

10-NA08

超並列計算によるマルチスケール・マルチフィジックス

心臓シミュレーション

岩村 尚（東京大学） 鷲尾 巧（東京大学） 岡田 純一（東京大学）

中島 研吾（東京大学） 久田 俊明（東京大学）

概要 当研究チームでは、マルチスケール・マルチフィジックス心臓シミュレータを開発している。これは有限要素法で内部構造まで再現した心筋細胞モデルを心臓モデルの各要素に割り当て、均質化法に基づきその力学特性を定義し、両者を連成させ同時に解くことで、細胞レベルのメカニズムと心室の拍動・血液拍出の巨視的現象の関連を明らかにするという世界に類を見ないものである。本共同研究プロジェクトにおいては、均質化法理論に基づいた 66 万心筋細胞のマルチスケールシミュレーションを次世代スーパーコンピュータで実行するために必要な技術開発、問題点抽出をする。これまでに、6144 コアでの実行および良好なスケーラビリティの達成、心筋細胞用ソルバ・前処理手法の開発、ハイブリッド並列のコーディング、性能分析と問題点抽出を行っている。

1. 研究の目的と意義

「ヒトゲノムプロジェクト」の完成をうけて、生命科学の焦点は遺伝情報の解明からタンパク、細胞内小器官、細胞、組織、臓器、そして個体という多階層に亘る複雑な生命現象の理解へと移行しつつある。第三の実験(in-silico 実験)として期待されている計算機シミュレーションにおいても、マルチスケール解析の重要性が増している。当研究チームが開発を進める UT-Heart は、細胞イオンチャンネルや収縮タンパクの数理モデルから出発し有限要素法でモデル化された心室の収縮、血液の拍出に至る現象を一貫して再現できる国際的にも突出した心臓シミュレータである。一方、有限要素法に基づく心筋細胞も並行して開発しており、今後これを組み込むことで、世界でも前例のないシームレスなマルチスケール・マルチフィジックス心臓シミュレータの完成を目指している。

UT-Heart の両心室・心房（マクロモデル）は約 66 万個の有限要素から構成され、各々の力学的特性は細胞のミクロモデルを均質化して定められる。従って原理的にはマクロモデルと 66 万個のミクロモデルを連成させた巨大な非線形方程式を刻々解く均質化解析に帰着する。当研究チームはこのシ

ミュレーションをペタフロップス級の次世代スパコンで実現することを目指とし、その準備を理論開発と実装技術の両面から進めている。現在のところ 10 万自由度のミクロモデルからなるマルチスケール解析に対し、心臓の 1 拍動に対する計算時間が 1～2 日になるものと予測している。本公募型共同研究においては、計算科学分野の研究者との密接な協力によって、大規模化に対応した様々な最適化を特に重点的に実施する。

この心臓シミュレータを完成することにより、近年の分子生物学の進歩により集積されつつある遺伝子、分子レベルの膨大な知見とマクロな臨床医学との間のギャップを埋めることが可能となり、新たな医学が創出されると共に、医療や創薬への画期的な貢献ができる。

2. 当拠点公募型共同研究として実施した意義

(1) 共同研究を実施した大学名

東京大学

(2) 共同研究分野

超大規模数値計算系応用分野

(3) 当公募型共同研究ならではの事項など

ペタフロップス級スパコンでの実行に向けた準

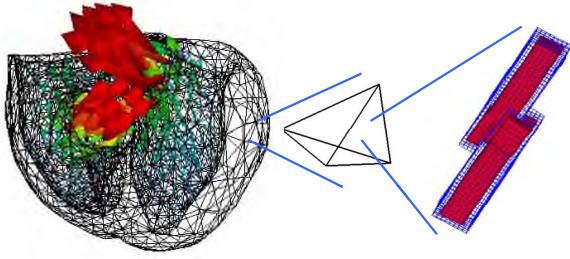


Fig. 1 マクロモデル有限要素への
ミクロモデル割り当て

備を計算科学分野の研究者との密接な協力によって実施する。

3. 研究成果の詳細

本章では、まず第1回シンポジウムおよび Super Computing 2010 (SC10) で発表したマルチスケール心臓シミュレータの概要とスケーラビリティについて述べる。次に、シミュレータの負荷不均衡に関する分析の現状報告を行う。

