

鈴木智博 (山梨大学)

行列分解のタイルアルゴリズムの高並列環境における最適化

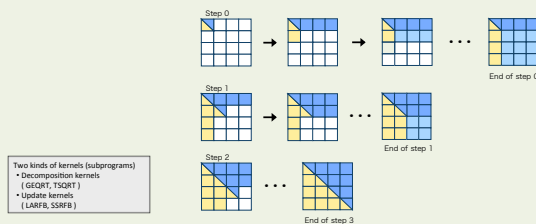


目的

行列分解の**タイルアルゴリズム**は細粒度のタスクを大量に生成可能
本研究ではタイルアルゴリズムの、

- ・タイルサイズチューニング
- ・ノード間通信の最適化

を行う



研究の意義

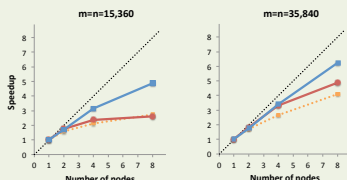
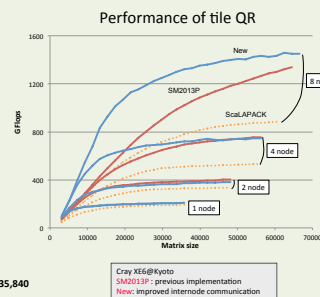
- ・大規模並列環境における高速な行列分解ルーチンの提供：
数値線形代数計算の大規模化、高速化の要求が高まり、高い並列性を持つ近年のハードウェアの性能を最大限に引き出す行列分解アルゴリズムが求められている
- ・国内最高クラスのスーパーコンピュータによる性能評価：
ScaLAPACKに代わるクラスタシステム向けの新しい数値計算ライブラリの開発をFX-10で行う
- ・未着手の研究課題への取り組み：
タイルアルゴリズムのタイルサイズチューニングは負荷分散、通信最適化、誤差への影響の面から 重要性が高い
→ タイルサイズチューニングに関する研究報告ない

これまでの研究成果

マルチコアクラスタ向け実装

実装の特徴

- ・1D block cyclic データ分散
- ・1 MPIプロセス/ノード
- ・OpenMP マルチスレッド
- ・1通信スレッド、他計算スレッド
- ・動的タスクスケジューリング
- ・通信の動的スケジューリング (左図のNew)

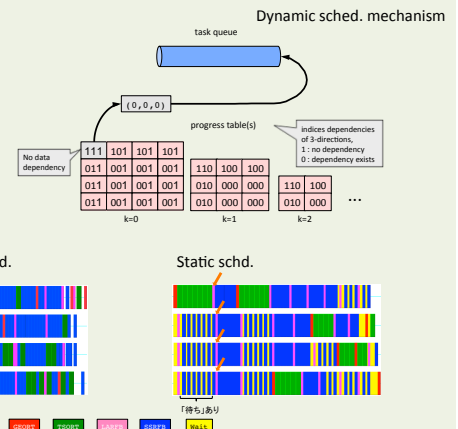


Strong scaling

動的タスクスケジューリング

動的タスクスケジューリングの構成要素

- ・タスクキュー
- ・プログレステーブル
- ・OpenMP (task parallel)



タイルサイズチューニング

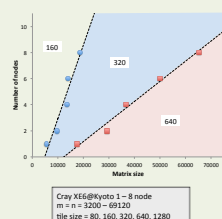
タイルサイズとタスク数

- ・タイルサイズ小 → タスク数多 → 負荷分散
- ・タイルサイズ大 → L3 BLAS 性能

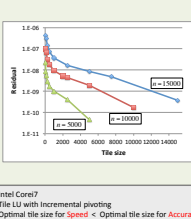
タイルサイズと誤差

- ・タイルサイズ大 → 誤差小

Optimal tile size shift point (QR)



Residual ||PA-LU|| (LU)



今後の研究計画

タイルサイズチューニング

- ・性能モデル (QR, LU)、誤差との関係 (LU)

ノード間通信の最適化

- ・高並列環境での性能評価 (QR)
- ・カーネル最適化による通信量削減 (QR, LU)