秒で実行している。1時間積分ごとに通信し、合計7回のNICAM側とAI側で通信を行い、通信に約20秒、エポック数を100として学習に約45秒程度かかっている。NICAMによる計算は通信の時間内で終わる。

(b) NHM-LETKFおよび asuca-LETKF

気象モデルによる予報とそこへの観測データの統合を繰り返す、気象予報の標準的なワークフローである。NHM、asucaはそれぞれ気象庁の旧版、新版の領域気象モデルであり、LETKFはアンサンブルカルマンフィルターの一種である。

富岳による計算では、1000アンサンブルメンバで1サイクル3時間積分を行うと、計算量は2500ノード時間程度で、気象モデルとLETKFで1対5程度の割合で資源を使う。典型的には900ノードを利用する。データは、ほとんどが $273 \times 221 \times 40$ 自由度のモデル出力で、合計200GB程度利用する。気象モデル（NHM、asuca）とLETKFの間の通信はI/Oを用いて行う。LETKFは、モデル出力である200GBとデータ同化のために用いられる観測データを受け取る。通信には、1分から最大10分程度の時間を要する。

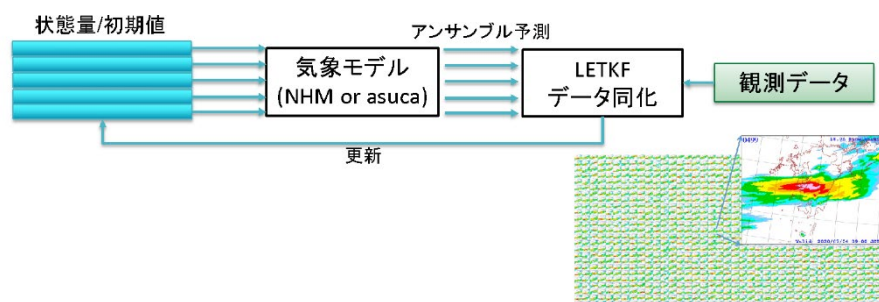


図 4.9.2.2

(c) WRF-UQ

気象モデルWRFの物理パラメタリゼーション選択や、モデルパラメータを観測データに基づいて最適化する。UQは Uncertainty Quantificationの略であり、パラメータの事後確率分布をベイズ推定するところが特徴である。計算コスト削減のため、パラメータと評価関数の関係を機械学習で近似し、また感度解析によりパラメータの次元削減を行っている。

富岳を用い、一度に生成する学習データを4000アンサンブルメンバとし、72時間先の降水量分布を評価関数として35パラメータを求める問題では、計算量は感度解析およびMCMC部分を除いて4000ノード時間程度必要とする。感度解析およびMCMCは、現時点では数ノード時間程度必要とするが、まだ試行錯誤している段階であり、今後、増減する可能性がある。特にデータ量が大きくなれば必要とする計算資源は増加する。データは、ほとんどが $273 \times 221 \times 40$ 自由度のモデル出力で、合計8000GB程度出力する。通信は、学習データのための受け渡しが生主であり、これは出力データのごく一部のみである。

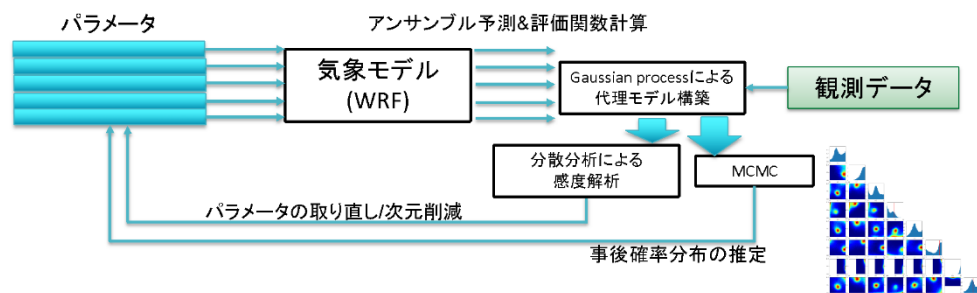


図 4.9.2.3 WRF-UQ の計算フロー

(d) CityLBM

メソスケール気象データを同化した局所風況シミュレーションである。シミュレーションを格子ボルツマン法（LBM）で実行し、これにデータ同化を併用する。2m解像度の計算の場合、典型的な利用ノード数とGPU数は、Wisteria-Aquarius（NVIDIA A100搭載）を用いると、アンサンブルメンバあたり1ノード（8 GPU）、8アンサンブルメンバで合計8ノード（64 GPU）を用いる。JAEA SGI8600（NVIDIA V100搭載）を用いると、アンサンブルメンバあたり4ノード（16 GPU）、16 アンサンブルメンバで合計64ノード（256 GPU）を用いる。

HPE SGI8600搭載のV100利用において、1解析あたりの計算時間は24時間で、1日の風況シミュレーションが可能である。1プロセスで利用している計算資源は次の通りである。LBMでは演算数が多いCumulantカーネルとメモリバンド幅を要求するstreamingから構成されているため、V100の演算性能 15.7 TFlops（単精度）、メモリバンド幅 0.9 TB/s により、高速な計算が初めて実現している。データ同化は、アンサンブル数の大きさの密行列（～642）の固有値を格子点数だけ計算する必要があり、演算性能 125 TFlops（テンソル演算）を利用した高速な行列演算積が用いられている。全体の計算コストの割合は、LBM対LETKFで3対1である。また、AlltoAll通信が用いられ、InfiniBand EDRではプロセスあたり通信時間のデータ同化処理に占める割合は10～50%（8～64アンサンブル）である。計算、通信、I/Oのコーデザインに取り組んでおり、メモリを使い切ることで通信領域の割合を減らしている。ノード内4GPU（8GPU）の計算・通信のバランスでも問題ない。また、I/Oに関連したデータはCPUメモリに退避し、計算とI/Oをオーバーラップを実現し、出力データは1プロセスあたり16GBで30分程度要する。

1プロセッサあたりの計算性能の実績（単位：MLUPS=Mega lattice update per second）は、498（NVIDIA V100）、732（NVIDIA A100）、432（AMD MI100）、10（Fujitsu A64FX、4CMG）である。富岳型プロセッサでは大きな性能劣化があり、コンパイラの不備およびレジスタ不足によるものと考えている。

4.9.3 2030 年頃のアプリ像とそれを実現するために必要な計算環境と解決すべき課題

ベンチマークとして整備した (a)、および(d)を簡易的にした2次元マイクロベンチマークコードの2種類について述べる。

(3) 次世代計算機で期待されるブレークスルー

(a)については、シミュレーションと機械学習の連携による、データ駆動型のパラメタリゼーションモデルの構築と、シミュレーションの一部をそのモデルで代替することによるシミュレーションの高解像度化と超高速化による現行の100倍～1000倍高速な100年から1000年といった長期間の全球気象計算を実現できることが期待される。

