

複数電源ドメインの実行時電圧制御による CMOS LSIの消費エネルギー最小化

塩見 準^{1,a)} 石原 亨¹ 小野寺 秀俊¹

概要：電源電圧 (V_{DD}) としきい値電圧 (V_{th}) の動的な調節は集積回路のエネルギー効率を改善する有効な手法の 1 つである。本稿では、与えられた要求動作速度の下、同一クリティカルパス上に存在する 2 つの電源ドメインの消費エネルギーを最適化する問題に関する議論を行う。従来手法ではこれらドメインの V_{DD} のみが調整されていた一方、本稿では V_{DD} と V_{th} を同時に最適化する。単純な CMOS LSI の性能モデルに基づき、この最適化問題をランタイムに解くアルゴリズムを提案する。65-nm SOTB プロセスを用いたシミュレーションにより提案手法の検証を行い、 V_{DD} のみ調整する従来手法と比べて提案手法により最大 18% 消費エネルギーが改善することを述べる。

1. 序論

情報通信技術の急速な発展を背景に、LSI のエネルギー効率改善が強く求められている。LSI のエネルギー効率改善は、CO₂ 排出力削減やモバイル機器のバッテリー駆動時間の延長に繋がるため、エネルギー効率の良い LSI の設計手法が長年研究されてきた。

電源電圧 (V_{DD}) としきい値電圧 (V_{th}) の動的調節技術は、LSI のエネルギー効率を改善する最も有効な手法の 1 つである [1–3]。 V_{DD} を下げることで対象回路の動的消費エネルギーを削減でき、 V_{th} を大きくすることでサブスレッショルドリーク電流に起因する静的消費エネルギーを削減することができる。他方、LSI の動作速度は V_{DD} と V_{th} に強く依存するため、同一の動作速度を達成する V_{DD} と V_{th} の組が無数に存在する。これら無数の組の中で、回路の総消費エネルギー (動的消費エネルギーと静的消費エネルギーの和) を最小化する V_{DD} と V_{th} の最適な組が存在する。

本稿では、クリティカルパス上に 2 つの電源ドメインが存在する LSI の消費エネルギーを最小化する電圧制御手法を提案する。文献 [4] によると、与えられた要求同速度の下、消費エネルギーを最小化する V_{DD} と V_{th} の組は回路の活性化率に強く依存して変化する。したがって、もし活性化率が大きく異なる複数のマクロが LSI に搭載されている場合、最適な V_{DD} と V_{th} は各マクロで大きく異なる。例えば、この関係はプロセッサにおいて典型的に見られる。オンチップメモリの活性化率は一般的なロジック回路の活性化率と比較して非常に小さく、静的消費エネルギーがオンチップメモリの総消費エネルギーの大部分を占めている。したがって、メモリ回路の V_{th} と V_{DD} をともに高く設定することで、回路の動作速度を悪化させることなくオンチップメモリの総消費エネルギーを削減できる。以上の事実は、プロセッサのロジック回路とメモリ回路に独立して電圧制御技術を適用すべきであることを示唆している。

しかし、一般的にプロセッサのメモリ回路やロジック回路は同一のクリティカルパス上に存在するため、プロセッサの要求動作速度を守りながらすべての電源ドメインの

V_{DD} と V_{th} を最適化する技術は自明ではない。また、これら最適な V_{DD} と V_{th} の組は、回路の活性化率、チップ温度、経年劣化、要求動作速度に依存して動的に変動する [4]。常に LSI が最小の消費エネルギーで動作するためには、 V_{DD} と V_{th} を動的に最適化する技術が必須である。本稿では、クリティカルパス上に存在する 2 つの電源ドメインの消費エネルギーを最小化するランタイム V_{DD} , V_{th} スケーリング技術を提案する。

本稿の構成は以下の通りである。まず、第 2 章で関連研究と簡単な Motivational example を示す。第 3 章で 2 つの電源ドメインをランタイムに最適化する電圧制御アルゴリズムを述べる。第 4 章では、SOTB 65-nm プロセスを用いた検証を行い、第 5 章で結論を述べる。

2. 関連研究と Motivational Example

2.1 関連研究

単一電源ドメインに対する V_{DD} と V_{th} のランタイム最適化技術に関する初期の研究の 1 つは Nomura らによる研究である [5]。文献 [4] では、要求動作速度を満たしながら消費エネルギーを最小化する V_{DD} と V_{th} で回路が動作した場合、静的消費電力総消費電力の約 30% になることを解析的に示している。この事実に基づき、[5] は V_{DD} と V_{th} をランタイムに最適化する単純な電圧制御アルゴリズムを提案している。静的消費電力、総消費電力のモニタ回路が対象チップに組み込まれ、モニタした電力の比が一定の比 (30%) になるよう V_{DD} と V_{th} をデルタ変調のように逐次的に調節する技術を提案している。文献 [5] と類似の技術が [6] により提案されている。静的消費エネルギー、動的消費エネルギー、チップ温度およびクリティカルパス遅延を動的にモニタすることで、与えられた要求動作速度の下、最適な V_{DD} と V_{th} をランタイムに導出する電圧制御技術が提案されている。文献 [4] が述べているように、最適な V_{DD} と V_{th} の組は回路の活性化率、チップ温度、要求動作速度によって動的に変動する。文献 [5, 6] で提案された電圧制御アルゴリズムを用いることで、動的に変動する最適な電圧組を追跡でき、LSI が常に最小の消費エネルギーで稼働することが可能になる。しかし、これらのアルゴリズムは複数電源ドメインの消費エネルギー最適化を対象とし

