

動的リコンフィギャラブルプロセッサの低電力化

天野 英 晴^{†1} 木村 優 之^{†1} 弘 中 和 衛^{†1}

動的リコンフィギャラブルプロセッサ MuCCRA の電力分析をシミュレーションおよび実チップで行い、動的再構成電力とスタンバイ電力の割合が大きいことを示す。動的再構成電力を削減するためには、PE アレイ上のデータパスをなるべく変更しないことが重要である。必要以外の部分の再構成を抑制する差分再構成法により 15% 程度の削減が可能であり、また、マッピングの工夫のみでも 10% 程度の改善が可能である。さらに、大域的にループ構造を変換することで、データパス変更の頻度を抑えることができる。また、リーク電力の削減のために、Dual-Vth およびパワーゲーティングの適用を試みる。温度が高い状態でなければ、Dual-Vth の利用で、リーク電力はかなりの程度抑えることが可能である。

Low Power Techniques for Dynamically Reconfigurable Processors

HIDEHARU AMANO,^{†1} MASAYUKI KIMURA^{†1}
and KAZUEI HIRONAKA^{†1}

The power consumption of a dynamically reconfigurable processors MuCCRA is analyzed both with simulation and a real chip, and it appears that the reduction target should be power for dynamic reconfiguration and stand-by power. In order to reduce the dynamic reconfiguration power, the datapath on the PE array should be kept as possible. The differential configuration which separates the configuration data and changes only required part can reduce the power by 15%. Only a sophisticated mapping algorithm which keeps the datapath as possible can reduce the power by 10% Global loop optimization further increases the time of keeping the data path. For reducing leakage power, the dual Vth and power gating have been tried. If the temperature is not so high, the dual Vth design can reduce the leakage power satisfactory.

^{†1} 慶應義塾大学情報工学科

Dept. of Information and Computer Science, Keio University

1. はじめに

動的リコンフィギャラブルプロセッサ¹⁾ は、8bit から 32bit の粗粒度の演算器 (Processing Element, PE) およびメモリモジュールを、アレイ状に多数配置した構成をとり、データ並列性の高いアプリケーションにおいて高い処理能力を発揮する。FPGA (Field Programmable Gate Array) を始めとする他のプログラマブルデバイス同様、開発後に回路構成情報を追加、変更するだけで、様々な処理に対応することのできる拡張性、柔軟性を持ち、ハードウェアの開発期間やコストを削減することができる。さらに、動的再構成機能を利用して、高速に処理対象を切り換えることにより、従来の FPGA に代表される細粒度のリコンフィギャラブルデバイスの課題である、面積効率も改善される。

動的リコンフィギャラブルプロセッサの利点の一つは、その高いエネルギー効率である。問題を解くためのデータパスを直接 PE アレイ上に実現することによりエネルギー効率の良い処理が可能である。しかし、動的再構成のエネルギーオーバーヘッド、多数の PE を持つことによるリーク電力の増大など動的リコンフィギャラブルプロセッサに特有の問題点がある。本稿では、我々が今まで動的リコンフィギャラブルプロセッサの電力オーバーヘッドの削減に向けて提案した様々な手法を紹介する。

2. MuCCRA-3 アーキテクチャ

ここでは、電力解析の対象として用いたアーキテクチャである MuCCRA について簡単に紹介する。MuCCRA は MuCCRA-1,²⁾ を経て 3 世代目のプロトタイプ MuCCRA-3 が動作している。それぞれに特徴を持つが、ほとんど同じ構造を持つため、ここでは MuCCRA-3 を簡単に紹介する。

2.1 PE アレイの構成

MuCCRA-3 は、先行して開発された MuCCRA-1,2 同様、4x4 の PE によるアレイ構造を持つ。データ幅は 16bit とした。図 1 に PE アレイの概観を示す。演算データを保持する MEM は、これまでアレイ下部に 4 つ配置していたものを、アレイ上部にも配置し合計 8 つとした。これは、MuCCRA でのアプリケーションマッピング時、メモリの読み書きがボトルネックの 1 つとなっていたためである。ただし、MEM の深さは MuCCRA-1 の 512 の半分である 256 ワードとした。また、MEM はチップ外とのデータ交換に利用されるが、ダブルバンク構造になっており、IO 時間の削減、隠蔽が可能である。また、MuCCRA-1,2 では、メモリアドレスの生成のために PE を無駄に利用していたため、MuCCRA-3 では、MEM 内にカウンタを実装し、簡単なアドレス計算であれば MEM 単体でできるようになっている。