(d)については、東京都心部の10km四方に対して、IoT機器のデータをリアルタイムに同化をともなった25cm解像度のリアルタイム風況デジタルツインを実現できることが期待される。

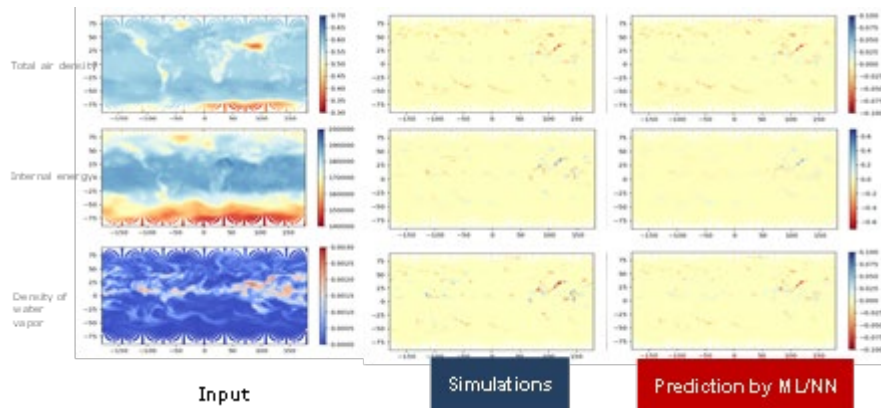


図 4.9.3.1

(4) ブレークスルーに必要な計算環境、解決すべき課題

(1)を実現するためには、次のような計算環境を必要とし、解決すべき課題があると考えられる。

(a) は、メモリ性能を求める気象モデルNICAMと機械学習という特性が異なるアプリケーションの組み合わせで構成される。前後の処理を含み、これらをシームレスに実行できるということが重要である。必要な計算環境としてホモジニアスなシステムとする場合、気象モデル、機械学習、データ処理という特性の異なるアプリケーション（異なるバイナリとして提供される）がそれなりに高い性能で動く必要がある。ネットワーク的には複数バイナリが密に連携しやすいのは利点である。一方で、ヘテロジニアスなシステムの場合は、異なるサブシステム間のネットワークの性能を高くすることが重要となる。浮動小数点計算は、精度を保ちながら性能を上げるためには混合精度があることが望ましい。気象モデルの一部の計算で、単精度までは利用可能であるが、半精度の利用は難しい。AI部分では半精度を利用する。

AIの利用は、現状は導入の初期段階であるが、今後は他の要素もAIに代替していくことが予想される。2030年頃はAIの利用は必須になると考えられる。AIはPythonフレームワークで実装されている。Python-Fortran 間でMPI通信を行えることが望ましい。量子計算の利用は想定していない。

コード開発・維持にかけられるコストの観点では、NICAMは100万行程度のコードを持つため、特定のアーキテクチャが長期間使われることが見通せない、そのアーキテクチャへの移植や最適化は難しいと考えている。

(d) は、メモリバンド幅律速のアプリケーションである。デジタルツインのためデータ同化が必要で密行列計算、固有値計算を多用しており、これらの高速化が必要である。アプリケーション本体はGPUに実装されており、IOの効率化のため、IO処理はGPUとCPU間のユニファイドメモリを使って実装している。IO処理はGPU計算で隠蔽している。

計算ノードは、ノードごとにMPIランクを立てる場合、表面積と体積の関係から、なるべく1ノードで多く計算ができることが好ましく、ファットノードが好ましい。また、アプリケーション開発では仮想化技術を利用したい。メモリバンド幅、メモリ容量ともに2028年頃であればNVIDIA H100の8倍程度あると望ましい。浮動小数点計算の一部では、単精度まで利用している。関連研究の報告によると半精度の利用は難しいと考えている。

現状では、AIは利用していないが、データ同化部分にAIを導入したいと考えており、今後もAIの利用は高まると想定している。

コード開発には、CUDA、HIP、Kokkosなどを利用しており、コード開発・維持にかけられるコストの観点では、既存の言語や指示文で対応できるアーキテクチャとなっていることが望ましい。

4.9.4 ベンチマークの整備計画と状況

2022年度末までに、ベンチマークとして、(a)、および(d)を簡易的にした2次元マイクロベンチマークコードの2種類を整備した。2023年度では、下記の更新を行った。

(a) では、大気モデルNICAMとPyTorchを用いた機械学習プログラムを連成実行する。1) Wisteria-Odyssey内でMPIを用いて連成する場合、2) Wisteria-Odyssey内で異機種間通信ライブラリh3-Open-SYS/WaitIO(以下WaitIO)を用いて連成する場合、3) Wisteria-OdysseyとWisteria-Aquarius間でWaitIOを用いて連成する場合、の3通り実行可能となっている。今年度は、WaitIOについて、新たな通信手法としてVerbs(InfiniBand, RoCE)とTofu(uTofu)に対応した。今後、両者を融合するライブラリを開発し、そのライブラリを用いたベンチマークコードの動作検証を行う予定である。

(d) については、データ同化の性能を測る目的でCityLBMを簡略化した2次元のマイクロベンチマークコードを整備した(<https://github.com/hasegawa-yuta-jaea/LBM2D-LETKF>)。2次元格子ボルツマン法 (LBM) と局所アンサンブル変換カルマンフィルター (LETKF) を用いた乱流のアンサンブルデータ同化を行う。今年度は、C++並列アルゴリズム (stdpar) 、thrust、senders/receiversを用いて実装し、CPU、NVIDIA GPUにおける性能可搬性について調査した。加えて、SYCLによる実装もCPUについては完了した。

4.10 科学技術計算・機械学習アルゴリズム調査研究サブグループ

4.10.1 調査研究の成果

科学技術計算アルゴリズムの調査研究として、次世代計算基盤の将来アーキテクチャ候補を想定した科学技術計算アルゴリズムの調査に基づき性能解析を行った。その上で、次世代計算基盤における科学技術計算アルゴリズムの改善の可能性について検討を行った。

アプリケーション調査研究グループの純アプリサブグループ（生命科学分野、新物質・エネルギー分野、気象・気候分野、地震・津波防災分野、ものづくりアプリ分野、基礎科学分野、社会科学分野、デジタルツイン・Society 5.0分野）とのヒアリングが2023年6月～7月に行われた。その結果、線形代数演算ライブラリのBasic Linear Algebra Subprograms (BLAS) またはLinear Algebra PACKage (LAPACK) を使っているとの回答があったのは気象・気候分野のGSM/asuca、デジタルツイン・Society 5.0分野のNICAM-AIおよび都市気流計算、生命科学分野のGENESIS、新物質・エネルギー分野のmVMC、HΦおよびSALMONであった。また、高速フーリエ変換（Fast Fourier Transform、以下FFT）を使っているとの回答があったのは気象・気候分野のGSM/asuca、生命科学分野のGENESISおよびテンプレートマッチング、新物質・エネルギー分野のSALMONであった。

