

パイプラインステージ統合：将来のモバイルプロセッサのための消費エネルギー削減技術

嶋田 創[†] 安藤 秀樹[†] 島田 俊夫[†]

近年のモバイルプロセッサにおいては、低消費エネルギーと高性能の両立が求められている。これらの要求に応えるため、現在では、DVS (dynamic voltage scaling) と呼ばれる方式が導入されている。DVS は、低クロック周波数での動作時に、電源電圧を低下させ、消費エネルギーを削減する。DVS は現在では有効な方式であるが、将来のプロセス技術においては、電源電圧の可変範囲が縮小し、有効性が低下する。これに対して我々は、低クロック周波数での動作時に、電源電圧を最大値に保ったまま複数のパイプラインステージを統合するパイプラインステージ統合 (PSU: pipeline stage unification) と呼ぶ方式を提案した。本論文では、現在および将来のプロセス世代における DVS と PSU の効果を比較する。評価の結果、現在において、PSU は DVS に対して 11~14% 程度消費エネルギーを削減できることが分かった。しかし、将来では DVS は大幅にその効果を落とすのに対し、PSU はその効果を維持し、その結果、約 10 年後には PSU は DVS に対して 27~34% と大きく消費エネルギーを削減できることが分かった。

Pipeline Stage Unification: A Low-Energy Consumption Technique for Future Mobile Processors

HAJIME SHIMADA,[†] HIDEKI ANDO[†] and TOSHIO SHIMADA[†]

Recent mobile processors are required to exhibit low-energy consumption as well as high performance. To satisfy these requirements, a method called dynamic voltage scaling or DVS is currently employed. DVS reduces energy consumption by decreasing the supply voltage when a processor runs at a low clock frequency. Although DVS is an effective method for reducing energy consumption, its effectiveness will be limited in future process generations because the variable supply voltage range will become small. As an alternative, we previously proposed a method called pipeline stage unification or PSU, which unifies multiple pipeline stages when the processor runs at a low clock frequency, leaving the supply voltage at its maximum level. This paper compares PSU to DVS in terms of their effectiveness in current and future process generations. Our evaluations show that currently PSU reduces energy consumption only moderately (11-14%) more than DVS. However, in the future, DVS will significantly decrease its effectiveness, whereas PSU will maintain its effectiveness. As a result, PSU will reduce energy consumption by 27-34% more than DVS after about 10 years.

1. はじめに

近年のモバイルプロセッサでは、低消費エネルギーと高性能の両方が要求されている。この要求を満たすために、現在 DVS (dynamic voltage scaling) と呼ばれる方式が導入されている (例えば、Transmeta Crusoe の LongRun^{(1),(2)}, Intel Mobile Pentium III,4 の SpeedStep^{(3)~(5)}, AMD Mobile Athlon 4 の PowerNow!⁽⁶⁾)。DVS はバッテリー持続時間要求やプロセッサ負荷に応じて、動的にクロック周波数と電源電圧を変更するものである。バッテリー持続時間要求が強いか、与えられた負荷が低ければ、クロック周波数を低下させ、消費電力を削減する。さらに、延びたクロックサイクル時間に信号の遅延を合わせ、電源電圧を低下させる。これにより、プログラム実行に要する消費エネルギーを削減する。

このように、現在 DVS は消費エネルギーを削減する有効な手法であるが、プロセス技術の進歩に応じてこの有効性は減少する。その理由は以下の通りである。プロセス技術の進歩に応じて、これまで、電源電圧とトランジスタの閾値電圧

はプロセス技術の進歩に応じてスケールアップされてきた。しかし、一般に閾値電圧のスケールアップは電源電圧のスケールアップに比べて緩やかである。加えて、閾値電圧を現在より大きく低下させると、サブスレッショルドリーク電流が劇的に増加するため、閾値電圧のスケールアップ速度はさらに緩やかになると考えられる。トランジスタが正しく動作する電源電圧は閾値電圧によって下限が与えられるため、DVS における電源電圧の可変の範囲 (正確には、最大の電源電圧に対する可変の範囲の割合) は減少する。その結果、DVS の有効性は減少することになる。

これに対して、我々はパイプラインステージ統合 (PSU: pipeline stage unification) と呼ぶ手法を提案した^{(7),(8)}。PSU は単純だが効果的に消費エネルギーを削減することができる。PSU では DVS と同様に、プロセッサの消費エネルギーを削減するためにクロック周波数を低下させるが、DVS と異なり、電源電圧を低下させるのではなく、パイプラインレジスタをバイパスさせることによって複数のパイプラインステージを統合する。PSU によって消費エネルギーを削減できる理由は次の 2 点である。まず第 1 に、バイパスされるパイプラインレジスタへのクロックの供給を止めることにより、クロックドライバの総負荷を減少させることができる。これに

[†] 名古屋大学大学院工学研究科
Graduate School of Engineering, Nagoya University

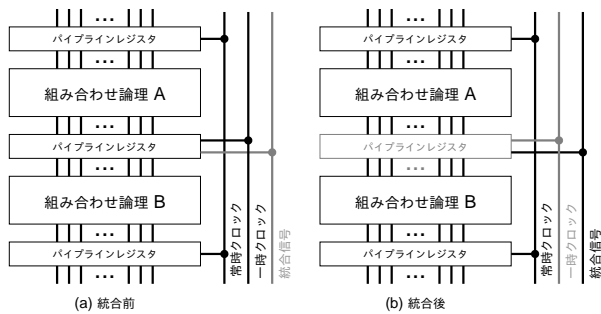


図1 クロック分配線の変更と統合信号の追加

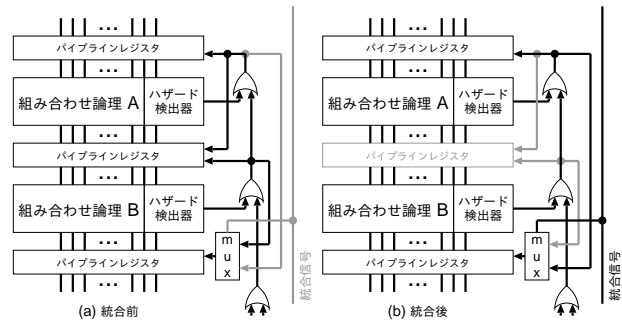


図2 インターロック回路

より、消費電力が削減される。第2に、パイプラインステージの統合によりプロセッサのパイプラインが短くなる。これにより、プログラムの実行に必要なサイクル数が削減され、電力を消費する時間を短くすることができる。例えば、フロントエンドパイプラインを短縮させることにより、分岐予測ミスペナルティは削減され、実行サイクル数は削減される。

