2007－ARC－174 (7)
2007／8／1

演算加速機構を持つオンチップメモリプロセッサの電力性能評価

高橋睦史[†1] 佐藤三久[†1] 高橋大介[†1]
朴泰祐[†1] 宇川彰[†1] 中村宏[†2,†1]
青木秀貴[†3] 澤本英雄[†4] 助川直伸[†3]

本稿では電力性能の向上に有効であるオンチップメモリプロセッサアーキテクチャSCIMAに，演算あたりのハードウェアおよび電力コストに有利な演算加速機構を導入することとし，その電力性能を評価する．演算加速機構としてベクトル型およびSIMD型の2種の方式を提案し，シミュレーションにより評価を行った結果，DAXPYおよびLivermore kernel 1,3では同じFMA数であれば電力性能に大きな差はなかったが，行列積演算においてはレジスタの要素数の差などの要因によりベクトル型がSIMD型の電力性能を上回った．Livermore kernel 7ではレジスタの再利用ができるSIMD型の電力性能がベクトル型を上回った．今後は電力モデルの詳細な検討や多様なアプリケーションを用いた評価を行う必要がある．

Power performance evaluation of on-chip memory processor with arithmetic accelerators

Chikafumi Takahashi,[†1] Mitsuhisa Sato,[†1]
Daisuke Takahashi,[†1] Taisuke Boku,[†1] Akira Ukawa,[†1]
Hiroshi Nakamura,[†2,†1] Hidetaka Aoki,[†3]
Hideo Sawamoto[†4] and Naonobu Sukegawa[†3]

In this paper, we evaluate power performance of arithmetic accelerators into on-chip memory processor, which is expected to be effective to improve power consumption. We propose vector-type arithmetic accelerators and SIMD-type arithmetic accelerators in to on-chip memory processor. The results of evaluation on our simulator indicates that it is almost same performance between the SIMD-type and the Vector-type on DAXPY, Livermore kernel 1 and 3. However, the vector-type accelerator's performance exceeds the SIMD-type accelerators's on matrix multiplication because of difference of the elements size of registers and so forth. On Livermore kernel 7, the power performance of the SIMD-type accelerators exceeds the vector-type because of register reuses. We needs further investigation about the power model and another simulations on various applications.

1. はじめに

マイクロプロセッサはデバイス技術やアーキテクチャの改良により，その動作周波数を向上させてきた．その結果，現在では3GHzを上回る動作周波数のプロセッサが現れるに至り，その高い動作周波数により高い性能を達成してきた．しかし代償としてプロセッサの消費電力は高くなり，集積度や実装に関して問題があることが明らかになってきた．最近では，比較的周波数を抑えたコアを1チップ内に集積したマルチコアプロセッサが使われるようになり，高性能システムにおいても低電力プロセッサを集積したBlueGene/L[1]が注目されている．

†1 筑波大学 計算科学研究センター
Center for Computational Sciences, University of Tsukuba
†2 東京大学 先端科学技術研究センター
Research Center for Advanced Science and Technology, The University of Tokyo
†3 （株）日立製作所 中央研究所
Central Research Laboratory, Hitachi Ltd.
†4 （株）日立製作所 エンタープライズサーバ事業部
Enterprise Server Division, Hitachi Ltd.

本稿では高性能な演算処理が必要とされる科学技術計算を対象に，オンチップメモリとそれを効率的に活用する演算加速機構により演算当たりの電力効率を高めるプロセッサアーキテクチャについて検討し，その電力性能の評価を行う．

オンチップメモリプロセッサはプロセッサコアと同一チップ上にメモリを持つプロセッサである．チップ外の主記憶にアクセスするには相対的に大きな電力が必要となるが，オンチップメモリを利用することでオンチップメモリ・主記憶間のデータ転送をソフトウェアで効率的に制御でき，電力性能を改善することができる．

他方，電力性能改善のアプローチとしてプロセッサあたりの演算器を増やすことが考えられる．プロセッサコアと比較して構成が単純な演算器を増やすことで単位時間当たりの演算量を増やし，消費電力の増加を最小限に抑えつつ性能を向上させることができる．そのためには効率的に演算器を利用するための機構が必要となる．

