

11-MD05

計測融合オペレーション実現のための 大規模計算機空気冷却風速場の実時間解析

研究課題責任者 松岡 浩 (理化学研究所)

メンバー: 松岡浩, 峯尾真一, 横川三津夫, 瀧塚博之, 伊賀崎誠 (以上、理研)、渡辺正 (原子力機構)、板倉憲一 (海洋機構)、福田正大 (計算科学振興財団)、菊池範子 (カスタムシステム株) / 東北大)、**拠点メンバー:** 小林広明, 江川隆輔 (以上、東北大)、竹村治雄, 菊池誠, 東田学 (以上、阪大)、青柳睦, 高見利也, 小林泰三 (以上、九大)

概要 “京コンピュータ”施設を解析対象にして、格子ガス法面心超立方体モデルによる計算機空気冷却風速場のシミュレーションの試計算を行った。定常状態の解析では、ナビエ・ストークス方程式解法に基づく Open-FOAM による計算との比較を行い、格子ガス法計算の場合に仮定した冷却空気の吹き出し量や流体粘性の条件が不適切であった可能性を認識した。また、ベクトル計算機 SX-9 の 32CPU を用いて、全空調機と全計算ラックのファンを同時に起動させた場合に、計算機冷却空気の流れが定常状態に達するまでの過渡変化シミュレーションを行った。この結果、1024×1024×256 ≒ 2.6 億個程度の微細格子点分割でも、実時間とほぼ同じオーダーの時間で計算できることを確認した。このことから、計算機空気冷却風速場の実時間シミュレーションを実現するには、10 テラフロップス程度の計算機が必要であること、及び、これは最新のスーパーコンピュータ 1 計算ラック程度の計算能力に相当することから、十分確保可能な規模であると評価した。

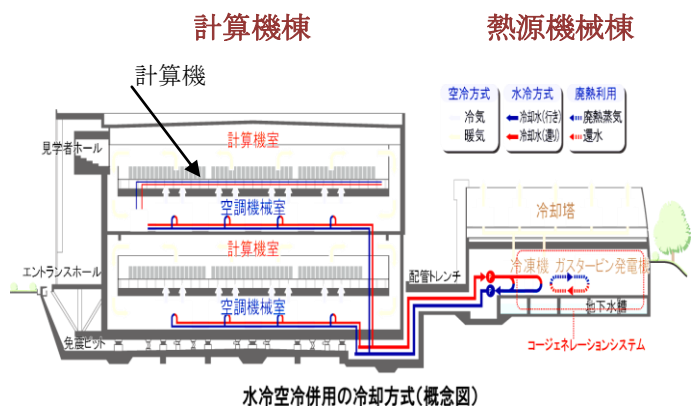
1. 研究の目的と意義

(1) 目的

格子ガス法超並列計算手法を利用して、大規模計算機の空気冷却風速場について実時間シミュレーションの実現が可能であることを示す。具体的には、現実の大規模計算機室(例：“京コンピュータ”施設)を解析対象にして、ナビエ・ストークス方程式に基づく流体シミュレーション手法との比較計算を行い、定常状態シミュレーションの妥当性を確認するとともに、異なる定常状態間の過渡変化を実時間で計算できる計算機能力が容易に確保可能な規模であることを示す。また、計算機センターの異常時にも過渡変化シミュレーションに基づく適切なオペレーション対応が可能になることを目指して、当該計算機能力を計算機の広域連携環境によって確保する方法を検討する。

以上の成果は、ペタフロップスを超える大規模計算機の空気冷却に関する計測融合オペレーションシステムの構築に資するものであり、これによって、本研究は、①筐体内の電子部品にやさしい過渡変化を実現して高信頼性・長寿命化を目指す運転と、②冷却・空調・電源系を含めた計算機システムプラント全体のエネルギー効率向上を目指

す運転の両立に貢献できるものになる。



(2) 意義

ペタフロップスを超える超並列高速計算機システムでは、通常数十メガワット以上の巨大な発熱を伴うばかりでなく、その発熱分布は広範囲に広がり、計算機システムのジョブスケジューラによ

る計算ノードの割当て・解除、各種プログラムのエラー発生、計算機ハードウェアの故障のほか、これらに伴う筐体ファン自動オンオフや、故障の修理保守後の起動、さらには、計算機システムの冷却・空調・電源系におけるトラブル発生で予測できない形状変化をする。他方、従来型の冷却制御は、液冷システム、空冷システムともに、計算機冷却場の入口温度と入口流量を確保し、各筐体では、温度や湿度がある制限値を超えた場合に電源オフにすることで計算ボード等の電子装置の保護を行っている。今後の大規模計算機システムにおいては、上述のとおり、異常な過渡変化の過程が多様化するため、従来の制御方式では、定格温度や定格湿度から大きなオーバーシュートやアンダーシュートが生じる可能性が高い。これらは、制限値に達すれば、稼働率の低下、達しない場合でも、電子部品の信頼性や寿命の低下を招く。このため、各種センサー情報とシミュレーションから過渡変化の全体像を的確に把握し、冷却システムの制御を適応的に実行する“計測融合オペレーション”を可能にする技術の構築がきわめて重要である。