3.1. マルチスケール心臓シミュレータの概要

マルチスケール心臓シミュレータでは、有限要素法で内部構造まで再現した心筋細胞モデル(ミクロモデル)を心臓モデル(マクロモデル)の各有限要素に割り当て (Fig. 1)、均質化法に基づきその力学特性を定義し、両者を連成させ同時に解く。これにより、細胞下レベルのメカニズムが心臓の拍動、血液の拍出などの巨視的現象にどのような影響を及ぼすか、また逆にその結果として定まる心臓全体としての運動が各心筋細胞にどのような環境をもたらすかを明らかにする。

マルチスケール解析は、各時刻での非線形方程式を線形化し、Newton-Raphson 法で反復的に解く問題に帰着する。行列の構成を Fig. 2 に示す。図の右上が解くべき全体の行列を表している。 A_{w_k} は各ミクロモデル k の変位増分 Δw_k に対応し、 $\bar{A}$ は通常の有限要素法で得られる行列で、マクロモデルの変位増分 Δu に対応する。なお以降、上線を付けた記号はマクロモデルを示すものとする。

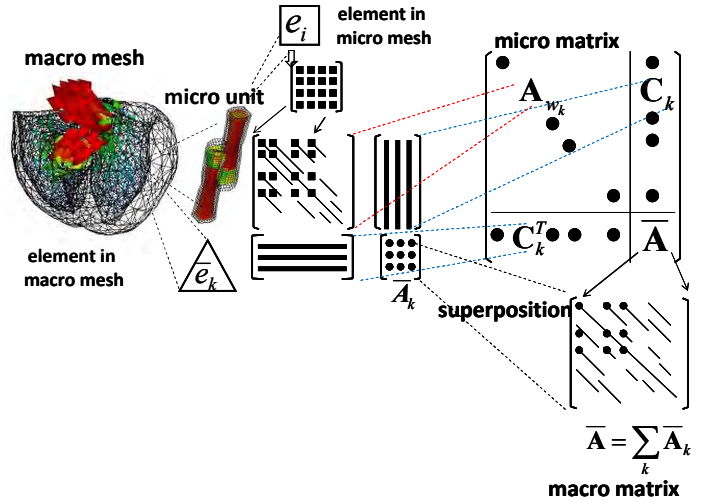


Fig. 2 線形方程式の行列の構造

C_k はミクロモデル・マクロモデル連成を表す。このうち、ミクロモデルに対応する自由度が非常に大きいので巨大な行列方程式となるが、その一方で、ミクロモデルの未知数は互いに直接の関与はない。このような方程式をブロック LU 分解によって解くために、Schur complement を計算する必要がある。この Schur complement の計算には、各ミクロモデルにおいて 9 成分の特性変位モード χ_k を利用する。 χ_k は

$$A_{w_k} \chi_k = G_k$$

により求められる。ただし、 G_k はマクロ変位勾配テンソルの 9 成分に対応する 9 個のベクトルからなる行列であり、 $C_k = G_k \bar{B}_k$ を満たす。 $\bar{B}_k$ はマクロモデルにおける勾配-変位マトリクスである。そして全体の Schur complement $\bar{S}$ は

$$\bar{S} = \sum_k \bar{S}_k$$

ここで

$$\begin{aligned} \bar{S}_k &= \bar{A}_k - \bar{B}_k^T \bar{G}_k^T A_{w_k}^{-1} G_k \bar{B}_k \\ &= \bar{A}_k - \bar{B}_k^T \chi_k^T A_{w_k} \chi_k \bar{B}_k \end{aligned}$$

となる。上記を踏まえ、Newton-Raphson 反復中の線型方程式の解法は下記ようになる。

1. Solve $A_{w_k} \Delta w_k = r_{w_k}, 1 \leq k \leq n$
2. Solve $\bar{S} \Delta u = \sum_k (\bar{r}_k - \bar{B}_k^T \chi_k^T r_{w_k})$

```

for  $istep = 1, \dots$  (time step)
  for  $iter = 1, \dots$  (Newton - Raphson step)
    comm1: Macro  $\rightarrow$  Micro
    for  $k = 1, \dots$  (Computation for each Micro Unit)
      assemble  $A_{w_k}, \{\bar{r}_k\}, \{\bar{r}_{w_k}\}, G_k, \bar{A}_k$ 
      Micro Solver( $\Delta w_k, \chi_k : N_t$ )
      compute  $\bar{S}_k = \bar{A}_k - \bar{B}_k^T \chi_k^T A_{w_k} \chi_k \bar{B}_k$ 
      compute  $\{\bar{g}_k\} = \{\bar{r}_k\} - \bar{B}_k \chi_k^T \{\bar{r}_{w_k}\}$ 
    end for
    comm1: Micro  $\rightarrow$  Macro
    compute  $\bar{S} = \sum_k \bar{S}_k, \{\bar{g}\} = \{\bar{g}_k\}$ 
    if converge return
      Solve  $\bar{S} \{\Delta \bar{u}\} = \{\bar{g}\}$ 
      update  $\{\bar{u}\}, \{w_k\}$ 
    end for
  end for

