

「京」の持つ8万ノード・64万コアのすべてを効 率よく利用するためには、ユーザはシステムに導入 された新機能や強化された機能を十分に発揮・活用 し、加えてアプリケーションに内在する固有の超並 列性を最大限に引き出さなければならず、そこには 高度な利用技術が必要とされる. 我々はプロジェク トの当初から、アプリケーションを用いたシステム の性能の検証と実証を行ってきた. 本稿では、その 経験から得られたアプリケーションの最適化とその 性能について具体的な事例を交えて解説する.

実証アプリケーションの選定

10101

特集スーパーコンピュータ「京」

「京」のアプリケーション実行性能の検証と実証 を目的に、スーパーコンピュータにおいて高い性 能での動作実績を持つ6本のアプリケーションを さまざまな応用分野から選定した.地球科学分野 のNICAM¹⁾, Seism3D^{2), 3)}, ナノ分野のPHASE⁴⁾, RSDFT⁵⁾, 工学分野のFrontFlow/blue⁶⁾, 物理分野 のLatticeQCD⁷⁾である.これらは, 表 -1 に示すよ うな要求 B/F 値 (プログラムが要求するメモリバ ンド幅と浮動小数点演算速度の比)や並列化の実装 手法などの計算機科学的特性の違いが考慮され, 偏 りがないように選ばれている.たとえば, NICAM, Seism3D, FrontFlow/blue, LatticeQCDの要求 B/ F 値は高いが, PHASE, RSDFT は低い.NICAM, Seism3D, LatticeQCD は隣接通信が中心であるが, FrontFlow/blue, PHASE, RSDFT は大域通信が中心 となる.このような6本のアプリケーションの特徴 にあわせて,プロダクション・ランも見据えながら, 後述するさまざまな改良が加えられた.

| アプリケー | - 分野 | 概要 | 物理モデル/手法 | 計算機科学的特徴 | | | | | |
|--------------------|------|--|---------------------|--|----------|-------------|-------------|-----|-------------|
| | | | | 演算 | | | 通信 | | |
| ション名 | | | | 演算量 | B/F 比 | メモリ アクセス | タイプ | 回数 | データ 量 |
| NICAM | 地球科学 | 全球解像大気大循環シミュ レーション | 大気大循環 有限差分法 | <i>O</i> (N) N : メッシュ数 | 高 | 連続 | 隣接 | 少ない | 小 |
| Seism3D | 地球科学 | 地震伝播・強振動シミュレ ーション | 地震波動 有限差分法 | <i>O</i> (N) N:メッシュ数 | 高 | 連続 | 隣接 | 少ない | 小 |
| PHASE | ナノ | 平面波展開第一原理分子動 力学解析 | 密度汎関数法 波数空間法 | <i>O</i> (N ³) N : 原子数 | 低 | 連続 | 大域 全対全含む | 多い | 大 |
| RSDFT | ナノ | 実空間第一原理分子動力学 解析 | 密度汎関数法 実空間差分法 | <i>O</i> (N ³) N : 原子数 | 低 | 連続 | 大域 | 多い | 大 |
| FrontFlow/ blue | 工学 | Large Eddy Simulation (LES) による非定常流解析 | 非圧縮非定常流体 有限要素法 | <i>O</i> (N) N:メッシュ数 | 高 | リスト | 大域 | 少ない | 小 |
| LatticeQCD | 物理 | LatticeQCD シミュレーション による素粒子・原子核解析 | LatticeQCD 経路積分法 | <i>O</i> (<i>L⁵/a⁷</i>) <i>L</i> :格子サイズ a:格子間隔 | 高 | 連続 | 隣接 | 多い | <u>را</u> ر |

表-1 アプリケーションの概要と計算機科学的特徴

特集スーパーコンピュータ「京」

最適化の手順

プログラムの性能はさまざまな要因によって決ま るため、それらが混在した状態では最適化の効果を 正しく評価できない.そのため問題の分離を狙い、 以下で述べる(A)~(F)の手順で最適化作業を 実施した.また、我々は「京」のシステム開発と並 行してアプリケーションの最適化を実施する必要が あったため、対象をより限定して模擬できる計算機 (たとえば、ノード単体性能は京試作機、並列性能 ならば可能な限りノード数の大きな並列計算機)を 使って開発を進める方法を採用した.