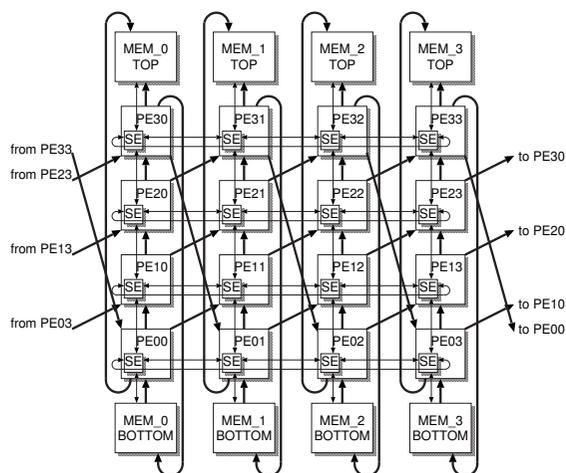


図 1 MuCCRA-3 の PE アレイ

PE 間を接続する結合網は、これまで MuCCRA-1,2 で採用してきた Switching Element(SE) によるアイランドスタイルに加え、近接の PE と接続する直結接続網との両方により構成される。図 1 中、太線が直接接続網、細線が SE による接続網を表している。これはアイランドスタイルを採用した MuCCRA-1 を直結接続網にて実装した MuCCRA-D の評価³⁾ から、直接接続網が動作周波数を向上させる一方で、データ移動の柔軟性を制限していたことが分かったためである。これにより、アプリケーションマッピング時に、近接の PE へデータを移動する場合には SE を再構成する必要のない直接接続網を利用し、遠い PE へデータを渡す場合には、SE を利用した接続といった使い分けが可能となり、データパスの形成がより容易である。SE によるチャンネル網は、A,B の 2 チャンネルが利用可能で、SE にて A から B, B から A へ乗り換えを可能とした。SE は配置配線のし易さから、PE 内のユニットとして持つようにした。

2.2 PE の構成

MuCCRA-3 を構成する PE は、2 項の四則演算、乗算、ビットシフト、比較演算を行う ALU、PE 内外からのデータ選択または定数生成を行い、1 項のシフト演算、ビット反転をする 2 つの ALU_SEL、8 エントリの 2 リードポート、1 ライトポートの RF からなっている。RF は、深さ 8 の FIFO として動作する FIFO モードを持つ。ALU_SEL_A,ALU_SEL_B, および RF の A_IN には、RF の出力、直接接続網、SE による接続網が接続されており、構

成情報により選択が可能である。ALU の機能は単純な 13 命令のみとし、ビット反転などは ALU_SEL が行う。これにより各ユニットの機能分離が明確になり、アプリケーションマッピングを容易にする。

2.3 MuCCRA-3 の再構成機構・動作制御

MuCCRA-3 は、コンテキストとよばれる、複数のコンフィギュレーションデータ (構成情報) のセットを切り替えて再構成を行う、マルチコンテキスト型の動的リコンフィギャブルプロセッサである。各 PE, MEM は自身を持つ CONTEXT MEMORY から、コンフィギュレーションデータを読み出して再構成する。したがって実行に先だて、CONTEXT MEMORY へコンフィギュレーションデータを転送しておく必要がある。この転送処理は PE アレイの外に配置されている Task Configuration Controller(TCC) が行う。

3. 電力分析

3.1 シミュレーションによる解析

動的リコンフィギャブルプロセッサの他のアーキテクチャに対する利点はエネルギー効率の良さである。これは、時間毎に必要なとされるデータパスをそのままの形で PE アレイ上に実現し、切り替えながら処理を実行できる点にある。一方で、コンテキスト切り替え時の電力が大きくなる問題が指摘されて来た。我々は小規模なマルチコンテキスト型リコンフィギャブルプロセッサ MuCCRA を用いて、電力の分析を行って来た。