¹ 京都大学大学院情報学研究科

^{a)} shiomi-jun@i.kyoto-u.ac.jp

ていない。本稿では [5,6] と異なり、与えられた要求動作速度の下、2つの電源ドメインの消費エネルギーをランタイムに最小化する電圧スケール技術を提案する。

複数電源ドメインの消費エネルギー最適化手法に関して述べられた初期の研究の1つは Usami らによる [7]。文献 [7] では、対象回路を2つの異なる電源電圧のドメインに分解し、動作速度を悪化させることなく消費エネルギーを削減する Clustered-Voltage-Scaling (CVS) に関して述べている。文献 [8,9] ではクリティカルパス上に存在する複数のドメインの電源電圧を最適化するスケール技術手法を提案している。単一の電源電圧で LSI を動作させた時と比較して、独立電圧制御により最大 57% の消費エネルギー削減に成功している。しかし、既存研究では、 V_{DD} のみの動的スケール技術が提案されている。既存研究と異なり、 V_{DD} と V_{th} の同時調節が省エネルギー動作の鍵であることを示す。SOTB 65-nm プロセスを用いたシミュレーションにより、 V_{DD} のみの動的スケールと比較して、 V_{DD} と V_{th} の動的な最適化により最大 18% の消費エネルギー削減できることを示す。

2.2 Motivational Example

単純なモデル回路を通して活性化率が異なる複数の電源ドメインに対して独立して電圧スケールする重要性を述べる。まず、与えられた要求動作速度の下、消費エネルギーを最小化する最適な V_{DD} と V_{th} が活性化率に依存することを示す。対象とする回路は、SOTB 65-nm プロセスを用いて設計されたファンアウト 4, 50 段インバータチェーンである。対象回路に与えられた遅延制約が 2 ns であるときの消費エネルギーの変化を図 1 (a) に示す。インバータチェーンの活性化率がそれぞれ 10% および 0.1% の時の結果を示している。10% および 0.1% の活性化率は、それぞれプロセッサにおけるロジック回路およびメモリ回路を模擬している。横軸はトランジスタのしきい値電圧 (V_{th}) である。ここで、2 ns の遅延制約を守るために、 V_{th} の増大とともに V_{DD} も増大することに注意。図の縦軸は回路の1クロックサイクルあたりの消費エネルギーである。消費エネルギーの値は、 V_{th} が 0.4 V の時の消費エネルギー値で正規化されている。図 1 (a) は、メモリ回路の最適な V_{th} はロジック回路の V_{th} より大きいことを示している。これは、メモリ回路の活性化率が極端に低いため、静的消費エネルギーがメモリ回路の消費エネルギーの大部分を占めているためである。したがって、 V_{th} を高くすることでメモリの総消費エネルギーを削減できる。以上の事実は、活性化率に応じて異なる V_{th} をマクロに印加すべきであることを示唆している。例えば、もし単純にロジック回路とメモリ回路の V_{th} を単一の 0.29 V に設定すると、ロジック回路は +85% のエネルギーオーバーヘッドに直面する。

次に、動的な V_{DD} と V_{th} の調節が LSI の消費エネルギー削減に重要であることを述べる。遅延制約が 2 ns から 10 ns に緩和され、 V_{DD} がスケールされた結果を図 1 (b) に示す。各マクロの最適な V_{th} が緩和に応じて増大している。これは、遅延制約の緩和により対象回路が低電圧領域で動作し、その結果静的消費エネルギーが総消費エネルギーの