DVSとPSUのどちらのアプローチが消費エネルギーの削減において有効であるかを直観的に判断することは難しい。以前発表の文献(7, 8)では、現在のプロセス技術のみを仮定した予備的な評価しか行われておらず、将来のプロセス技術において、電源電圧と閾値電圧が変わった時の評価が行われていない。この論文では現在と将来のプロセス技術において詳細な評価を行い、PSUとDVSを比較する。

残りの論文の構成は以下の通りである。2章ではPSUの実装について述べ、3章ではDVSとPSUによる消費エネルギーの削減について解析的に分析する。4章では評価における仮定について説明し、5章で評価結果を示す。6章では関連研究について述べ、最後に、7章でまとめる。

2. PSUの実装

本章では、PSUの実装について述べる。2.1節でPSUを実現するための基本的な回路の変更を示す。PSUの大部分は、この節で説明する方法で実装できるが、例外的に取り扱わなければならない部分が存在する。2.2節ではこれについて述べる。最後に、2.3節では、さらなるエネルギー削減のためのPSUとDVSの併用について述べる。

2.1 PSUのための回路

図1にPSUに関連する信号線とパイプラインレジスタとの結線関係を示す。説明を簡単にするために、2ステージの統合を例としている。図1に示すように、常時クロック、一時クロック、統合信号と呼ぶ3本の信号線が存在する。常時クロックは統合をするかしないかにかかわらず、常に動作しているクロックである。一方、一時クロックは統合を行った時には動作を停止させるクロックであり、統合を行わないときのみ動作する。統合信号は統合を指示する信号である。図1(a)はステージを統合していない状態を、図1(b)は統合した状態を示す。図中の黒い部分は動作、あるいは、アサートされている部分を示し、灰色の部分は動作していない、あるいは、アサートされていない部分を示す。図1(a)に示すように、通常のパイプラインでは、隣接する組み合わせ論理回路AとBは、それらの回路の間のパイプラインレジスタが動作しているため、異なったステージとして動作する。一方、

図1(b)ではこのパイプラインレジスタは動作せず、2つの組み合わせ論理回路は1つのステージとして動作する。

ここで、パイプラインレジスタには、フロントエンド、実行コア、バックエンド間にある命令ウィンドウなどデカップリング用の記憶素子を含まないことを注意しておく。これらは、パイプラインレジスタと同じくステージをつなぐ記憶素子であるが、複数の命令の状態を記憶する機能が必要なので、バイパスさせることはできない。

パイプラインレジスタをバイパスさせるには、2つの方法が考えられる。1つ目の方法は、統合時にクロックとは無関係に信号が通過するようパイプラインレジスタの論理を構成することである。透過型のラッチでパイプラインレジスタを実現する場合、この論理の変更は容易である。2つ目の方法は、パイプラインレジスタの後方にマルチプレクサを配置し、パイプラインレジスタの出力と前方のステージの出力を統合信号により選択する方法である。この方法は、パイプラインレジスタをどのような回路で構成しても適用可能である。

インターロック回路にもわずかな変更が必要である。図2に変更したインターロック回路を示す。図2(a)に統合していない状態での信号の流れを、図2(b)に統合した状態での信号の流れを示す。

一般に、あるステージのストール信号がアサートされるのは、そのステージでハザードを検出するか、または、後方のステージのストール信号が真の時である。図2に示した例では、 $S(A)$ と $H(A)$ を、それぞれ、ステージAのストール信号、ハザード信号とすると、 $S(A)$ は以下の式により与えられる：

$$S(A) = H(A) + S(B) \quad (1)$$

ステージAとBを統合する場合、統合後のステージのストール信号 $S(A+B)$ は、AまたはBがストールするときアサートされる必要があるため、以下ようになる：

$$S(A+B) = S(A) + S(B) \quad (2)$$

式(1)を式(2)に代入すると、

$$S(A+B) = H(A) + S(B) \quad (3)$$

となり、結局、 $S(A)$ と等しいことがわかる。

ところで、ストール信号によりステージを実際にストールさせるためには、そのステージの前方にあるパイプラインレジスタの更新を抑止し、後方のパイプラインレジスタへ渡す信号をNOPに変更する必要がある。組み合わせ回路Aの前方のパイプラインレジスタへ渡す信号は、統合前後で、それぞれ、 $S(A)$ 、 $S(A+B)$ なので、式(1)、(3)より変更の必要がないことがわかる。一方、組み合わせ回路Bの後方の

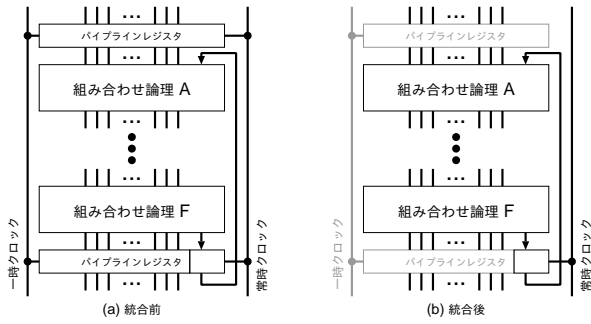


図 3 ループをなす信号バス中のパイプラインレジスタ

パイプラインレジスタへ渡す信号については、統合するかどうかに応じて $S(B)$ と $S(A+B)$ を切り替える必要がある。このためにマルチプレクサが必要である。

以上では 2 ステージ統合の場合についてのみ述べたが、一時クロック、統合信号、および、関連する回路を複数用意して適切に動作させることにより、さらに多くのステージの統合を行えるように簡単に拡張可能である。