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と研究体制

東北大学、大阪大学、九州大学。研究体制は、理研等が格子ガスモデルと模擬計算、東北大・阪大が広域連携計算、九大が Open-FOAM 計算を分担。

(2) 共同研究分野

超大規数値計算系応用分野、超大規模データ処理系応用分野、超大規模情報システム関連研究分野

(3) 当公募型共同研究ならではの事項など

拠点が関心をもつ計算機室を実験計測の現場にするとともに、計算機センターの異常時にも過渡変化シミュレーションに基づく適切な操作が可能のように、本計算をクラウドコンピューティングで実現することを目指して、SX-9 の広域連携計算基盤によるシミュレーションにも挑戦している。

3. 研究成果の詳細と当初計画の達成状況

(1) 定常状態シミュレーションの妥当性確認

(i) 計算体系の仮定

理化学研究所計算科学研究機構の“京コンピュータ”は、2012 年 6 月の完成を目指して神戸ポートアイランドに設置が進められている。既に、それを構成する 864 台全ての計算ラックは、計算機棟 3 階にある計算機室（約 50m×約 60m）への搬入・据え付けを完了している。計算ラック内の各種装置のうち、主記憶メモリボードや電源ユニットなどは、空気によって冷却される仕組みになっており、当該空気は、計算機室階下にある空調機室の空調機から供給される。下図に、計算ラック搬入前の計算機室と空調機室の写真を示す。

計算機室 (3 階)



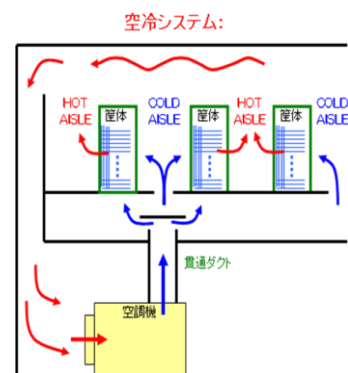
空調機室 (2 階)



計算機棟 2 階の空調機室と 3 階の計算機室の間には 2 重床があり、空調機室の空調機で冷却された空気は、2 重床の下穴につながっているダクトに送られ、2 重床の間の空間に注入される。その後、2 重床の上穴から 3 階の計算ラック配置スペースに吹き出される。(cf. 下図参照のこと。)

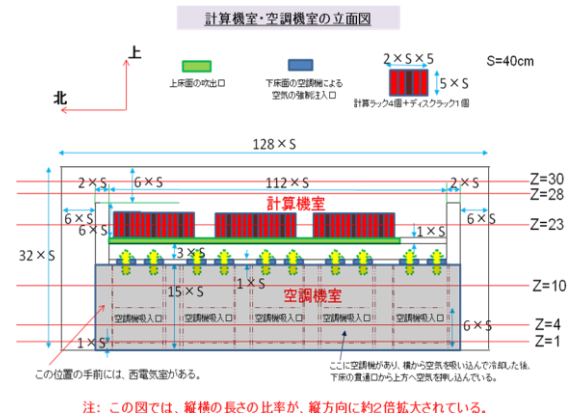
吹き出された冷却空気は、計算ラック内にあるファンの働きによって計算ラック前面から吸いこまれ、内部のメモリーボードや電源ユニットなどを冷却して、