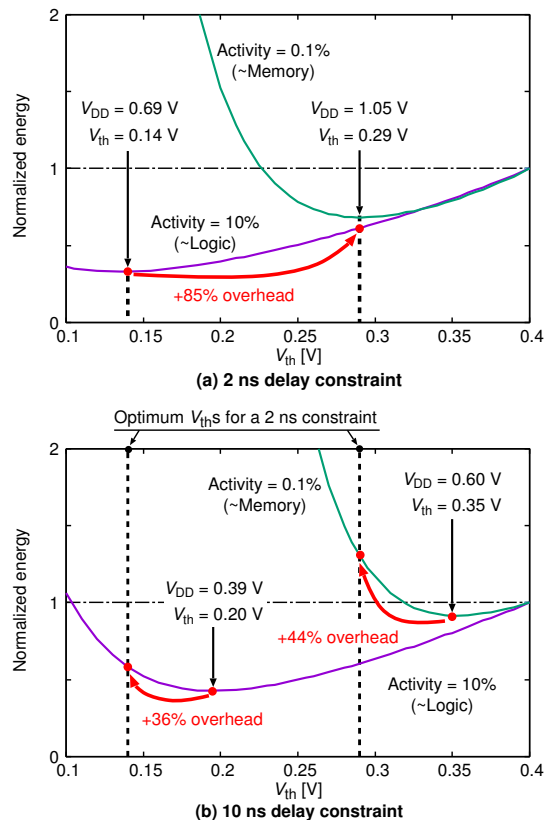


図 1 (a) 2 ns 遅延制約, (b) 10 ns 遅延制約 に対するモデル回路の消費エネルギー。

大半を占めるためである。したがって、各電圧ドメインの V_{th} の調節が消費エネルギー改善に重要な役割を果たす。 V_{th} を 2 ns 遅延制約で最適化した値で固定していると、図 1 (b) に示すように、ロジック回路とメモリ回路でそれぞれ +36% と +44% のエネルギーオーバーヘッドが発生する。したがって図 1 (b) は、既存の V_{DD} のみを用いた多電源ドメイン最適化手法と異なり、 V_{th} の動的な制御も省エネルギー動作に重要であることを示している。しかし、ロジック回路とメモリ回路は典型的には同一のクリティカルパスに存在するため、プロセッサの要求動作速度を守りながら2つの回路の V_{DD} と V_{th} を最適化する手法は自明ではない。本稿では、与えられた要求動作速度の下、ランタイムに2つの電源ドメインの消費エネルギーを最適化する電圧制御技術を提案する。

3. 複数電源ドメインのエネルギー最小化に向けたランタイム電圧制御アルゴリズム

3.1 問題設定

図 2 に示された回路の消費エネルギーを削減する問題を考える。対象回路には2つの電源ドメインが存在する。それぞれのドメインはプロセッサのロジック回路とオンチップメモリ回路に対応する。一般的なプロセッサのロジック回路には、エネルギー効率を改善するため、DC-DC コンバータが実装されている。一般的に、メモリ回路の電源はロジック回路の電源と分離されている。本稿では、オンチップメモリのエネルギー効率を改善するため、異なる DC-DC コンバータがメモリ回路に取り付けられていると

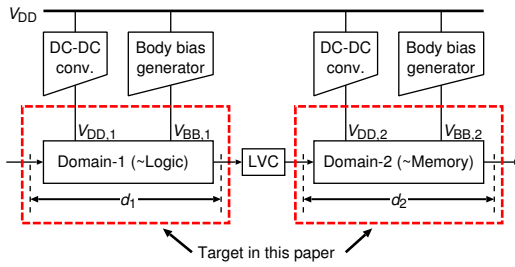


図 2 2つの電源ドメインを有する対象回路。

仮定する．2つの回路のしきい値電圧 ($V_{th,1}$, $V_{th,2}$) は基板バイアス ($V_{BB,1}$, $V_{BB,2}$) を調節することで変更可能である．例えば，[10] のような文献で単一電源で動作する低コスト基板バイアス生成回路が提案されている．レベルコンバータ (LVC) が2つのドメイン間に挿入されている．以降の解析では，簡単のため，図2に示した“Domain-1”および“Domain-2”のみを対象とし，周辺回路により生じるエネルギー・遅延オーバーヘッドは考慮しない．

本稿では，対象回路の2つの電源ドメインの遅延をそれぞれ d_i とする．ここで，添え字 i は電源ドメインを識別するパラメータである ($i = 1, 2$)．簡単のため，クリティカルパス遅延が d_i の和で表されると仮定する． $e_{t,i}$ を Domain- i で消費される消費エネルギーとする． $e_{t,i}$ は，動的消費エネルギー ($e_{d,i}$) と静的消費エネルギー ($e_{s,i}$) の和で表現される．動的消費エネルギーは電源電圧の2乗に比例し，静的消費エネルギーはしきい値電圧に対して指数関数的に変化する．消費エネルギーモデルを以下にまとめる：

$$e_{t,i} = e_{d,i} + e_{s,i}, \quad (1)$$

$$e_{d,i} = k_{1,i} V_{DD,i}^2, \quad (2)$$

$$e_{s,i} = k_{2,i} T_0 V_{DD,i} \exp(-V_{th,i}/N_{s,i}). \quad (3)$$

ここで， $k_{1,i}$ と $k_{2,i}$ はプロセステクノロジー，回路構造， i に依存して変化するフィッティング係数である． $N_{s,i}$ は理想係数と熱電圧の積であり， T_0 は対象回路のクロック周期 (遅延制約) である．もし電源電圧がしきい値電圧より十分高い場合，回路の伝搬遅延は α -乗則モデル [11] を用いて以下のように表現することができる：

$$d_i = k_{3,i} V_{DD,i} (V_{DD,i} - V_{th,i})^{-\alpha_i}. \quad (V_{DD,i} \gg V_{th,i}) \quad (4)$$