線形代数演算において、行列サイズを $n \times n$ とすると密行列積の計算量は $O(n^3)$ であることから演算密度が高く、アクセラレータに適していることが知られている。純アプリサブグループの多くのアプリケーションにおいて密行列積が使われており、密行列が使われている計算に対しては加速が期待できる。一方で、3次元FFTの問題サイズを n^3 とすると計算量は $O(n \log n)$ であり、密行列積に比べると演算密度が低いことから計算の加速はそれほど期待できないことが予想される。また、3次元FFTにおいて問題サイズを n^3 とすると、 n^2 組の n 点FFTを並列に実行することができるため、並列度は n^2 となる。これまでのプロセッサやアクセラレータでは演算コア数よりも n^2 の並列度が大きくなる場合が多かったが、次世代計算基盤の将来アーキテクチャ候補では演算コア数をさらに増やすことで性能向上を図ることが予想されるため、問題サイズが小さい場合にはすべての演算コアを有効に活用することができないという問題がある。

そこで、3次元FFTにおいて二重ループを一重化することで並列度を n^2 から $n^3/2$ に増やす手法を検討したところ、問題サイズが小さい場合において性能を改善できることが分かった。

4.10.2 機械学習アルゴリズム調査研究の成果

機械学習分野は急速に発展しており、ニューラルネットアーキテクチャやその中で用いられるアルゴリズムも日々変化している。特に最近ではCNNやRNNからtransformerにアーキテクチャの流れが遷移しており、これからも数年単位で変化していく可能性がある。このような状況の中で、MLPerfは幅広い機械学習分野のアプリケーションから主要なものを継続的に抽出しており、ベンチマークとして広く用いられている。計画当初は、ベンチマークの選定に相当な時間をかける予定であったが、MLPerfの中から代表的なものを抽出する方向に昨年度の時点で方針を転換した。今年度はChatGPTの登場以降、大規模言語モデルの重要性が高まりMLPerfにもGPT3のベンチマークが採用された。そこで、今年度はGPT3の事前学習における性能を「富岳」とABCIで評価を行った。

「富岳」上で大規模言語モデルを学習する際にはMegatron-DeepSpeedをA64FXに移植する作業を行った。A64FXにはAI用のFP16の演算器が搭載されているが、NVIDIAのTensor Coreなどとは異なり、足し込みがFP16で行われるため誤差があまりにも大きく学習が発散することが確認された。このため、「富岳」での学習はFP32で行わざるを得なかった。また、A64FXのFP32の理論ピーク性能は6.8TFLOP/sであるのに対して、PyTorchから呼び出す行列積の実測値は最大でも2.5TFLOP/s程度であった。さらに、多くの行列積は縦長の行列に対して行われ、行列の次元も大きくないことから中には0.4TFLOP/s程度の性能になる行列積も多数存在する。本調査研究では、行列積を実行する際の並列形状などのパラメータを調整し、スレッド並列化の工夫を行うことで6倍程度の高速化を実現した。大規模言語モデルで用いられるTransformerアーキテクチャは99%の演算が行列積によ

て占められるため、このような行列積の高速化はベンチマークを正しく行う上で重要である。

「富岳」の強みはTofu-Dネットワークの持つネットワーク性能であるが、これを活かすためのTrinaryアルゴリズムはグローバルコミュニケータでしか動作しない仕様になっている。大規模言語モデルを学習するためには、データ並列、パイプライン並列、テンソル並列を併用した3-D Parallelismを駆使する必要があるが、このうちデータ並列とテンソル並列はサブコミュニケータ上でAllReduceの集団通信を行う。このサブコミュニケータ上でのAllReduceにTofuネットワークを最大限に活かすためのTrinaryアルゴリズムを独自実装した。3軸に対して双方向BCastと双方向Reduceを実装し、Rank mapを用いてノードの配置を最適化した。また、uTofuを用いたRDMA実装も行い、高速化前と比べて約3倍の高速化を達成した。

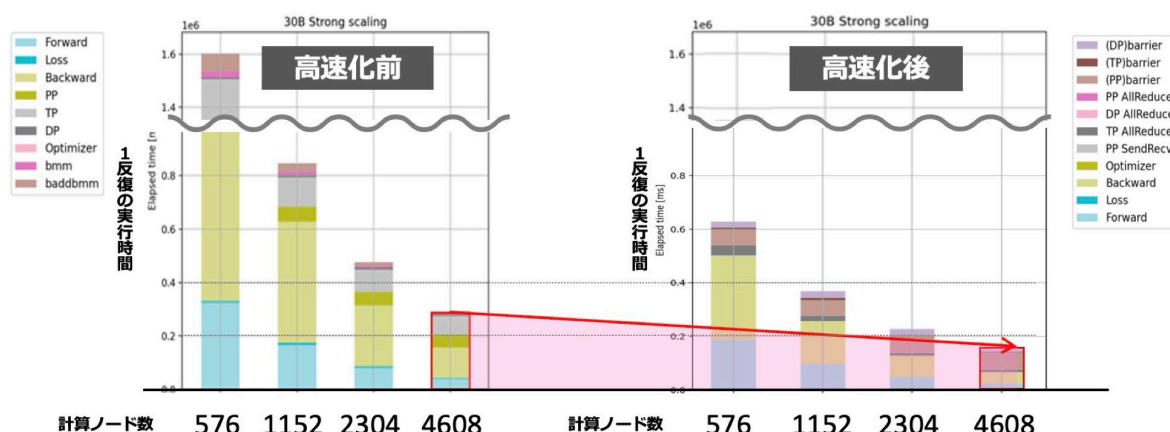


図 4.10.2.1 「富岳」における大規模言語モデルの高速化前後の実行時間

このように、演算と通信が最適化されたMegatron-DeepSpeedを用いて30BパラメータのGPTモデルの性能を評価した。「富岳」における大規模言語モデルの高速化前後の1反復にかかる実行時間を図 4.10.2.1に示す。

4.11 ベンチマーク構築サブグループ

4.11.1 調査研究の成果

4.11.1.1 ベンチマークの収集

昨年度に収集済みのベンチマークコードの加え、さらに各サブグループからの提供を受け、以下のベンチマークコードをGitHubリポジトリ https://github.com/RIKEN-RCCS/FS_Benchmarks で公開した。

- CUBE
- ChemTS
- EbE-method
- FFVHC-ACE
- FrontFlow_blue
- GENESIS
- Gromacs
- ISPACK
- LLM
- LQCD-DWF-HMC
- LQCD-Measurement
- LQCD-Wilson-HMC
- NICAM
- NICAM-AI
- QWS
- README.md
- RMCSM
- SALMON
- SCALE-DG
- SCALE-LETKF
- SCALE-LT
- SCALE-SDM
- asura-fdps
- athena-benchmark
- mVMC
- qNET_kernel
- vlasov_refact

