

第8回 自動チューニング技術の現状と応用に関するシンポジウム(ATTA2016)

日時:2016年12月26日(月)13:30~14:00

場所:東京大学 山上会館

通信回避・削減アルゴリズムのための 自動チューニング技術の新展開

名古屋大学情報基盤センター 片桐孝洋

科研基盤B「通信回避・削減アルゴリズムのための
自動チューニング技術の新展開」2016-2018



背景

- ▶ 次世代スパコンでは、ノード数が10万～100万ノード級と想定され、高い並列性の確保と通信時間の削減が強く要請される。
- ▶ 特に通信時間については、通信レイテンシの削減がハードウェア上困難であることが知られている。
- ▶ そのため、**通信回避 (Communication Avoiding (CA))**、および、**通信削減 (Communication Reducing (CR))** アルゴリズムによる数値計算ライブラリの再構築が必須となる。
- ▶ 利便性の向上と多様な入力や計算機環境での高性能化のため、アプリケーションの性能パラメタを対象計算機のキャッシュサイズ、コア数、通信性能などの計算機アーキテクチャの特性に加えて、数値アルゴリズム選択に至る広範な要因を自動的にチューニングする**ソフトウェア自動チューニング技術 (AT 技術)**が、国内外から注目を集めている。

HPCソフトウェア開発の流れ図

コア数の増大、
プログラミング
モデルおよび
コード最適化の
複雑化
(多様な計算機
アーキテクチャ)

コンパイルと実行

3. 最適化フェーズ

結果の解析

2. プログラミング
フェーズ

4. チューニングに
関する知識探索と
データベース化
のフェーズ

対象計算機

チューニング知識
データベース

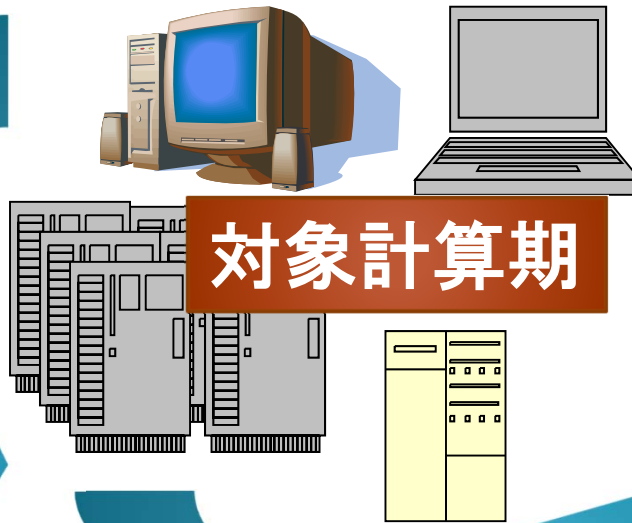
1. 仕様策定フェーズ

コード
生成

```
do i=1, n
do j=1, n
do k=1, n
C(i, j) = C(i, j) + A(i, k) * B(k, j)
enddo
enddo
enddo

do i=1, n, 2
do j=1, n, 2
do k=1, n, 2
Ctmp1 = C(i, j)
Ctmp2 = C(i+1, j)
do k=1, n, 2
Btmp1 = B(k, j)
Btmp2 = B(k+1, j)
Ctmp1 = Ctmp1 + A(i, k) * Btmp1
+ A(i, k+1) * Btmp2
Ctmp2 = Ctmp2 + A(i+1, k) * Btmp1
+ A(i+1, k+1) * Btmp2
enddo
C(i, j)=Ctmp1
C(i+1, j)=Ctmp2
enddo
enddo
```

```
!ABCLib$ install unroll (i,k) region start
!ABCLib$ name MyMatMul
!ABCLib$ varied (i,k) from 1 to 8
do i=1, n
do j=1, n
do k=1, n
C(i, j) = C(i, j) + A(i, k) * B(k, j)
enddo
enddo
enddo
!ABCLib$ install unroll (i,k) region end
```



研究期間内に何をどこまで明らかにしようとするのか

1. CA/CR アルゴリズムで現れる性能パラメタと、そのAT 実装のフレームワーク
2. CA/CR アルゴリズムのためのハイブリッド MPI/OpenMP 実行における高性能実装方式
3. 1、2 の処理で利用可能なAT 時間削減方式
4. プラズマシミュレーションおよびデータ同化処理の「実用コード」において、開発したAT を用いて達成できる高性能化の原理