2.2 ループをなす信号バスにおけるパイプラインレジスタ

前節で述べたように、統合するステージ間のパイプラインレジスタは基本的にバイパスさせるが、例外がある。ループをなす信号バスにおいて、後方のステージへ信号を出力するパイプラインレジスタは、統合するステージ間であってもバイパスさせない。図 3 に例を示す。図 3(a) は、ステージ F の出力がパイプラインレジスタを通り、後方のステージ A に入力され、再びステージ F に戻るといふ、ループをなす信号バスを持つパイプラインを示している。ステージ F とその次のステージを統合する際、図 3(b) に示すように、統合されるステージ間のパイプラインレジスタは基本的にバイパスさせるが、ループをなす信号バス上にある部分はバイパスさせない。パイプラインレジスタをバイパスさせないことにより、ループをなす信号バスは正しいタイミングを維持できる。このようなループをなす信号バスの例として、命令発行とそれに伴う命令の wakeup、実行結果のフォワーディング等があげられる。

図 4 に、命令の実行結果のフォワーディングのタイミングを例として示す。単純な DLX 型の 5 段のパイプライン⁹⁾において、実行ステージ (Exec) とメモリアクセスステージ (Mem) を統合した場合を考える。Exec と Mem ステージ間のパイプラインレジスタは基本的にバイパスされるが、実行結果のフォワーディングの部分ではバイパスされない。図 4(a) および (b) は、それぞれ統合前および統合後のタイミングを示している。図 4(b) に示すように、Exec ステージで得られた add 命令の結果は、バイパスさせない Exec と Mem ステージ間のパイプラインレジスタで保持され、次のサイクルに、正しいタイミングで Exec ステージの lw 命令に渡される。一方、Mem ステージで得られる lw 命令の結果は、通常のタイミングで Exec ステージの sub 命令に渡される。統合後も、add 命令のレイテンシは 1 サイクルであるが、lw 命令のレイテンシは統合により、1 サイクルに短縮されることに注意されたい。

2.3 PSU と DVS の組み合わせ

PSU は f_{max} を最大のクロック周波数、 $U(\geq 2)$ を統合さ

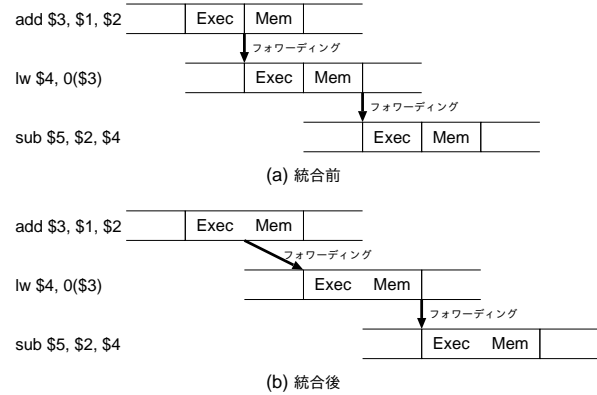


図 4 実行結果のフォワーディングのタイミング

れるステージ数とした時、 f_{max}/U の点 (以下、スイッチングポイントと呼ぶ) ごとに段階的に消費エネルギーを削減する。このことは、PSU は DVS と異なり、クロック周波数のスケールングに対して連続的にエネルギー消費を削減することはできないことを意味している。この欠点を補うためには、PSU と DVS を組み合わせれば良い。つまり、各スイッチングポイントではステージを統合し、スイッチングポイントの間は DVS を適用する。より具体的には、まず、電源電圧とクロック周波数を、最大値から最初のスイッチングポイント ($f_{max}/2$) まで減少させていく。次に、電源電圧を最大値に戻し、2 ステージを統合する。同様に、次のスイッチングポイント ($f_{max}/3$) まで電源電圧は最大値から減少させていく。このプロセスをクロック周波数を減少させている間、繰り返す。クロック周波数を上昇させる時はこの反対を行う。

なお、本論文では、PSU と DVS を組み合わせた場合の評価は行っていない。

3. 消費エネルギー削減量の解析

本章では、DVS と PSU による消費エネルギーの削減について解析的に説明する。一般に、プログラムの実行の際に消費するエネルギー E は以下の式で表される：

$$E = P \times T_{ex} \quad (4)$$

ここで、 P は消費電力、 T_{ex} は実行時間である。 P と T_{ex} は以下の式で与えられる：

$$P = a \times f \times C \times V_{DD}^2 \quad (5)$$

$$T_{ex} = \frac{N}{IPC \times f} \quad (6)$$

ここで、 f はクロック周波数、 C はスイッチするノードの全容量、 a はアクティビティファクタ、つまり、各ノードの平均スイッチング確率、 N は実行命令数、 IPC は 1 サイクルあたりの実行命令数の平均である。

3.1 DVS による消費エネルギー削減

式 (5) と (6) を式 (4) に代入し、以下に示す消費エネルギー $E(f, V_{DD})$ が得られる：

$$E(f, V_{DD}) = \frac{a \times N \times C \times V_{DD}^2}{IPC} \quad (7)$$

通常モードでは最大のクロック周波数 f_{max} と最大の電源電圧 V_{DDmax} で動作し、低消費電力モードでは f_{low} と V_{DDlow} で動作する DVS プロセッサを考える。通常モードの消費エネルギーで正規化した低消費電力モードにおける消費エネ

ギーは以下の式となる：

$$E_{DVS,n}(f_{low}, V_{DDlow}) = \frac{IPC_{max}}{IPC_{low}} \times \left(\frac{V_{DDlow}}{V_{DDmax}} \right)^2 \quad (8)$$

ここで、 IPC_{max} と IPC_{low} はそれぞれ通常モードと低消費電力モードでの IPC である。プロセッサ以外のデバイスも、プロセッサのクロック周波数の低下に応じて動作速度を低下させるという単純な仮定をおくと、 IPC_{low} は IPC_{max} に等しい。よって、次の式が導かれる：

$$E_{DVS,n}(f_{low}, V_{DDlow}) = \left(\frac{V_{DDlow}}{V_{DDmax}} \right)^2 \quad (9)$$

上式より、消費エネルギーの削減は電源電圧の低下によるみ達成されることが分かる。また、電源電圧の低下によるエネルギーの削減は、電源電圧の比の 2 乗の関数となっていることが分かる。2 乗の関数のため、消費エネルギーは劇的に削減される。同時に、式 (9) は、将来のプロセス技術では DVS の効果は急速に減少することを意味している。将来のプロセス技術では、 V_{DDlow}/V_{DDmax} は現在のプロセス技術ほど小さくできない。なぜなら、トランジスタが確実に動作するためには、 V_{DDlow} は、閾値電圧の 2 倍以上が必要であるが¹⁰⁾、閾値電圧は電源電圧より緩やかにしかスケールされないからである。