自らの温度を上昇させ、計算ラックの背面から反対側に排出される。計算ラックは、互いに背中を向かい合わせるように配置されているため、暖め

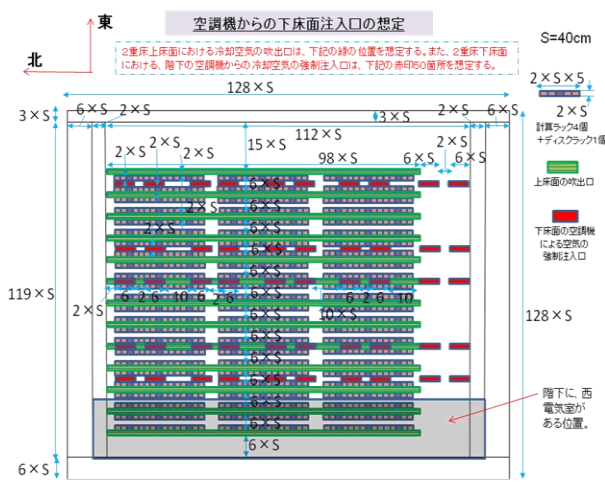
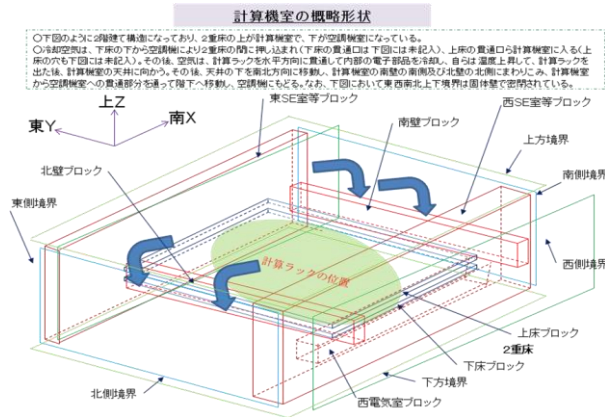
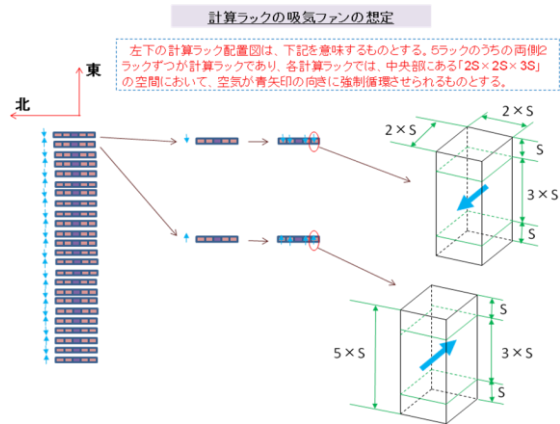


られた空気は、計算ラックの背中どうしが囲む“ホットアイル”と呼ばれる空間に出て、計算機室天井方向に上昇する。上昇した空気は、2階にある空調機の吸引力に引かれて、計算機室の上部空間を南北方向に水平に移動し、南北の両端にある空気の下降路を通過して2階の空調機室に下がり、それぞれの空調機の吸気口に到達して一巡する。

冷却空気の風速場を数値シミュレーションで求めるためには、上に示した空気の流動経路を境界条件として入力する必要がある。しかしながら、実際には、例えば、計算機室下の2重床の間には、計算ラック中のCPU等を冷やすための冷水供給配管、信号通信ケーブル、電源ライン、ケーブルトレイなどが存在する。大規模な体系では、これらの入力が非常に大変な作業になる。このため、風速場の概略を知ることが目的に、近似計算として、下図のような簡略化した体系を仮定した。すなわち、2重床内部は何もない空間であるとし、3階計算機室には直方体の計算ラックとディスクラックのみが縦横に整然と配列されているものとした。



また、前述のとおり、空気の流動経路の途中には、空気を加速する空調機と計算ラック内のファンがある。前者については2重床下面の穴の部分に、後者については下図に示すとおり計算ラック中央部分の直方体領域に仮想粒子を加速する粒子衝突規則を適用した。



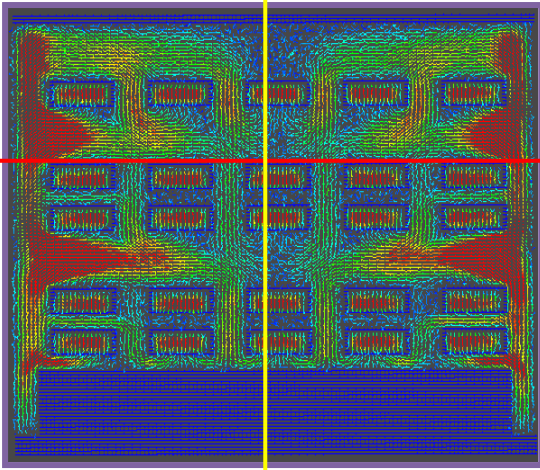
(ii) 格子ガス法超並列計算手法

流体シミュレーション計算には、格子ガス法“面心超立方体 (FCHC) モデル”を使用した。このモデルの詳細は、JHPCN 平成 22 年度共同研究最終報告書(2011 年 5 月)「計測融合オペレーション実現のための大規模計算機空気冷却風速場の高解像度過渡変化解析」及び 東北大学サイバーサイエンスセンター大規模科学計算システム広報 SENAC Vol. 44 No. 3(2011 年 7 月)pp21-28「“多段 2 体衝突法”による流体粘性の制御効果」に記載している。

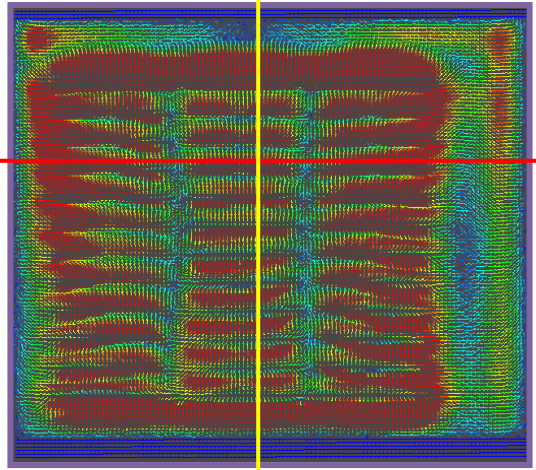
(iii) 定常状態のシミュレーション結果

計算機室に全計算ラック 864 台と全ディスクラック 216 台を配置し、空調機室の全空調機 50 台を運転するという仮想的な状況のもとで、定常状態のシミュレーション計算を行った。(cf. 次ページ)