(A)並列特性分析

ソースコードの反復・分岐等の論理構造を静的に 解析し^{☆1}実行時間の測定結果をもとに、プログラ ムから処理ブロック(計算ブロックと通信ブロック) を抽出する.計算ブロックごとに、非並列/完全並 列/部分並列等の計算特性、計算量の問題規模(N) に対する依存特性(たとえば、Nに比例、N²に比例 等)を把握する.通信ブロックごとには、隣接通信 (隣のプロセスとの1対1通信)、大域通信(全プロ セスもしくは一部のプロセス間の集団通信)等の通 信特性,通信データ量の問題規模に対する依存性(た とえば、分割された領域の隣接面に比例、領域体積 に比例等)を把握する.また、処理ブロックごとに スケーラビリティを調査し、高並列時の挙動を分析 する.

(B) 計算・通信カーネル評価

(A) で得た情報をもとに、計算・通信ブロック 群を同種の計算と通信に分類する.それらの中から ターゲット問題実行時に主要部となるブロックを計 算カーネル、通信カーネルとして洗い出す.

(C) ノード単体性能向上

計算カーネルをテストプログラムとして切り出し, 要求 B/F 値とキャッシュの利用効率から推定性能の 見通しをたてる.その上で配列次元の入れ替え,配 列融合,ループの組替え等さまざまな性能向上策を

☆1 プログラムの論理構造を静的に解析し最適化を支援するツール K-scopeを開発した. 2012 年度中の公開を予定している. 試行し、それらの効果を評価する.

(D)高並列化

(A) の高並列時の挙動の評価結果に基づき,通 信カーネルの高並列最適化を試行し,その効果を評 価する.大規模なターゲット問題を想定した次の4 点の評価が重要である.(1)非並列部の有無とそ のコスト,(2)プロセス間のロードバランス,(3) 隣接通信のコスト,(4)大域通信のコスト,である. 演算ブロックごとの計算特性,通信ブロックごとの 通信特性から評価モデルを作成し,演算および通信 の挙動を予測・評価する.

0101070

(E) 高性能化策の実装

(C) と(D)の試行結果を合わせて実装し、並列計算機で評価する.

(F)「京」での最適化

(E) まで実施されたプログラムの「京」の上での性能を確認し、さらに最適化を施す.実機での測定結果と(C)(D)で実施した評価結果と比較し、著しい差異がある場合はその原因を究明し、対策を実施する.

NICAM

●アプリケーション概要

NICAM (Nohydrostatic ICosahedral Atomospheric Model) は、正二十面体非静力学大気モデルを採用 した、全球を対象とする大気大循環モデルの1つで ある.地球シミュレータ (ES) 上で開発されてき た経緯により、ベクトルプロセッサを意識しており、 ES において最高性能が得られるようにコーディン グされている.その一方で、T2K-Tsukuba、T2K東 大、Cray XT4、RICC等の異なるアーキテクチャで の大規模実行の動作実績があり、ポータビリティの 高いソースコードとなっている.

ノード単体性能向上

ノード単体性能の最適化では,流体計算を行う 「力学過程」と外力項を計算する「物理過程」を把 握した上でカーネルの抽出を行った.力学過程から 6 アプリケーションの性能と最適化-アプリケーションによるペタ FLOPS 性能の実証-

は、水平発散項(divergence)、水平拡散項(diffusion)、 鉛直方向の音波・重力波の陰解法(divdamp3d)、 単調性を保存したリミッタ付き風上差分の発散 項(divergence2_rev)、水平勾配項(gradient)、鉛直 方向のフラックス収束計算(src_flux_convergence)、 物理過程からは雲微物理(NSW6)についてカーネ ルを抽出した.