4.11.1.2 優先ベンチマークの整備

まず、各サブグループにヒアリングを行い、今年度のアーキテクチャ評価において優先すべきものとして、前節で収集したベンチマークコードの中から以下の8本を選定した。原則として、各アプリケーション分野から1本を選定し、それぞれカーネル版およびフルアプリ版を整備した (AI分野のみ、シングルノード用のHugging Face GPT-2 XLとマルチノード用のMegatron-LM DeepSpeedを選定)。

- GENESIS
- SALMON

- SCALE-LETKF
- EbE-method
- FrontFlow/blue
- LQCD-DWF-HMC
- Hugging Face GPT-2 XL
- Megatron-LM DeepSpeed

ここで、選定の基準としては、

- 各分野において、代表的・重要なアプリケーションであること。
- 全体として、代表的な計算種別を網羅していること。
- 富岳 NEXT が運用される見込みの 2030 年頃においても、重要であり続けると目されること。

を用いた。

選定した各ベンチマークのカーネル/シングルノード版およびフルアプリ/マルチノード版のそれぞれに対し、ベンチマーク指標 (e.g. 特定区間の実行時間) とベンチマーク規則 (e.g. コード改変の可否) を定めた。

2023年3月末における、各コードの整備状況を表 4.11.1.1に示す。

表 4.11.1.1 各コードの整備状況

名称	分野	種類	計算種別	言語	GPU対応
GENESIS	生命科学	カーネル/フルアプリ	MD (粒子)	Fortran	CUDA
SALMON	新物質・エネルギー	カーネル/フルアプリ	DFT, Stencil, FFT	Fortran	OpenACC
SCALE	気象・気候	カーネル/フルアプリ	CFD (構造格子)	Fortran	OpenACC
EbE-method	地震・防災	カーネル	FEM (非構造格子)	C++	OpenACC
FrontFlow/blue	ものづくり	カーネル/フルアプリ	FEM (非構造格子)	Fortran	OpenACC
LQCD-DWF-HMC	基礎科学	カーネル/フルアプリ	Stencil, 疎行列ベクトル積	C++	OpenACC
Hugging Face GPT-2 XL	AI	シングルノード	Transformer	PyTorch	n/a
Megatron-LM DeepSpeed	AI	マルチノード	Transformer	PyTorch	n/a

4.11.1.3 フルシステム性能推定手法の策定

前節で整備したベンチマークコードのうちGENESIS、FrontFlow/blue、LQCD-DWF-HMCおよびMegatron-LM DeepSpeedを用いて、アーキテクチャ調査研究グループは提案システムのフルシステム性能推定を行う。そのため、アーキテクチャ調査研究グループとも協議の上、次のような性能推定手法を定めた (以下に現れるxxx/yyyyは、リポジトリのソースツリーにおけるディレクトリ名またはファイル名を示す)。さらに、性能推定のためのベースラインとして、「富岳」における性能プロファイルおよびMPI統計情報を提供した。

- (1) For each of the applications of the three target sizes, estimate the benchmark index if it were executed on your proposed systems, based on the provided performance profile and MPI statistics information.
 - The benchmark index (i.e. the elapsed time of a certain region) of an application is defined in XXX/README.md.
 - For the target sizes, called x1, x4, and x10, and the performance profile and MPI statistics information, as the "baseline" performance measured on Fugaku, refer to XXX/Full_System_Estimation/README.md of each application.
 - The target sizes are stored in XXX/Full_System_Estimation/inputs.
 - The performance profile information is stored in XXX/Full_System_Estimation/profiles/cpu_pa.
 - The MPI statistics information is stored in XXX/Full_System_Estimation/profiles/mpistat.
- (2) Put the benchmark indices in *Performance Estimation Report*.
- (3) Rationalize the indices by explaining the method used to estimate them in *Estimation Method Report*.

ただし、GENESISおよびLQCD-HMC-DWFに関しては、その特徴を考慮して、以下のように別途に性能推定手法を定めた。

● GENESIS の性能推定手法

GENESISでは、ベンチマークとしてはMDのメインループ部分 (6,000ステップ, 21ps分) の実行時間を指定した。ベンチマークセットの入力ファイルはGitHubレポジトリ内 (Full_System_Estimation/inputs) に置いている。ここでは実行時間は演算時間と通信時間の和として考える。ターゲットシステムに関しては、原子数の大きさにより基準値 (= 1 倍)、4 倍、8 倍として設定した。具体的には108,739,584原子、434,958,336原子、869,916,672原子である。これらの原子は35,397原子 (solと呼ぶ) からなるシステムを3,072倍、12,288倍、24,576倍したものである。

基準として提示したCPU性能値は、sol1個分を富岳2ノード計算した値と、sol2個分を富岳2ノードと4ノードで計算した値である。

MPI性能値は、24個のsolと48個のsolを富岳48ノード、96ノードで計算した値である。これらの性能値は1ノード当たり4MPI、12スレッドで計算したものである。

演算時間の見積もり方法に関して指定していないが、「非結合性相互作用」のみを考慮に入れる事は明確に禁止し、結合性相互作用、ペアリスト生成部分、数値積分部分も考慮するように明記した。

通信時間に関しては、長距離相互作用計算で利用される3D FFTの(1D alltoall)の計算だけを考慮するように明記し、更にMPIプロセス数から与えられる計算式を与えた、単純にプロセス辺りのデータ量をBandwidthで割ったものとなっている (レイテンシ等のスタートアップに係る時間は計算式には含まない)。

計算式の詳細はFull_System_Estimation/README.mdに記載している。

● LQCD-DWF-HMC の性能推定手法

LQCD-DWF-HMCでは、4次元時空の格子サイズが $128^3 \times 256$ のクォークの伝搬関数を計算する区間の実行時間をベンチマークの指標として指定し、これを「Problem size x1-s」と呼ぶことにした。他のベンチマークとは異なり、x4やx10といったターゲットサイズは指定していない。その代わり、フルシステムを使って、同じ問題サイズで、独立して走行可能な1000個のジョブによるクォーク伝搬関数の計算に必要な

時間をフルシステムのキャパシティを評価するための指標として指定し、これを「Problem size x1-w」と呼ぶことにした。「Problem size x1-w」においてはベンダが自由にジョブ当たりのノード数を調整してもよいとした。例えば、フルシステムのノード数がN、ジョブ当たりのノード数がM、ノード数Mでのクォーク伝搬関数の計算に必要な時間をTとした時、「Problem size x1-w」は、 $1000 * T / (N/M)$ となる。

LQCD-DWF-HMC/Full_System_Estimation2 では、問題設定の詳細、様々な問題サイズで富岳において採取されたプロファイル情報に加え、プロファイル情報から抽出された計算時間、演算数、袖領域交換通信の回数や通信量、MPI_Allreduceの回数などを表にしてわかりやすく掲載している。また、富岳全系を利用したときの「Problem size x1-s」「Problem size x1-w」の推定の実演や、提案システムでの推定の助けとなる演算数や、通信量、通信回数の理論式が掲載されている。