さらに，次世代のプロセッサで用いられると想定される45nmプロセスではさらに多くのトランジスタを利用できるため，多数のコアを搭載するマルチコアとし，ク

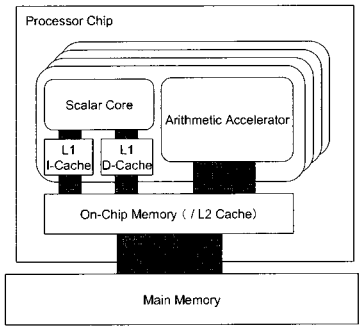

図 1 演算加速機構を持つオンチップメモリプロセッサの基本構成（4 コア時）

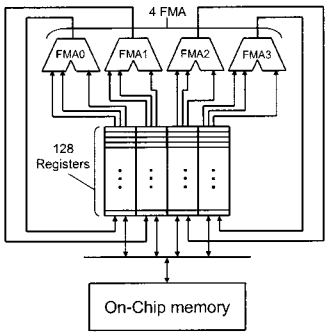

図 2 SIMD 型演算加速機構

ロック周波数を抑えて比較的低電圧で駆動し並列処理を行うことにより電力効率を高めることができる．

我々はすでに演算加速機構を搭載したオンチップメモリプロセッサの構成について検討を行い，シミュレータを用いた予備評価により基本性能に関する問題点と改善方法を考察した[2]．本稿ではそれを踏まえてアーキテクチャを再度検討し，検討したアーキテクチャに対してシミュレータを用いた実効性能と電力性能の評価を行う．評価では，考えうる演算器方式とパラメータに着目してシミュレーションを行い，パラメータが各構成の実効性能と電力性能に与える影響を考察する．

2. 演算加速機構を持つオンチップメモリプロセッサ

本稿では演算加速機構を付加したオンチップメモリプロセッサについての検討と評価を行う．検討するプロセッサの基本構成を**図 1** に示す[2]．

このプロセッサは単層のオンチップメモリと演算加速機構を持つ．演算加速機構は既存のプロセッサコア（スカラコア）に付加される形で実装されるものし，このスカラコアと演算加速機構で 1 つのコアを構成する．このコアをプロセッサチップ内に複数配置し，L2 キャッシュおよびオンチップメモリを共有する．なお，メモリ階層はオンチップメモリと演算加速機構，およびオンチップメモリと主記憶を直接接続する方法を想定する．

以降，本プロセッサの特徴であるオンチップメモリと検討する 2 種類の演算加速機構について，詳細を述べる．

2.1 オンチップメモリ

オンチップメモリを用いたプロセッサアーキテクチャは数多く提案されている．我々は SCIMA と呼ぶアーキテクチャを提案している[3),4)]．我々はこの SCIMA を基本的なオンチップメモリアーキテクチャとして考える．

SCIMA はハイパフォーマンスコンピューティング向けのオンチップメモリプロセッサアーキテクチャであり，キャッシュと同じメモリ階層に，SRAM を用いたオンチップメモリを実装する．オンチップメモリはメモリウェイごとに，すべてのデータ転送がソフトウェア制御可能な一時記憶か従来型のキャッシュとして使い分けることが可能である．

オンチップメモリの利用には以下のような利点がある．

(1) 演算に必要なデータを明示的に指定することで，従来のキャッシュで生じていたような意図しないキャッシュミスや固定されたキャッシュラインサイズでのデータ転送で発生する不要なデータ転送を抑制する．

(2) 演算に必要となるより前にデータをオンチップメモリに転送しておくことでデータプリフェッチが明示的に行える．これはキャッシュプリフェッチとほぼ同等の機能である．しかしキャッシュプリフェッチとは異なり，プリフェッチしたデータが意図せずキャッシュから追い出されることはない．オンチップメモリが主記憶との DMA 転送をサポートする場合はキャッシュプリフェッチと同様にデータ転送時間の隠蔽が可能になる．

(3) (2) によりオンチップメモリにあらかじめデータを転送しておくことで，オンチップメモリへのアクセスに関して，各コア間でアクセスするアドレスがコンフリクトしなければ，アクセスレイテンシが一定であることを保証できる．