NICAM の力学過程カーネルの特徴を整理すると, (1)最内ループのループ長が長く、ベクトルパイプ ラインを意識したコーディングであり、メモリバン ド幅を要求すること、(2)ループ内に出現する配 列数が多い一方、キャッシュが効く配列もあること、 が挙げられる.物理過程カーネルの特徴を整理する と、(3)1つのループボディの中に大量のスカラ変 数が中間変数として使われておりレジスタスピル等 による演算待ちが発生しやすいこと、が挙げられる. NICAM 全体の特徴は、(4)実行時に問題サイズが 与えられるため、コンパイル時にループ長が不明な こと、(5)ソースコードが巨大でボトルネックが比 較的満遍なく分布しているため最適化個所が多いこ と、(6)IF文による実行処理の分岐が多いコーデ ィングであることが挙げられる.

(1),(2) については、配列添字の入れ替えと配
列融合を軸としたデータのプリフェッチとキャッシュのラインアクセスの有効利用,(3) については、コンパイラの改善を前提とした効率の良いスケジューリング,(4),(5) については、コンパイラが提供する自動並列化とディレクティブの有効利用,(6) については、マスク付き SIMD (Single Instruction Multiple Data) またはリストベクトルの 採用,等を検討した.

結果, divergence では実行効率が2.0%から 16.7%まで向上した.その他のカーネルについても おおむね10~15%に達した.

約8万ノードを用いる最終ターゲット問題では, 初期条件,境界条件等の入力ファイルが 28TB に達

☆2 1ノードで実行する問題規模を一定にして並列数を増やす方法. 一方,全体の問題規模を一定にして並列を増やす方法をストロン グ・スケーリングと言う. する.現在はウィーク・スケーリング^{☆2}の測定と 評価を進めている.

Seism3D

●アプリケーション概要

Seism3Dは、有限差分法を用いて数値的に粘弾性 方程式を時間発展させることにより、地震伝播と津 波を連動して解く、大規模な並列化に対応したプロ グラムである. Seism3Dは以下の6つの計算カーネ ルから構成される.

- (a) 応力空間微分計算
- (b) 速度空間微分計算
- (c) 応力時間積分計算
- (d) 応力時間積分吸収計算
- (e) 速度時間積分計算
- (f) 速度時間積分吸収計算

●ノード単体性能向上

Seism3Dの主要な6つの計算カーネルは,メモリ バンド幅を要求し,要求B/F値が高いことが特徴 として挙げられる.「京」の実効メモリバンド幅の 実測値が46GB/sであり,ハードウェアのB/F値が 0.36であることを踏まえると,プログラムの要求 B/F値から各カーネルの性能値を推定できる.ノー ド単体性能は,(1)各カーネルが推定性能値まで達 しているか,(2)推定性能値に達していない原因は 何か,(3)さらに性能向上は可能か,の3つのス テップを踏むことにより検証と最適化を進めた.

応力,速度積分部は,応力成分,速度成分がスト リーム配列(再利用性のない配列)であり,最内に 10 配列程度のアクセスが発生する.その場合,L1 データキャッシュでのキャッシュ競合が発生しやす い状況であるため,推定値に達していないケースが 散見された.このようなループの場合はループ分割 や配列融合など,ストリーム配列の局所的なアクセ ス数を減らすことにより推定性能値まで性能を改善 できた.