4.12 性能モデリング調査研究サブグループ

4.12.1 調査研究の成果

すべてのモデリングとシミュレーション手法がすべてのアプリケーション分野に簡単に適用できるわけではないため、目標はアプリケーション分野に適したモデリングアプローチを決定することである。したがって、性能モデリング研究開発は、ターゲットとするアプリケーション分野、ベンチマーク、および提案されているアーキテクチャと整合させ、実現可能でありながら正確なスーパーコンピュータ設計の予測を行う必要がある。目標は、確立された指標と技術予測に基づき、候補設計の「what-if」分析を迅速に実行する方法と、アプリケーションが将来的に高速化を得るために提案されたハードウェア機能を利用できるかどうかのベースラインアプローチを確立することである。

4.12.2 将来の HPC アーキテクチャにおけるデザイン空間探索のための性能投影アプローチ

ハードウェア、ソフトウェアスタック、およびアプリケーションの変更が性能に与える影響を評価する可能性は、コード設計環境において最も重要である。ここでは、マルチコアノードおよびマルチスレッドアプリケーションの最初の設計空間探索フェーズを可能にすることを目的とした性能予測ワークフローを提案する。この目的のために、ソースアクセス可能なマシン上での解析により、仮想的なターゲットマシンの持続可能な最大性能を特徴付ける。最後に、これらの特徴を利用して、ソースアクセス可能なマシンからターゲットマシンまでの性能の深化を予測する。

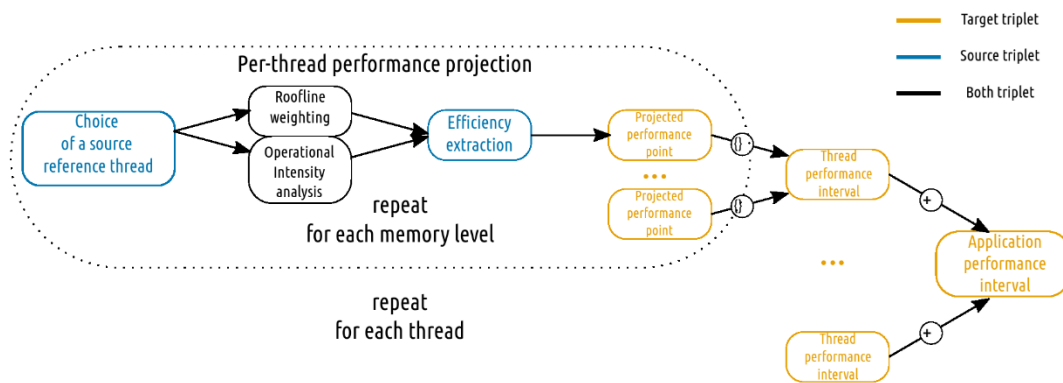


図 4.12.2.1 スレッドごとのパフォーマンス予測

NAS並列ベンチマークとccs-qcdにおけるgem5のサイクルレベルシミュレーションとMachine Code Analyzerに基づく高速探索手法の精度とオーバーヘッドを比較することにより、我々の手法の長所と短所を示す。最後に、QWSとgenesisにおいて、gem5よりも大規模で代表的なアプリケーションを研究するための我々のアプローチの能力を示す。我々のアプローチは、MCAの高速な探索とgem5の精度の中間に位置する新しい手法であるため、通常gem5ではシミュレーションできないような、より広範なアプリケーションや入力/ワークロードに適用できる。

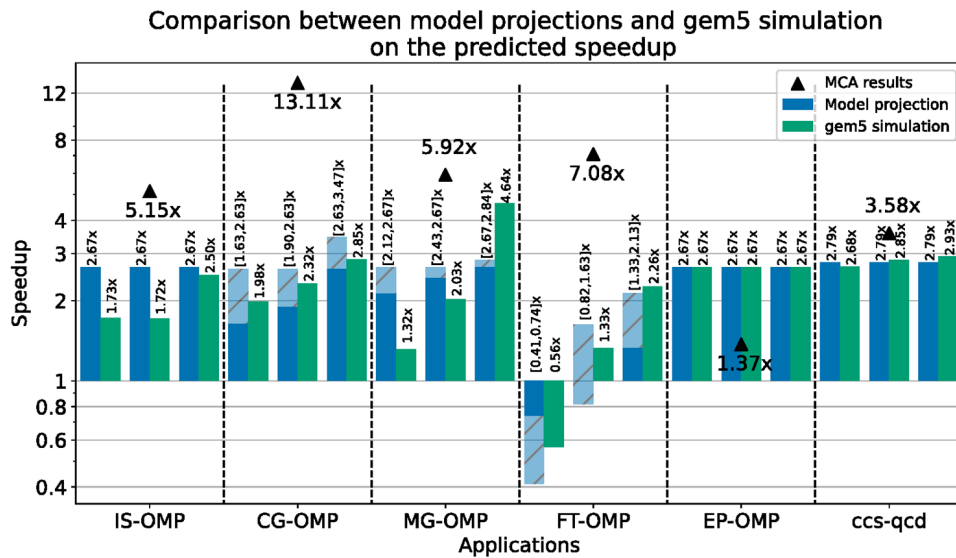


図 4.12.2.2 高速化予測に関するモデル予測と gem5 シミュレーションの比較

4.12.3 精度操作と近似計算のための解析とツールチェーン

様々な精度を実装するための既存のアプローチは、HPC開発者にとって課題となる特定のアプローチに限られていた。BITFLEXは、HPCアプリケーションにおける浮動小数点（FP）の動作をフロントからバックまで完全に把握できるため、開発者の作業を軽減することを目的としている。コンパイル時に、MPFRを採用してユーザが選択したさまざまな精度を可能にし、Serialbox 2を採用して任意のFP演算とオペランド値のマルチ出力をサポートするようにカーネルを装飾する。その結果、関心のあるデータ構造をダンプして比較し、切り捨てや丸めなどの「quiet」エラーを発見して追跡できるバイナリが完成した。今後数カ月の間にこのアプローチの能力を向上させる予定である。このプロジェクトに関連して、我々は現在、R-CCSの複数のアプリケーションチームやCSチームと協力して、彼らのコードに縮小精度を使用することの実現可能性を検討している。しかし、残念なことに、シミュレーションや数値計算に誤差が生じることもある。そこで、数値誤差を回避するための高精度エミュレーションスキーム（Ozaki-methodなど）をアプリケーションチームと積極的に検討している。