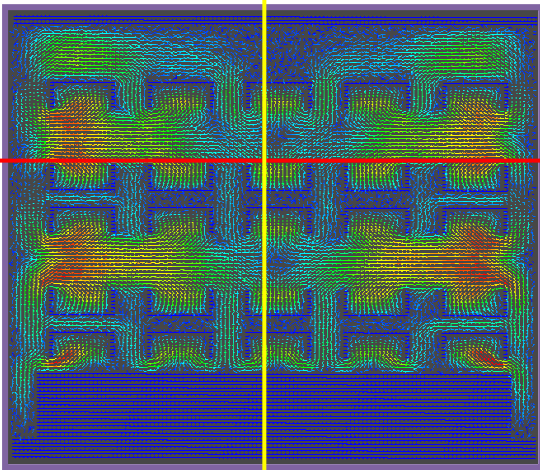
空調機室床上 4.0m 水平面 (Z=10) 上の風速場



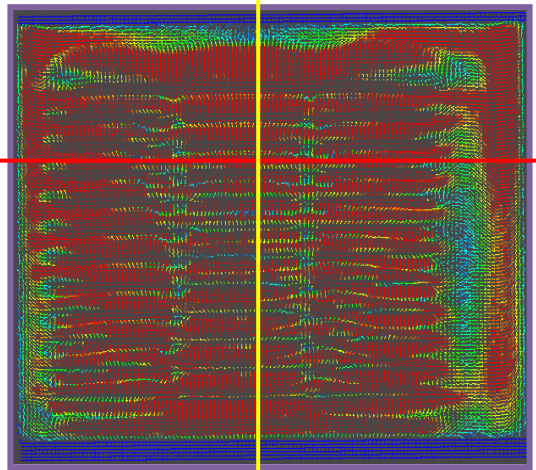
計算機室床上 4.0m 水平面 (Z=30) における風速場



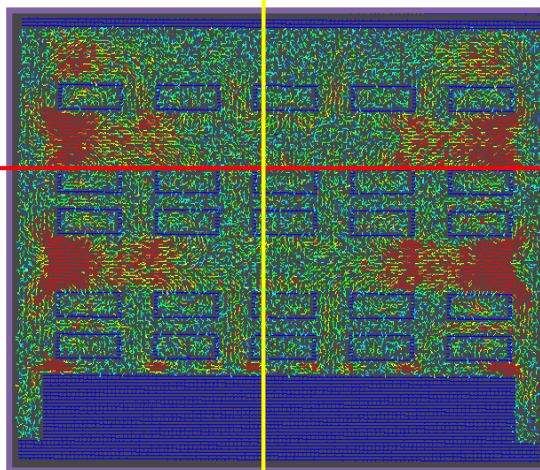
空調機室床上 1.6m 水平面 (Z=4) 上の風速場



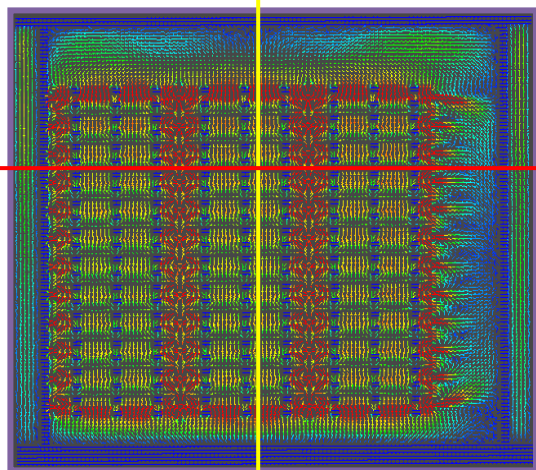
計算機室床上 3.2m 水平面 (Z=28) における風速場



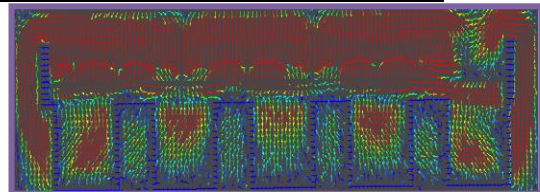
空調機室床上 40cm 水平面 (Z=1) 上の風速場



計算機床上 1.2m 水平面 (Z=23) における風速場

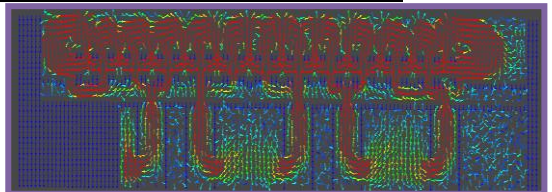


上記赤線位置 (Y=85) 鉛直面上の風速場



(↑ 上半分が計算機室、下半分が空調機室)