まず、32bit の PE を 16 個持つ MuCCRA-P について、合成結果に基づくシミュレーションによる解析を行った⁴⁾。利用したプロセスは ASPLA 90nm である。この結果を図 2 に示す。分析の結果 (MuCCRA-P)、動的再構成に要する電力が 20%程度、クロックツリーなどのスタンバイ電力が 15%程度、結合網が 10%程度で残りは PE アレイにより消費されている。これは結合網による電力が半分程度を占める FPGA とはかなり異なった結果であり、電力の削減は、動的再構成のオーバーヘッドを減らすこと、クロックツリーやリークなどのスタンバイ電力を減らすことが重要であることがわかる。図中には、電力削減のための簡単な工夫の結果も示してある。まず、コンテキスト切り替え電力を削減するために、必要時のみコンテキストメモリから読み出しを行う選択的コンテキストフェッチ (MuCCRA-SF) を試した。また、PE 内と結合網のレベルの変化を抑えるオペランドアイソレーション (MuCCRA-OI) を適用した⁵⁾。結果として前者はあまり大きな効果は得られなかったが、後者は 10%程度電力を削減する効果があることがわかった。

3.2 実機による解析

シミュレーションによる解析は、消費電力の詳細な分析が可能であるが、実レイアウトに基づいていない場合には、配線容量の影響が含まれず、リーク電流の精度にも疑問がある。

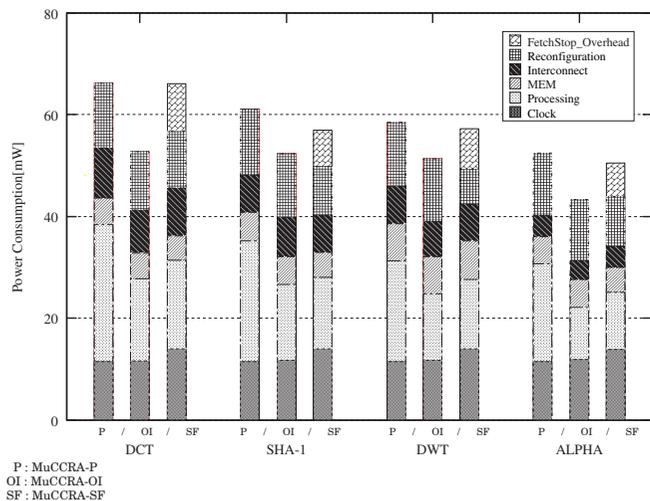


図2 MuCCRA-2の消費電力分析

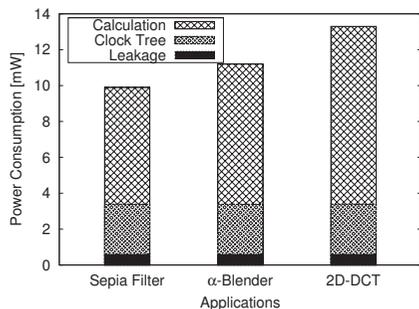


図3 MuCCRA-3の消費電力

そこで、e-shuttle 65nm プロセスで実装した MuCCRA-3 の実チップを用いて評価を行った⁶⁾。MuCCRA-3 は、MuCCRA-2 同様 16PE で 32 コンテキストを持つマルチコンテキスト型であるが、PE のデータ幅は 16bit である。また、アクセスのボトルネックを減らすために、アレイの上下にメモリを持つ点、PE の構造を単純化し、命令数を削減している点などに相違点がある。

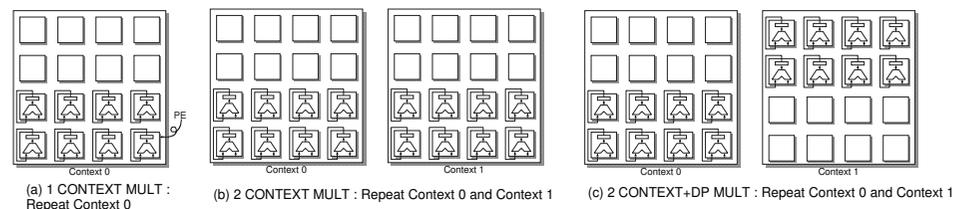


図5 3種類の実行方式

図3にはアプリケーションの電力を示す。シミュレーション同様、スタンバイ電力(図3における Clock Tree)は大きいですが、リーク電力はスレッシュホールドの高いライブラリを利用した影響で予想よりも少ない割合であった。図4は、コンフィギュレーションデータの転送および演算に利用するデータ転送に要する電力を測定した結果を示す。MuCCRA-3はコンフィギュレーションデータの転送には RoMulTiC と呼ぶマルチキャスト手法を用いているため、若干電力が大きいですが、両方共、演算に対する割合は小さいことが分かる。マルチコンテキスト型動的リコンフィギュラブルプロセッサでは、コンテキストメモリからコンフィギュレーションデータの読み出し、これを設定することによるデータパスの変更、データパス上で行われる演算が切れ目なく実行されるので、実機における評価では、これらを分離することが難しい。

