

3. Power Wall 問題への ブレークスルーを目指して ～リーク電力削減への試み～

中村 宏 (東京大学)

Power Wall 問題

携帯電話・端末、ゲーム機器、デジタルハイビジョンテレビといったフロントエンド系から、基幹サーバ、スーパーコンピュータといったバックエンド系に至るまで、高性能なコンピュータシステムが我々の生活の中で多く利用されており、より高度で快適な情報化社会を実現するために、コンピューティングのさらなる高性能・高機能化が求められている。しかし、高性能化・高機能化には VLSI 内部のトランジスタのスイッチング回数の増加が必然的に要求されるため、消費電力の増大を引き起こす。

文献 1) によると、2006 年の時点で IT 機器の消費電力はすでに日本全体の消費電力の約 5% を占め、2025 年にはそれが 20% を超えると予測されている。これは低炭素社会を目指す上で容認されるものではなく、より快適な高度情報化社会の実現には、IT 機器の低電力化が必須となっている。

高性能化と低電力化の相反する要求に対し、用途に応じてどちらかをより優先させる、すなわち、電源や冷却環境が厳しいフロントエンド系では低電力化を、性能要求が厳しいバックエンド系では高性能化をより重視することで、コンピュータシステムの設計・開発は行われてきた。しかし、今では性能が最も重要視されるスーパーコンピュータでさえその高性能化は消費電力によって抑えられるようになってきている。電力消費によって発生する熱量を冷却能力より低く抑えないとシステムの動作温度が上昇し正常に動作させることができなくなるからである。図-1 に、文献 2) で報告されている、2009 年 11 月

時点での、世界最高速 200 位までのスーパーコンピュータの性能と消費電力を示す。図中の直線は、200 位までの電力/性能の平均値を傾きとするものであり、性能が 1 位と 2 位のものは○で囲んである。この図から分かるように、きわめて高い性能を達成するコンピュータは電力/性能の値が小さい。これは、スーパーコンピュータの高性能化が、消費電力によって抑えられていることを意味する。

このように、消費電力の壁 (Power Wall) 問題は、より高度で快適な情報化社会を実現するために解決せねばならない重要な課題となっている。

ブレークスルーを目指して

▶ 消費電力と性能の関係

VLSI における消費電力には、スイッチング動作に伴って消費されるダイナミック電力と、漏れ電流 (リーク電流) に伴って消費されるリーク電力^{☆1} がある。前者のダイナミック電力は、半導体素子の微細化に伴い、トランジスタ当たりの値としては小さくなるもののチップに搭載されるトランジスタ数の増大により増加する。一方、後者のリーク電力に関しては、以前は無視できるほど小さかったものの、半導体素子の微細化に伴いトランジスタ当たりの値としても大きくなるため、搭載トランジスタ数の増大とあいまって、より深刻な問題となっている³⁾。

電力を消費するのは半導体素子であるから、低消費電力化の唯一の手段は、トレードオフ関係にある

^{☆1} スイッチング動作とは無関係であることから、ダイナミック電力の対語としてスタティック電力とも言う。

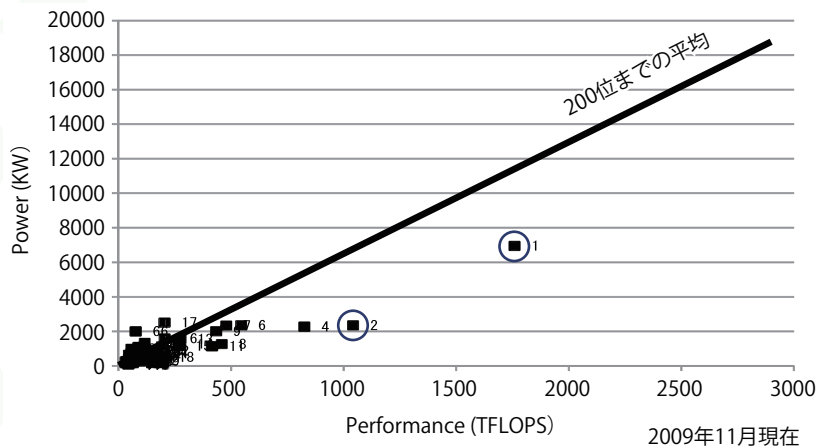


図-1 スーパーコンピュータの性能と消費電力の関係

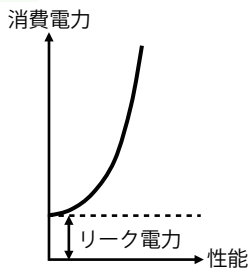


図-2 スイッチング素子の性能と電力の関係

各素子の性能と電力を制御することである。たとえば、DVFS (Dynamic Voltage Frequency Scaling : 動的周波数・電源電圧制御) は、電源電圧レベルを変動させることで、性能と消費電力を調節する機構であり近年のVLSIでは広く用いられている。電源電圧レベルを変化させると、スイッチング素子の性能と電力は図-2のような関係を示す(関係式導出の詳細は文献3)などを参照されたい)。したがって、ある半導体素子に必要とされる性能が与えられれば、その性能を達成可能な最低電圧レベルを設定することで、その素子の消費電力を最低限に抑えることができる。