研究計画・方法

- ▶ <AT 方式設計グループ>、<AT 実装方式開発グループ>、<AT アプリケーション適用グループ>の3グループを編成し、CA/CR のためのAT 機能のアルゴリズム、API 仕様策定、実装、および、性能評価を機能的に実施する。
- ▶ 開発工程を5つのフェーズに分け、CA/CR のためのAT 方式を開発する。
平成28年度の計画は、主に <CA/CR 用高性能実装とCA/CR 用AT 方式> の技術開発を行う。

AT要素技術の適用

AT基盤技術
(片桐、櫻井)

ppOpen-ATのAT技術

CAの数理
(須田)

BCBCGの数理

AT時間削減、
CA/CRアルゴリズム
実装 (田中)

d-Spline、
CBCG実装

MPI実装のAT
(黒田)

MPI最適化技術

SpMV高性能実装
メニーコア最適化
(大島)

SpMV、前処理
実装最適化技術

新OpenATLib設計

AT機能のAPI

次世代スパコン環境
に向けたCA/CR用
AT技術の新展開

データ同化
データ同化処理へHPC・
通信削減技術の適用

アプリケーション適用
AT有効性検証
(佐藤、岩下、長尾)

●統括 片桐孝洋(名大 教授)

(★はグループ代表者)

●AT方式設計グループ(ATMD-G)

- ★片桐孝洋(名大 教授、代表): 新OpenATLib設計、CA/CR向けAT方式の設計
- 黒田久泰(愛媛大 准教授、連携): MPI最適化方式開発と通信ライブラリATの設計
- 今村俊幸(理研、連携): GPU向けAT方式の知識提供とメニーコア適用の検討
- 須田礼仁(東大 教授、連携): CA/CR数理の提供

AT仕様の提示、AT機能要求

AT仕様変更、ATの実装

AT方式適用
性能評価

●AT実装方式開発グループ(ATIM-G)

- ★田中輝雄(工学院大 教授、分担): AT時間削減技術、CA/CRの実装
- 大島聡史(東大 助教、連携): GPUおよびメニーコア向け高性能SpMV
- 櫻井隆雄(日立、連携): OpenATLib、Xabclib実装方式の知識提供

AT仕様・実装変更
実性能の提示

AT技術の適用評価

実性能の提示

●ATアプリケーション適用グループ(ATAA-G)

■ステンシルアプリ適用サブグループ(SAA-SG)

- ★佐藤雅彦(核融合研 助教、分担):
陰解法MHDコード適用・検証
- 岩下武史(北大 教授、連携):
電磁気シミュレーション(FDTD)の知識提供

■データ同化適用サブグループ(DAA-SG)

- ★長尾大道(地震研 准教授、分担):
データ同化コード適用・検証
- 福田淳一(地震研 助教、連携):
データ同化コードの知識提供

研究計画

フェーズ	平成 28 年度	平成 29 年度	平成 30 年度
(1)新 OpenATLib 設計、 CA/CR 用 AT 開発	→		
(2)CA/CR 用高性能実装 開発	←	↓ (1)反映	
(3)CA/CR 用 AT 機能 高度化(AT 高度化)		↓ (1)(2)(4)反映	↓ (2)反映
(4)アプリケーション 適用評価	←		
(5)開発コードの 公開・保守			パッケージ化↓

研究事例：アジョイント法における 時空間ブロッキングの適用

- ▶ データ同化における新しいアルゴリズム
(東大長尾研と共同研究)
- ▶ 時空間ブロッキング
 - ▶ 新しい軸(時空間)の並列性の活用
 1. 階層メモリ(階層キャッシュ)最適化
 2. 通信回避可能(通信回避アルゴリズム)
 - ▶ 複数の時空間ステップの通信を1回に集約し、
通信レイテンシを隠蔽なアルゴリズム

第157回ハイパフォーマンスコンピューティング研究発表会

日程：2016年12月21日(水)～12月22日(木)

場所：沖縄産業支援センター で発表済み

アジョイント法におけるForward model への 階層ブロッキング適用による高性能化

名古屋大学大学院 情報科学研究科

東京大学 地震研究所

東京大学 地震研究所/東京大学大学院 情報理工学系研究科

名古屋大学 情報基盤センター 大規模計算支援環境研究部門

//

//

池田 朋哉

伊藤 伸一

長尾 大道

片桐 孝洋

永井 亨

荻野 正雄

アウトライン

- ・ データ同化
 - ・ 非逐次型データ同化：アジョイント法
- ・ テストモデル：フェーズフィールドモデル
- ・ Forward modelの計算式
- ・ 階層ブロッキング
 - ・ Forward modelのステンシル計算への時空間ブロッキング
 - ・ 複数Forward modelの計算のブロッキング
- ・ 性能評価
- ・ まとめと今後の課題