微分計算部は各軸方向の差分計算が発生するため,

特集スーパーコンピュータ「気」

010100

| $\frac{1}{OMP}$ DO SCHEDULE(static,1) do J = 1, NY | |
|---|----|
| do $K = 3$, NZ-1 | |
| DYV (k,I,J) = (V(k,I,J) -V(k,I,J-1))*R40 & - (V(k,I,J+1)-V(k,I,J-2))*R41 | ম |
| end do end do | Se |
| end do | 度割 |

』-1 eism3D y方向速 ξ差分項(cyclic 分 リスレッド並列化)

キャッシュの利用効率が各差分軸によって変化する. この場合,ループ融合により元の複数ループよりも 相対的に要求 B/F 値を下げることが可能となり,性 能を改善できた.また,要求 B/F 値を下げる他の手 法には,図-1 のような OpenMP ディレクティブに よる cyclic 分割を指定する方法がある⁸⁾.これによ り最外軸において 1 つのスレッドが L2 キャッシュ にロードしたデータを他のスレッドが再利用可能と なる.要求 B/F 値が 4.8 から 2.4 まで下がった結果, 実行効率は 7.5% から 13.6% まで向上した.本手法 はコア間で L2 キャッシュを共有する「京」の特性 を活用している.

●高並列化

Seism3Dは「京」への移植に伴い,地震動に加え て津波のシミュレーションに対応する改良が行われ ている.その改良に伴い,Seism3Dは3次元分割か ら水平方向のみの2次元分割に変更されたため,通 信部は隣接4ノードにおける境界データの交換の みとなった.2次元分割モデルをTofuインターコ ネクト上にマッピングした場合,1ホップによる隣 接通信が保証され,ウィーク・スケーリングによる 測定でも通信コストの増加は見られなかった.現在, 5PFLOPSの規模においてもプログラム全体の実行 効率は17%を超えており,要求 B/F値が高いプロ グラムでも「京」で高い実行効率を確保できた事例 である.

PHASE

●アプリケーション概要

PHASE は、擬ポテンシャルと密度汎関数法による

ナノ材料第一原理分子動力学プログラムである.平 面波基底を用いることで、多様な物質に対して高精 度な電子状態計算が可能である.なかでも、結晶・ アモルファスなどの解析を得意とし、ミクロな観点 から量子論効果を厳密に解析し、新規材料開発につ ながる計算を目指している.「京」でターゲットとす る計算対象は、半導体のデバイス特性や燃料電池の 輸送特性など、新材料やエネルギー問題が挙げられる.

010107

●ノード単体性能向上

電子状態計算とは、離散化されたハミルトニアン 演算子に対する固有値問題である.その大部分は演 算密度が高く、行列-行列積(DGEMM)の形に書 きくだすことで高効率が見込める.直交化について も依存関係を分離することで、DGEMMへの書換 えが可能となり、高性能が期待できる.

●高並列化

PHASE のオリジナル並列手法は,エネルギーバ ンド並列を採用しており,一部の処理では波数空間 並列を用いていた.この両並列手法間には,データ 持ち替えに伴う全体通信が発生し,並列特性分析に より,「京」の持つ8万ノードという高並列性への 対応が不可能であることが分かった.したがって, 最適化における最大の作業は,両並列手法間の全体 通信を削減するために,全面的にエネルギーバンド と波数空間の二軸並列を実装することであった.

PHASE では、電子の状態を波数空間でフーリエ級 数展開することで離散化するため FFT (Fast Fourier Transform)を用いる.一般に FFT は、転置転送に伴 う全体通信のため高並列化が困難とされる.上記の 二軸並列を採用すると、波数方向に新たな通信は発 生するが、転置転送に伴う全体通信は、小規模な部 分空間内にパックされ、多重実行も可能となる.大 規模計算では、FFT の演算量は DGEMM に比べて 増大しないため、通信性能が向上すれば、相対的 に FFT の比率が下がり、並列性能の向上が見込める. PHASE の最適化では、これらの利点と欠点を総合し て高性能が得られるかがポイントであった.性能見

B18191

6 アプリケーションの性能と最適化-アプリケーションによるペタ FLOPS 性能の実証-



図 -2 PHASE の並列性能(SiC 3800 原子 SCF1 回)

積りの結果,二軸並列の方が高い性能が見込めると 判断し作業を実施した.