しかしながら、コンピュータシステム全体の性能と電力の関係は決してこのように単純には表せない。電力は構成要素の電力の総和となるのに対し、性能はボトルネックとなる構成要素のみの性能で決まるからである。したがって、システム全体の消費電力を低減するには、性能上のボトルネックになっていない構成要素を低速・低電力で動作させることが重要となる。

▶ 設計階層間協調の重要性

システム全体の低消費電力化は、各設計階層がそれぞれ独立の役割を担うことで推し進められてきた。すなわち、回路技術は構成要素の性能と電力を制御するための調節弁を提供し、アーキテクチャあるいはそれより上位階層はシステム全体の中で不要不急の動作が「いつ」「どこに」あるのかを見つけ、回路技術によって提供される調節弁を適切に活用する役割を担ってきた。

低電力化の回路技術としては、クロック信号を停止させるクロックゲーティング、先に述べたDVFS、電源遮断を行うパワーゲーティング、トランジスタの基板にバイアス電圧を印加しトランジスタの閾値電圧を制御する基板バイアス技術等がある。これらを電力・性能の調節弁として考えた場合、その特性としては、電力削減効果が大きいことだけでなく空間的にも時間的にも細粒度な調節が可能であることが望ましい。粒度が小さいと調節の適用範囲を広げられるからである。たとえば先に述べたクロックゲーティングは、クロック信号ごとにクロックを抑制するANDゲートを付加することで容易に実現できるためきわめて細粒度な制御が可能である。そのため、適用可能個所を回路レベルで判断できる扱いやすい技術であり、クロックゲーティングを適用した論理合成も可能となっている。一方、他の回路技術はいずれも制御に時間を要し、空間的粒度を細かくすることも容易ではない。半導体素子の微細化に

伴って単位時間・単位空間あたりのアクティビティがこれからも指数的に増大することを考えると、不急の動作が存在する時間的・空間的粒度はより小さくなる。したがって、前述の回路技術を有効に活用するためには、システムにおいて不要不急の動作が「いつ」「どこに」あるのかを単に見つけるだけではなく、アーキテクチャレベルにおいてシステム内のアクティビティを回路技術が提供する調節弁の制御粒度に積極的に合わせる必要がある。今後さらなる低消費電力化を実現するためには、このような設計階層間の協調が必須となる。

▶ データレジデントコンピューティング： アーキテクチャの新パラダイム

このように考えると、低電力化を実現するためには、コンピュータシステムの設計思想を従来とは変える必要性が生じる。従来の設計思想は、集積度向上を頼りに多くのトランジスタを投入し、投入するトランジスタをいかに性能向上に結びつけるかがコンピュータアーキテクト(コンピュータの設計者)の大きな仕事であり、その後で、性能重視の設計思想に基づいて構成されたシステムにおいて、不要不急の動作が「いつ」「どこに」あるのかを見つけることが低電力化アーキテクチャ技術であった。これに対し、低電力化においてブレークスルーを達成するには、設計思想を大転換し、動作すべき真に必要なトランジスタを必要時に必要部だけ動作させるような設計を最初からすべきであろう。

このためには、アーキテクチャの新しいパラダイムが必要となると私は考えている(図-3)。アーキテクチャは設計階層においてソフトウェアとハードウェアのインタフェースとなる重要な抽象化レベルであるが、その捉え方は時代とともに変わってきている。メインフレームが主流の時代にはハードウェアが実現する機能としての命令セットを考慮してソフトウェアを実現していた。そのため命令セットアーキテクチャがインタフェースレイヤとしての抽象化レベルであった。その後1980年代にRISCが登場し、命令パイプライン処理に代表される命令実