3.2 PSU による消費エネルギーの削減

2 章で述べたように、PSU は一時クロックのドライバを停止することによって消費電力を削減する。U ステージ統合 (以下、これを統合度 U という) でプロセッサが動作しているとき、理想的にはクロックドライバが消費する電力は $1/U$ となる。また、通常のプロセッサと同じく、総消費電力はクロック周波数の低下率に比例して削減する。従って、クロック周波数 f_{low} で動作する、統合度 U の PSU プロセッサの消費電力は以下の式で表される：

$$P_{PSU}(f_{low}, U) = \left(P_{total} - P_{clock} + \frac{P_{clock}}{U} \right) \times \frac{f_{low}}{f_{max}} \quad (10)$$

ここで、 P_{total} と P_{clock} はそれぞれ、通常モードにおけるプロセッサの総消費電力とクロックドライバの消費電力である。

式 (4) を用い、通常モードで正規化した消費エネルギーは以下の式で表される：

$$E_{PSU,n}(f_{low}, U) = \frac{P_{PSU}(f_{low}, U) \times T_{ex}(f_{low}, U)}{P_{total} \times T_{ex}(f_{max}, 1)} \quad (11)$$

ここで、 $T_{ex}(f, U)$ はクロック周波数 f と統合度 U における実行時間である。 $T_{ex}(f_{max}, 1)$ は通常モードにおける実行時間であることを注意されたい。式 (6) と (10) を式 (11) に代入することにより、以下の式を得ることができる：

$$E_{PSU,n}(f_{low}, U) = \frac{IPC_{max}}{IPC_{low}} \times \left\{ 1 - k \times \left(1 - \frac{1}{U} \right) \right\} \quad (12)$$

ただし、

$$k = \frac{P_{clock}}{P_{total}} \quad (13)$$

である。

式 (12) から分かるように、消費エネルギーは IPC の向上に反比例して削減される (パイプラインが短縮されるため、 $IPC_{max} < IPC_{low}$ であることに注意)。また、消費エネルギーの削減は、クロックドライバにより消費される電力が、プロセッサの総消費電力のどれだけの割合を占めているかということに依存する。この割合は、近年の高速なプロセッサ

表 1 ベンチマーク

ベンチマーク	入力	実行命令数
compress95	bigtest.in	95M
gcc	genoutput.i	84M
go	2stone9.in	75M
ijpeg	specmun.ppm	450M
li	train.lsp	183M
m88ksim	ctl.in	100M
perl	scrabbl.in	80M
vortex	vortex.in	80M

表 2 プロセッサの構成

命令発行幅		8
RUU		64 エントリ
LSQ		32 エントリ
メモリポート		8
機能ユニット	整数 ALU	8
	整数乗除算	4
	浮動小数点 ALU	8
	浮動小数点乗除算	4
分岐予測	予測方式	gshare
	履歴	6 ビット
	インデックス	13 ビット
	BTB	2048 エントリ/4-way
	RAS	16 エントリ
キャッシュ	L1 命令	64KB/32B ライン/1-way
	L1 データ	64KB/32B ライン/1-way
	L2 統合	2MB/64B ライン/4-way
メモリ	初期参照レイテンシ	64
	バースト転送間隔	2
TLB	命令	16 エントリ/4-way
	データ	32 エントリ/4-way
	ミスレイテンシ	128

表 3 PSU において仮定する実行レイテンシ、キャッシュヒットレイテンシ、分岐予測ミスペナルティ

統合されるステージ数	1	2	4	
クロック周波数	100%	50%	25%	
実行レイテンシ	整数乗算	3	2	1
	浮動小数点 ALU	2	1	1
	浮動小数点乗算	4	2	1
L1 キャッシュヒットレイテンシ	4	2	1	
L2 キャッシュヒットレイテンシ	16	8	4	
分岐予測ミスペナルティ	20	10	5	

では非常に大きく (例えば、Alpha 21264 では 32%¹¹⁾), この傾向は深いパイプラインや小さいクロックスキュー等の達成のために将来も続き、PSU は消費エネルギーを大きく削減できると期待できる。

4. 評価環境

3 章の式 (12) から分かるように、PSU の消費エネルギーには IPC と統合度 U とクロックドライバの消費電力の割合 k が関係する。本章では最初に IPC の評価環境を示す。次に、評価に用いたパイプラインの仮定を示す。最後に、クロックドライバの消費電力の割合 k に関する仮定について説明する。

4.1 シミュレーション環境

SimpleScalar Tool Set¹²⁾ 中の out-of-order 実行シミュレータを用いて、パイプラインのステージ数を変化させ、IPC を測定した。命令セットは SimpleScalar PISA である。表 1 に示すように、ベンチマークプログラムとして、SPECint95 の 8 本を用いた。ベンチマークプログラムのバイナリは gcc ver.2.7.2.3 を用い、-O6 -funroll-loops のオプションでコン

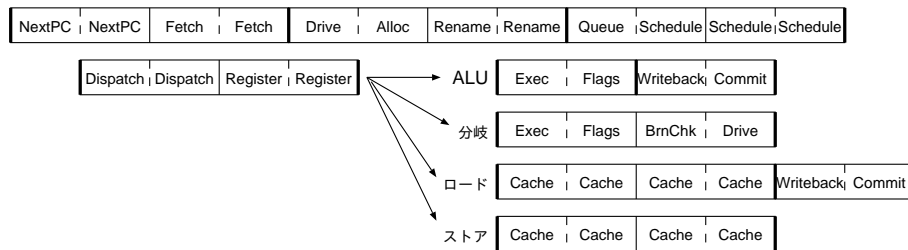


図5 仮定する PSU のパイプライン

パイルし作成した。シミュレーション時間が過大にならないようにするために、命令ミックス、関数の出現頻度など、特徴をほぼ維持しつつ、それぞれのベンチマークプログラムへの入力を調整した。