結果として、パワーデバイスとして用いられるシ リコンカーバイド (SiC) 16,384 原子の螺旋転移の 電子状態の計算を行い、12,288 並列では 28.8% の 実行効率、36,864 並列では 21.3% もの実行効率を 達成した.二軸並列による FFT の影響も、性能悪 化要因となっていないことが確認できた(図-2).

また当初の見積り通り,二軸並列の効果として並 列数の上限が拡大した.1万原子ほどの電子状態計 算を電子数より遙かに多い「京」の全ノードを用い ての計算が可能となり,計算科学的に画期的な成果 が得られた.

RSDFT

●アプリケーション概要

RSDFT (Real Space Density Functional Theory code) は密度汎関数法による電子状態計算を行うプ ログラムである.ナノスケールでの量子論的諸現象 を第一原理に立脚して解明し,新機能を有するナノ 物資・構造の予測を行うことを目的としている.実 空間に格子を導入し各格子上に離散化された物性値 を差分法で計算する.並列処理は,空間格子をいく つかの小領域に分け,それぞれの小領域をそれぞれ のノードに対応させることで実装されている. 演算 の主要部は,前述の PHASE と同様に行列 - 行列積 (DGEMM) で処理される.一方,通信は大域的な ものが多く高並列で通信ボトルネットとなることが 予想される.本章では最適化のポイントになる高並 列化について解説する.

●高並列化

並列特性分析から、オリジナルのプログラムに実 装されている空間並列のみでは、「京」の持つ8万ノ ードの高並列性に対応できないことが分かった.その ため、前述のPHASEと同様に二軸並列を採用し、空 間並列に加えてエネルギーバンド並列を実装した⁹⁾. これによりエネルギーバンドに関する通信が新たに 発生するが、空間並列の全プロセスを対象とする大 域通信が、グルーピングされたプロセス群単位とな り、通信コストを削減できる. RSDFT における最適 化のポイントは、このように増加と減少の両面を持 つ通信コストをいかに最小化できるかである.

通信コストの大きいプログラムの場合、性能を向 上させる一番効果的な方法は MPI (Message passing Interface) 関数の呼び出し回数を削減することであ る. 関数呼び出し回数の削減は複数回の呼び出しを まとめることで実装できるが、それによって1回 あたりの通信データ量は大きくなる. このタイプ の通信は「京」の MPI に実装されている Tofu 向け 集団通信アルゴリズム Trinaryx3^{☆3}を利用できる. RSDFT では、二軸並列、Trinaryx3の適用、さらに Tofu インターコネクトへのプロセスマッピングの 組合せによって、通信コストの大きな削減が見込め る. 二軸並列化されたプロセス群(サブコミュニケ ータ)の形状と並び(データの連続性)を意識して、 プロセス全体のマッピングを決める必要があり、プ ログラムの実行時にマッピング情報をテキストファ イルで与えることで実現している.図-3の例では 二軸並列化された3つのプロセス群(サブコミュ

☆3 Tofu インターコネクトの性能を引き出すため、3 つのネットワー クインタフェースを用いて通信経路の重なりがないように実装さ れたアルゴリズム.

特集スーパーコンピュータ「京」



図 -3 二軸並列化されたプロセスの Tofu インターコネクトへのマッピング

ニケータ)を3つのサブメッシュ/トーラスにマッ ピングしている.これらによって,(1)1プロセス あたりの MPI 関数呼び出し回数が削減される,(2) バンド分割でグルーピングされたプロセス内で通 信が閉じられるため通信の輻輳が起こらない,(3) Trinaryx3 が適用され通信時間が短縮される,等の 効果が得られるので,Allreduce, Bcast, Reduce の通信 時間が 1/5 ~ 1/10 に短縮された.