上記黄線位置 (X=60) 鉛直面上の風速場



(↑ 上半分が計算機室、下半分が空調機室)

格子ガス法計算に用いた 3 次元微細格子点の数は、1024(南北方向)×1024(東西方向)×256(上下方向)=2.6 億個であり、これを 8×8×8 個の 3 次元微細格子点ごとに合計して 128×128×32 個の運動量ベクトルを算出している。前ページに示した計算結果では、これをいろいろな水平面上または鉛直面上の 2 次元断面上で表示しており、図の色は、運動量ベクトルの大きさを、小さい順に、青色→空色→黄緑色→黄色→赤色と変化する色で表現している。

前記(1)(i)で述べた空気流動経路が、定性的には正しく反映された結果になっている。また、現時点における実際の計算機室及び空調機室の状況は、本シミュレーションの計算条件と一致はしていないがある程度近い。少なくとも現時点の空調機室における風速実測との比較では、似たような傾向の風速分布になっていることを確認できた。

(iv) Open-FOAM による計算との比較

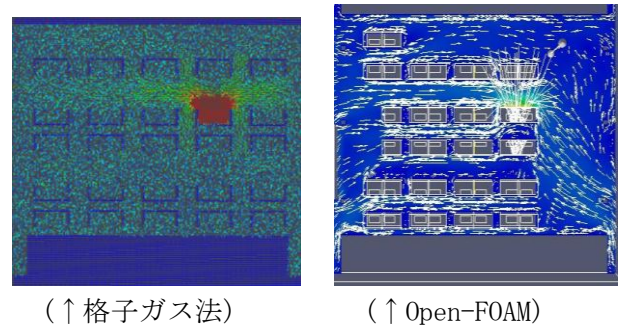
ナビエ・ストークス方程式に基づく通常の流体シミュレーション手法による計算と比較するため公開コードである Open-FOAM による計算を行った。大規模複雑形状の入力や格子生成に長時間を要するため、上記(iii)より簡単な計算条件として、計算機室にラックをひとつも置かず、空調機は、下図の 2 か所のみを運転するものと仮定した。



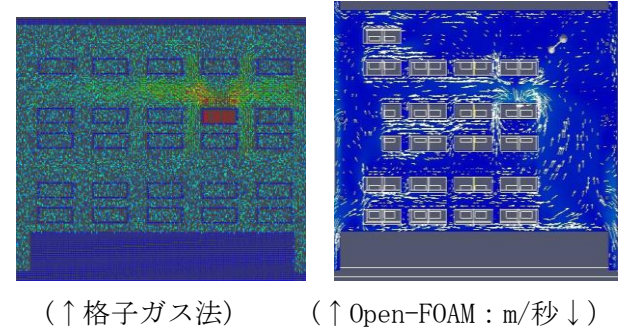
右ページに空調機室の床上①1.6mと②4.0mの水平面における風速場、及び計算機室の床上③1.2mと④3.2mの水平面における風速場の計算結果を示す。左が格子ガス法、右が Open-FOAM によるものである。

①及び②の風速場の比較では、空調機の入り口付近の流速場は似ているが遠方は定かではない。

①空調機室床上 1.6m 水平面 (Z=4) 上の風速場

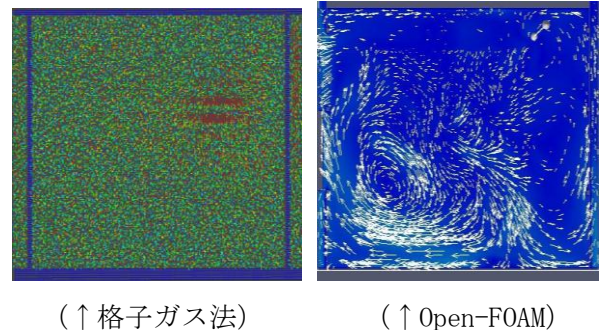


②空調機室床上 4.0m 水平面 (Z=10) 上の風速場

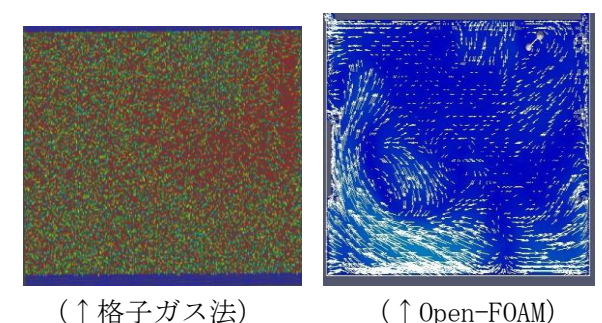


また、③及び④の風速場の比較では、流れ場の様子かなり異なり、これは入力した境界条件の差のみでは説明できそうにない。発熱条件は、ラックなしのゼロ発熱で差はないため、空調機吹き出し量と流体粘性の差だと思われる。