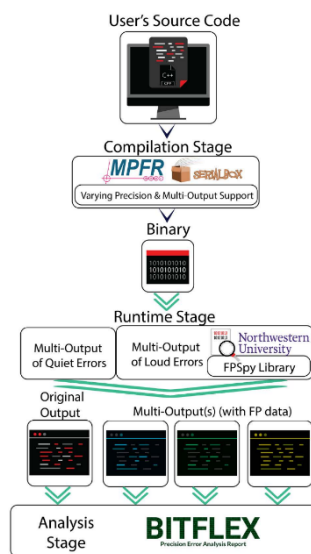


図 4.12.3.1 精度操作と近似計算のための解析とツールチェーン

4.12.4 SimEngとSST gem5の代替品

サイクルレベルシミュレータgem5を使用した結果、開発期間が非常に長くなり、現実的なワークロードに対して不当に長いシミュレーションが行われた経験から、代替案を検討する必要があると感じている。我々のMCAベースのアプローチは、非常に高速な一次近似を得るために必要なツールの一つに過ぎない。さらに、我々はStructural Simulator Toolkit (SST) との共同研究を続けており、特にネットワーク・シミュレーションに力を入れている（ただし、現時点では公表可能な結果は得られていない）。さらに、SimEngの開発者とも積極的に協力し、主にHPC用のロングベクトルやArm SME（行列ユニット）の研究を行っている。SimEngは英国の大学発の新しいシミュレータで、gem5よりも精度が高く、大幅なスピードアップが期待できる。SimEngはまだ完成には程遠いが、英国との協力により、将来のパフォーマンス・モデリングに役立つツールがまた一つ増えることを期待している。

4.13 気象庁モデル評価・解析サブグループ

4.13.1 調査研究の成果

気象・気候モデルの中でも特に気象庁モデルのベンチマーキング評価とその解析を行い、性能情報等の知見を基に次世代計算基盤に求められる性能・機能について調査を行うとともに、次世代計算基盤を対象にしたグランドチャレンジ級の成果が期待できる気象・気候モデルを選定し、そのベンチマーク候補を検討することを目的として、本サブグループの調査を行った。

具体的には、気象庁気象研究所から、気象庁モデルおよび関連する情報を提供していただくとともに、気象庁で実施した各種最適化の評価について議論・検討を行った。気象庁で運用中の数値予報モデルは、集中豪雨や突風など局地的な激しい現象から、暖・寒候期予報といった長期の現象まで、時間・空間スケールが多岐にわたる現象を対象としている。これらの予測対象に合わせて、モデル領域、解像度、予報期時間等多岐にわたっており、対象に適した解析手法が選択されている。それらの中から、大規模計算が必要な問題規模にスケールアップできることや、気象庁側の将来的な取組も考慮したベンチマークとして、GSM、asuca、asuca-Varの三つのモデルを提供して頂いた。気象庁でも「富岳」の商用機であるFX1000が導入され、アプリケーションの移植や最適化が行われており、これらの取組状況と評価結果について情報提供を受けた。また、将来の計算機システムを考慮したGPUなどのアクセラレータへの対応に向けた検討状況も情報提供を受けた。

理化学研究所では、提供された複数のベンチマークを「富岳」上で実施するとともに、詳細プロファイラを用いた評価を行い、そのベンチマークの特性や、ボトルネックとなっている箇所などの評価を行った。特に、ボトルネックの可能性の高い箇所については、コンパイル結果のメッセージを確認し、SIMD化やスレッド並列化、ソフトウェアパイプラインの適用などの最適化の詳細について、検討を行った。その結果、ハードウェアバリアによるスレッド間同期の高速化が、スレッド実行にばらつきがある場合などに効果があることが明らかとなった。また、今後のアーキテクチャ候補について、気象庁気象研究所へフィードバックを行った。

今後は、IO処理の影響などについて、ネットワークポロジとの関連を含めて調査していきたいと考えており、両者の評価結果をもとに、さらなる最適化を進める予定である。

5. まとめと今後の課題

2023年度も、次世代計算基盤に係る調査研究（システム調査研究）理研チームの調査研究活動としてアーキテクチャ調査研究グループ、システムソフトウェア・ライブラリ調査研究グループ、アプリケーション調査研究グループのそれぞれで、様々な項目での技術動向や利用動向調査、次世代のアプリケーション検討を行った。特に、次世代計算基盤に向けたアーキテクチャの絞り込みを行い、候補アーキテクチャを提案した。また、システムソフトウェア開発戦略検討のための各種調査や、重要開発項目策定に向けた各システムソフトウェアにおける類似ソフトウェアに関する調査研究も実施したほか、重要システムソフトウェアの富岳へのポーティングも引き続き行い、一部評価を実施した。さらに、アーキテクチャの絞り込みに向けたベンチマークセットの更新やシステム性能を見積るためのベンチマーク作成なども重点的に実施し、また次世代計算基盤の実現により可能となるターゲットサイエンスに関してもより深く検討・調査を行った。

次年度は、絞り込まれたアーキテクチャ候補に基づき、引き続きベンチマーク評価による性能推定を行いつつ、さらに詳細なアーキテクチャパラメータの検討を行うほか、候補アーキテクチャに向けたシステムソフトウェアの開発ロードマップ作成を実施する。さらに、ベンチマークセット更新とともに、候補アーキテクチャ向けのアルゴリズム検討やアプリケーションの最適化に関する調査研究を実施する予定である。

以 上

学 会 等 発 表 実 績

委託業務題目 「次世代計算基盤に係る調査研究（システム調査研究）」

機関名 国立研究開発法人 理化学研究所

1. 学会等における口頭・ポスター発表

発表した成果（発表題目、口頭・ポスター発表の別）	発表者氏名	発表した場所（学会等名）	発表した時期	国内・外の別
次世代計算基盤に向けたアクセラレータの調査と初期評価(口頭発表)	小田嶋 哲哉、伊藤 真紀子、依田 勝洋、中村 洋介、Nguyen Anh、吉川 隆英	第 191 回 HPC 研究発表会	2023/9/26	国内
次世代計算基盤に向けたノードアーキテクチャの初期検討結果と課題分析(口頭発表)	安島 雄一郎、本車田 強、秋谷 定則、小田嶋 哲哉、吉川 隆英、水谷 康志	第 191 回 HPC 研究発表会	2023/9/26	国内
Fujitsu's AI & Computing Technology Vision(ブース展示)	Fujitsu Limited	2023 International Conference for High Performance Computing, Networking, Storage and Analysis (SC23)	2023/11/14～16	国外
アクセラレータのデザイン・スペース探索に向けた RTL Architect 適用事例(オンライン講演)	中村 洋介	SNUG Japan 2023 Online	2023/9/20	国内
ISC High Performance 2023 に現地参加しました(ブログ記事公開)	小田嶋 哲哉	Fujitsu TECH BLOG、 https://blog.fltech.tech.dev/entry/2023/07/07/isc2023-ja	2023/7/7	国内
Design Automation Conference に参加しました(ブログ記事公開)	萩原 汐	Fujitsu TECH BLOG、 https://blog.fltech.tech.dev/entry/2023/11/10/Dac60-ja	2023/11/10	国内