7PFLOPS 規模の 55,296 ノードを用いて,シリ コンナノワイヤ 107,292 原子(格子数:576×576 × 192, エネルギーバンド数:229,824)の1 反復 計算を測定し,56.3%の実行効率を得た.全実行時 間に対する通信時間の割合は 29.7% であり,通信 コストの削減が大きな効果を発揮した.

FrontFlow/blue

●アプリケーション概要

FrontFlow/blue は非圧縮性流体の非定常流動を高精 度に予測可能な LES(Large Eddy Simulation)に基づ いた汎用流体解析プログラムである.形状適合性に 優れた有限要素法による離散化を採用し,ファンや ポンプ等の流体機械や複雑形状周りの非定常乱流解 析および流れから発生する騒音の予測が可能である.

●ノード単体性能向上

10101070

FrontFlow/blue で採用して いる行列ソルバは BiCGSTAB 法(Bi-Conjugate Gradient Stabilized method)であり、疎 行列とベクトル積計算が大き なコストを占める.FrontFlow/ blueの計算カーネルの課題とし て,(1)ベクトル部はリストア クセスのため、リストの非局所 性により性能が左右されること, (2)行列部は CRS (Compressed Row Storage)形式により記述さ

れており,最内のループ回転数が可変かつ高々30 回転程度と比較的小さいため,コンパイラによるル ープ最適化が難しいことが挙げられる.(1)の解決 方法として,図-4のように解析空間を小領域に分 割し,さらに節点番号を小領域の内側(A),外側(B) の順にリナンバリングすることにより,ベクトルの 局所性を高めた⁸⁾.

(2)の解決方法として、すべての節点の隣接節点 数を0埋めによって最大隣接節点数に合わせること により回転数を一定化し、フルアンローリングによ り回転数不足を解消した.(1),(2)の手法を適用 することにより、本カーネルの要求 B/F 値から求め られる推定性能値 9% に迫る 8.5% の実行効率が確 認された.

●高並列化

FrontFlow/blue の高並列化の際に留意しなければ ならない点は、大域通信である Allreduce の実行時 間が並列数の増加に伴い、どのように推移するかで ある.1要素データの Allreduce の場合、Tofu インタ ーコネクトの持つハードウェアバリア機能が用いら れるため、並列数が増えても高速な通信が可能であ る.FrontFlow/blue では、10PFLOPS 規模において 44マイクロ秒の通信性能が確認できていることか ら、数万ノード規模の高並列実行でも、通信はボト ルネックにならないと考えられる.ただし、基礎性

6 アプリケーションの性能と最適化-アプリケーションによるペタ FLOPS 性能の実証-



図-4 節点番号のリナンバリング

能は 10PFLOPS 構成の場合に 21 マイクロ秒であり, FrontFlow/blue で用いた場合との性能差が確認でき ている.今後はこの性能差の検証を予定している.

LatticeQCD

●アプリケーション概要

LatticeQCDは、素粒子クォークおよびグルオン の基本法則である量子色力学(QCD)を4次元時 空間格子上に定式化した第一原理計算である。今 回用いたLatticeQCDは、領域分割されたハイブリ ッドモンテカルロ(DD-HMC)に基づく手法⁷⁾で LDDHMCと呼ばれる。

●ノード単体性能向上

LatticeQCDでは、モンテカルロ積分において逆 行列を求める部分が計算時間の大部分を占めるため、 最適化は大規模疎行列ソルバの高速化に帰着する.

最適化の流れとしては、ソルバの選択、並列性 と収束性を考慮した前処理の選択を行い、次に SPARC64TMVIIIfx のコア数を意識した OpenMP に よるスレッド並列化、最後に SIMD 演算の最適化 を実施した. なお、オリジナルの LDDHMC の 前処理部分は、インテル・アーキテクチャの SSE (Streaming SIMD Extensions)を用いることで単精 度浮動小数点演算による4命令同時実行が実装され ている. これを SPARC64TMVIIIfx の SIMD 命令に 置き換える作業を行った.