③計算機室床上 1.2m 水平面 (Z=23) での風速場



④計算機室床上 3.2m 水平面 (Z=28) での風速場



以上のことから、当初計画との比較では、格子ガス法と Open-FOAM による計算比較は実行できたが、妥当性確認までは達成できなかった。

(2) 過渡変化シミュレーションの実時間計算達成評価

格子ガス法計算手法を用いて、それまで静止していた計算ラック内の冷却ファンと空調機が、同時に時刻ゼロで運転を開始するという仮想的な初期条件のもとに、計算機室及び空調機室の風速場が定常状態に達するまでの過渡変化シミュレーションを行った。なお、本計算では、主に、東北大学サイバーサイエンスセンターの SX-9 の 2 ノード(32CPU)を利用した。

本ページ右側及び次ページの左側の図では、1024(南北方向)×1024(東西方向)×256(上下方向)=2.6 億個の 3 次元微細格子点を設定し、1024 タイムステップ進むごとに、8×8×8 個の 3 次元微細格子点ずつを合計して 128×128×32 個の運動量ベクトルをひとつのスナップショットとして算出し、はじめの 4 つのスナップショットを描いた。それぞれ、空調機室の床上①1.2m と②4.0m の水平面における風速場の時間変化、及び計算機室の床上③1.2m と④4.0m の水平面における風速場の時間変化を表している。

時刻ゼロで空調機が起動し、静止していた空気が循環しはじめる。次のスナップショット時点では、空気の流れが、既に最終的な定常状態にかなり近い状態になっている。そして、3 回目と 4 回目のスナップショットになると、風速場の間にはほとんど差が見られず、定常状態に到達した。

他方、計算時間は、SX-9 の 2 ノード(32CPU)で約 5 分/スナップショットだったので、運転している空調機の位置や台数の変更による定常状態間移行は、数分～10 分間のオーダーだと推定できる。

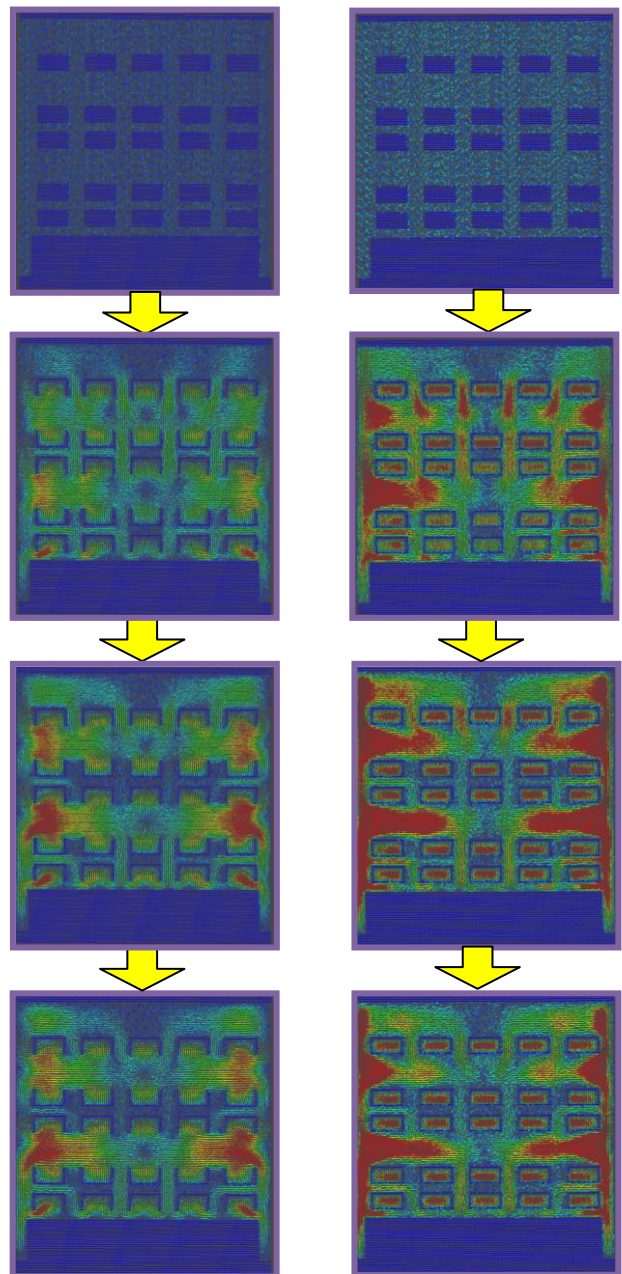
この計算時間は、現実の計算機の冷却空気循環システムが同様の定常状態間を移行するのに要する時間の実測値とほぼ同じオーダーであった。すなわち、SX-9 の 2 ノードの計算能力=約 3Tflops に余裕をみて 10Tflops あれば、空気冷却シミュレ