"Compiler and System Software Feasibility Study and Benchmark for Next Generation Supercomputer" 口頭発表	Miwako Tsuji	ADAC13 SYMPOSIUM	2023/9/26	国内
"次期フラグシップマシンのストレージシステム調査研究の中間報告" 口頭発表	建部 修見	秋葉原（Gfarm シンポジウム 2023）	2023/9/8	国内
"利用者が望む HPC ストレージ" 口頭発表	建部 修見	宮古島（Gfarm シンポジウム 2024）	2024/1/29	国内
"次世代 HPC ストレージへの提案" パネル討論	建部 修見（モデレータ）	宮古島（Gfarm シンポジウム 2024）	2024/1/30	国内
Challenges of Scaling Deep Learning on HPC Systems	Mohamed Wahib	SIAM Conference on Parallel Processing (SIAM PP'24)	2024/3/08	国外
"量子関連技術への自動チューニングの新展開" 口頭発表	片桐孝洋	東京都新宿区（ATTA2023）	2023/12/22	国内
"Auto-Tuning for Quantum Computing Related Technology on Supercomputers" 口頭発表	Takahiro Katagiri	米国、メリーランド州ボルチモア（SIAM PP24）	2024/3/7	国外
"Adaptation of Auto-tuning for Quantum Annealer" 口頭発表	Takahiro Katagiri, Makoto Morishita	台湾、新竹市（ATAT2024）	2024/3/22	国外
"System Software Research – The FugakuNext Supercomputer and Synchrotron Radiation Facilities" 口頭発表	Kento Sato	WSSP36 (36th Workshop on Sustained Simulation Perfor	2023/12/11	国内
"スーパーコンピュータがもたらす AI の革命" 口頭発表	佐藤 賢斗	2023 年度 スパコンセミナーのぞいてみようスパコンの世界 - 「富岳」を見て、聞いて、知って -	2023/10/14	国内

"複数の特性予測モデルの信頼性を考慮した分子の多目的最適化" 口頭発表	吉澤竜哉、石田祥一、佐藤朋広、大田雅照、本間光貴、寺山慧	第 51 回構造活性相関シンポジウム	2023/11/21	国内
"Multiple ligands docking を用いた STING を標的とした新規ヒット化合物の探索" ポスター	戸板太陽、石田祥一、浴本亨、池口満徳、出水庸介、辻巖一郎、寺山慧	第 51 回構造活性相関シンポジウム	2023/11/21	国内
"Multi-objective Molecular Structure Generation Using Dynamic Applicability Domains Adaptation" 口頭発表	Tatsuya Yoshizawa, Shoich Ishida, Tomohiro Sato, Masateru Ohta, Teruki Honma, Kei Terayama	情報計算科学生物学会 2023 年大会(CBI2023)	2023/10/23-26	国内
"RNA とタンパク質を用いた液-液相分離実験に関するデータベース構築と予測モデル開発" ポスター	陳佳盈、石田祥一、寺山慧	第 23 回蛋白質科学会年会	2023/7/5-7	国内
"Massively Parallel Computing Approaches to Entangled Many-Body Electrons—Applications to Cuprates and Frustrated Quantum Magnets" 口頭発表	Youhei Yamaji	2024 International Workshop on Massively Parallel Programming in Quantum Chemistry and Physics (MPQCP 2024)	2024/1/9	国外
"Parallelized remapping algorithms for km-scale global weather and climate simulations with icosahedral grid system" 口頭発表	C. Kodama, H. Yashiro, T. Arakawa, D. Takasuka, S. Matsugishi, and H. Tomita	International Conference on High Performance Computing in Asia-Pacific Region (HPCAsia 2024)	2024/1/26	国内

"Global cloud-resolving simulations with NICAM on the supercomputer Fugaku" ポスター発表	C. Kodama, H. Yashiro, T. Suematsu, D. Takasuka, Y. Takano, R. Masunaga, Y. Yamada, S. Matsugishi, T. Arakawa, M. Nakano, H. Miura, T. Seiki, A. T. Noda, M. Satoh, T. Ohno, T. Nasuno, T. Miyakawa,	AOGS2023	2023/8/1	国外
"Initiatives in the field of weather and climate for the HPC environment around 2030" ポスター発表	C. Kodama, S. A. Adachi, M. Nakano, T. Seiki, H. Yashiro, R. Yoshida	JpGU Meeting 2023	2023/5/22	国内
"Validation of a Global Nonhydrostatic Atmospheric Dynamical Core using Discontinuous Galerkin method" ポスター発表	Y. Kawai, H. Tomita	AMS 104th Annual Meeting	2024/1	国外
"不連続ガラーキン法を用いた非静力学大気力学コアの開発: 地形の考慮" 口頭発表	河合佑太、富田浩文	JpGU Meeting 2023	2023/5	国内
"不連続ガラーキン法に基づく流体計算コードの富岳における計算効率について" 口頭発表	河合佑太、任軒正博、片桐孝洋、富田浩文	富岳 NEXT FS 気象気候研究会	2024/1	国内

"Compression of Scientific Simulation Data by Stochastic Basis Expansion – Example on Multiple Computer Systems" 口頭発表	Kohei Fujita, Tsuyoshi Ichimura, Lalith Maddegadara, Muneo Hori	PDSW23: 8th International Parallel Data Systems Workshop (held in conjunction with SC23)	2023/11/12	国外
“Sequential Monte Carlo sampling を用いた断層位置とすべり分布の同時ベイズ推定手法の拡張” 口頭発表	中尾魁, 市村強, 藤田航平, 小林知勝, 宗包浩志, 堀高峰	第 43 回地震工学研究発表会	2023/9/8	国内
"Mixed Precision Solvers for Lattice QCD on Supercomputer Fugaku" ポスター発表	Issaku Kanamori, Tatsumi Aoyama, Kazuyuki Kanaya, Hideo Matsufuru, Yusuke Namekawa, Hidekatsu Nemura and Keigo Nitadori	the 6 th R-CCS International Symposium	2024/1/29	国内
"Multigrid Solver for Lattice QCD on GPU Machines" 口頭発表	Tatsumi Aoyama, Issaku Kanamori, Kazuyuki Kanaya, Hideo Matsufuru, Yusuke Namekawa, Hidekatsu Nemura and Keigo Nitadori	the 34 th IUPAP Conference on Computational Physics (CCP2023)	2023/8/8	国内

"Multigrid solvers in general purpose lattice QCD code set Bridge++" 口頭発表	青山龍美、金森逸作、金谷知至、松古栄夫、滑川裕介、根村英克、似鳥啓吾	日本物理学会第 78 回年次大会	2023/09/16	国内
"Development of a Heterogeneous Coupling Library h3-Open-UTIL/MP" 口頭発表	Takashi Arakawa, Shinji Sumimoto, Hisashi Yashiro, Kengo Nakajima	Asia Oceania Geoscience Society	2023/8/2	国外
"h3-open-UTIL/MP, a coupling library for heterogeneous computing" 口頭発表	Takashi Arakawa, Shinji Sumimoto, Hisashi Yashiro, Kengo Nakajima	International Congress on Industrial and Applied Mathematics	2023/8/22	国外
"異機種間連携ライブラリの性能評価" 口頭発表	荒川隆、住元真司、八代尚、中島研吾	FIT2023、大阪公立大学	2023/9/8	国内
"oneAPI を用いた様々なデバイス上でのステンシル計算の実装" 口頭発表	佐久間 大我、下川辺 隆史、大森 拓郎	第 28 回計算工学講演会	2023/6/2	国内
Top500 上位スパコンを用いた大規模深層学習の状況について	横田理央	CloudWeek2023	2023/08/30	国内
「富岳」政策対応枠における大規模言語モデル分散並列学習手法の開発と AI for Science に向けた国際連携	横田理央	第 6 回 HPCI コンソーシアムシンポジウム	2023/10/26	国内
国内外の大規模言語モデルに関する取り組みについて	横田理央	PCCC23「HPC 基盤技術と生成 AI」	2023/12/08	国内
Overview of Japanese Efforts to Train LLMs	Rio Yokota	Multilingual Workshop at NeurIPS 2023	2023/12/15	国外

