

jh180023-NAH

中島研吾(東京大学情報基盤センター)

高性能・変動精度・高信頼性数値解析手法とその応用

Large-scale Simulations by Advanced Numerical Methods with High-Performance/Adaptive-Precision/High-Reliability



岩下武史<sup>1</sup>, 深谷猛<sup>1</sup>, 中島研吾(代表)<sup>2,7</sup>, 市村強(副代表)<sup>2</sup>, 奥田洋司<sup>2</sup>, 古村孝志<sup>2</sup>, 塙敏博<sup>2</sup>, 伊田明弘<sup>2</sup>, 近藤正章<sup>2</sup>, 坂本龍一<sup>2</sup>, 藤田航平<sup>2</sup>, 星野哲也<sup>2</sup>, 有間英志<sup>2</sup>, 森田直樹<sup>2</sup>, 横田理央(副代表)<sup>3</sup>, 荻田武史<sup>4</sup>, 尾崎克久<sup>5</sup>, 片桐孝洋<sup>6</sup>, 八代尚<sup>7</sup>, 河合直聡<sup>7</sup>, 井上弘士<sup>8</sup>, 大島聡史<sup>8</sup>, 荒川隆<sup>9</sup>, 堀越将司<sup>10</sup>, 成瀬彰<sup>11</sup>, G. Wellein<sup>12</sup>, A. Basermann<sup>13</sup>

1: 北大, 2: 東大, 3: 東工大, 4: 東大女, 5: 芝工大, 6: 名大, 7: 理研 R-CCS, 8: 九大, 9: RIST, 10: Intel, 11: NVIDIA, 12: FAU, Germany, 13: DLR, Germany



本研究の背景: Exascaleシステムにむけての技術的課題

高性能数値アルゴリズム実現には通信最適化が必須

- メモリ上のデータ移動(Serial通信), ノード間ネットワークを介したParallel通信
技術革新により, メモリ・ネットワークの大容量化, 広帯域化が進む一方で, 階層がより深化し, 局所化やレイテンシ対応が必要となる
Serial・Parallel通信最適化のための様々な手法が提案, 実装されている

もう一つの技術的課題は消費電力, エネルギー(消費電力)

- ハードウェアの技術革新と共に, 省電力・省エネルギー(省電力)のためのアルゴリズムの開発
実用化により, 実計算時の消費電力の抑制が期待される
Approximate Computing (S. Mittal, 2016)は低精度演算の積極的活用により計算時間短縮, 消費電力削減を図る試み
混合精度演算はその一種であり, 既に多くの研究事例があるが, Approximate Computingでは, 半精度から四倍精度まで演算精度を動的に変動させる変動精度(Transprecision)の研究が進められている
数値計算による近似解(数値解)は様々な計算誤差を含み, 計算結果の信頼性の観点から, 数値解の正しさを数学的に保証する必要がある, 低精度・変動精度使用時, 悪条件問題には重要である
昨今は, スパコンによる大規模計算向け精度保証の研究も実施されているが, 実問題で現れる大規模疎行列・H行列系への応用例はほとんどない

全球大気循環モデルにおける精度変更(64bit => 20bit)の影響

C/O Peter D. Düben (ECMWF) To Reduce Numerical Precision to Achieve Higher Accuracy in Weather and Climate Modelling, MS5 Approximate Computing Towards Exascale in SIAM PP18 (March 7-10, 2018, Tokyo)

Table with 4 columns: Resolution, Precision in number of bits, Normalised Energy Demand, Mean error Z500 at day 2. Rows show 235 km, 315 km, and 235 km with precision changes from 64 to 20 bits.

To save power a reduction in precision is much more efficient when compared to a reduction in resolution.

Studies with real hardware (such as FPGAs) confirm this result.

Düben et al. MWR 2015; Düben et al. DATE 2015; Düben et al. JAMES 2015; Russel, Düben et al. FCCM 2015.