ソルバは, LatticeQCD で一般的な BiCGSTAB 法 を採用している. オリジナルは, SSOR法 (Symmetric Successive Over Relaxation method) が前処理として 採用されており、ノード間は Red-Black (Even-Odd) Ordering によって並列化されている.

最適化は、ノード内の並列性と収束性を考慮し Red-Black Ordering と Block-Natural Ordering につい て検討を行った. Red-Black Ordering は、スレッド 内およびスレッド間にデータ依存がないため並列性 は高いが収束性は悪い.一方、Natural Ordering は、 スレッド内およびスレッド間にデータ依存を残して いるため収束性は良い. Block-Natural Ordering は、 Natural Ordering と Red-Black Ordering の中間的な 方法であり、スレッド内はデータ依存を残しつつも、 スレッド間のデータ依存をなくし並列化を可能とす る方法である.ここではスレッド並列が可能とな る利点と、Natural Ordering に比べた場合、収束性 能の低下がそれほど大きくなかったため、Natural Block Ordering を採用した.

作業の効率とプログラムの見通しを良くするため に、計算主要部分から3つのカーネルを抽出した. カーネル1はSSORの前進代入に関する部分、カー ネル2は後退代入に関する部分、カーネル3は両方 向の代入部分である.なお、ノードあたりの問題サ イズはL2データキャッシュに載るように設定した.

その結果、オリジナルではカーネル全体で15% 程度の実行効率が得られたが、L1 データキャッシ ュミス率が高く, 整数ロードキャッシュアクセス待 ち,バリア同期待ちが大きい.バリア同期待ちにつ いては、アルゴリズム由来のもので、境界部分の計 算量がスレッドごとに偏りがあることが原因である が,L1 データキャッシュミス率については、イン ライン展開された場合に適切なプリフェッチが生成 されていなかったことが要因であった. これについ ては、コンパイラの改善により解決を図った. 整数 ロードキャッシュアクセス待ちについては、不要な インデックス計算を削除することで改善された. 有 効総命令数に対する SIMD 命令数の割合については、 C 言語の SIMD 組込み関数を用いることで 45% 程 度だったものが 75% 程度まで向上した. 最適化の結 果,カーネル全体で実行効率33%まで向上した.

●高並列化

大規模並列としては、5PFLOPS と 10PFLOPS 規 模の測定を行った.それぞれ、通信を含めた主要計 算部分の実行効率は 15%程度で、ウィーク・スケ ーリングしていることが確認された.

まとめ

「京」のアプリケーション性能の実証を目的に、6 本のアプリケーションのノード単体性能の向上と高 並列化を実施してきた. ノード単体性能に関しては、 スカラ型計算機が不得意とする要求 B/F 値が大きい プログラムについても,要求 B/F 値とキャッシュ の利用効率から見積もられる推定性能と同程度の性 能が得られた. さらに、性能向上のためのキャッシ ュの有効利用法等の最適化手法を適用し、Seism3D はトータル性能で17%以上の高い実行効率が得ら れた. PHASE と RSDFT では、「京」の 8 万ノード を活かすために適用した二軸並列化手法等が十分に 機能し、高並列に対応できることが確認された.高 並列時の実行で重要な通信性能については、隣接通 信、大域通信のどちらでも高速な通信が実現できて いることが確認された. ここで述べたことは「京」 が汎用システムとして実際のアプリケーションにお いて十分に高い性能を出すことができることを示し ている. 演算部, 通信部ともに性能を引き出しやす く、「京」の供用開始後に動き出す多くのアプリケ ーションにおいても高い実効性能が得られることが 期待できる.