そこで、図5に示すように、(a)では半分のPEに同一コンテキストで乗算を実行させ、(b)では同じ位置でコンテキストを切り替えて実行した。(c)はさらに、乗算の実行を行うPEをコンテキストの切り替え毎に変更して実行した。実行した乗算自体は同じであるので、(b)と(a)の差はコンテキストの切り替えによるもの、(b)と(c)の差はコンテキストの切り替えによるデータパスの変更の電力を示すと考えられる。

図6に、三つの場合の電力消費の測定結果を示す。(b)と(a)の差がさほど大きくないのに対して、(c)と(b)では倍程度異なることがわかる。

以上の解析により以下の点が明らかになった。

- 動的リコンフィギュラブルプロセッサの電力消費を抑えるためには、コンテキスト切り替えの電力の削減とスタンバイ電力の削減を行うべきである。
- コンテキスト切り替えの電力を減らすためには、データパスの変化を抑えるのが有効である。

4. 電力削減手法

ここでは、前章の結果に基づき、動的リコンフィギュラブルプロセッサの消費電力を削減

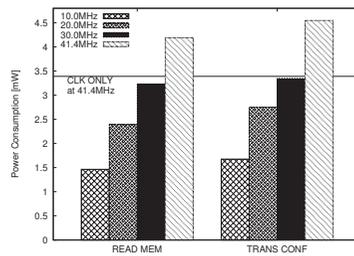


図4 データ/構成情報転送の電力

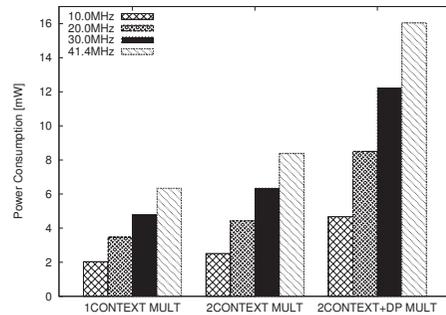


図 6 コンテキスト切り替えの電力消費

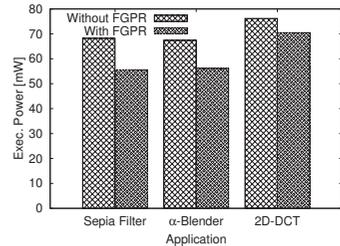


図 7 差分再構成による電力削減

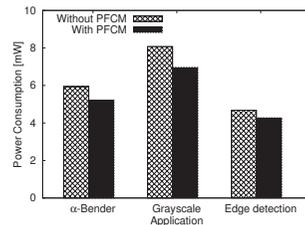


図 8 PFCM による電力削減

するための提案手法について紹介を行う。まず、コンテキスト切り替え時のデータバス変更を抑制手法について紹介し、次に各 PE のリーク電力を削減する手法について紹介する。

4.1 データバス変更の抑制

データバスの変更を抑制するためには、構成情報をいくつかのフィールドに分割し、変化させない部分の構成情報を転送しない手法が考えられる。これを実現したのが差分再構成法⁷⁾である。これを実現するには、再構成を行うかどうかを示す情報を他の構成情報に比べて読み出し周期を半クロックずらして読み出す必要があり、比較的複雑なハードウェアが必要である。しかし、図 7 に示すように、電力を削減することが可能である。この評価結果は MuCCRA-3 に改造を施してシミュレーションによって求めたものである。

もう一つは、アプリケーションを PE アレイにマッピングする際にできる限りデータバスを変更しないように割りつける手法が考えられる。この手法を PFCM(Partial Fixed Con-

figuration Mapping) と呼び⁸⁾、これを MuCCRA-3 に施した結果を図 8 に示す。これは実チップによる評価である。PFCM の効果はアプリケーションによって差があるが、この方法の利点は全くハードウェアのオーバーヘッドなしに電力効率を改善できることである。