(1), (2) の利点により性能が向上し，加えて主記憶へのアクセスを削減することによって電力性能の向上も見込まれる．また, (3) によりアプリケーションプログラムの実装時に，オンチップメモリへのアクセスレイテンシの隠蔽するための最適化が容易になる．

2.2 演算加速機構

2.2.1 SIMD 型演算加速機構

本稿で検討する SIMD 型演算加速機構を**図 2** に示す．この演算加速機構は 4 個の倍精度浮動小数点数積和演算器（Fused Multiply-Add, 以降 FMA）を持ち，128 個の SIMD レジスタを持つ．1SIMD レジスタあたり 4 要素を持ち，演算時には 1 サイクルで 4 要素がそれぞれ対応する演算器へと送られる．

SIMD 型演算加速機構では既研究[2]において，スカラコアでのアドレス演算に起因するストールが発生することと，オンチップメモリへのアクセスレイテンシ隠蔽のために必要なループアンローリングを行う際にレジスタ数が不足することが分かっている．そこで本稿では SIMD 型演算加速機構のロード・ストア命令用にインデックスレジスタを導入し，レジスタ数は従来の 32 本より多い 128 本とする改良を行った．

SIMD 型演算加速機構をオンチップメモリプロセッサに導入した場合は，転送したデータは必ずオンチップメモリに保持されるのでアクセスレイテンシがほぼ一定であることを保証できる．プリロード命令を用いて，演算

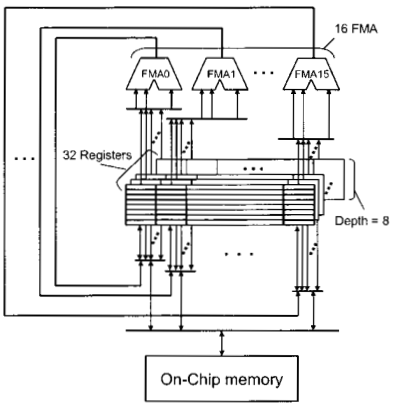

図 3 ベクトル型演算加速機構（16FMA 時）

とオーバーラップする効率的な命令スケジューリングも可能となる．また，実効性能および電力性能はキャッシュを利用する場合よりも向上することが予想される．

SIMD 型演算加速機構の場合，比較的少ない要素を SIMD レジスタにロードして処理を行う．再利用される値をレジスタに保持しておき，最適化を図るレジスタブロッキングやレジスタ内要素間の交換命令を用いるなど細かい最適化が可能である．ただし，連続したベクトルを処理する場合，SIMD 型では 1 命令あたりの処理要素数がベクトル型より少ないため，ベクトル型演算加速機構と比較して単位演算あたりの命令発行数が最大で 32 倍になり，ベクトル型よりも不利となる．

電力の観点からは，ベクトル型に比べレジスタなどの機構が簡単なため，消費電力を抑えられる可能性がある．加速機構に L1 キャッシュを用いないことやレジスタの再利用によりオンチップメモリと演算加速機構間のトラフィックを削減することで電力削減の効果が見込まれる．

2.2.2 ベクトル型演算加速機構

本稿で検討するベクトル型演算加速機構を**図 3** に示す．この演算加速機構では 16 個の FMA を持ち，32 個のベクトルレジスタを持つ．1 ベクトルレジスタは通常のベクトルプロセッサよりも少ない 16 要素 ×8=128 要素で構成される．1 サイクルに 1 ベクトルレジスタあたり 16 要素ずつ演算器に送り込まれ，これが 8 サイクル繰り返される．チェイニング機構を持つものとし，後続のベクトル演算のオペランドとして用いられる場合には，バイパスして後続の演算に供給される．

