

研究背景・目的

- メモリ階層の深化, CPUのメニーコア化, GPUの導入など, 計算機が多様化・複雑化
- アプリケーションの性能に影響を及ぼすパラメータも複雑化
 - ピーク性能やメモリ性能はもちろん, ベクトル長と命令レイテンシの大きさ, キャッシュの速度やサイズ, コア間の通信レイテンシ, ノード間の通信レイテンシなど
 - アプリケーションの最適化は専門家でも難しくなっている
- 計算機の性能モデリングはアプリケーションの手動・自動最適化, 計算機の開発・導入において重要
 - マイクロベンチマークレベルでの評価はよく行われているが, 幅広くより実用的・先端的なアプリケーションを用いた, 様々なアーキテクチャにおける性能モデリングは十分でない
- **目的: アプリケーションの性能理解や自動最適化に有用な性能モデルの開発**
 - ハードウェアをメモリ性能と演算性能で単純化したルーフラインモデルは, ルーフラインに至らない最適化途上のアプリケーションの性能律速原因の理解に適さない
 - ハードウェアの複雑性やアプリケーションの特性を考慮した, 人間にわかりやすい性能モデルが必要
 - ハードウェアの複雑化に伴い, 自動最適化におけるパラメータ探索空間が増加
 - 探索空間を狭めるための, 自動最適化向けの性能モデルが必要

研究実施項目

- マイクロベンチマークによる性能評価
 - メモリ性能, キャッシュ性能, コア間通信レイテンシ等の計測
 - 実アプリケーションでの性能モデリングに活用
- ステンシル計算の時空間ブロッキング
 - ステンシルカーネルのパラメータ (次元数, 近傍セルの参照点数, 各セルの物理量など)
 - 時空間ブロッキングのパラメータ (空間ブロック形状・サイズ, 時間ブロッキングサイズなど)
 - プロセッサのパラメータ (キャッシュサイズ・速度, メモリ性能, 演算性能, コア間レイテンシなど)
 - 上記を踏まえた性能モデルの構築
- 低精度演算を含む非線形ソルバ
 - 精度影響が小さい計算カーネルの一部の低精度化は, 速度の向上に変換によるオーバーヘッド (変換そのもののコスト, 変換による最適化の阻害) が伴う
 - 低精度計算適用による性能の予測モデル
- 階層型行列演算の性能モデル
 - 階層型行列のパラメータ (低ランク部分行列のサイズ・ランク数, 小密行列の数など) は解析対象の形状によって大きく異なり, 構築するまでわからない
 - 階層型行列パラメータや実行するプロセッサによって適用すべき最適化手法が異なる
- 自動最適化への応用
 - 性能モデルの自動最適化への応用手法の検討
 - 自動最適化ツールであるppOpen-ATへの取り込み