ここで， α_i は1以上2以下のフィッティング係数である．与えられた要求動作速度の下， $V_{DD,i}$ と $V_{th,i}$ を動的に調節することで消費エネルギーを最小化する問題を考える．この最適化問題を以下に示す：

$$\begin{aligned} \min \quad & e_{t,1} + e_{t,2} \\ \text{s.t.} \quad & d_1 + d_2 = T_0 \\ & V_{DD,1}, V_{th,1}, V_{DD,2}, V_{th,2} \in \mathbb{R}. \end{aligned} \quad (5)$$

本稿の目的は (5) の最適解をランタイムに導出することである．文献 [5, 6] のような既存研究で既に，与えられた要求動作速度の下，単一の電源ドメインに対して V_{DD} と V_{th} をランタイムに最適化する技術が提案されている．しかし，以下の問題が単一電源ドメインの最適化問題と比較して複数電源ドメインの最適化問題を複雑にしている．

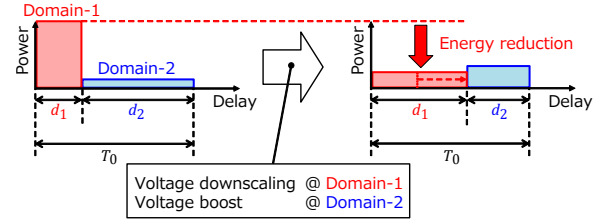


図 3 遅延割り当ての最適化による消費エネルギーの改善。

電源ドメイン間の遅延割り当て問題： d_1 と d_2 の和が T_0 を超えない限り， d_1 と d_2 の値を変更することができる．本稿では，各ドメインに割り当てる遅延の値を遅延割り当てと呼ぶ．例えば，もし $d_1 < d_2$ かつ $e_{t,1} \gg e_{t,2}$ が成立する場合，図3に示すように遅延割り当てを変更することで $e_{t,2}$ の微増と引き替えに $e_{t,1} + e_{t,2}$ を大幅に削減できる．

次章で，複数電源ドメインの消費エネルギーを最小化するために，遅延割り当て問題をランタイムに解く電圧制御アルゴリズムを提案する．

3.2 ランタイム電圧制御アルゴリズム

3.2.1 提案アルゴリズムの概要

提案アルゴリズムを単純化したフローチャートを図4に示す．提案アルゴリズムは3つのステップ (Step1, Step2, Step3) に分けられる．Step1 では，遅延割り当て (d_1, d_2) が固定され，各ドメインの $V_{DD,i}$ と $V_{th,i}$ が最適化される．本稿では，[6] で述べられた手法のみを用いて，ランタイムに Step1 を実行できることを示す．Step2 では，遅延割り当ての最適性を判定する．本稿では，遅延割り当ての最適性をランタイムに判定可能であることを示す．もし Step2 において遅延割り当てが最適でない場合，Step3 で遅延割り当ての更新を行う．提案手法ではすべてのステップを反復実行することでランタイムに複数電源ドメインの消費エネルギー最適化を実現する．

3.2.2 単一電源ドメインに対するランタイム電圧制御手法

Step1 のキーアイデアは，以下の事実に基づいている：Domain-1 に対する遅延割り当て (d_1) が固定されている限り，Domain-1 での電圧制御 ($V_{DD,1}$, $V_{th,1}$) は Domain-2 の性能 ($e_{t,2}$, d_2) に一切の影響を与えず，その逆も成立する．したがって遅延制約 (d_1) が固定されている場合，Domain-1 の消費エネルギー ($e_{t,1}$) を $V_{DD,1}$ と $V_{th,1}$ の制御のみを用いて最小化することができる．このとき， $V_{DD,2}$ と $V_{th,2}$ の調節は Domain-1 に一切の影響を与えない．したがって， d_1 や d_2 の値が固定されると，単一電源ドメインに対して提案されている既存のランタイム V_{DD} , V_{th} スケーリング

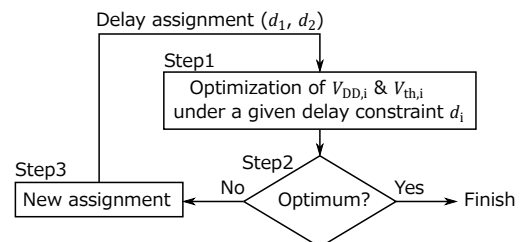


図 4 提案するアルゴリズムを単純化したフローチャート。

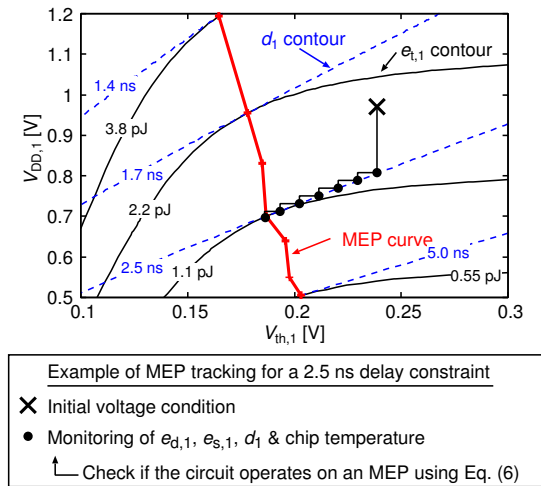


図5 Minimum Energy Point (MEP) と MEP 追跡手法の例 [6].