本稿で想定する演算加速機構では L2 キャッシュ階層と同等のメモリ階層に置いたオンチップメモリにアクセスする．したがって従来のベクトル機とは異なりオンチップメモリへのデータ転送を行っておき，そこからのロード命令を発行する．また，オンチップメモリ上でのデータの再利用も考えるため，従来のベクトルプロセッサと比較して主記憶へのアクセスを削減することができる．また，再利用性を十分に活用できた場合は主記憶へのバスバンド幅がある程度限られた状態でも効率的に演算できることが見込まれる．しかし，ベクトル長が大きい場合にはオンチップメモリに保持できるデータセットの数が少なくなり再利用性を十分生かせない可能性がある．このため，オンチップメモリを組み合わせる場合にはベクトル長を大きくとることが難しい．この点で本稿の演算加速機構で想定するベクトル型演算加速機構は通常のベクトルプロセッサとは異なる．

2.3 オンチップメモリと演算加速機構を用いるプログラム最適化

演算加速機構で演算を行う際には，いったんオンチップメモリに主記憶のデータを転送する必要があるが，そのデータ転送はソフトウェアにて制御される．この部分については基本的な仕組みが同じであるので既存研究[3),4)] の手法を襲用することができる．オンチップメモリ上に転送し終わったデータに対しては演算加速機構による従来手法と同じ最適化が行える．

オンチップメモリを利用したプログラムは

(1) 再利用性のあるデータを選択して主記憶からオンチップメモリにデータを転送
(2) オンチップメモリにあるデータを利用して演算
(3) 書き戻す必要のあるデータを主記憶に書き戻す

といった手順で実行される．このとき，データセットがオンチップメモリよりも大きい場合は何らかのブロッキングを行う．この手順についてはベクトル型および SIMD 型の演算加速機構で共通の動作となる．データ転送は完了するまでにレイテンシがあるために，あらかじめ主記憶・オンチップメモリ間のデータ転送命令を発行しておくことで，このレイテンシを隠蔽できる．

なお，2.2.1 節でも述べたが，演算加速機構とオンチップメモリ間のデータ転送におけるレイテンシについても，隠蔽するために先行してロード命令を発行する必要があり，その実現にはループアンローリングが基本的な実装方法となる．

オンチップメモリの利用はデータのアクセスパターンが予測可能であることが前提となる．予測困難なアクセスパターンのデータは，プログラミングもしくはコンパイル時にデータ転送の制御を明示することが難しく，基本的にオンチップメモリを用いた最適化は行えない．この場合は予測可能なデータのみをオンチップメモリを用いて最適化を行い，それ以外のデータは従来どおりにキャッシュを利用することで解決する．

オンチップメモリプロセッサ向けのプログラム最適化にはキャッシュと同様に時間的局所性を利用する方法と，オンチップメモリをストリームバッファ的に使用する方法がある．ストリームバッファとして利用する場合は各コアがオンチップメモリの異なる領域をそれぞれ利用してプリロードなどを行うことにより最適化を行う．時間的局所性を利用する場合は，あるコアがロードしたデータを各コア間で共有する方法と各コアが個別にデータをロードする方法が考えられる．前者の最適化では各コア間でデータを共有できた場合，後者と比較してメインメモリへのアクセスを削減することが可能である．しかし，同じオンチップメモリアドレスにアクセスが集中した場合は性能が低下する恐れがあるので，これを防ぐためにはアクセスタイミングの最適化が重要になる．

3. 性能評価環境

3.1 評価環境と評価用プログラム

本稿では想定するプロセッサアーキテクチャの実行性能

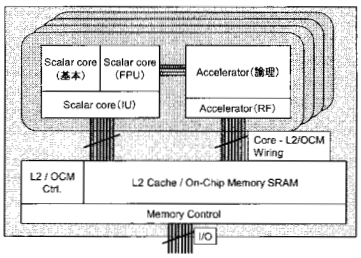

図 4 電力モデルにおける機能ブロック

と電力性能を評価するためにシミュレーションを行った．シミュレータは演算加速機構とオンチップメモリを付加するスカラコアとして低消費電力プロセッサである SH-4A プロセッサを想定し，GDB5.0 付属の SH エミュレータにクロックシミュレーション機能および演算加速機構を駆動するためのベクトル型および SIMD 型の命令セットを拡張することで開発した．なお，命令ストリームはすべて SH-4A コアで処理されることとする．