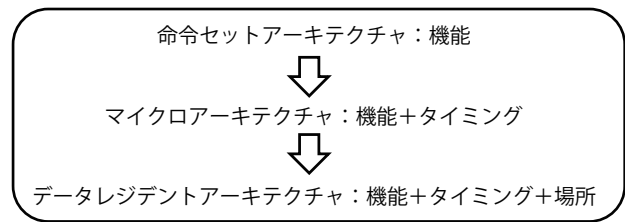


図-3 アーキテクチャのパラダイムシフト

行制御方式が用いられるようになり、そのハードウェア機構を活用する命令スケジューリングも重要な技術となった。これは、ハードウェアによって実現される機能だけでなく、処理のタイミングをも考慮してソフトウェアが実現されるべきであることを意味し、命令実行制御方式までも含むマイクロアーキテクチャがインタフェースレイヤとしての重要な抽象化レベルとなった。しかるに現在、低消費電力という新しい目標に対しては、機能とタイミングだけでなくデータに対する処理の場所をも含む概念をインタフェースレイヤとして考えるべきであろう。この概念をデータレジデントアーキテクチャ(data resident architecture)と呼び、回路技術が提供する調節弁の制御粒度に、システム内のアクティビティをハードウェアからもソフトウェアからも可制御にすることで画期的な低消費電力化を実現することを提唱している⁴⁾。この方針のもと、科学技術振興機構の戦略的創造研究推進事業(CRESTタイプ)において、研究課題名「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」を平成18年度より実施している。次章では、その中で取り組んでいるリーク電力削減技術について説明する。

リーク電力削減への試み

▶ 細粒度パワーゲーティング

パワーゲーティングは、図-4に示すようにパワースイッチと呼ばれるリーク電流の少ないトランジスタをVDD(電源)と対象回路の間、または、GND(接地)と対象回路の間に挿入し(図-4は後者)、対象回路が利用されていないときにこのパワースイッチを切り、電源を遮断するスリープ状態とするこ

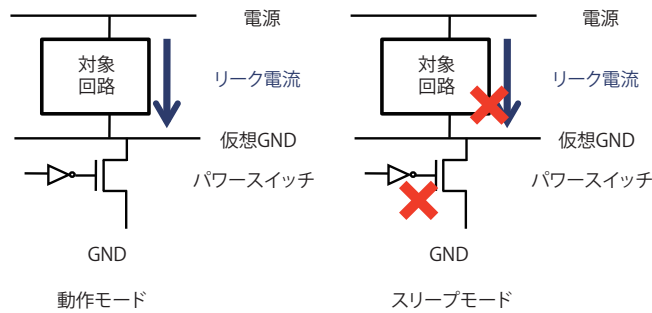


図-4 パワーゲーティングの原理

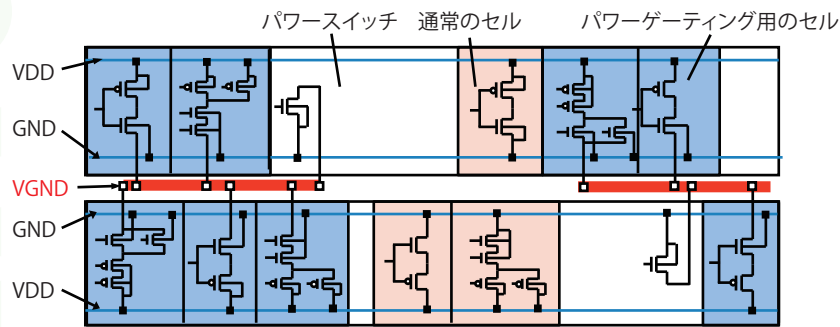


図-5 細粒度パワーゲーティング回路

とによりリーク電力を削減する手法である。

パワーゲーティング手法自体は新しい回路技術ではないが、これまでは、通常のゲートを用いて設計した対象回路の周囲に仮想VDD用の電源リングを設け、この仮想VDDとVDDの間にパワースイッチを挿入していた。しかし、この方法では電源制御の空間的粒度が大きく、またパワースイッチがオフであるスリープモードとオンである動作モードの間の遷移にはマイクロ秒単位の時間を要した。そのため、これまでは長い期間スリープ状態を保つことが期待され空間的にも大きい領域、たとえばアクセラレータ、I/Oモジュール、メモリなどが対象であった。Intel社は文献5)においてコア部に対してパワーゲーティングを適用したNehalemプロセッサの発表を行ったが、時間的・空間的粒度はやはり大きい。

これに対し、細粒度なパワーゲーティングを実現できればさらなる低電力化が可能となる。たとえば、動作中においてもプロセッサコアの演算ユニットの稼働率はそれほど高いわけではない。高速性が要求される演算ユニットではリーク電流の大きい高速ト

ランジスタが用いられているため、これらの部分にパワーゲーティング手法を適用できればその効果は大きいと考えられる。しかし、演算ユニットは定常的に使用されるため、その適用には時間的に細粒度な制御が必要となる。