```

```

procedure Micro Solver( $\Delta w_k, \chi_k : N_t$ )
  if ( $\text{mod}(istep, N_t) == 0 \ \&\& \ iter == 1$ )
    LU factorization  $A_{w_k} = L_{w_k} U_{w_k}$ 
    Solve  $A_{w_k} \{\Delta w_k\} = \{\bar{r}_{w_k}\}$  by preconditioned GMRES
      using  $L_{w_k} U_{w_k}$  as the preconditioner
  if ( $\text{mod}(istep, N_t) == 0 \ \&\& \ iter == 1$ )
     $L_{w_k} U_{w_k} \chi_k = G_k$ 
  end procedure

```

Fig. 3 マルチスケールシミュレーションのアルゴリズム

$$3. \quad \Delta w_k = \Delta w_k - \chi_k^T \bar{B}_k \Delta \bar{u}, 1 \leq k \leq n$$

1. のマイクロモデルの変位増分を求める箇所が計算の大部分を占めるが、これはマイクロモデルごとに独立に解くことが可能なため、良好なスケーラビリティを期待できる。

次に並列計算のフローを Fig. 3 に示す。Newton-Raphson のループ内における Macro->Micro 通信および Micro->Macro 通信には含まれた箇所がマイクロモデルの計算であり、それぞれの MPI プロセスが独立に計算可能な箇所である。procedure Micro Solver では完全 LU 分解した行列を前処理行列として GMRES 法によりマイクロモデルの変位増分および特性変位モードを求める。実際には LU 分解と特性変位モードの計算は 10 time step ごと、かつ Newton-Raphson 法の最初の反復のみで行い、それぞれの結果を使い続ける。このような近似を行っても Newton-Raphson 反復において収束解を厳密に求めているので問題はない。

最後に MPI プロセスの割り当てについて述べる。全ての MPI プロセスは 1 個以上のマイクロモデルの計算を担当するが、マクロモデルの計算も行うマ

クロ・マイクロ担当とマイクロモデルのみ計算を行うマイクロ担当が存在する。実際には前者は MPI rank 0-63、後者が残りである。マクロ・マイクロ担当はマクロモデル(流体・心筋の計算)を行い、Fig. 3 の Macro->Micro 通信により、マイクロモデルの計算に必要なデータ(マクロモデルの変位情報など)をマクロ・マイクロ担当およびマイクロ担当のプロセスに送信する。逆に Micro->Macro 通信では各マイクロモデルの計算結果(Schur complement など)を送信する。

3.2. 東大 T2K スパコン 6144 コアでの計算性能

本節では、マルチスケール心臓シミュレーションのスケーラビリティ確認のために行った東大 T2K スパコンでの計算結果を述べる。

計算環境として、東京大学の T2K オープンスパコン 6144 コアを使用し、Flat MPI で実行した。Numa policy は default とした。また、計算の time step は心臓の弛緩期から開始し、1.5 拍分に相当する 600 ステップの計算を行った。220 ステップ付近から歪の大きな収縮期が開始する。

Table 1 は細胞数(マイクロモデル数)とシミュレーションの自由度(NDOF)についてまとめたもので

ある。

Table 1 細胞数および自由度

Number of cells	6,144
NDOF per cell	26,121
Total NDOF of cells	1.6×10^8
NDOF of the organ	46,165
Total NDOF of a heart	1.6×10^8

まず、スケーラビリティの測定では1プロセスが担当する細胞の数を1から4個まで変更した。例えば1536coreでの実行では各MPIプロセスは細胞を4個、6144coreではそれぞれ1個を計算する。Fig.4は1536coreを1としたときの速度向上率である。6144coreまで良好なスケーラビリティを示している。

次に、6144core使用時におけるマイクロモデル(Fig. 3の2つの通信で挟まれた部分)とマクロモデルの累積計算時間の推移をFig.5に示す。これまで述べてきたように、マイクロモデルの計算が大部分(96%以上)を占めていることがわかる。

最後にマイクロモデルの計算時間の内訳をFig.6に示す。これはtime step 200から250までの変位増分計算(LU分解+GMRES)、特性変位モード計算、行列アセンブルそれぞれの計算時間を表している。LU分解を行うステップではLU分解の時間が大部分を占める一方、非LU分解ステップではGMRESの時間が、特に収縮期が始まるtime step 220以降で大きな割合を占めるようになる。