データの依存関係は浮動小数点数レジスタの依存関係のみをシミュレーションし，それ以外の汎用レジスタなどは依存なく実行できるものと仮定する．また，命令分岐予測は 100%成立するとする．命令発行は SH-4A と同じくインオーダー型の Dual Issue スーパスケーラとし，演算加速機構の演算命令とロードストア命令は同時に発行できることとする．演算加速機構にはロードストアユニットを 1 つ持つこととし，アクセスレイテンシについては演算加速機構からオンチップメモリへは 20 サイクル，オンチップメモリから主記憶へは 100 サイクルとする．

シミュレーションはベクトル型と SIMD 型について，それぞれの FMA 数を 16 個と 4 個としてシミュレーションを行う（以降，“Vector-16FMA” および “SIMD”）．またベクトル型と SIMD 型の演算器数以外の要素の比較を行うために一部評価では 4 個の演算器を持つベクトル型のシミュレーションもあわせて行う（以降，“Vector-4FMA”）．

評価用プログラムには DAXPY ループ，Livermore kernel 1,3,7[5] および $C = A \times B$ で表される N 次の倍精度行列積演算を使用した．それぞれのプログラムは C 言語により実装し，演算加速機構を利用する部分についてはベクトルもしくは SIMD 命令セットをインラインアセンブラでバイナリとして挿入した．

DAXPY ループおよび Livermore kernel については各ベクトルを分割し，行列積演算については，最内側ループ内でのデータセットがオンチップメモリ容量を上回ることから，タイリング[6] を用いて行列について 2 次元分割を行ってオンチップメモリにロードした後，再び 2 次元分割を行って各コアにスレッドとして割り当てた．この方法では 2.3 節で述べたように主記憶へのメモリアクセス量を減少させられるがコア間でアドレスコンフリクトが発生する可能性がある．今回はコーディングの都合上，コンフリクトを回避する最適化は行えていない．

3.2 電力モデル

消費電力については現段階では本研究はアーキテクチャレベルの検討を行っているため，RTL 設計のデータを利用するなどによる詳細な見積もりはできない．そこで本稿では想定するプロセッサについて，機能ブロックごとに分割し，その機能ブロック駆動回数あたりの消費電力単価を概算し，駆動率を掛け合わせることで消費電力を求めることとする．

分割した機能ブロックのうちクロックツリーを除いた図を**図 4** に示す．まず各コアのスカラコアについて命令発行ユニット (IU)，基本論理，浮動小数点数演算ユニットの 3 つに，演算加速機構は論理とレジスタファイルに，オンチップメモリ部分は SRAM とコントロール部分とメモリインタフェースの 3 ブロックに分ける．その他，配線としてクロックツリー，演算加速機構／オンチップメモリ間配線，外部 I/O を考える．演算加速機構については乗算命令・ロードストア命令・それ以外の 3 種別に分けて駆動率を考える．スカラコア／オンチップメモリ間の配線は特に分けて考慮せずスカラコア部分の消費電力に含まれると仮定する．以上，11 ブロックについてシミュレーションにより駆動率を求める．

電力は各機能ブロックについて想定より見積もったトランジスタ数，ゲート数や配線長を元に相対値とした電力単価を導出し，電力単価に機能ブロックの駆動率を掛け合わせることで求める．機能ブロックあたりの電力単価を C_i，機能ブロックの駆動率を n_i とすると，全体の電力 P は $P = \sum C_i \times n_i$ で求められる．

なお，本稿ではプロセッサチップで生じる消費電力をシミュレートし，主記憶自体の消費電力は含まない．また，今回の評価で用いる電力単価は相対値とし，0/1 のスイッチングで発生するアクティブ電力についてのみ評価を行う．リーク電流による消費電力については，想定するプロセスルールや温度などの条件により著しく値が変化し，電力値の導出が困難であるため，現時点では評価には含めないこととする．