さらに積極的にデータバスを保持するためにコンパイラを使う方法も検討している。動的リコンフィギュラブルプロセッサのアプリケーションはマルチメディア処理など、一定の大きさのデータに繰り返し処理を施すものが多い。これらの処理は、ループ内の処理は複雑でも、ループ構造自体は単純なものが多く、多数のデータに対して同一処理が繰り返して施される。図 9 は、このタイプの処理を非常に小さい 2×2 の PE アレイで実行する例を示す。マルチコンテキスト型の DRPA では、通常、PE アレイ上のデータバスを毎回切り替えて図 9a に示すようにこの処理を実行する。

しかし、この処理はループキャリア依存性がないので、図 9b に示すようにそれぞれのステップにおけるループに置き換えることが可能である。ここで、次のハードウェアの助けがあれば同じコンフィギュレーションデータをそれぞれのループで利用することができる。

- ループを数える専用のループカウンタ
- ループカウンタによって切り替わる中間結果格納用レジスタアレイ
- ループカウンタに依存して適切なアドレスを発生してデータメモリを読み出すアドレスカウンタ。処理の幅を広げるには単純なカウントアップだけではなく、スライドアクセスなどが必要である。

上記のハードウェアの助けを前提に行うループの再構成をここでは Loop Separating for Keeping Datapath (LSKD) と呼ぶ⁹⁾。LSKD は中間結果を格納するレジスタを必要とする点で、ループアンローリングと似ているが、LSKD は並列性を高めるためではなく、データバスを連続利用するためにループ構造を分離する点で異なる。

LSKD と、コンフィギュレーションデータの高速転送手法を併用すると、コンテキストメモリ自体をなくしてしまうことが可能である。この場合、場合によっては実行時間が延びてしまうが、消費エネルギーを節約することが可能である。図 10 は、MuCCRA-3 の元の構成を 1 とした場合の LSKD を用いた場合の相対的な消費エネルギーを示す。

図中に示す通り、ループ依存性のないアプリケーションでは、消費エネルギーは大きく低減されていることがわかる。

4.2 リーク電力の削減

MuCCRA-3 の評価では、スタンバイ電力の中でリーク電力はさほど大きな割合を占めていないが、特に将来のプロセスではリーク電力の削減は重要である。我々は Dual-Vth およびパワーゲーティングの利用でこの削減を試みた。動的リコンフィギュラブルプロセッサは、通常のプロセッサと FPGA の中間的な性格を持つため、Dual-Vth の適用手法も両者

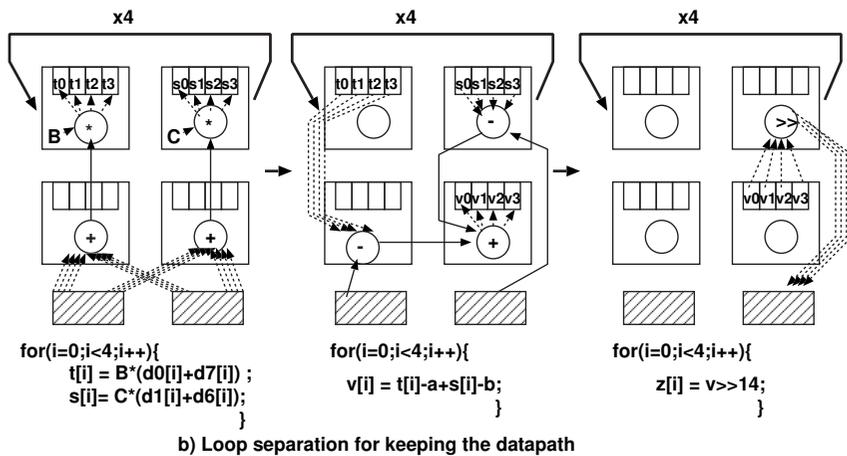
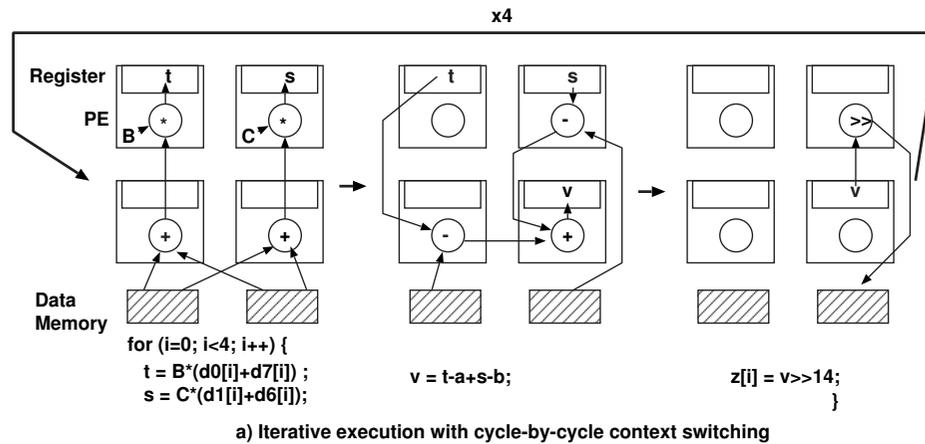