3.3. ミクロモデル計算の負荷不均衡

前節で述べたように、速度向上率としては、現状のマルチスケール心臓シミュレータで良好なスケーラビリティが得られている。しかし速度向上率の基準とした1536コアの計算においてマイクロモデル(細胞)間の計算時間不均衡が見られる。次世代スパコン上でより大規模な並列計算を行う場合、そのばらつきはさらに大きくなると考えられ、効率の良い高速なシミュレーションを実行するには、そのばらつきの原因を特定し、対策を施すことが必須である。本節では、マイクロモデル計算時間のばらつきの分析について現状報告を行う。な

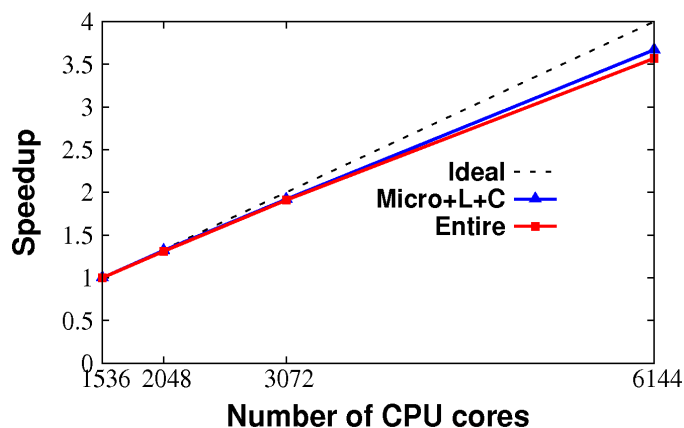


Fig. 4 速度向上率

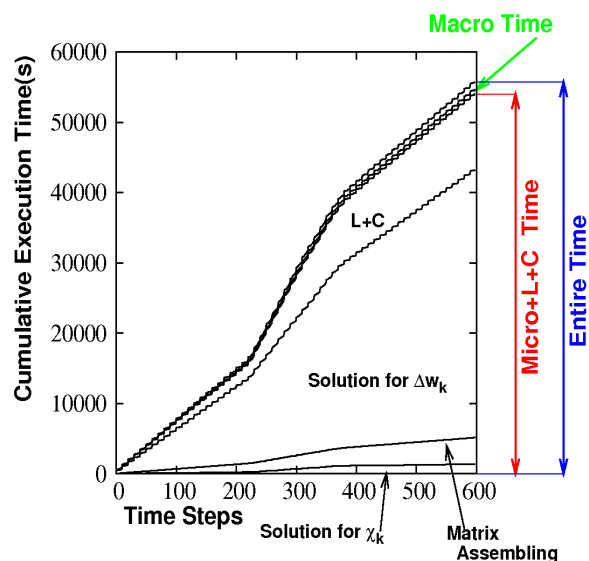


Fig. 5 ミクロモデル、マクロモデルおよび全体の計算時間推移

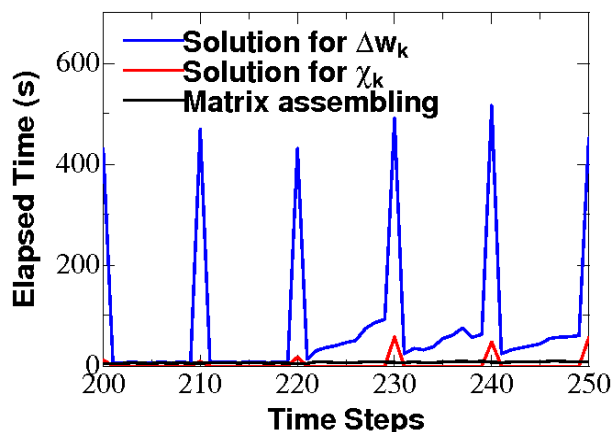


Fig. 6 ミクロモデル計算中の3種の計算時間

お、分析はT2K オープンスパコン(東京大学)の4,096コア・4,096細胞を用いた計算により行った。細胞が割り当てられた要素をFig.7に示す。これは心基部側から心室を見た図であるが、左心室自由壁と中隔の一部に細胞を割り当てている。

3.3.1. ミクロモデル計算時間のばらつきの概要

我々は、上述の 6144core 計算の段階から、マクロ・ミクロ担当プロセスとミクロ担当プロセスを比べた場合、同一の計算を行っている筈のミクロモデルの計算時間が前者において大となる傾向にあることを確認していた。今回の分析はその傾向を確認することから開始した。

