

エクサスケール時代の数値計算手法に対する性能予測技術

深谷 猛(代表・北大), 星 健夫(副代表・鳥取大), 山本有作(電通大), 福島 孝治(東大), 中西 義典(東大), 望月 出海(KEK), 高山 あかり(早大), 田中 和幸(鳥取大), 塩谷 明美(電通大)

概要

本研究課題では、応用分野の研究者のニーズに基づき、エクサスケール時代の超並列型アプリケーションの性能(実行時間)を事前予測する技術を開発することを目指す。プログラムの実行前に、性能情報を予測することで、効率的な実験計画の立案や適切な計算リソースの選択が可能となる。さらに、予測結果に基づき、計算プロセスを自動最適化することも可能となる。これにより、スパコン上での超並列計算の敷居を下げる。近年、データ駆動系アプリケーションの需要が増加しており、今後、スパコンにあまり馴染みのない実験系研究者が、実験データとシミュレーションデータを組み合わせたデータ解析を、スパコン上で手軽に実行できるようになることが期待される。

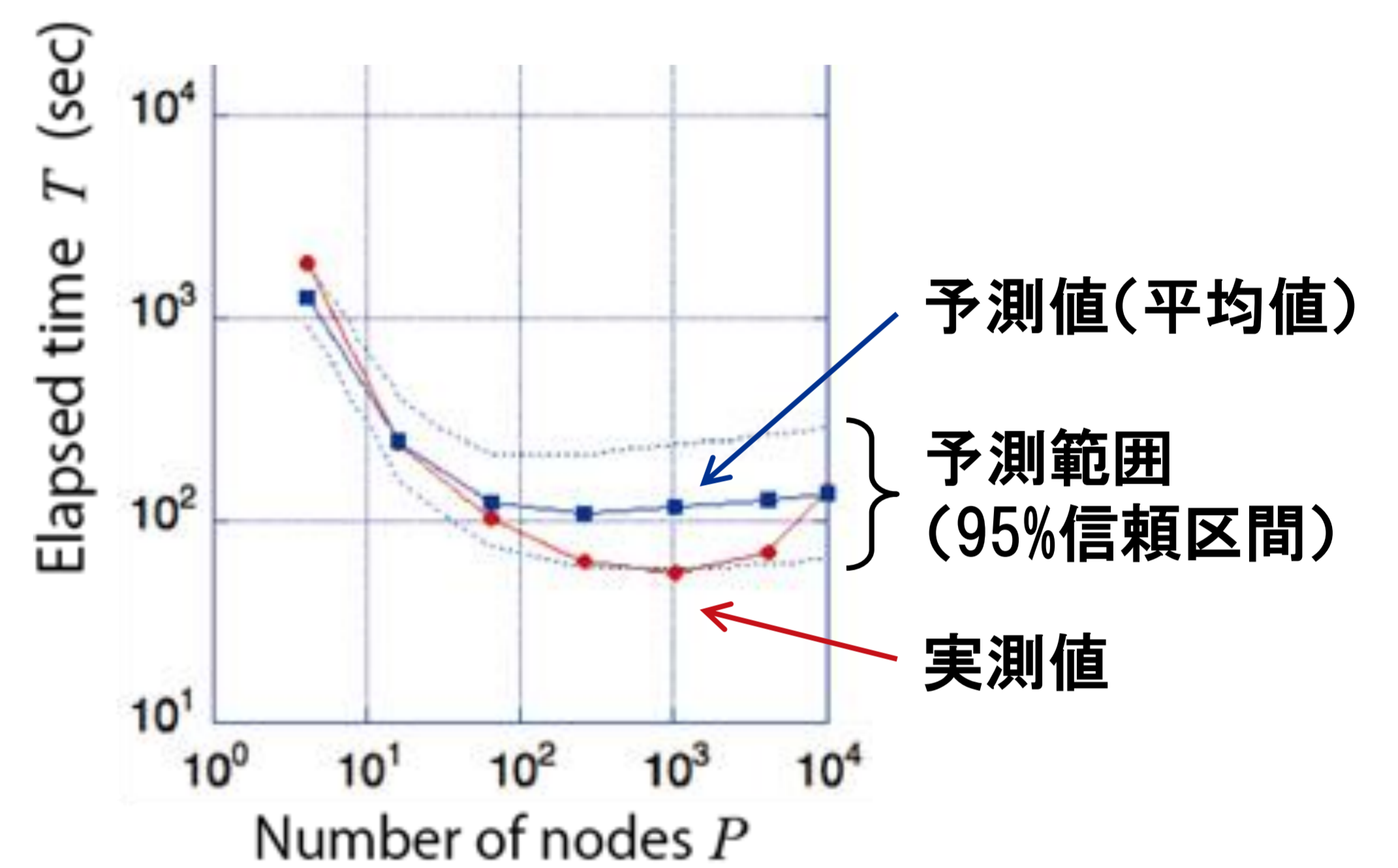
◆ 実施項目1: 実行時間を外挿する性能予測技術の開発

実行が容易なベンチマークの結果
(問題サイズ・並列数: 小~中規模)