4. 評価結果

4.1 演算加速機構の基本特性

まず，各方式の基本性能を確認するためにすべてのデータがオンチップメモリにあると仮定して，演算加速機構とオンチップメモリ間のバスバンド幅をピーク演算性能あたりのデータ転送量（Byte/flop）を基準に変化させて評価を行った．演算コアは 4 コアと仮定し，DAXPY ループ，Livermore kernel 1,3,7 ではデータサイズ N=200000 を 10 回，行列積演算では行列サイズ 576×576 を 1 回実行した．これらのベンチマークのうち，DAXPY，Livermore 7 および行列積演算について結果を**図 5** に示す．実効性能は 1 サイクルあたりの演算数 (flop/cycle) を，電力性能比は SIMD 4FMA 0.5B/flop 時の性能を 1 として，各ベンチマークごとに正規化したものである．本稿では電力性能の指標としてエネルギー遅延積 (Energy Delay Product: 以降，EDP) を使用する．EDP は消費電力量に時間を乗じたものであり，電力あたりの性能を表す指標として広く用いられている．

実効性能について DAXPY における SIMD-4FMA と Vector-4FMA を比較すると，各バンド幅においてほぼ同じ性能を示した．この性能は，バンド幅より求められる各ベンチマークプログラムの理想性能にほとんど等しい値を示しており，これは Livermore1，3 においてもほぼ同様の結果となった．Vector-16FMA については Vector-4FMA のほぼ 4 倍となる性能を示している．これらより基本的

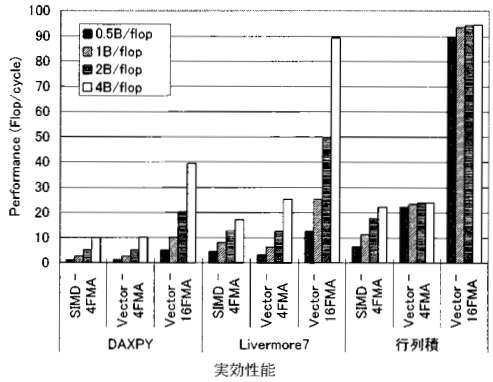

図 5 加速機構・オンチップメモリ間バンド幅による評価結果（4 コア）

なベクトル処理においては演算加速機構によらず，高い性能を発揮できることが分かる．

Livermore7 ではバンド幅が 2B/flop 以下の場合に SIMD-4FMA の性能が Vector-4FMA の性能を上回っている．これは SIMD 型はレジスタあたりの要素数が少ないためにレジスタにロードしたデータをループ間で再利用でき，演算あたりのロード数を削減できるためである．

これに対して行列積では，SIMD 型ではバンド幅に従って徐々に性能が向上するのと比較して，ベクトル型では 4FMA・16FMA とも低いバンド幅から性能を発揮している．これは，SIMD 型のレジスタが 4×128 要素しかデータを保持できないのと比較して，ベクトル型は 128×32 要素と，4 倍のデータをレジスタに保持できるため，行列積では演算量あたりのロード数を少なくできることから，ベクトル型では低バンド幅でも高い性能を発揮できていると考えられる．

EDP による電力性能を見ると，同じ FMA 数では SIMD 型がベクトル型よりも命令発行数が多くなるために同じ実効性能であっても SIMD 型にとって不利となるが，DAXPY における EDP を比較すると SIMD 型とベクトル型の電力性能はほぼ同じであった．Livermore7 では実効性能と同じく，EDP でも Vector-4FMA を下回っている．行列積では実効性能と同様，SIMD 型はベクトル型よりも EDP が大きい．また，各アーキテクチャでバンド幅増加による EDP 削減率は DAXPY が最も大きい．これは消費電力のうちクロックが占める割合が多いため，実効時間削減の効果をもっとも受けやすいからである．