図9 Loop Separating for Keeping Datapath

の組み合わせが有効である。図11は、MuCCRA-3と類似したシステムを対象とし、富士通 e-shuttle 65nm プロセスの Low-Vth ライブラリと High-Vth ライブラリを組み合わせた場合のリーク電力と遅延時間積の相対値を示す。

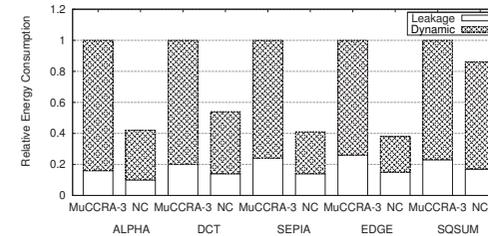


図10 相対消費エネルギー

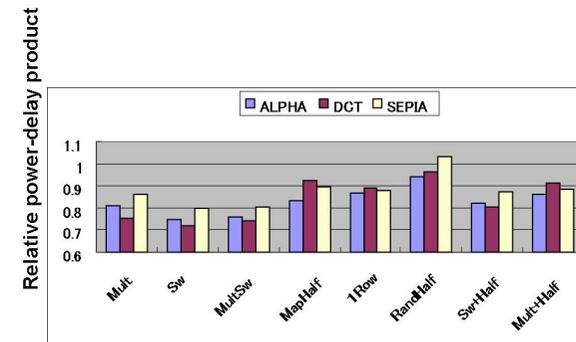


図11 Dual-Vth を用いた電力時間積 (ALL-LVt を 1 とする)

様々な High-Vth ライブラリと Low-Vth ライブラリの組み合わせ方を試しているが、中でも結合網とマルチプレクサ部分にのみ Low-Vth を使い、他を High-Vth とした Sw が優れた電力時間積を実現し、電力時間積を 20% から 28% 改善した。ただし、この手法では、実行遅延が大きくなるため、性能が不足する場合は、Sw+Half が優れている。この方法は、結合網部分および、PE の半分を Low-Vth のライブラリを用いると共に、アプリケーションのマッピングを最適化する方法である。この方法は全てのセルに Low-Vth を使った場合に比べて遅延の増大は 5%-14% で、28% のリーク電力を削減できる¹⁰⁾。

もう一つの手法は、パワーゲーティングの活用である¹¹⁾。コンテキスト毎に PE をスリープ状態にしてリーク電力を節約する細粒度パワーゲーティングを試みた。ここで、適用の対

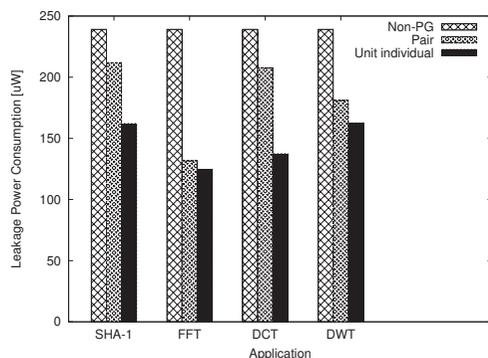


図 12 パワーゲーティングによるリーク電力の削減

象を PE 全体とした場合および PE の個別の演算モジュール単位に行った場合を評価した結果を図 12 に示す。

ここでは、チップの温度を 80 度として仮定しているため、リーク電力の割合が大きく、また削減効果も大きめに評価され、最大 48%削減している。

5. まとめと今後の課題

本稿では、動的リコンフィギャラブルプロセッサの電力を分析し、様々な削減手法を紹介した。動的再構成に要する電力を抑えるためには、PE 上に実現されるデータパスの変更を必要以上に行わないことが重要であり、コンパイラによるマッピングの工夫が有効である。また、スタンバイ電力についてはリーク電力の削減手法を示した。実際上、Dual-Vth の活用によりリーク電力はかなりの程度抑え込めると考えられる。