大規模言語モデルにおける分散並列学習	横田理央	スパコンコロキウム（第 10 回）	2024/01/30	国内
Training a Large Language Model on ABCI	横田理央	Centra7	2024/2/08	国内
生成 AI を支える HPC 技術	横田理央	第 3 回「富岳」成果創出加速プログラム研究交流会	2024/3/12	国内
国内における大規模言語モデルの開発状況について	横田理央	スーパーコンピューティング・ジャパン	2024/3/12	国内

2. 学会誌・雑誌等における論文掲載

掲載した論文(発表題目)	発表者氏名	発表した場所 (学会誌・雑誌等名)	発表した時期	国内・ 外の別
次世代計算基盤に向けた アクセラレータの調査と初期 評価	小田嶋 哲哉、伊藤 真紀子、依田 勝 洋、中村 洋介、 Nguyen Anh、吉 川 隆英	情報処理学会研究報告 ハイパフォーマンス・コンピュ ーティング(HPC)	2023/9/19	国内
次世代計算基盤に向けた ノードアーキテクチャの初期 検討結果と課題分析	安島 雄一郎、本車 田 強、秋谷 定則、 小田嶋 哲哉、吉川 隆英、水谷 康志	情報処理学会研究報告 ハイパフォーマンス・コンピュ ーティング(HPC)	2023/9/19	国内
At the Locus of Performance: Quantifying the Effects of Copious 3D- Stacked Cache on HPC Workloads	Jens Domke, Emil Vatai, Balazs Gerofi, Yuetsu Kodama, Mohamed Wahib, Artur Podobas, Sparsh Mittal, Miquel Pericàs, Lingqi Zhang, Peng Chen, Aleksandr Drozd, Satoshi Matsuoka	ACM Transactions on Architecture and Code Optimization	2023/12	国外
Myths and legends in high-performance computing	Satoshi Matsuoka, Jens Domke, Mohamed Wahib, Aleksandr Drozd, Torsten Hoefler	The International Journal of High Performance Computing Applications	2023/4	国外

Koopmans' Theorem-Compliant Long-Range Corrected (KTLC) Density Functional Mediated by Black-Box Optimization and Data-Driven Prediction for Organic Molecules	K. Terayama, Y. Osaki, T. Fujita, R. Tamura, M. Naito, K. Tsuda,* T. Matsui,* M. Sumita,*	Journal of Chemical Theory and Computation	2023/9	国外
Design of antimicrobial peptides containing non-proteinogenic amino acids using multi-objective Bayesian optimisation	Y. Murakami, S. Ishida, Y. Demizu, K. Terayama*	Digital Discovery	2024/8	国外
ChemTSv2: Functional molecular design using de novo molecule generator	S. Ishida, T. Aasawat, M. Sumita, M, Katouda, T. Yoshizawa, K. Yoshizoe, K. Tsuda, K. Terayama*	WIREs Computational Molecular Science	2023/7	国外
Parallelized remapping algorithms for km-scale global weather and climate simulations with icosahedral grid system	C. Kodama, H. Yashiro, T. Arakawa, D. Takasuka, S. Matsugishi, and H. Tomita	HPCAsia '24: Proceedings of the International Conference on High Performance Computing in Asia-Pacific Region	2024/1	国内

Earth Virtualization Engines (EVE)	B. Stevens, S. Adami, T. Ali, H. Anzt, Z. Aslan, S. Attinger, J. Baack, J. Baehr, P. Bauer, N. Bernier, B. Bishop, H. Bockelmann, S. Bony, V. Bouchet, G. Brasseur, D. N. Bresch, S. Breyer, G. Brunet, P. L. Buttigieg, J. Cao, C. Castet, Y. Cheng, A. D. Choudhury, D. Coen, S. Crewell, A. Dabholkar, Q. Dai, F. Doblas-Reyes, D. Durran, A. El Gaidi, C. Ewen, E. Exarchou, V. Eyring, F. Falkinhoff, D. Farrell, P. M. Forster, A. Frassoni, C. Frauen, O. Fuhrer, S. Gani, E. Gerber, D. Goldfarb, J. Grieger, N. Gruber, W. Hazeleger, R. Herken, C. Hewitt, T. Hoefler, H.-H. Hsu, D. Jacob, A. Jahn, C. Jakob, T. Jung, C. Kadow, I.-S. Kang, S. Kang, K. Kashinath, K. K.-von Konigslow, D. Klocke, U. Kloenne, M. Klower, C. Kodama, S. Kollet, T. Koling, J. Kontkanen, S. Kopp, M. Koran, M. Kulmala, H. Lappalainen, F. Latifi, B. Lawrence, J. Y. Lee, Q. Lejeun, C. Lessig, C. Li, T. Lippert, J. Luterbacher, P. Manninen, J. Marotzke, S. Matsouoka, C. Merchant, P. Messmer, G. Michel, K. Michielsen, T. Miyakawa, J. Muller, R. Munir, S. Narayanasetti, O. Ndiaye, C. Nobre, A. Oberg, R. Oki, T. Ozkan-Haller, T. Palmer, S. Posey, A. Prein, O. Primus, M. Pritchard, J. Pullen, D. Putrasahan, J. Quaas, K. Raghavan, V. Ramaswamy, M. Rapp, F. Rauser, M. Reichstein, A. Revi, S. Saluja, M. Satoh, V. Schemann, S. Schemm, C. S. Poberaj, T. Schulthess, C. Senior, J. Shukla, M. Singh, J. Slingo, A. Sobel, S. Solman, J. Spitzer, D. Stammer, P. Stier, T. Stocker, S. Strock, H. Su, P. Taalas, J. Taylor, S. Tegtmeier, G. Teutsch, A. Tompkins, U. Ulbrich, P.-L. Vidale, C.-M. Wu, H. Xu, N. Zaki, L. Zanna, T. Zhou, and F. Ziemer	Earth System Science Data	2024/2	国外
Numerical Accuracy Necessary for Large-Eddy Simulation of Planetary Boundary Layer Turbulence using Discontinuous Galerkin Method	Y. Kawai, H. Tomita	Monthly Weather Review, 151, 6	2023/6	国外

（注）発表者氏名は、連名による発表の場合には、筆頭者を先頭にして全員を記載すること。

以 上