表2に、シミュレーションにおいて仮定したプロセッサの構成を示す。プロセッサは近年のプロセッサと同様に、深いパイプラインを持つと仮定した。メモリアクセス時間は、プロセッサのクロック周波数の低下に比例して遅くなると仮定した。そのため、プロセッサのクロック周波数によらず、メモリアクセスのサイクル数は一定である。

4.2 パイプラインの仮定

図5にベース、2ステージ統合、4ステージ統合のパイプラインを示す。Pentium 4⁽¹³⁾ のパイプラインとほぼ等しいパイプラインをベースとしている。2ステージ統合時には、パイプラインを区切る破線部のパイプラインレジスタをバイパスさせる。4ステージ統合時には、さらに細線部のパイプラインレジスタをバイパスさせる。表3に、これらのパイプラインにおける命令の実行レイテンシ、分岐予測ミスペナルティ、キャッシュヒットレイテンシを示す。表の2行目に示すように、1, 2, 4ステージを統合した場合、それぞれ最大クロック周波数の100%, 50%, 25%で動作する。

なお、整数/浮動小数点除算と平方根演算については、同一資源を繰り返し使用し完全なパイプライン化はされておらず、ステージの統合はできないと仮定した。レイテンシはそれぞれ、20, 12, 24サイクルとした。

4.3 クロックの消費電力の割合の仮定

プロセッサの総消費電力に対するクロックの消費電力の割合 k は、プロセッサの設計によって異なる。文献(11), (14)~(16)によれば、その範囲は18%~40%である。そこで、5章の評価においては、特に断わりのない限り、これらの値のほぼ中央値である30%と仮定した。また、クロックドライバが消費する電力は、駆動するパイプラインレジスタ数に比例すると仮定し、単純に統合度 U に反比例するとした。この仮定はおおまかではあるが、以下の理由で我々の評価においては妥当であると考えられる。一般に、クロックは多段のネットワークにより分配するが、クロックドライバの消費電力のほとんどは、最終段のドライバで消費される(例えば、Intel Itanium 2の階層化クロックネットワークの場合、88%を最終段のドライバが消費している⁽¹⁴⁾)。また、最終段のドライバの負荷は、およそ、そのファンアウトであるパイプラインレジスタの数に比例すると考えられる。

5. 評価結果

本章では、まず、現在の180nmのプロセス技術におけるエネルギー削減の評価結果を示す。次に、将来のプロセス技術

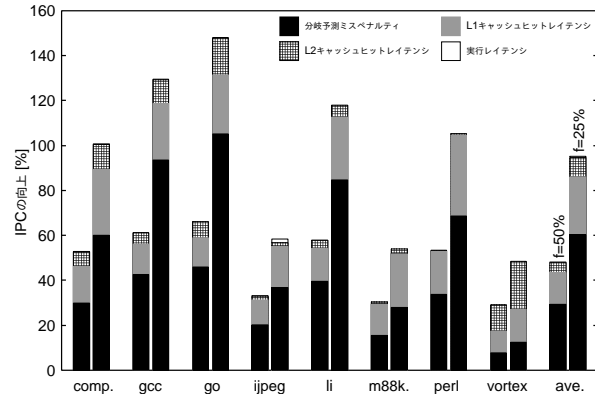


図6 PSUによるIPCの向上

におけるDVSとPSUの消費エネルギーの比較結果を示す。

5.1 現在のプロセス技術における消費エネルギーの比較

本節では、現在の180nmのプロセス技術におけるPSUによるIPCの向上を評価し、次にPSUとDVSの消費エネルギーを比較する。表4に評価におけるクロック周波数と電源電圧の関係の仮定を示す。DVSにおける各仮定は、Transmeta Crusoe TM5400⁽¹⁷⁾の対応するデータをもとにしている。なお、PSUでの電源電圧はクロック周波数によらず一定である。

図6に、各ベンチマークにおける、PSUによるIPCの向上率とその平均を示す。各ベンチマークの2本の棒はそれぞれ2ステージ統合($f=50%$)と4ステージ統合($f=25%$)の場合のIPC向上率である。積層された棒は下からそれぞれ分岐予測ミスペナルティ、L1キャッシュヒットレイテンシ、L2キャッシュヒットレイテンシ、命令の実行レイテンシの短縮による寄与である。IPCは $f=50%$ 時に平均約48%、 $f=25%$ 時に平均約95%向上した。特に、jpeg, m88ksim, vortexを除くベンチマークでは、IPCが大幅に向上することが分かる。IPCが大きく向上すれば、3.2節で示した式(12)から分かるように、消費エネルギーの大幅な削減が期待できる。

図7に、消費エネルギーの評価結果を示す。縦軸は、通常モードにおける消費エネルギーで正規化した消費エネルギーである。各ベンチマークにおける2本の棒グラフは共に、PSUにおける正規化消費エネルギーである。左の棒グラフは $f=50%$ における正規化消費エネルギーを、右の棒グラフは $f=25%$ における正規化消費エネルギーを示す。また、2本の横線(破線)は、DVSにおける正規化消費エネルギーを示している。上下の破線は、それぞれ $f=50%$ と $f=25%$ の場合である。なお、3.1節の式(9)からわかるように、DVSにおいては、エネルギーの削減は電源電圧の減少のみによ