そこで、我々はそれぞれのセルごとに仮想GNDを設けてこれらを一定数接続し、この仮想GNDとGNDとの間にパワースイッチを設ける回路構成を採用した。図-5に、この手法に基づく実際のセル配置を示す。従来のパワーゲーティングとは異なり、1つの仮想GNDを数個程度のセルが共有している。この方法は仮想GNDを持つ新しいセルを必要とするが、スリープトランジスタ数を調整することで、スリープ状態への遷移やウェイクアップ状態への復帰に要する時間をナノ秒単位の短縮することができる。

この回路構成により正しく動作する細粒度パワーゲーティングは実現できる。しかし、状態遷移時にエネルギーを消費するという問題が残されているため、遷移頻度が大きいと逆に消費電力は増大する。この様子を図-6に示す。状態遷移によるエネ

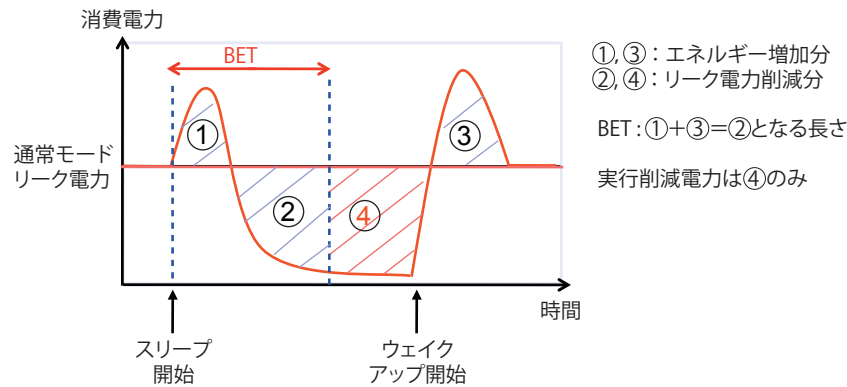


図-6 状態遷移時の消費エネルギー

ルギー損失よりパワーゲーティングによる削減エネルギーが大きくなるまでの時間を、ブレイクイーブンタイム (BET: Break Even Time) で表す。BET はパワーゲーティングの損益分岐点であり、回路技術が電力・性能調節弁として提供可能な最小時間粒度と考えることができる。したがって、演算ユニットのアクティビティをこの最小時間粒度でまとめるデータレジデントアーキテクチャが低電力化には必要である。

▶ 細粒度パワーゲーティングマイクロプロセッサ Geyser

細粒度パワーゲーティングによるリーク電力の削減を実証するプロトタイプとして、我々は汎用マイクロプロセッサ Geyser を開発している。Geyser は MIPS アーキテクチャ互換で、単純な 5 段階命令パイプラインを採用し in-order で命令を実行する。基本的なパワーゲーティング動作は、命令フェッチ時に命令をプリデコードすることで必要な演算ユニットのみをウェイクアップし、演算の実行後は速やかにスリープ状態に戻すものである。しかし、前節で述べたようにスリープとウェイクアップを頻りに繰り返すとむしろ電力を増加させてしまうため、演算ユニットの非使用期間が BET より長い時のみスリープ状態に戻すべきである。この問題に対し、各演算ユニットの非使用期間は命令スケジューリングを行うコンパイラが判断できることに着目し、スリープすべきか否かを指定できる特殊ビットを持つ演算命

令を用意することで、コンパイラによるパワーゲーティング戦略の指示を可能とした。これは、データレジデントアーキテクチャの設計思想に基づくものである。しかしこれだけでは十分ではない。BET はリーク電流の値に依存し、リーク電流の値は実行状況で変化する温度に強く依存するからである。後述の我々の設計では、温度が 100℃ の場合には BET は 50 ナノ秒程度となるが 25℃ では 500 ナノ秒程度にもなる。数サイクル程度の非使用期間はコンパイラが正確に把握できるものの、100 サイクル以上の非使用期間は分岐やメモリアクセスなどの動的状況で変化するため、コンパイラで正確に把握することは難しい。このため、キャッシュミス時には長期間演算ユニットが非使用状態になることに着目し、キャッシュミス時にスリープ状態にする、という別のパワーゲーティング戦略も用意する。さらに、動的に変化する BET を OS が実行時に正確に把握し、状況に応じて OS がパワーゲーティング戦略を切り替えるようにする。このように、Geyser では、システム設計階層間の協調に基づいてパワーゲーティング方式を最適に制御することを目指している。