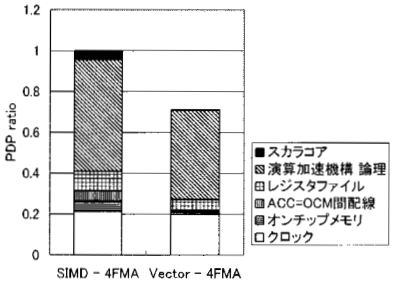

図 6 行列積演算 4 コア時における消費電力量内訳（4B/flop 時）

ここで SIMD 型とベクトル型の消費電力の違いを見るため，行列積 4B/flop 時の SIMD-4FMA と Vector-4FMA の消費電力量内訳を**図 6** に示す．これは各ブロックの消費電力量を SIMD-4FMA の消費電力量全体を 1 として正規化したグラフである．これを見ると演算加速機構・オンチップメモリ間配線とオンチップメモリの消費電力量が大きく減っていることが分かる．これはベクトル型では SIMD 型と比較して多くの要素をレジスタ内に保持できるためレジスタブロッキングのブロックサイズが大きくなり，演算あたりのロードストア回数が少なくてすむためである．これに伴い，演算加速機構の論理部分の消費電力やレジスタファイルの電力も減少している．また，スカラコアの消費電力量が減少しているのは，前述の理由でロードストア命令が少なくてすむことに加え，同じ演算量であってもベクトル型では 1 命令あたりの演算量が多いため，相対的に命令発行数が少なくてすむためであると考えられる．

SIMD-4FMA と Vector-16FMA の比較では Vector-16FMA の方が電力性能が圧倒的に良いが，これは前述のとおり同 FMA 数でも有利であったベクトル型で，FMA 数がさらに増加したため実行時間が短縮され，EDP が大幅に減少したためだと考えられる．

4.2 主記憶バンド幅とスケーラビリティ

次に，オンチップメモリ・メインメモリ間データ転送を含めたシミュレーションで，コア数のスケーラビリティに関する評価を行った．行列サイズ $N = 1728 \times 1728$ の倍精度行列積演算において，プロセッサコアの駆動周波数が 2.0GHz のときに 1ch 25.6GB/s の DDR3 インタフェースを実装したと仮定して，SIMD と Vector-16FMA の演算加速機構でチャネル数を 1 から 4ch と変化させてシミュレーションを行った．なお演算加速機構とオンチップメモリ間のバンド幅は 4B/flop とした．実効性能は 1 サイクルあたりの演算数を，電力性能比は SIMD 1 コア 25.6GB/s 時の性能を 1 として，各 EDP を正規化したものである．

実効性能を見ると，一部を除いてほぼコア数に比例して性能が向上している．1 コア時と 16 コア時の性能を比較すると，SIMD 102.4GB/s では 11.05 倍，Vector 102.4GB/s では 9.45 倍に性能が向上した．しかし，ベクトル型の 4 コア以下と SIMD 型ではバンド幅増加に伴う性能向上はほとんど見られなかった．これは FMA 数が一定以下では演算器律速となり，バスバンド幅の増加による効果が現れなかったと考えられる．また，電力性能とバンド幅の関係を見ると，バンド幅を増加させるこ

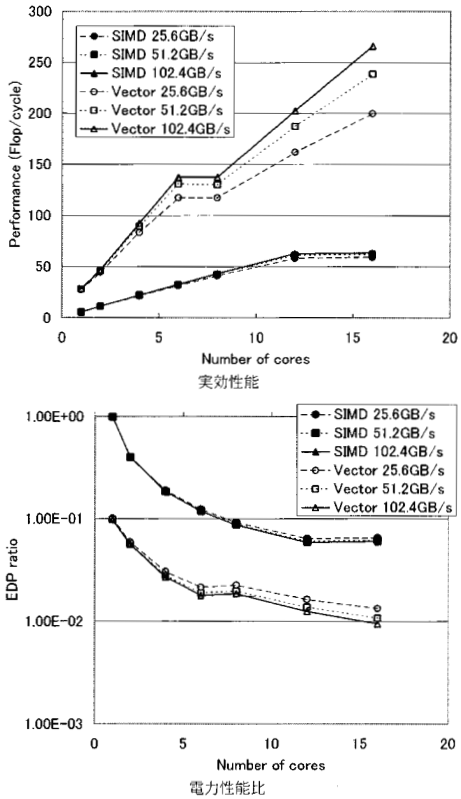

図 7 メモリバンド幅別 スケーラビリティ評価結果

とにより消費電力は低下する傾向があることが分かった.