表 4 180nm プロセス技術における DVS と PSU のクロック周波数と電源電圧の仮定

クロック周波数	電源電圧	
	DVS	PSU
100%	1.65V	1.65V
50%	1.35V	1.65V
25%	1.10V	1.65V

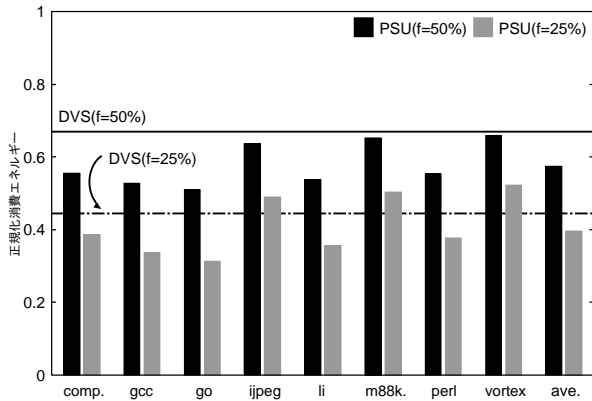


図 7 180nm のプロセス技術における 50%と 25%のクロック周波数における正規化消費エネルギー

て達成されるため、正規化消費エネルギーはどのベンチマークにおいても等しい。

図より、 $f=50\%$ における PSU の消費エネルギーは、すべてのベンチマークにおいて DVS よりも少ないことが分かる。正規化消費エネルギーは DVS で平均 0.67 である一方 PSU では平均 0.57 であり、14%削減された。 $f=25\%$ においては、PSU の DVS に対する優位性は小さくなり、PSU は一部のベンチマーク (jpeg, m88ksim, vortex) で DVS より多くのエネルギーを消費している。これは、DVS のエネルギー消費は電源電圧の減少の 2 乗で削減されるが、PSU による IPC の増加はそれほど急速ではないからである。図 6 から分かるように、上記の 3 本のベンチマークでは IPC の向上率が少ない。しかしながら、平均ではエネルギー消費は 0.44(DVS) から 0.39(PSU) へと、11%削減されている。現在のプロセス技術における PSU の優位性はこの程度にとどまっているが、5.2.2 節ではこの優位性は将来のプロセス技術ではより大きくなることを示す。

これまで k を 30%と仮定してきた。しかし、 k はプロセッサの設計により変化する。3.2 節式 (12) から分かるように、クロックの消費電力の割合が減少すると、DVS に対する PSU の優位性は減少する。文献 11), 14)~16) の例では、 k は 18%から 40%と幅がある。そこで、PSU が最も不利となる 20%の場合について評価した。その結果、PSU は DVS よりも、 $f=50\%$ において 9%、 $f=25\%$ において 2%消費エネルギーを削減可能なことが確認できた。このことは、プロセッサの多くの設計において、PSU が有効であることを意味している。

5.2 将来のプロセス技術における消費エネルギー

本節では、将来のプロセス技術における DVS と PSU の消費エネルギーの削減量を比較する。

5.2.1 将来の DVS プロセッサにおける電源電圧のスケールリングの見積もり

将来における DVS による消費電力の削減量を見積もるには、プロセス技術の進歩に応じて、どのように電源電圧が変化するか、また、低下させたクロック周波数に対して電源電圧をどこまで減少させることができるかを調べる必要がある。

最大クロック周波数で動作する通常モードでの電源電圧 V_{DDmax} については、これまで発表されたプロセッサの電源電圧のトレンド¹⁸⁾、および、TSMC と UMC が発表している電源電圧^{19),20)} より求めた。その結果、以下の式が得られた：

$$V_{DDmax} = 0.0381 \times Technology^{0.7171} \quad (14)$$

ここで、 $Technology[nm]$ は、プロセス技術の最小加工寸法である。

一方、低消費電力モードでの電源電圧に関しては、 V_{DDmax} と同様の方法で見積もるだけの十分なデータはなかった。そこで、以下のようにして求めた。

一般に、電源電圧と閾値電圧が与えられたとき、それらとゲートが動作する最大クロック周波数 f との間に、以下に示す関係がある (例えば、文献 21))：

$$f \propto \frac{(V_{DD} - V_{th})^2}{V_{DD}} \quad (15)$$

ここで、 f を通常モードでのクロック周波数で正規化した低消費電力モードでのクロック周波数とし、比例定数を a とおくと、上式は次のように表すことができる：

$$f = a \times \frac{(V_{DD} - V_{th})^2}{V_{DD}} \quad (16)$$

比例定数 a は $f = 100\%$ の時の電源電圧 V_{DDmax} (式 (14)) と V_{th} から得ることができる：

$$a = \frac{V_{DDmax}}{(V_{DDmax} - V_{th})^2} \quad (17)$$

一方 V_{th} は V_{DDmax} と同様の方法で見積もり、以下の式を得た：

$$V_{th} = 0.0226 \times Technology^{0.5111} \quad (18)$$

比例定数 a を用い、低消費電力モードにおける V_{DD} を式 (16) より求めることができる：

$$V_{DD} = \frac{(2aV_{th} + f) + \sqrt{(2aV_{th} + f)^2 - 4a^2V_{th}^2}}{2a} \quad (19)$$

実際の電源電圧は、上式で定まる V_{DD} にある程度のマージンが必要と思われる。種々の理由で、電源電圧が低いほどこのマージンを大きくとる必要がある。たとえば、式 (15) からわかるように、プロセスのばらつきによる V_{th} のばらつきのクロック周波数への影響は、 V_{DD} が低いほど大きい。また、 V_{DD} が低いほどノイズの影響を受けやすい。そこでこのマージンを M とし、式 (19) を以下のように書き直す：

$$V_{DD} = M \times \frac{(2aV_{th} + f) + \sqrt{(2aV_{th} + f)^2 - 4a^2V_{th}^2}}{2a} \quad (20)$$

ここで、 M は通常モードの時 1、低消費電力モードの時 $M > 1$ であり、プロセス世代によらず一定とする。

本研究で関心のある $f=50\%$ 、 25% の M を計算するため、Crusoe TM5400 (180nm プロセス) と TM5800 (130nm プロセス) の f と V_{DD} の関係のデータを用いた。これらのデータと式 (17), (18) より与えられる a と V_{th} を式 (20) に代入することにより、 M を計算することができる。表 5 に計算

表5 式(20)のマージン M の計算結果

正規化した クロック周波数	TM5400		TM5800	
	V_{DD}	M	V_{DD}	M
100%	1.65V	1.00	1.30V	1.00
50%	1.35V	1.24	—	—
25%	1.10V	1.39	0.90V	1.41

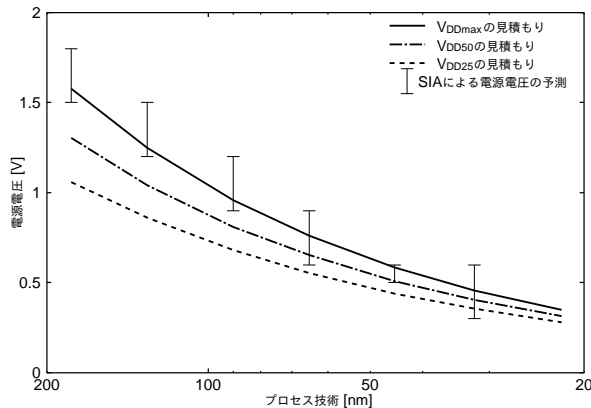
図8 将来の DVS プロセッサの V_{DDmax} と V_{DDmin} の見積もり結果