実際に興味がある計算の性能予測値
(問題サイズ・並列数: 大)

- 小・中規模のベンチマークデータから、問題サイズや並列数(ノード数・コア数)を大きくした場合の実行時間を外挿する技術が必要。
- 一般的に外挿が困難であることを踏まえて、結果の信頼性を付与した予測技術を開発。
- プログラムの実行時間を確率変数としてモデル化し、ベンチマーク結果から、ベイズの定理を用いて、事後確率分布として、実行時間を予測。
- LAPACK・ScaLAPACK, 課題参加者が過去に開発した数値計算プログラムを対象として、研究開発を実施予定。

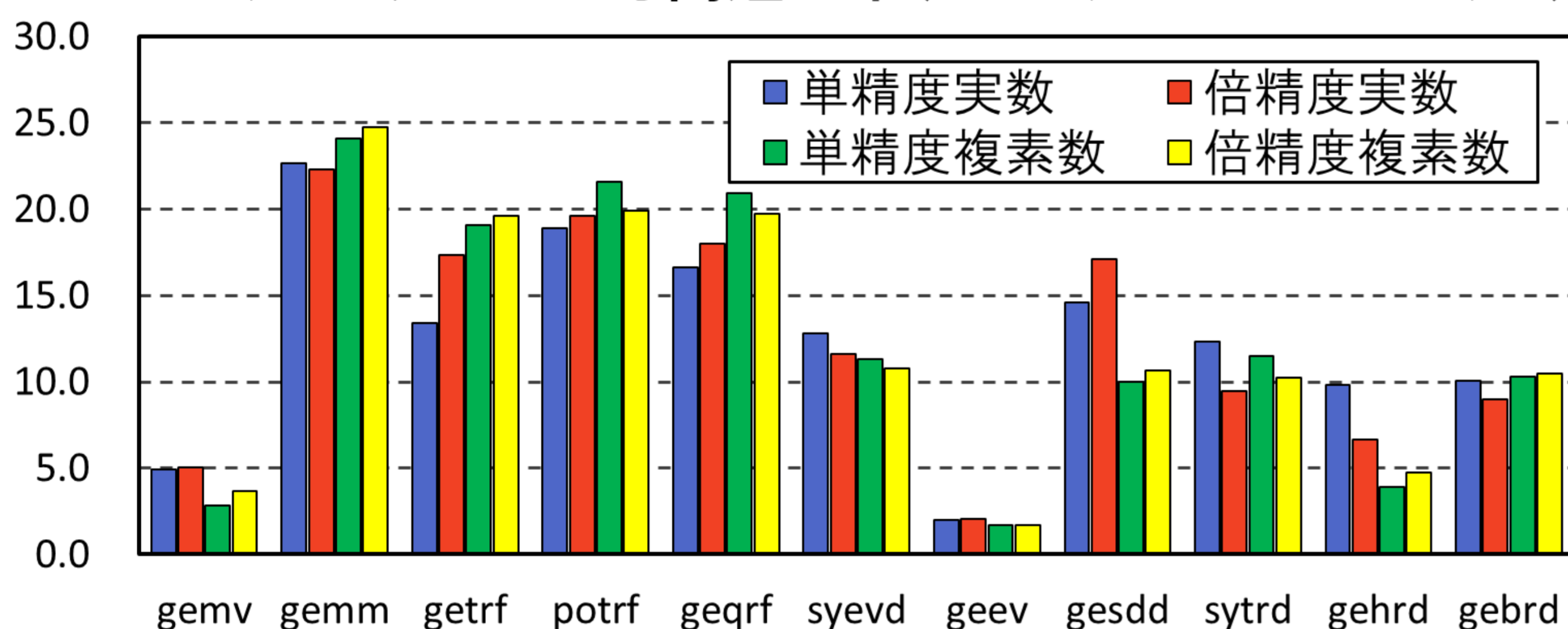


固有値計算プログラムに対する確率分布としての性能予測事例 (K. Tanaka et al., JJIAM Vol. 36(2), 719-742, 2019.)