Fig.8 は time step ごとのミクロモデル計算時間をプロセス種別で、すなわちミクロ担当プロセスとマクロ・ミクロ担当プロセスに分けてプロットしたものである。赤の点がミクロ担当プロセス、緑の点がマクロ・ミクロ担当プロセスを表す。ただし計算時間に大きな差がある3グループ、すなわち、非LU分解 time step での弛緩期(～time step 220)、興奮期(time step 220～)およびLU分解 time step に着目したグラフとした(それぞれ上・中・下段)。上段の弛緩期の非LU分解 time step では、マクロ・ミクロ担当プロセスが計算時間のばらつきを大きくしている箇所が存在する。一方、中段の興奮期の非LU分解 time step ではプロセス種別間で顕著な差は無い。下段のLU分解 time step では1つのプロセスだけ計算時間が大きくなっている。これはrank 0のプロセスである。

これらの結果から、弛緩期・非LU分解 time step でのばらつきの一因としてマクロ・ミクロ担当プロセスがばらつきを大きくしていると考えられる。一方、収縮期・非LU分解 time step であるが、こ

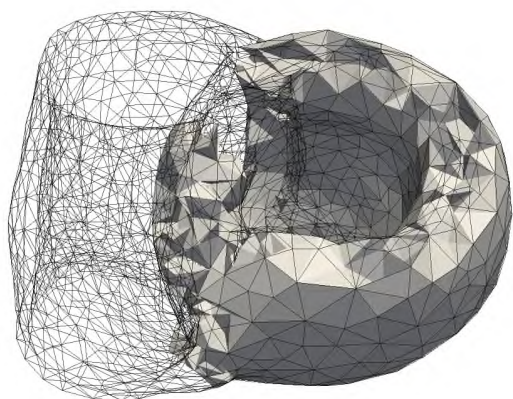


Fig. 7 細胞が割り当てられたマクロモデル心筋要素

のばらつきの主たる要因はマクロ・ミクロ担当プロセスとは異なると考えられる。そこで、弛緩期と収縮期での計算時間の違いを考えると、収縮期では心臓壁の運動により歪が増大して行列の性質