表6 将来のプロセス技術における DVS と PSU のクロック周波数と電源電圧の関係の仮定

プロセス	DVS			PSU		
	100%	50%	25%	100%	50%	25%
180nm	1.58V	1.29V	1.06V	1.58V	1.58V	1.58V
130nm	1.25V	1.04V	0.86V	1.25V	1.25V	1.25V
90nm	0.96V	0.81V	0.68V	0.96V	0.96V	0.96V
65nm	0.76V	0.65V	0.56V	0.76V	0.76V	0.76V
45nm	0.58V	0.51V	0.44V	0.58V	0.58V	0.58V
32nm	0.46V	0.41V	0.36V	0.46V	0.46V	0.46V
22nm	0.35V	0.32V	0.28V	0.35V	0.35V	0.35V

結果をまとめる*. これより, M を以下のように決定した:

$$M = \begin{cases} 1.00, & f = 100\% \\ 1.24, & f = 50\% \\ 1.40, & f = 25\% \end{cases} \quad (21)$$

以上まとめると, 通常モードの電源電圧 V_{DDmax} は, 式(14)より得られる. 低消費電力モードでクロック周波数が最大の50%, 25%での電源電圧は, 式(17), (18), (21)を式(20)に代入して得られる.

図8に V_{DDmax} , V_{DD50} , V_{DD25} の見積もり結果を示す. ここで, V_{DD50} , V_{DD25} は, それぞれ, $f=50\%$, 25% の時の電源電圧である. 図には, 同時に SIA 予測²²⁾の電源電圧の範囲も示している. 図からわかるように, V_{DD50}/V_{DDmax} と V_{DD25}/V_{DDmax} はプロセスが進むに応じて増加していく. このことは, 将来 DVS の有効性が減少していくことを意味している.

5.2.2 将来のプロセス技術における消費エネルギーの比較

将来における DVS と PSU の消費エネルギーを比較する. 表6に前節で見積もった各プロセス世代におけるクロック周波数と電源電圧の関係をまとめる.

図9に, 各プロセス世代において, 通常モードの消費エネ

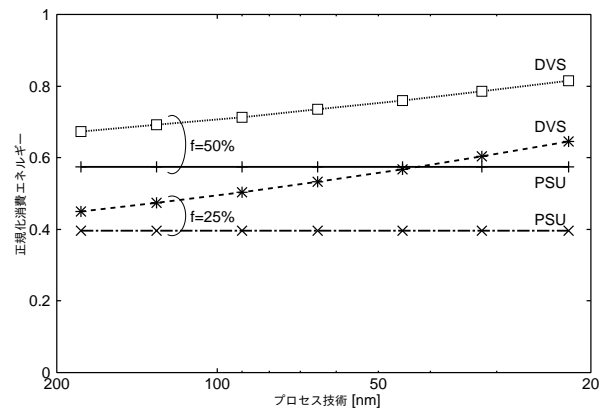


図9 将来のプロセス技術における50%と25%のクロック周波数における正規化消費エネルギー

ルギーで正規化した消費エネルギー(ベンチマーク平均)を示す. まず, DVS の有効性はプロセス世代の進歩に応じて低下していくことがわかる. $f=50\%$ の場合はおよそ3~4%/世代, 25%の場合は5~7%/世代の割合で消費エネルギーは増えていく. この結果, プロセス世代が現在の180nmからおよそ10年後の32nmの世代になると, 消費エネルギーの削減率は, $f=50\%$ の場合は33%から21%, $f=25\%$ の場合は55%から40%に減少する. 削減率が単調に減少していく理由は, 前節で述べたように, 通常モードにおける電源電圧に対する低消費電力モードでの電源電圧の割合が増加していくためである.

これに対して, PSU の消費エネルギーはプロセス世代の進歩に関係なく一定である. その結果, 32nmのプロセス技術においては, PSU は DVS よりも $f=50\%$ において27%, $f=25\%$ において34%多くの消費エネルギーを削減できる. この理由は, PSU の消費エネルギーの削減が, 電源電圧の低下に依存しているのではなく, 動的なアーキテクチャの変更によって達成されているためである.

6. 関連研究

Albonesi は, 動的にプロセッサの資源のサイズとクロック周波数を変更し, 実行時間を削減する方法を提案した²³⁾. 一般に, 資源サイズと実現可能なクロック周波数は, トレードオフの関係にある. このトレードオフの最適点を求め, プログラムの実行時間を削減する.

Bahar らは, プログラムの実行中に同時発行命令数が大きく変化することに着目し, 必要に応じて動的に命令発行論理と機能ユニットの一部を停止させ, 消費電力を削減することを提案した²⁴⁾.

Manne らは分岐予測ミスしたパスからフェッチした命令を減らして消費電力を削減することを提案した²⁵⁾. 具体的には, 分岐予測の信頼性が低い分岐命令が連続してフェッチされた場合, その後の命令のフェッチを停止する.

Canal らは, データ, アドレス, 命令をある特別なエンコーディングによって圧縮し, パイプラインの活動量を減少させ, 消費電力を削減する方法を提案した²⁶⁾.

7. まとめ

本論文では, PSU と呼ぶ消費エネルギーを削減する手法

* V_{DDmax} には, データの一貫性を保つため式(14)から得られる値ではなく, 実際の値を用いた. また, TM5800 の $f=50\%$ における電源電圧は残念ながら入手できなかった.