技術 [6] を直接適用できる。本節では、[6] で提案されているランタイム電圧制御アルゴリズムの概要を述べる。

遅延制約 d_1 の下、Domain-1 の消費エネルギー $e_{t,1}$ を最小化する問題を考える。与えられた遅延制約 d_1 を守りながら、 $e_{t,1}$ を最小化する $V_{DD,1}$ と $V_{th,1}$ の組が存在する。文献 [6] では、この最適な電圧組を Minimum Energy Point (MEP) と定義している。図5は、Domain-1 の MEP を示した図である。Domain-1 として、ファンアウト 4, 50 段インバータチェーンを使用している。図5では、電源電圧としきい値電圧が制御変数として調整されている。黒色の実線と青色の破線はそれぞれ対象回路の消費エネルギー等高線およびクリティカルパス遅延等高線である。赤色の太い実線は様々な遅延制約 d_1 に対する MEP の集合である。文献 [6] で述べられているように、MEP はクリティカルパス遅延等高線と消費エネルギー等高線の接点として発見される。したがって、Domain-1 の MEP は以下の方程式の解として求めることができる：

$$\nabla_i e_{t,i} \parallel \nabla_i d_i \Leftrightarrow \nabla_i e_{t,i} = -G_i \nabla_i d_i, \quad (6)$$

ここで、 ∇_i は微分演算子 $[\partial/\partial V_{DD,i}, \partial/\partial V_{th,i}]^T$ である。 $\nabla_i e_{t,i}$ と $\nabla_i d_i$ はそれぞれ消費エネルギー等高線とクリティカルパス遅延等高線の法線ベクトルであることに注意。したがって、MEP は2つの等高線の接点であることを(6)は述べている。式(1)–(4)より、(6)はチップ温度 ($N_{s,i}$), $e_{d,i}$ そして $e_{s,i}$ を含む方程式に変換することができる。したがって、(i) $e_{d,i}$, $e_{s,i}$, d_i およびチップ温度をランタイムにモニタし、(ii) (6)を満たす $V_{DD,i}$ と $V_{th,i}$ の組を探す、ことにより MEP をランタイムに発見できることを [6] が述べている。文献 [6] が提案した電圧制御アルゴリズムでは、図5の“Example of MEP tracking for a 2.5 ns delay constraint”のように、 $V_{DD,i}$ と $V_{th,i}$ をデルタ変調のように少しずつ変更し、徐々に MEP に接近する。本稿では、(6)に現れるパラメータ G_i が遅延割り当ての最適化において重要な役割を果たすことを次節で述べる。

3.2.3 最適な遅延割り当てのための条件

Step2 と Step3 で解決すべき問題は、遅延割り当て d_i の最適化である。文献 [1] で定義された Energy gradient

を用いて遅延割り当て問題をランタイムに解決する。一般的に、遅延制約を緩くすることで、LSI の消費エネルギーを改善できる。Energy gradient は、遅延制約を緩めることで削減可能な消費エネルギーの割合を示す。例えば、 G'_i を Domain- i に対する Energy gradient とすると、もし d_i が Δd だけ緩められたとき、Domain- i の消費エネルギーが $G'_i \Delta d$ だけ改善することを意味する。文献 [1] では、Energy gradient を数値計算により導出しているが、本稿では Energy gradient が閉形式関数で導出できることを示す。具体的には、Energy gradient G'_i は(6)に現れる G_i と等しい。詳細な証明を付録1に示す。

もし d_1 が Δd だけ緩められた場合、Domain-1 の消費エネルギーは $G_1 \Delta d$ だけ改善する。他方、 d_2 は Δd だけ厳しくなるため Domain-2 の消費エネルギーは $G_2 \Delta d$ だけ悪化する。もし G_1 が G_2 より大きい場合、この遅延緩和は全体回路の消費エネルギー削減に繋がる。しかし、[1] に示されているように、 G_1 の値は d_1 の遅延緩和とともに徐々に悪化する。結果として、以下の式が最適な遅延割り当てのための条件として成立する：

$$G_1 \Delta d = G_2 \Delta d \Leftrightarrow G_1 = G_2. \quad (7)$$

式(7)は、遅延割り当てが最適な場合、遅延割り当ての変更により消費エネルギーの改善が得られないことを示している。式(6)に(1)–(4)を代入することで、 G_i を次式に変換できる：

$$\begin{aligned} \frac{\partial e_{t,i}}{\partial V_{DD,i}} &= -G_i \frac{\partial d_i}{\partial V_{DD,i}} \\ \Leftrightarrow G_i &= \frac{2e_{d,i} + e_{s,i}}{\left(\frac{\alpha_i}{1 - V'_{th,i}/V_{DD,i}} - 1 \right) d_i}. \end{aligned} \quad (8)$$