本研究の概要と意義

本研究の概要

- 疎行列演算, H行列演算, ステンシル演算等の代表的数値アルゴリズムについて, 各参加者によって開発されたSerial・Parallel通信の最適化に着目した下記の高性能手法を各システムにおいて実装, 最適化し, 低精度演算・変動精度演算についての検討を実施する
SELL-C-σ疎行列格納法(SIMD演算に適した疎行列格納法)(Wellein等)
時空間タイリング(ステンシル演算の特性により, 通信削減・メモリアクセス最適化実施)
Pipeline法(MPI Allreduceの通信オーバーヘッドを隠蔽する手法)
Dynamic Loop Scheduling(疎行列ベクトル積における通信と計算のオーバーラップ)
hCGA法(並列多重格点法におけるオーバーヘッド削減する手法)
Adaptive CG(要素単位多重格点法前処理を適用した共役勾配法)
pK-Open-SOL(悪条件問題向け並列前処理ソルバー)
HACApK(世界初のオープンソースOpenMP/MPIハイブリッドによるH行列ライブラリ)
各アルゴリズムを各参加者によって開発された地震(Seism3D, GHYDRA), 大気科学(NICAM), 量子科学, 工学(FrontISTR)等の実シミュレーションに適用し, 従来手法と性能・精度を比較, 一部のアルゴリズム, シミュレーションについては計算時の消費電力測定を実施
疎行列演算やH行列演算を対象として, 悪条件問題における実用的な精度保証法を確立, 反復改良法(Iterative Refinement)を変動精度演算により実現し, 精度保証法と融合
所望の結果精度達成という条件下で, 計算時間や消費電力を最小化する最適な演算精度を自動チューニング技術によって動的に制御する手法を確立
本研究では, 単精度・倍精度の変動精度を前提とするが, 一部アルゴリズム, アプリケーションについては, 半精度, 四倍精度についても検討
本研究によって開発された高性能・変動精度・高信頼性数値解析法を東京情報基盤センターで開発中の自動チューニング機構を有するアプリケーション開発・実行環境ppOpen-HPC及びpK-Open-HPC(ppOpen-HPCの発展形)(http://ppopenhpc.cc.u-tokyo.ac.jp/)に実装, Oakforest-PACS(OPF), Reedbush(RBU, RBH, RBL), Tsubame-3等で公開, 将来的には他のセンターのスパコンへの導入も視野に

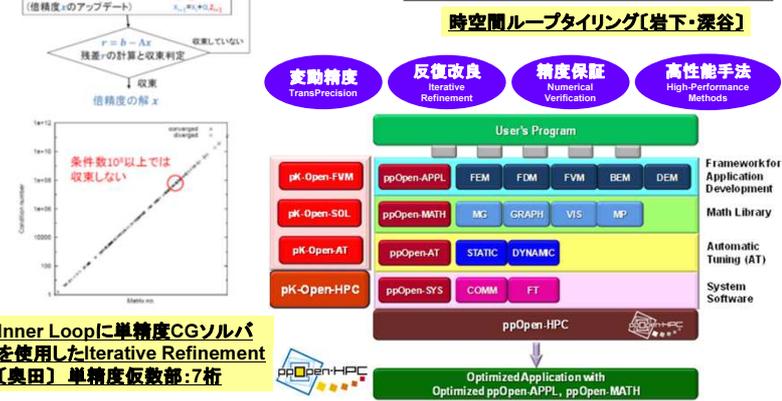
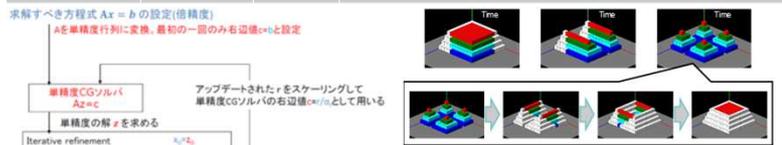
研究の意義

- 本研究は, 最先端のスパコン向け高性能数値アルゴリズムに対して, 半精度から倍精度, 倍々精度までの広範囲をカバーする変動精度演算を適用し, 精度保証, そのための自動チューニング手法を開発する試みとしては初めてのもの
開発された手法を様々なアプリケーションに適用することで, 低精度を中心とした変動精度演算の科学技術シミュレーションへの有効性を検証できる
開発したアルゴリズム, アプリケーションの消費電力の直接測定によって, 各計算の特性と低精度演算の有効性を消費電力の観点から検討可能となる

実施内容: 3年計画

- JHPCNの中でも消費電力当たり計算性能の高いシステム(Oakforest-PACS(OPF), Reedbush(U/H/L), Tsubame-3等)を使用
1年目: Intel Xeon Phi(KNL)・NVIDIA P100向け最適化, 実装及び精度評価
2年目: 精度保証手法の確立
3年目: 自動チューニング手法の確立, ppOpen-HPC等へ実装・公開

Table with 4 columns: 実施項目, 担当者, 目標精度, 概要. Rows include stencil calculations, sparse matrix operations, adaptive CG, and precision refinement methods.



Inner Loopに単精度CGソルバを使用したIterative Refinement 【奥田】単精度仮数部: 7桁