を、現在と将来のプロセス技術において評価し、現在の手法である DVS と比較した。PSU はクロック周波数が減少した時に、電源電圧を最大値に保ったまま、複数のパイプラインステージを統合することによって消費エネルギーを削減する手法である。将来のプロセス技術では電源電圧の可変範囲が減少するため、DVS の効果は減少する一方、電源電圧に依存せずにアーキテクチャの変更によって消費エネルギーを削減する PSU は、その効果を維持することができる。

評価の結果、現在では PSU は DVS よりも最大の 50% のクロック周波数のとき 14%、25% のクロック周波数のとき 11% 多く消費エネルギーを削減できる。この改善率はプロセス技術の進歩に伴って増加していく。例えば、10 年後には、改善率は 50% のクロック周波数のとき 27%、25% のクロック周波数のとき 34% に増加する。このように、PSU は将来のモバイルプロセッサにおいては、より魅力的なエネルギー削減手法となる。

より深いパイプラインがプロセッサの性能をさらに向上させるといふ予測がいくつかの文献で述べられている^{27)~29)}。パイプラインがより深くなれば、全消費電力に対するクロックの消費電力の割合は増加する。この結果、統合できるステージ数も増加する。このため、PSU の有効性はこの文献で述べたよりもさらに大きくなると考えられる。

参 考 文 献

- 1) Deifendroff, K.: Transmeta Unveils Crusoe, *Microprocessor Report*, Vol. 14, Archive 1, pp. 15-16 (2000).
- 2) Transmeta Corporation: *Crusoe Processor Model TM5800 Product Brief* (2001).
- 3) Intel Corporation: *Mobile Pentium III Processor in BGA2 and Micro-PGA2 Packages Datasheet* (2001).
- 4) Intel Corporation: *Mobile Intel Pentium III Processor-M Datasheet* (2002).
- 5) Intel Corporation: *Mobile Intel Pentium 4 Processor-M Datasheet* (2002).
- 6) Advanced Micro Devices, Inc.: *Mobile AMD Athlon 4 Processor Model 6 CPGA Data Sheet* (2001).
- 7) 嶋田創, 安藤秀樹, 島田俊夫: 低消費電力化のための可変パイプライン, 情報処理学会研究報告, 2001-ARC-145, pp. 57-62 (2001).
- 8) Shimada, H., Ando, H. and Shimada, T.: Pipeline Stage Unification for Low-Power Consumption, *Int. Symp. on Low-Power and High-Speed Chips (COOL Chips V) Presentation Slides*, pp. 194-200 (2002).
- 9) Hennessy, J. L. and Patterson, D. A.: *Computer Architecture: A Quantitative Approach, Second Edition*, Morgan Kaufmann Publishers Inc. (1996).
- 10) Semiconductor Industry Association: *International Technology Roadmap for Semiconductors 2001 Process Integration, Devices, Structures* (2001).
- 11) Gowan, M. K., Biro, L. L. and Jackson, D. B.: Power Considerations in the Design of the Alpha 21264 Microprocessor, *Proc. of the 35th Design Automation Conf.*, pp. 726-731 (1998).
- 12) Burger, D. and Austin, T. M.: The SimpleScalar Tool Set, Version 2.0, Technical Report CS-TR-97-1342, University of Wisconsin-Madison Computer Sciences Dept. (1997).
- 13) Glaskowsky, P. N.: Pentium 4 (Partially) Previewed, Vol. 14, Archive 8., pp. 1-4 (2000).
- 14) Anderson, F. E., Wells, J. S. and Berta, E. Z.: The Core Clock System on the Next-Generation Itanium Microprocessor, *2002 IEEE Int. Solid-State Circuits Conf. Visual Supplement to the Digest of Technical Papers*, pp. 110-111 (2002).
- 15) Clark, L. T., Hoffman, E. J., Miller, J., Biyani, M., Liao, Y., Strazdus, S., Morrow, M., Velarde, K. E. and Yarch, M. A.: An Embedded 32-b Microprocessor Core for Low-Power and High-Performance Applications, *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 11, pp. 1599-1608 (2001).
- 16) Gronowski, P. E., Bowhill, W. J., Preston, R. P., Gowan, M. K. and Allmon, R. L.: High-Performance Microprocessor Design, *IEEE Journal of Solid-State Circuits*, Vol. 33, No. 5, pp. 677-686 (1998).
- 17) Laird, D.: *Crusoe Processor Products and Technology*, Transmeta Corporation (2000).
- 18) Butts, J. A. and Sohi, G. S.: A Static Power Model for Architecture, *Proc. of the 33rd Annual Int. Symp. on Microarchitecture*, pp. 191-201 (2000).
- 19) Diefendorff, K.: TSMC Sets Sights on #1, *Microprocessor Report*, Vol. 14, Archive 6, pp. 17-21 (2000).
- 20) <http://www.umc.com/>.
- 21) Mudge, T.: Power: A First-Class Architectural Design Constraint, *IEEE Computer*, Vol. 34, No. 4, pp. 52-58 (2001).
- 22) Semiconductor Industry Association: *International Technology Roadmap for Semiconductors 2000 Update Process Integration, Devices, Structures* (2000).
- 23) Albonesi, D. H.: Dynamic IPC/Clock Rate Optimization, *Proc. of the 25th Annual Int. Symp. on Computer Architecture*, pp. 282-292 (1998).
- 24) Bahar, R. I. and Manne, S.: Power and Energy Reduction Via Pipeline Balancing, *Proc. of the 28th Annual Int. Symp. on Computer Architecture*, pp. 218-229 (2001).
- 25) Manne, S., Klauser, A. and Grunwald, D.: Pipeline Gating: Speculation Control For Energy Reduction, *Proc. of the 25th Annual Int. Symp. on Computer Architecture*, pp. 132-141 (1998).
- 26) Canal, R., Gonzalez, A. and Smith, J. E.: Very Low Power Pipelines using Significance Compression, *Proc. of the 33rd Annual Int. Symp. on Microarchitecture*, pp. 181-190 (2000).
- 27) Hartstein, A. and Puzak, T. R.: The Optimum Pipeline Depth for a Microprocessor, *Proc. of the 29th Annual Int. Symp. on Computer Architecture*, pp. 7-13 (2002).
- 28) Hrishikesh, M. S., Jouppi, N. P., Farkas, K. I., Burger, D., Keckler, S. W. and Shivakumar, P.: The Optimal Useful Logic Depth Per Pipeline Stage is 6-8 FO4, *Proc. of the 29th Annual Int. Symp. on Computer Architecture*, pp. 14-24 (2002).
- 29) Sprangle, E. and Carmean, D.: Increasing Processor Performance by Implementing Deeper Pipelines, *Proc. of the 29th Annual Int. Symp. on Computer Architecture*, pp. 25-34 (2002).