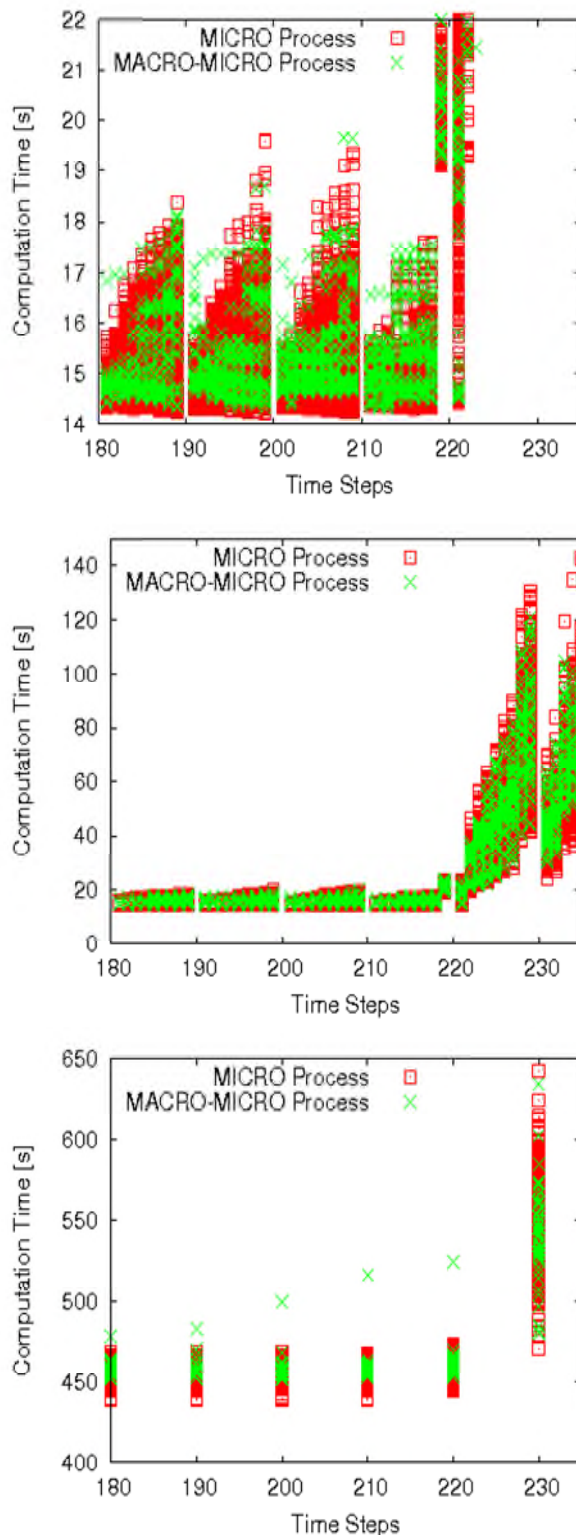


Fig. 8 MPI プロセス種類ごとのミクロモデル計算時間

が悪化し、それにより GMRES の収束に要する反復回数ひいては計算時間が増大すると考えられる。

ところが心臓の特徴として、心臓壁の内側の歪が大きくなる傾向にある。したがって、マイクロモデルの心室に占める空間的位置の違いにより、ばらつきが発生しているのではないかと考えた。

そこで、今回の分析では以下の要因によるばらつきの発生を調査・確認する。

- MPI プロセス種別
- ミクロモデルの空間的位置の違い

3.3.2. MPI プロセス種別による負荷不均衡

弛緩期・非 LU 分解 time step における計算時間は、GMRES によるものが支配的である。一方で、弛緩期ではマイクロモデル間の運動に差がないため、GMRES 反復数の差は少なく、負荷不均衡の原因は数理モデルやアルゴリズムよりも計算機システムによるものではないかと考えた。ここで Fig. 3 で述べたアルゴリズムを見ると、10 time step ごとに行う LU 分解結果を前処理行列として、GMRES の反復処理中に前進後退代入をしている。T2K スパコンは ccNUMA つまり、CPU ソケットとメモリの距離に差がつけられているが、仮に LU 分解結果が置かれたメモリが、それを使用する MPI プロセスが置かれた CPU ソケットのローカルメモリでない場合、メモリ参照時間が増大すると考えられる。

そこで実際に LU 分解結果を格納する配列を初期化したときと GMRES 実行時における、各 MPI プロセスが存在する CPU ソケットを調査した。その結果、

- 全 236 time step の計算で、マクロ・マイクロ担当プロセスが GMRES 試行回数の 20%、マイクロ担当プロセスが 3%の割合で LU 分解結果が置かれたメモリと離れた CPU ソケットで実行していた
- 計算時間が他と比較して大きくなっているプロセスは、上記の MPI プロセスのソケット移動を起こしていた

このようなソケット移動が起きた原因のひとつとして NUMA policy を default 値で計算を実行していたことが考えられる。そこで numactl で -physcpubind と -membind を設定し、MPI プロセス