様々な手法を提案してきたが、動的リコンフィギャラブルプロセッサの基本的なアーキテクチャを維持した上での可能な電力削減は限界に達しつつある。さらに、電力を削減する場合、動的リコンフィギャラブルプロセッサの柔軟性のある程度犠牲にする必要がある。我々は、実行中の再構成を行わず、PE アレイを組みあわせ回路で構成してクロックツリーをなくしてしまう SMA について検討を行っている。

謝辞 本研究は科学技術振興機構「JST」の戦略的創造研究推進事業「CREST」における研究領域「情報システムの超低消費電力化を目指した技術革新と統合化技術」の研究課題「革新的電源制御による次世代超低電力高性能システム LSI の研究」による。また、東京大学大規模集積システム設計教育研究センターを通し、株式会社半導体理工学研究センター、(株)イー・シャトル、富士通株式会社の協力で半導体設計ライブラリを利用させていただいた。

参考文献

- 1) 末吉敏則, 天野英晴: リコンフィギャラブルシステム, オーム社 (2005). (編著).
- 2) H.Amano, Y.Hasegawa, S.Tsutsumi, T.Nakamura, T.Nishimura, V.Tanbunheng, A.Parimala, T.Sano, M.Kato: MuCCRA chips: Configurable dynamically-reconfigurable processors, Proc. of the ASSCC '07. IEEE Asian (Nov. 2007).
- 3) M.Kato, Y.Hasegawa, H.Amano: Evaluation of MuCCRA-D: A Dynamically Reconfigurable Processor with Directly Interconnected PEs, Proc. of The 2008 International Conference on Engineering of Reconfigurable Systems and Algorithms (ERSA'08) (Aug. 2008).
- 4) T.Nishimura and K.Hirai and Y.Saito and T.Nakamura and Y.Hasegawa and S.Tsutsumi and V.Tunbunheng and H.Amano: Power Reduction Techniques for Dynamically Reconfigurable Processor Arrays, Proc. of Int'l Conf. on Field Programmable Logic and Application (FPL) (2008).
- 5) 西村隆, 平井啓一郎, 斎藤貴樹, 中村拓郎, 堤 聡, 長谷川揚平, 天野英晴: 動的リコンフィギャラブルデバイスにおける電力分析と低電力化手法の検討, 電子情報通信学会論文誌, Vol.J92-D, No.10, pp.1763-1771 (2009).
- 6) Y.Saito, T.Sano, M.Kato, V.Tunbunheng, Y. Yasuda, H.Amano: A Real Chip Evaluation of MuCCRA-3: A Low Power Dynamically Reconfigurable Processor Array, Proc. of ERSA 2009, pp.283-286 (2009).
- 7) T.Sano, Y.Saito, M.Kato and H. Amano: Fine Grain Partial Reconfiguration for Energy Saving in Dynamically Reconfigurable Processors, Proc. of FPL, pp.530-533 (2009).
- 8) Hironaka, K., Kimura, M., Saito, Y., Sano, T., Kato, M., Tunbunheng, V., Yasuda, Y. and Amano, H.: Reducing power consumption for Dynamically Reconfigurable Processor Array with partially fixed configuration mapping, Proc. of the IEEE Int'l Conf. on Field Programmable Technology (ICFPT2010) (2010).
- 9) 天野、木村、小崎: マルチコンテキスト型動的リコンフィギャラブルプロセッサからコンテキストメモリをなす方法の提案, 信学技報 RECONF-2010 9 月 (2010).
- 10) Kei'ichiro Hirai, Masaru Kato, Yoshiki Saito and Hideharu Amano: Leakage Power Reduction for Coarse-Grained Dynamically Reconfigurable Processor Arrays Using Dual VT Cells, In Proc. of International Conference on Field Programmable Technology (to be appeared) (2009).
- 11) Yoshiki Saito, et. al.: Leakage Power Reduction for Coarse Grain Dynamically Reconfigurable Processor Arrays With Fine-Grained Power Gating Technique, ICFPT, pp.329-332 (2008).