ーションを実時間で実行できると概略評価できる。

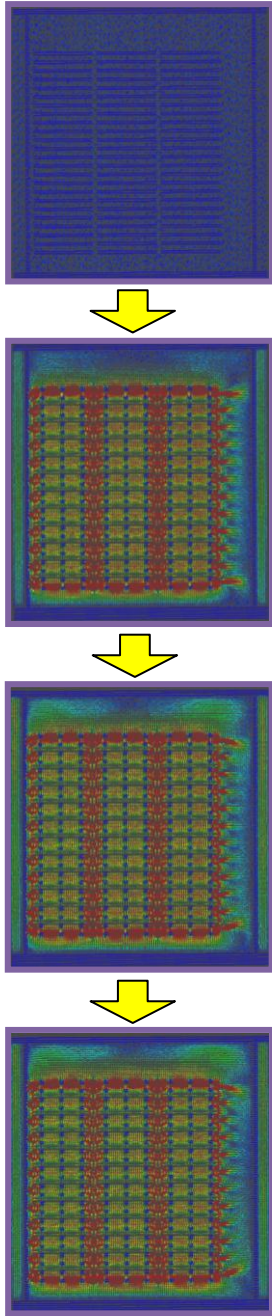
(注：ここで、CPU の論理演算速度は、浮動小数点表示による加減乗除演算の速度に大体比例するものと仮定。) また、10Tflops という値は、“京コンピュータ”の場合、計算機全体システム(ピーク性能で約 11Pflops)の 1/1000 程度、すなわち、計算ラック 1 台を計算機空気冷却風速場の運用制御専用に使えば確保可能な能力である。従って、費用対効果の面からも、空気冷却の実時間シミュレーションは十分実現可能であろう。

①空調機室床上 1.2m
平面上の風速過渡変化
(1024 タイムステップごと)

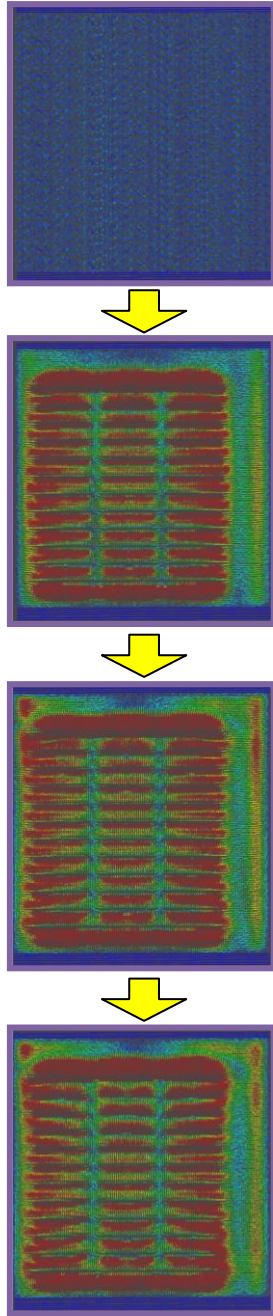
②空調機室床上 4.0m
平面上の風速過渡変化
(1024 タイムステップごと)



③ 計算機室床上 1.2m
 平面上の風速過渡変化
 (1024 タイムステップごと)



④ 計算機室床上 4.0m
 平面上の風速過渡変化
 (1024 タイムステップごと)



以上のことから、当初計画との比較では、実時間シミュレーションの達成に必要な計算機能力の推定評価を達成できた。

(3) 計算機の広域連携環境による高速計算性能評価

計算機センターの異常時にも過渡変化シミュレーションに基づく適切なオペレーション対応が可能になることを目指して、格子ガス法計算手法に

よる円柱後流渦の流体計算を計算機の広域連携環境で実行した場合に要する計算時間を推定した。

具体的には、東北大学サイバーサイエンスセンター、及び同センター担当の日本電気スタッフの協力を得て、SX-9、2 ノードを **ケース 1** として専用高速ノード間スイッチ (IXS 64GB/秒×双方向) で、**ケース 2** として Gigabit Ethernet で接続し、3 次元格子ガス法流体シミュレーションプログラムを実行し計算時間を測定した。(注: 並列計算は、OpenMP で 8 並列×MPI で 2 並列×2 ノード=32CPU)

○ **ケース 1 の実測: (右図)**

(SX 演算時間 + IXS 通信時間)

ジョブ完了時間 = 1235 秒



○ **ケース 2 の実測: (右図)**

(SX 演算時間 + GbE 通信時間)

ジョブ完了時間 = 2520 秒



以上の結果を利用して、東北大学サイバーサイエンスセンターの SX-9 と大阪大学サイバーメディアセンターの SX-9 を JGN2 (東北大と阪大の通信速度は、ほぼ 8.9MB/秒) で接続して同じ計算を実行した場合における SX 演算時間と JGN2 通信時間の推定を行い、