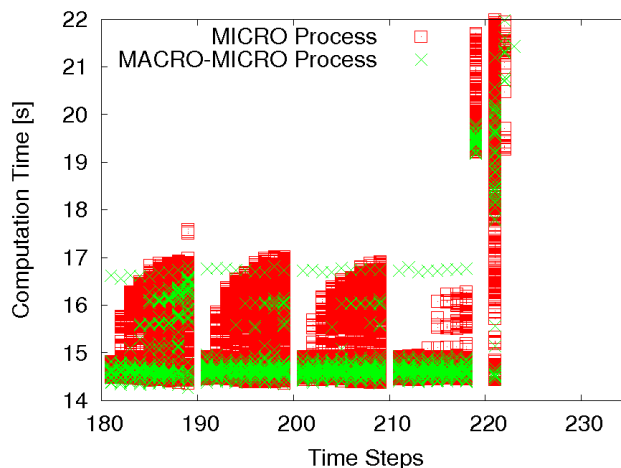


Fig. 9 CPU とメモリの bind 時での弛緩期・非 LU 分解 step におけるマイクロモデル計算時間

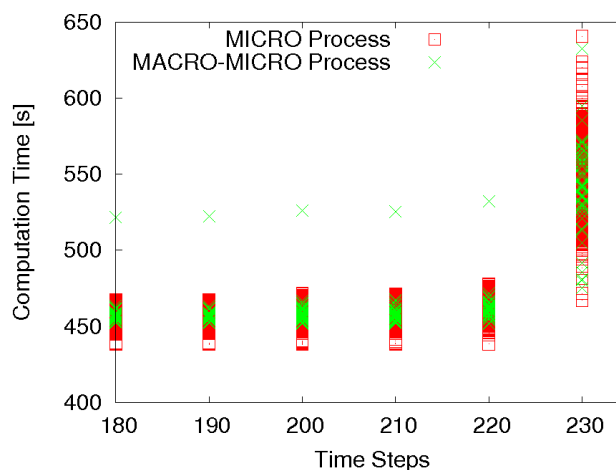


Fig. 10 CPU とメモリの bind 時での LU 分解 step におけるマイクロモデル計算時間

を CPU コアに固定し、必ずローカル(一番近い)メモリを使用する計算を行った。弛緩期の非 LU 分解 time step における計算時間を Fig. 9 に示す。Fig. 8 上段と比較するとばらつきが抑えられていることがわかる。なお各ステップの 16.6s 付近に離れて存在するマクロ・マイクロ担当プロセスは rank0 である。このグラフの範囲では rank0 が NUMA policy default のときよりもばらつきを大きくしている箇所もあるが、他の time step ではこのような影響は無かった。

しかしながら、Fig. 10 に示す LU 分解 time step では rank0 の計算時間が bind 無しよりも、最大で 50 秒ほど伸びていることがわかる。また他の time step でも rank 0 の計算時間が増大していた。こ