我々は、富士通 e-shuttle 65nm プロセスを用いてプロトタイプチップ Geyser-1 を試作した^{☆2}。図-7 に、そのレイアウト図とチップ写真を示す。チップサイズは 2.1mm × 4.2mm である。レイアウト図中のリークモニタは動作中にリーク電流値を測

☆2 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)・日本ケイデンス(株)の協力で行われた。

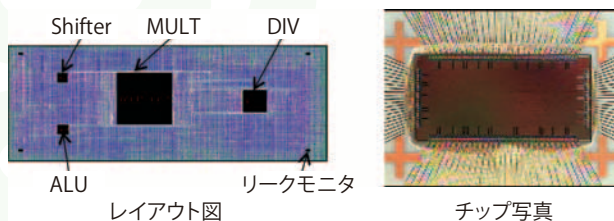


図-7 Geysler-1 プロセッサ

定する回路であり、その測定結果から動的に変化する BET を OS が正確に把握し、BET に応じた最適なパワーゲーティング戦略を選択する。Geysler-1 は、細粒度パワーゲーティングの効果を実チップで示すことに目的を絞り、周波数方向に設計を最適化しなかったため動作周波数は 60MHz にとどまったが、サイクルごとの細粒度パワーゲーティングが正しく動作することを実証することに成功した。また、80°C で 10MHz 動作をさせた場合、細粒度パワーゲーティングを用いることで、MiBench ベンチマークの Quick Sort, Dijkstra プログラムの消費電力をそれぞれ 29% と 24% 削減でき、その有効性も合わせて示すことができた。Geysler-1 の詳細については文献 6) を参照されたい。

今後の展望

本稿では、Power Wall 問題解決のブレークスルーを実現するにはシステム設計階層間の協調が重要であることを述べ、近年深刻になっているリーク電力を削減する試みとして、ハードウェアとソフトウェアの新しいインタフェースレイヤとしてデータレジデントアーキテクチャを導入して研究を進めている Geysler プロセッサを紹介した。

コンピュータシステムに対してさらなる高性能・高機能化が求められている限り、Power Wall 問題は存在し続ける。さまざまな低電力化技術で壁を遠くへ動かすことはできても壁が消えることはなく、コンピュータの高性能・高機能化が進むと、再びその壁は我々の前に立ちはだかる。

電力は構成要素の電力の総和で決まり、性能はボトルネックとなる構成要素のみの性能で決まる、と

いうのは、システム全体の電力と性能の本質的な関係であり、性能上のボトルネックとなっていない構成要素の動作を低速・低電力にする、すなわち無駄な消費電力を省くことが、システム全体の性能/電力を向上させるために必要である。今後従来の半導体とはまったく異なるデバイスが利用されるかもしれないし、コンピュータシステムの利用形態も大きく変わるかもしれない。その場合も、デバイスからアプリケーションまで、あるいはさらに広範囲な階層間の協調の必要性は増えこそすれ、減ることはない。

謝辞 CREST 研究の共同研究者として日頃ご議論いただき、芝浦工業大学宇佐美公良教授、慶應義塾大学天野英晴教授、東京農工大学並木美太郎教授、電気通信大学近藤正章准教授に感謝いたします。

参考文献

- 1) 経済産業省:「グリーン IT イニシアティブ」の推進に向けて、http://www.meti.go.jp/discussion/topic_2008_06/main_03.htm (2008).
- 2) TOP500 : <http://www.top500.org>
- 3) Kim, N. S. et al. : Leakage Current : Moore's Law Meets Static Power, IEEE Computer, Vol.36, No.12, pp.68-75 (2003).
- 4) 中村 宏, 他: 革新的電源制御による超低消費電力高性能システム LSI の構想, 情報研報 ARC-173(14), pp.79-84 (2007).
- 5) Kumar, R. et al. : A Family of 45nm IA Processors, pp.58-59, ISSCC09 (2009).
- 6) Ikebuchi, D. et al. : Geysler-1, A MIPS R3000 CPU Core with Fine Grain Runtime Power Gating, Proc. of IEEE ASSCC2009, pp.281-284 (2009).

(平成 22 年 5 月 9 日受付)

中村 宏 (正会員) nakamura@fw.ipsj.or.jp

平成 2 年東京大学大学院工学系研究科博士課程修了, 工学博士. 筑波大学電子・情報工学系助手, 東京大学先端科学技術研究センターを経て, 平成 20 年より東京大学大学院情報理工学系研究科准教授. 高性能・低電力・高信頼コンピューティングの実現を目指す. 本会論文賞 (平成 5 年), 山下記念賞 (平成 6 年), 坂井記念特別賞 (平成 13 年), 各受賞. ACM, IEEE, IEICE 各会員.