($V'_{DD,i}$, $V'_{th,i}$) は Domain- i の MEP である。 $e_{d,i}$, $e_{s,i}$, d_i の値が判明した場合、式(8)により G_i の値を推定できる。したがって、前節と同じようにこれらパラメータの値をランタイムにモニタする回路をチップに搭載することで、 G_i をランタイムに推定することができる。

3.2.4 遅延割り当ての最適化

遅延割り当て最適化のコンセプトを図6に示す。文献 [1] で述べられているように、Energy gradient G_i は d_i に対して単調減少する。Domain-1 の遅延割り当て d_1 を大きくすると G_1 が減少する一方、Domain-2 の遅延割り当て d_2 が小さくなるため G_2 が増大する。結果として(7)を満たす遅延割り当てが唯一存在する。図6からわかるように、 $G_1 > G_2$ の場合、 d_1 を大きくすることで遅延割り当ての最適化を行うことができ、その逆も成立する。以上の事実

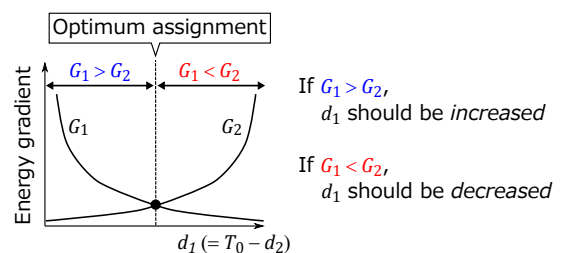


図6 遅延割り当ての最適化のコンセプト。

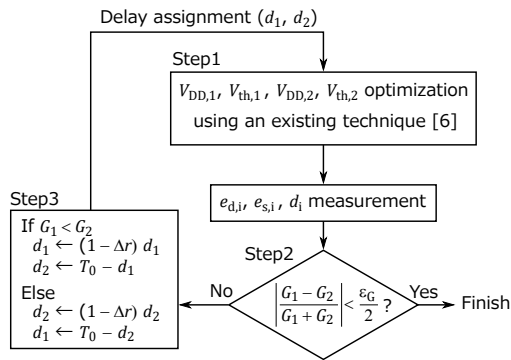


図 7 提案アルゴリズムの詳細。

に基づき、Step3 では遅延割り当てをデルタ変調のように逐次的に変更する。

提案アルゴリズムの詳細なフローチャートを図 7 に示す。Step1 では、[6] で提案されたランタイム電圧制御アルゴリズムを 2 つの電源ドメインに適用する。Step2 を開始する前に、 $e_{d,i}$, $e_{s,i}$ および d_i の値をモニタ回路を用いて取得する。例えば文献 [5] にこれらのパラメータのモニタ回路が述べられている。Step2 で遅延割り当ての最適性を評価する。もし G_1 と G_2 の相対的な差が ε_G である場合 (i.e., $G_1 \simeq G_2$ である場合)、提案アルゴリズムは終了する。もしそうでない場合、Step3 で遅延割り当てが Δr だけ変更される。逐次的にすべてのステップを繰り返すことで、電圧および遅延割り当ての最適化を実現する。

4. SOTB 65-nm プロセスを用いた検証

SOTB 65-nm 低しきい値電圧プロセスを用いたランジスタレベル回路シミュレーションを通して提案アルゴリズムの検証を行う。本稿では以下の回路を実験に用いる：

- Domain-1: 100 並列, 50 段, ファンアウト 4 インバータチェーン。活性化率は 10% である。
- Domain-2: 100 並列, 50 段, ファンアウト 4 インバータチェーン。活性化率は 1% である。

ここで、Domain-1 と Domain-2 はそれぞれプロセッサのロジック回路およびメモリ回路を模擬している。各回路の電源電圧と基板電圧を幅広く調節し、性能を評価する。提案アルゴリズム実装のために、 α_i や $N_{s,i}$ といった物理パラメータを事前に求める必要がある。これらの値はランジスタの DC シミュレーションにより求める。 $V_{th,i}$ の値を求めるため、本稿では定電流法を用いる。パラメータ ε_G , Δr の値をそれぞれ 15%, 5% とする。本稿では、理想的なモニタ回路を用い、 $e_{d,i}$, $e_{s,i}$, d_i , チップ温度の正確な値をランタイムに推定可能と仮定する。提案アルゴリズム実装にあたり、遅延割り当ての初期値を与える必要がある。本稿では d_1 と d_2 の初期値を共に $T_0/2$ とする。

遅延制約 5 ns (i.e., $T_0 = 5$ ns) に対する最適化結果を図 8 に示す。縦軸および横軸は各ドメインの電源電圧および基板バイアスである。ここで、図の右側に進むと回路のしきい値電圧が増大することに注意。黒い実線は、Step1 において既存の MEP 追跡技術を用いた際の電圧制御結果の軌跡であり、“+”が導出された MEP である。本稿では、電源電圧および基板電圧をそれぞれ 5 mV ずつ調節