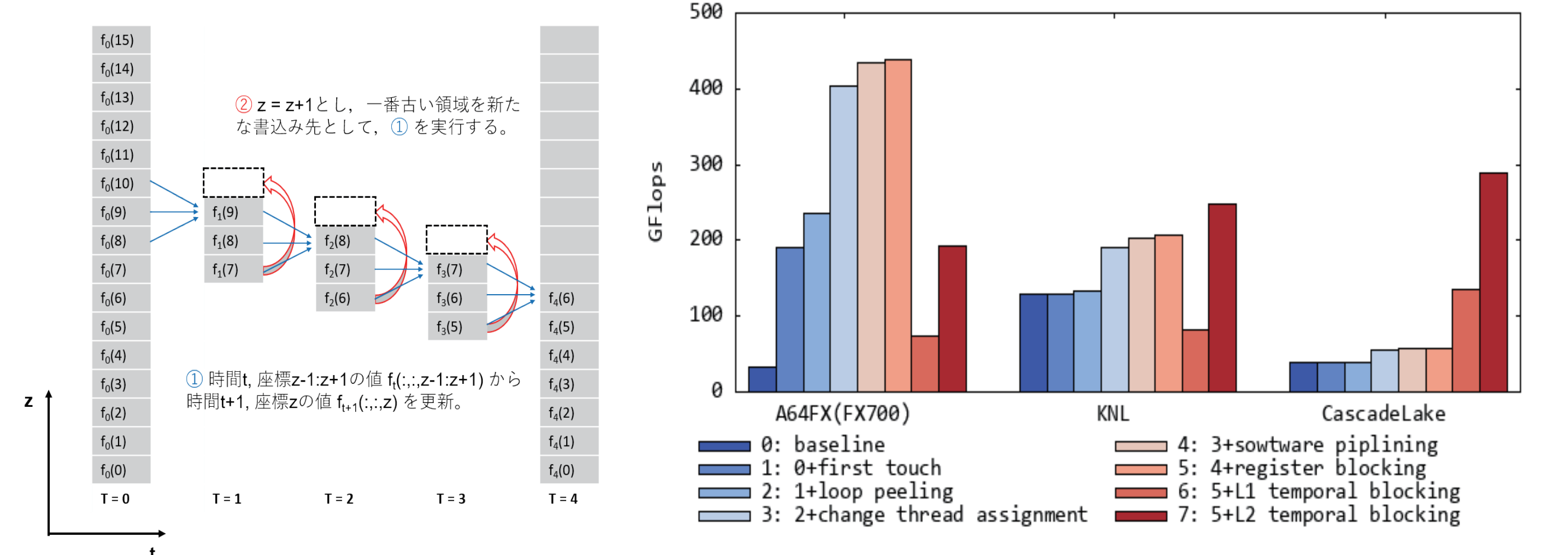


図1: 時空間ブロッキングの実装例。一度に進める時間ステップ数 (図の場合は4) や X, Y, Z 方向の空間ブロックサイズがパラメータとなる

図2: 倍精度3次元ステンシル計算の性能評価 [1]。右2つが時空間ブロッキングの適用によるもので, 空間ブロックサイズの違いで大きく性能が異なる。

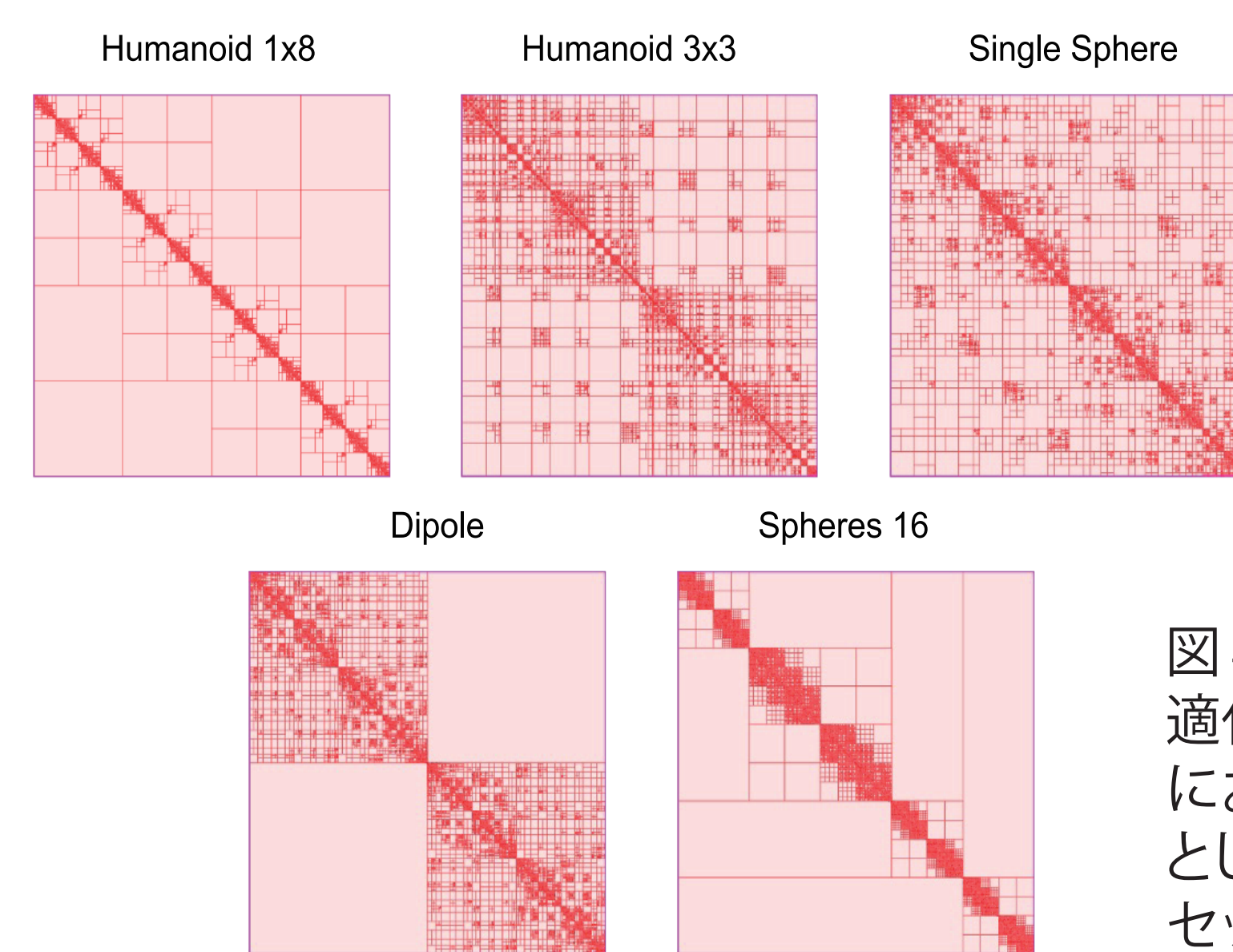


図3: 静電場解析で生じる階層型行列の例。薄いピンクで表現されている部分行列が低ランク近似化されており, 濃い赤で表現されている部分行列は密行列。解析対象の物体形状により, 階層型行列の構造が大きく異なる。

図4 (下): 階層型行列 - ベクトル積演算の最適化と A64FX, AMD EPYC, Intel CascadeLake における性能評価 [2]。一番左をベースラインとして, 順次最適化を追加している。対象プロセッサや対象とする階層型行列の違いによって最適化の効果度合いが異なっている。