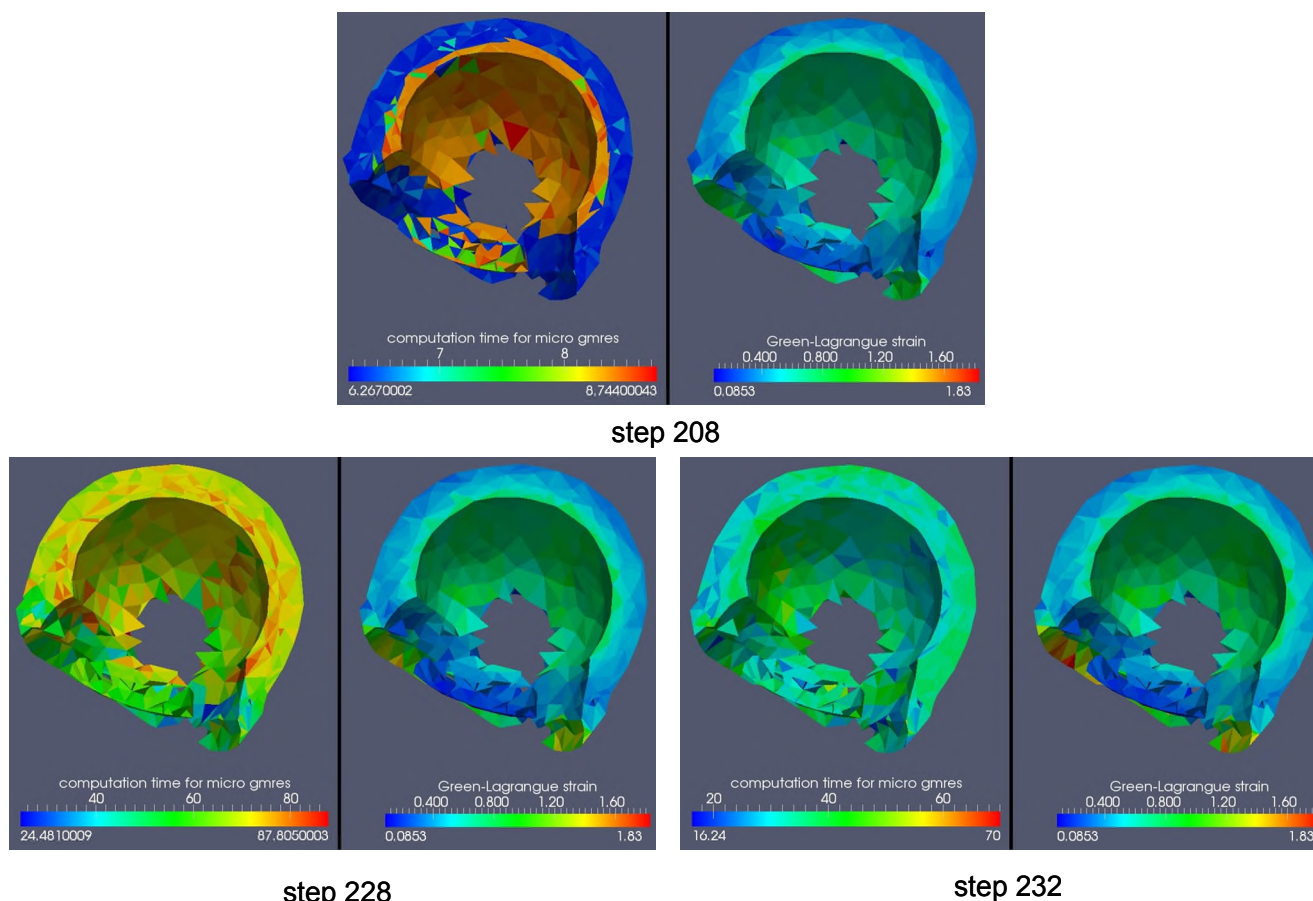


Fig. 11 GMRES 計算時間と Green-Lagrange 歪の大きさの分布. 弛緩期:step208, 収縮期:step228, 232. 各 step と
も左が計算時間, 右が歪の大きさ.

のような rank 0 の挙動によりシミュレーションの全実行時間、マイクロモデルの計算時間は bind あり/なしでほぼ変化がなかった。つまり、弛緩期・非 LU 分解 time step ではばらつきが減少したが rank 0 の伸びにより相殺された形となった。

この rank 0 の計算時間が大きくなる現象については、

- bind あり/なしに関わらず発生する
- 256 コア、1024 コアの計算(いずれも 1 細胞/1 プロセス)では現れなかった

という点が確認できた。つまりこの問題は並列度を増やしていくことで顕在化したものであり、将来のより大規模な並列計算においても現れ、更により大きな問題点になりうる可能性がある。今後はこの原因と対策を考える必要がある。

3.3.3. ミクロモデルの心室に占める空間的位置による負荷不均衡

ミクロモデルの心室に占める空間的位置と負荷

の関連を調べるために、非 LU 分解 time step の弛緩期・興奮期における GMRES 実行時間と、Green-Lagrange 歪テンソルの大きさの分布を確認した。歪テンソルの大きさには主ひずみの二乗和の平方根をとったものを用いた。計算結果には CPU とメモリの bind を行ったときのものを使用した。

Fig. 11 は分布を可視化したものである。弛緩期では計算時間大の領域と歪の大きい領域は比較的一致しているが、計算時間の差は小さい(6.2~8.7s)。興奮期のデータ 2 つに関しては、心室壁の内側(endo 側)の歪が高くなっているものの、計算時間との相関は見られない。

今回の分析では、弛緩期における歪と計算時間の相関が確認できたが、興奮期においては確認できなかった。弛緩期の小さい計算時間のばらつきを解消しても効果は小さく、今後は興奮期におけるばらつきの分析に力を入れる必要がある。

3.4. まとめと今後の予定

マルチスケール心臓シミュレータのマイクロモデル計算の負荷不均衡について分析を行い、以下のことがわかった。

- MPI プロセスの種別(マクロ・マイクロ担当プロセス、マイクロ担当プロセス)によるばらつきは、プロセスのCPUソケット移動に起因するGMRES計算時間の増大によるもの
- rank 0 の計算時間が並列度に伴い増大
- 興奮期におけるGMRES計算時間と歪の大きさの関連を調査したが、相関は見られず

今後の予定としてはrank 0 の計算増大の原因究明とその対策、興奮期における計算時間増大への対策を行っていく。

4. これまでの進捗状況と今後の展望

今年度の課題として以下3点を挙げた。

1. 10-20万自由度の心筋細胞用ソルバの開発
 - ・強力な前処理手法の開発
2. ハイブリッド並列化の試行
 - ・1細胞の計算をスレッド並列高速化
3. ペタスケールシミュレーションを目指した問題点抽出・最適化

これに対し進捗状況は下記のとおりで、概ね計画通りに進行している。

1. LU分解結果を前処理行列に利用することで収束性向上
2. 1細胞1スレッドのハイブリッド並列化は完成
3. rank 0 の増大、興奮期の計算時間のばらつきを継続して分析、対策する必要がある

今後の展望として、ハイブリッド並列化作業と性能分析を継続して行っていく。

5. 研究成果リスト

- (1) 学術論文
- (2) 国際会議プロシーディングス
- (3) 国際会議発表

Hosoi A, Washio T, Okada J, Kadooka Y, Nakajima K and Hisada T, "A Multi-Scale Heart Simulation on Massively Parallel Computers, to be presented," ACM/IEEE

Supercomputing Conference(SC10)

- (4) 国内会議発表

細井 聡, 鷲尾 巧, 門岡良昌, 久田俊明, 非圧縮拘束条件付き有限要素ソルバーのILU前処理におけるスレッドレベル並列化に関する検討, 第15回計算工学講演会論文集, pp. 182-186, 2010

- (5) その他