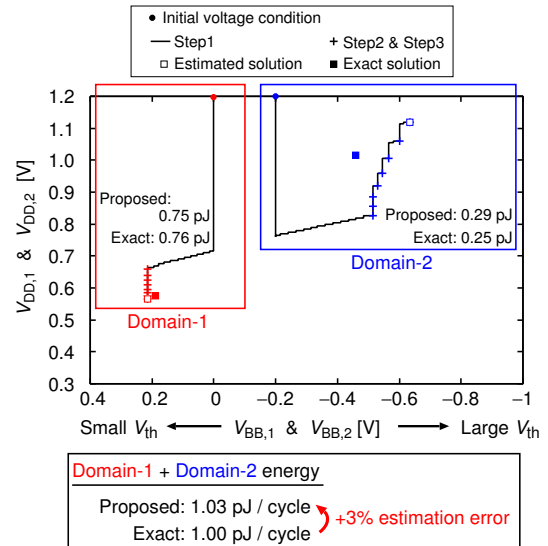


図 8 遅延制約 5 ns に対する電圧の最適化結果。

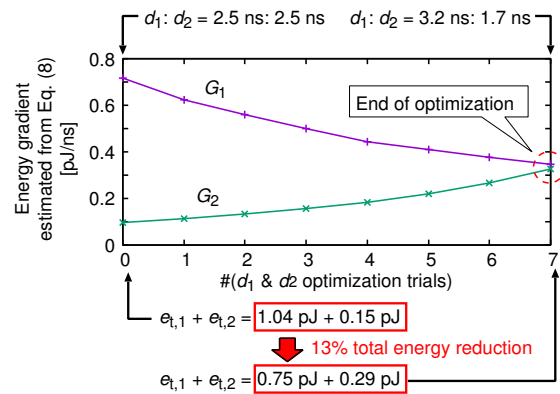


図 9 遅延制約 5 ns に対する G_1 と G_2 の最適化の過程。

し、MEP の探索を行っている。対象回路の総消費エネルギー ($e_{t,1} + e_{t,2}$) は提案手法により 1.03 pJ/cycle になる一方、全探索法で求めた厳密解は 1.00 pJ/cycle である。したがって、提案手法により +3% の以内の誤差で最適解を探索することができた。

T_0 を 5 ns としたときの、 G_1 と G_2 の最適化の過程を図 9 に示す。横軸は遅延割り当ての最適化の回数 (Step3 を実行した回数) を示す。Step3 は、図 8 の“+”で行われている。初期状態では G_1 は G_2 より大きいので、 d_2 より大きい遅延を d_1 に割り当てることで総消費エネルギーを削減できる。 G_1 の値は d_1 の増大と共に減少するため、 G_1 と G_2 は Step3 を繰り返す度に等しい値に近づく。図 8 が示すように、遅延割り当てを最適化することで、初期状態と仮定して最大 13% 消費エネルギーを削減できることを確認した。以上の事実は遅延の割り当てが複数電源ドメイン最適化において重要な役割を果たすことを示唆している。

遅延制約 (T_0) が 5 ns から 20 ns に緩和された時の最適化結果を図 10 に示す。図 8 の時と同様に、提案アルゴリズムにより発生する消費エネルギーの誤差は +2% である。文献 [8,9] のように、2 つのドメインの V_{DD} のみをスケールリングした時を考える。各ドメインの V_{th} は、遅延制約 (T_0) が 5 ns の時の最適値で固定されていると仮定する。

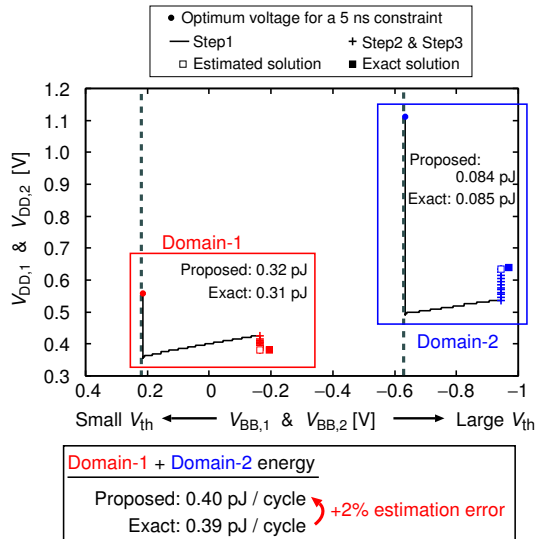


図 10 遅延制約 20 ns に対する電圧の最適化結果。

この場合、Domain-1 と Domain-2 は図 10 の破線部で稼働する。もし T_0 が 5 ns から 20 ns に緩和された場合、電源電圧のみ調節で $e_{t,1} + e_{t,2}$ は 0.49 pJ/cycle まで削減することができる。一方、提案手法で V_{DD} と V_{th} を同時に最適化することで消費エネルギーは 0.40 pJ/cycle まで削減され、この差は 18% に対応する。以上の結果は、複数電源ドメインの省エネルギー動作実現において、 V_{DD} だけでなく V_{th} の調節も重要であることを示している。