参考文献

- Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T. and Iga, S. : Nonhydrostatic Icosahedral Atmospheric Model (NICAM) for Global Cloudresolving Simulations., *Journal of Computational Physics, the Special Issue on Predicting Weather*, Climate and Extreme events, 227, pp.3486-3514 (2008).
- Furumura, T. and Chen, L. : Parallel Simulation of Strong Ground Motions during Recent and Historical Damaginge Earthquakes in Tokyo, Japan, *Parallel Computing*, 31, pp.149-165 (2005).
- 古村孝志:差分法による3次元不均質場での地震波伝播の大 規模計算,地震2,61巻,S83-S92 (2009).

- 4) 黒田明義,長谷川幸弘,寺井優晃,井上俊介,市川真一,小 松秀実,大井憲行,安藤琢也,山崎隆浩,大野隆央,南一 生:ナノ材料第一原理分子動力学プログラム PHASE の京速コ ンピュータ「京」上の計算性能最適化,ハイパフォーマンス コンピューティングと計算科学シンポジウム論文集, pp.144-152 (2012).
- 5) Iwata, J., Takahashi, D., Oshiyama, A., Boku, T., Shiraishi, K. and Okada, S. : A Massively-parallel Electronic-structure Calculations based on Real-space Density Functional Theory, *Journal of Computational Physics*, 229, pp.2339-2363 (2010).
- http://www.ciss.iis.u-tokyo.ac.jp/rss21/theme/multi/fluid/fluid_ softwareinfo.html
- 7) Aoki, S., Ishikawa, K.-I., Ishizuka, N., Izubuchi, T., Kadoh, D., Kanaya, K., Kuramashi, Y., Namekawa, Y., Okawa, M., Taniguchi, Y., Ukawa, A., Ukita, N. and Yoshie, T. : 2+1 Flavor Lattice QCD toward the Physical Point, Physical Review D, 79, 034503 (2009).
- 8)南 一生,井上俊介,堤 重信,前田拓人,長谷川幸弘,黒 田明義,寺井優晃,横川三津夫:「京」コンピュータにおける 疎行列とベクトル積の性能チューニングと性能評価,ハイパ フォーマンスコンピューティングと計算科学シンポジウム論 文集,pp.23-31 (2012).
- 9) Hasegawa, Y., Iwata, J. I., Tsuji, M., Takahashi, D., Oshiyama, A., Minami, K., Boku, T., Shoji, F., Uno, A., Kurokawa, M., Inoue, H., Miyoshi, I. and Yokokawa, M.: First-principles Calculations of Electron States of a Silicon Nanowire with 100,000 Atoms on the K Computer. In: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis. pp.1:1-1:11. SC '11, ACM, New York, NY, USA (2011).

(2012年4月27日受付)

0101070

南 一生 minami_kaz@riken.jp

(独) 理化学研究所次世代スーパーコンピュータ開発実施本部開発グ ループアプリケーション開発チームチームリーダー.

■井上俊介 inoue.shunsuke@riken.jp

(独) 理化学研究所次世代スーパーコンピュータ開発実施本部開発グ ループアプリケーション開発チーム所属.開発研究員. Seism3D およ び FrontFlow/blue の最適化に従事.

熊畑 清 kuma@riken.jp

(独) 理化学研究所計算科学研究機構運用技術部運転技術課所属. 研究員.博士(情報科学). FrontFlow/blueの最適化に従事.

黒田明義 kro@riken.jp

(独)理化学研究所次世代スーパーコンピュータ開発実施本部開発 グループアプリケーション開発チーム所属.開発研究員.博士(人間・ 環境学). PHASE の最適化に従事.

■寺井優晃(正会員) teraim@riken.jp

(独)理化学研究所次世代スーパーコンピュータ開発実施本部開発 グループアプリケーション開発チーム所属.開発研究員.博士(情 報科学).NICAM および LatticeQCD の最適化に従事.

■長谷川幸弘 y.hasegawa@riken.jp

(独)理化学研究所次世代スーパーコンピュータ開発実施本部開発 グループアプリケーション開発チーム所属.開発研究員. RSDFT の 最適化に従事.