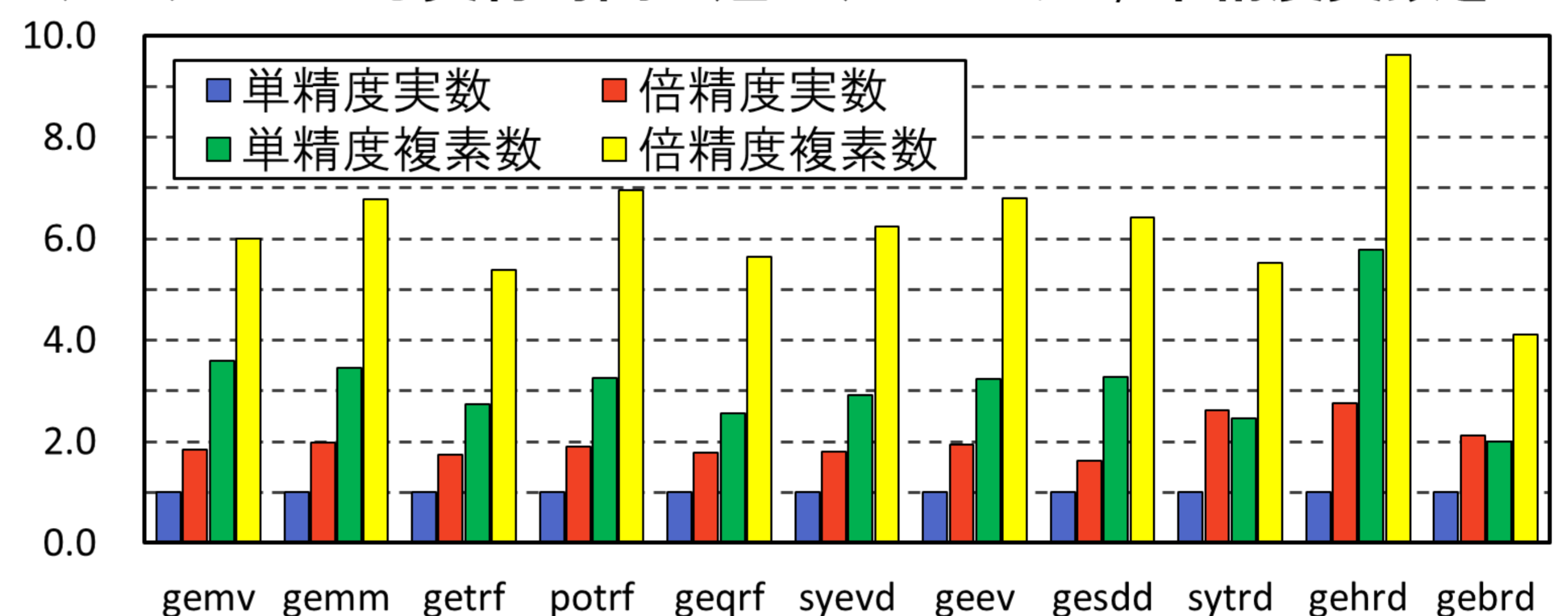
◆ 実施項目2: 演算精度・型が異なるプログラムに対する性能予測技術の開発

- アプリケーションにより、数学的に同じ問題に対して、データの精度(倍精度・単精度・半精度)と型(実数・複素数)が異なる。
- アルゴリズムの基本構造が同じであれば、精度や型が異なる場合でも、実行時間の挙動が共通点がある場合が多い。
- 各ケース(精度×型)に対して、個別に性能予測を行うのではなく、一体化による効率的な性能予測技術の開発を目指す。
- LAPACKで提供されている主要ルーチンを対象に、性能分析を行い、精度・型間の実行時間の関連性を明らかにする。

スレッド並列化による高速化率(1スレッド vs. 40スレッド)



データ型による実行時間の違い(40スレッド, 単精度実数を基準)



◆ 実施項目3: データ駆動型アプリケーションに対する性能予測技術

- 従来のシミュレーションとその結果に対するデータ処理を組み合わせた、階層的並列計算アプリケーションに対する性能予測技術を研究する。
- 並列データ処理部分の実行時間は、既存の事例が乏しいため、テストプログラムの整備とその性能評価・分析から行う。
- シミュレーション処理部分は、既存の知見(や実施項目1と2)を活用する。
- 具体的なアプリケーションとして、陽電子回折の実験データの解析を想定。(K. Tanaka et al., arXiv:1910.05743, 2019.)