演算時間が約 20 分、

通信時間が約 190

分との結果を得た。



すなわち、演算と通信の時間比は 2:19 であった。

以上のことから、当初計画との比較では、広域連携環境によるジョブ実行時間の把握結果から、

「通信時間を演算時間に隠せるように通信頻度を減らす必要があり、例えば、隣接領域の境界データを複数層まとめて通信するなどの方法を試していくべき」など今後の課題を抽出できた。

(4) 大規模計算機の空気冷却に関する計測融合オペレーションシステムの考案

風速計等のセンサーから得られる時系列計測データをもとに格子ガス法計算モデルを時々刻々補正することで、センサーの設置されていない場所も含めた全体の風速場について、現実に近いシミュレーション結果を得て、これを実際の空調機等

の操作に適応的に反映させる計測融合オペレーションシステムを実現したい。このために必要な実時間シミュレーションの実現方法を検討した。

前記 3. (2) に示したとおり、「格子ガス法 FCHC モデルを利用して、1024(南北方向)×1024(東西方向)×256(上下方向)=2.6 億個の 3 次元微細格子点を設定し、8×8×8 個の 3 次元微細格子点ずつを合計して 128×128×32 個の運動量ベクトルをひとつのスナップショットとして算出していく」という方法を用いれば、10Tflops 級のスパコン、すなわち、“京コンピュータ”の 1 ラック程度の計算機を用いることで実時間計算が可能になる。しかし、3 (1) (iv) に述べたとおり、まずは、非熱計算シミュレーションの範囲で、従来のナビエ・ストークス法との一致を確かなものにしておきたい。解決策は、格子ガス法による高レイノルズ数流体シミュレーションの実現である。次に、計算ラックにおける発熱や空調機における冷却を考慮できる熱計算を適切に実行したい。格子ガス法 FCHC モデルは、仮想粒子の速さを 1 種類しか仮定しないので発熱を伴う問題は通常扱えないとされている。

上記 2 つの問題は、1989 年に S. Chen と K. Molvig が提案した“多速さ LGA モデル”、及び、その後の C. M. Teixeira らの研究成果を利用して解決できると考える。彼らは、3 次元空間において仮想粒子の速さを 3 種類仮定するモデルにより運動量保存則を、4 種類仮定するモデルによりマッハ数 0.5 以下の範囲でエネルギー保存則を、5 種類仮定するモデルにより全ての保存則とナビエ・ストークス方程式までも正しく再現できることを示したとされている。(cf. C.M. Teixeira, Continuum Limit of Lattice Gas Fluid Dynamics, Ph. D. Thesis of MIT, 1992) (cf. H. Chen, C. Teixeira and K. Molvig, Int. J. Mod. Phys. C, 8, 675-684, 1997) また、“多速さ LGA モデル”を利用して、100 万オーダーのレイノルズ数まで達成できたとの報告がある。今後、大規模問題で“多速さ LGA モデル”を実証することが大切である。

なお、当チームがこれまで開発してきた格子ガス法超並列計算法では、繰り返し計算部分で if 文

を一切使用せず、ビット演算の完全なるパイプライン処理が可能である。“多速さ LGA モデル”への変更は、仮想粒子のもつ速度の種類数の増加と衝突規則の変更のみで対応できるので、計算速度を大きく遅らせる要因は特にないと考えられる。

さらに、格子ガス法モデルを全く変えなくても、計測データのフィードバックにより、適切なタイミングで適切な場所に適切な運動量やエネルギーを注入することで、現実に近いシミュレーション結果を得ることができると考えられる。

以上のことから、当初計画との比較では、実時間シミュレーションの実現方法を具体的に示すことで計測融合オペレーションシステムの考案を達成できたと考える。

4. 今後の展望

これまでの研究成果により、計算速度という点では、大規模計算機空気冷却風速場の実時間解析を実現する目途がついた。今後は、3. (4) に示した計算モデルの改良を行うとともに、必要な場所における風速等の実測データをフィードバックして現実の風速場全体を正確に把握する計測融合オペレーション基盤技術の実証を目指すこととしたい。

5. 研究成果リスト

- | | |
|--------------------------|------|
| (1) <u>学術論文</u> | : なし |
| (2) <u>国際会議プロシーディングス</u> | : なし |
| (3) <u>国際会議発表</u> | : なし |
| (4) <u>国内会議発表</u> | |

①松岡ほか，“計測融合オペレーション実現のための大規模計算機空気冷却風速場の実時間解析”，JHPCN 拠点第 3 回シンポジウムポスターセッション，July 14, 2011.

②峯尾，“計算機設備の最適運転のためのシミュレーション技術活用の試み”，MATLAB EXPO 2011, Nov. 2, 2011.

③峯尾，“OpenFOAM による「京」計算機施設のシミュレーション”，OpenCAE シンポジウム, Dec. 2, 2011.

(5) その他 (特許, プレス発表, 著書等): なし