5. 結論

クリティカルパス上に 2 つの電源ドメインが存在する LSI の消費エネルギー最適化手法を提案した。エネルギー効率の良い LSI を実現するためには、 V_{DD} だけでなく V_{th} の動的な調節が重要である。本稿では、与えられた要求動作速度を守りながら、2 つの電源ドメインの V_{DD} と V_{th} をランタイムに制御する電圧制御手法を提案した。SOTB 65-nm プロセスを用いたシミュレーションを通し、提案アルゴリズムが最大 3% 以下の誤差で消費エネルギーを最適化可能であることを示した。 V_{DD} のみをスケールリングする従来手法と比較し、 V_{DD} と V_{th} を同時最適化することで最大 18% 消費エネルギーを削減可能であることを示した。DC-DC コンバータや LVC のオーバーヘッドを考慮した電圧スケールリング手法の開発が今後の課題である。

付録 1: Energy Gradient の導出

Domain- i の遅延割り当てを Δd だけ緩めることで、Domain- i の消費エネルギーを $-\Delta e_i$ だけ改善する状況を考える。遅延割り当てを Δd だけ緩和することにより、Domain- i の MEP (i.e., Step1 の最適化結果) が $(\Delta V_{DD,i}, \Delta V_{th,i})$ だけ移動すると仮定する。ここで、 $-\Delta e_i$ は正の値であることに注意。 $\Delta V_{DD,i}$ と $\Delta V_{th,i}$ が十分小さい場合、以下が成立する：

$$\Delta d = \frac{\partial d_i}{\partial V_{DD,i}} \Delta V_{DD,i} + \frac{\partial d_i}{\partial V_{th,i}} \Delta V_{th,i}. \quad (9)$$

式 (6) が Domain- i 上の MEP で成立するため、(9) は次のように変換できる。

$$-G_i \Delta d = \frac{\partial e_{t,i}}{\partial V_{DD,i}} \Delta V_{DD,i} + \frac{\partial e_{t,i}}{\partial V_{th,i}} \Delta V_{th,i} \Leftrightarrow G_i = \frac{-\Delta e_i}{\Delta d}.$$

したがって、 G_i は Domain- i の Energy gradient である。

謝辞

本研究は JSPS 科研費 (16H01713, 17H01712) による支援によって行われた。本研究は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、シノプシス株式会社の協力で行われた。

参考文献

- [1] L. Yan, J. Luo, and N. Jha, "Joint Dynamic Voltage Scaling and Adaptive Body Biasing for Heterogeneous Distributed Real-Time Embedded Systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 7, pp. 1030–1041, July 2005.
- [2] S. Martin, K. Flautner, T. Mudge, and D. Blaauw, "Combined Dynamic Voltage Scaling and Adaptive Body Biasing for Lower Power Microprocessors under Dynamic Workloads," in *International Conference on Computer Aided Design*, Nov 2002, pp. 721–725.
- [3] A. Basu, S.-C. Lin, V. Wason, A. Mehrotra, and K. Banerjee, "Simultaneous Optimization of Supply and Threshold Voltages for Low-Power and High-Performance Circuits in the Leakage Dominant Era," in *Design Automation Conference*, July 2004, pp. 884–887.
- [4] K. Nose and T. Sakurai, "Optimization of VDD and VTH for Low-power and High Speed Applications," in *Asia and South Pacific Design Automation Conference*, Jan 2000, pp. 469–474.
- [5] M. Nomura, Y. Ikenaga, K. Takeda, Y. Nakazawa, Y. Aimoto, and Y. Hagihara, "Delay and Power Monitoring Schemes for Minimizing Power Consumption by Means of Supply and Threshold Voltage Control in Active and Standby Modes," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 4, pp. 805–814, April 2006.
- [6] S. Hokimoto, T. Ishihara, and H. Onodera, "Minimum Energy Point Tracking Using Combined Dynamic Voltage Scaling and Adaptive Body Biasing," in *International System-on-Chip Conference*, Sept 2016, pp. 1–6.
- [7] K. Usami, M. Igarashi, F. Minami, T. Ishikawa, M. Kanazawa, M. Ichida, and K. Nogami, "Automated Low-Power Technique Exploiting Multiple Supply Voltages Applied to a Media Processor," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 3, pp. 463–472, Mar 1998.
- [8] T. Kuroda and M. Hamada, "Low-power CMOS digital design with dual embedded adaptive power supplies," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 4, pp. 652–655, April 2000.
- [9] T. Kuroda, "Optimization and Control of VDD and VTH for Low-Power, High-Speed CMOS Design," in *International Conference on Computer Aided Design*, Nov 2002, pp. 28–34.
- [10] N. Kamae, A. K. M. M. Islam, A. Tsuchiya, and H. Onodera, "A Body Bias Generator with Wide Supply-Range down to Threshold Voltage for Within-die Variability Compensation," in *Asian Solid-State Circuits Conference*, Nov 2014, pp. 53–56.
- [11] T. Sakurai and A. Newton, "Alpha-Power Law MOSFET Model and its Applications to CMOS Inverter Delay and Other Formulas," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 2, pp. 584–594, Apr 1990.