5. 考　　察

今回の評価で用いた，DAXPY や Livermore1,3 のような典型的なベクトル処理では FMA 数が同じならば SIMD 型とベクトル型で実効性能や電力性能に大きな差はなかったが，行列積演算では電力性能ではベクトル型が SIMD 型を上回る結果となった．ベクトル型は 1 レジスタあたりの要素数が多いために発行命令数が少なくなる上に，レジスタブロッキング時にブロックサイズを大きく取れることが電力性能を考える上でメリットとなる．また，レジスタに保持できるデータ量の違いから行列積演算ではベクトル型よりも SIMD 型の方が演算加速機構・オンチップメモリ間のバンド幅を要求することが分かった．

しかしながら，Livermore7 のようにレジスタにロードしたデータを再利用できる場合や，今回のベンチマークには含まれなかったが複素数を扱うような場合でレジスタ内要素交換が生かせるような状況下であれば SIMD 型演算加速機構の利点が現れると考えられる．また本評価に用いた電力単価についてはベクトル型と SIMD 型の FMA あたりの論理数やレジスタファイルの 1 要素あたりの電力コストは違わないという見積りに基づいており，ハードウェアコストを正しく反映できていない可能性もある．4 章では B/flop というパラメータ下で比較したため，ベクトル型と SIMD 型ではオンチップメモリの絶対的なバンド幅も異なっていたが，電力コストを含めたハードウェアコストの見積りも不十分である．今後はこの点について検証する必要があると考えられる．

今回，後半の評価で用いた行列積演算プログラムでは，スケーラビリティは良好であり，オンチップメモリ・主記憶間のバスバンド幅の増加によって性能はあまり向上しなかった．これは行列積演算ではデータの再利用性によりバスバンド幅があまり必要とされず，通常マルチコアプロセッサで問題となるバスバンド幅のボトルネックに関する評価が十分行えなかった．今後はこの点に着目して異なったアプリケーションで評価を行う必要がある．

6. ま と め

本稿ではオンチップメモリプロセッサに多数の演算器を備えた演算加速機構を導入することを検討し，SIMD 型演算加速機構とベクトル型演算加速機構について実効性能と電力性能の指標において評価を行った．

シミュレーションによる評価の結果，典型的なベクトル処理では実効性能と電力性能において SIMD 型演算加速機構とベクトル型演算加速機構はほぼ同等の性能であったが，行列積演算ではレジスタ内の要素数の差などによりベクトル型は SIMD 型より実効および電力性能が勝る結果となった．しかし Livermore7 のような SIMD 型が有利となるプログラムの存在もあったことから，今後は実用に近い多様なアプリケーションでの評価が必要である．

今後，論理の詳細化と方式の改善により電力単価をより正確にしていく予定であり，想定するプロセスルールを定めた上でリーク電流を含めた評価を行いたい．また，本稿ではオンチップメモリに演算加速機構を付加したが，従来のキャッシュに付加した場合との比較も行いたい．

謝辞　本研究にあたりまして貴重なアドバイスを頂きました（株）日立製作所 小高雅則様，内田万亀男様に感謝いたします．なお本研究の一部は文部科学省「次世代 IT 基盤構築のための研究開発」プロジェクト「低電力高速デバイス・回路技術・理論方式の研究開発」による．

参 考 文 献

1) Adiga, N. et al.: An overview of the Blue Gene/L supercomputer, *Proc. SC2002*, pp. 1–22 (2002).
2) 高橋睦史ほか: オンチップメモリプロセッサでの演算加速機構の検討, 情報処理学会研究報告, 2007-ARC-172, pp. 263–268 (2007).
3) 中村宏ほか: ハイパフォーマンスコンピューティング向けアーキテクチャ*SCIMA*, 情報処理学会論文誌, Vol. 41, No. SIG 5(HPS 1), pp. 15–27 (2000).
4) Kondo, M. et al.: Software-controlled on-chip memory for high-performance and low-power computing, *ACM SIGARCH Computer Architecture News*, Vol. 30, pp. 7–8 (2002).
5) McMahon, F. H.: The Livermore Fortran Kernels: A Computer Test Of The Numerical Performance Range, Technical report, Lawrence Livermore National Laboratory (1986).
6) Lam, M. et al.: The cache performance and optimizations of Blocked Algorithms, *Proc. ASPLOS-IV*, pp. 63–74 (1991).