

# Heterogeneous Microarchitectures Trump Voltage Scaling for Low-Power Cores

著者: Lukefahr, A., Padmanabha, S., Das, R., Dreslinski Jr, R., Wenisch, T. F., & Mahlke, S.

出典: *Proceeding of the 23rd int'l Conference on Parallel Architectures and Compilation Techniques. 2014. pp. 237-250.*

発表者: 高性能コンピューティング学講座 本多・三輪研究室 1553007 澁谷俊憲

## 1 はじめに

プロセッサのエネルギー効率の向上には、処理性能とエネルギー効率にトレードオフが存在する。エネルギー効率の向上の技術には DVFS と HMs が知られている。これらの技術は、処理速度を低下させてもよい場合に、各手法でプロセッサの性能を低下させてエネルギーの節約を図るものである。しかし、実行状態を切り替える際に多大なオーバーヘッドを生じるため、切り替えを判断するまでの命令区間 (Quantum) が広くせざるを得ない問題があった。

### Dynamic Voltage and Frequency Scaling(DVFS)

DVFS は、実行時のプロセッサの消費電力が周波数と電圧に比例することを利用した技術である。低い電圧状態に遷移させることにより消費電力を下げる事が可能だが、同時に周波数も下げる必要が生じる。

### Heterogeneous Microarchitectures(HMs)

HMs は、異なる性能を持つ複数のコアを用意し、状況に応じて使い分ける技術である。高性能だが消費電力の高い big コアと低性能だが消費電力の低い little コアを併用し、big コアの高い性能が十分活用されない場合に、処理の実行コアを little コアに切り替えることでエネルギー削減を図る。

## 2 先行研究

各エネルギー削減方式は実行状態切り替えの際に大きな遷移レイテンシが生じていたが、近年では高速に実行状態の切り替えを行える技術も開発されてきた。DVFS では、Le らがレギュレータをオンチップ化し、ワイヤ長と大容量キャパシタを不要にしたことで、従来数十  $\mu$ s 要した電圧切替を 20ns 未満で行うことを可能にした [1]。HMs では Lukefahr らが、2 種類のコアのフロントエンドを共有することで遷移の際のデータ転送量を減らし、10M 命令程度の切り替え判断間隔を 1K 命令程度まで短縮することができた [2]。

## 3 目的

一定の性能低下を許容する性能制約下でエネルギーを可能な限り抑えるには、どの場所 (区間) で、どの技術を、どの程度適用するかを適切に割り当てる必要がある。また、表 1 は HMs における 3 区間で構成されたプログラムの遅延、エネルギーの例である。この組み合わせは  $2^3$  通

表 1: 各区間の遅延、消費エネルギー例 (HMs のみに単純化)

core	Quantum1		Quantum2		Quantum3	
	Delay	Energy	Delay	Energy	Delay	Energy
Big	10ms	50mJ	20ms	60mJ	30mJ	60mJ
Small	20ms	20mJ	40ms	50mJ	35mJ	40mJ

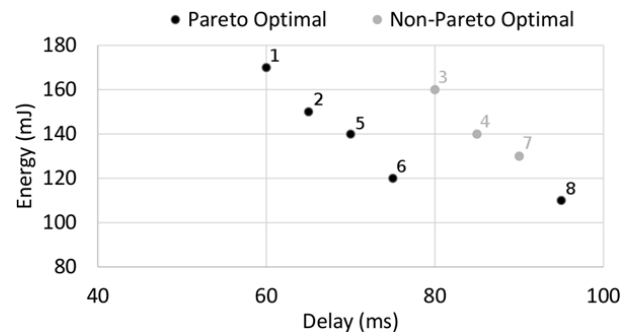


図 1: パレート最適スケジュール

りとなり、スケジューリング候補は 8 通りとなるが、高速な切り替えが可能な現在、実際の区間数は更に増加し、DVFS の状態数も考慮する必要がある。そのため、組み合わせは膨大な数となり、トレードオフの計算には、事前計算をするための高速なスケジューリング探索アルゴリズムが求められる。

本研究は、適切にスケジューリングされた各手法が、各々の程度エネルギー節約に貢献するかの相対的な評価を目的としている。そのため、高速にスケジューリング探索ができる計算ツール PaTH (Pareto-optimal Tradeoffs for Heterogeneity) を開発した。

## 4 パレート最適スケジュール

パレート最適とは、他のある要素を低下させない限り、すべてのいずれかの要素を向上できない状態のことを指す。今回の“要素”は、消費エネルギーと遅延時間である。スケジュール  $X, Y$  において、消費エネルギー  $Energy(X) > Energy(Y)$  かつ遅延  $Delay(X) > Delay(Y)$  となる  $Y$  が存在しない、またその場合にのみ、 $X$  をパレート最適スケジュールという。

表 1 の 8 通りの組み合わせをプロットしたのが図 1 であり、グラフ中の 5 つの黒点がパレート最適スケジュールである。この各点を曲線で補間するとパレート最適曲線となる。PaTH は分割統治法によりパレート最適曲線の近似曲線を高速に求める。

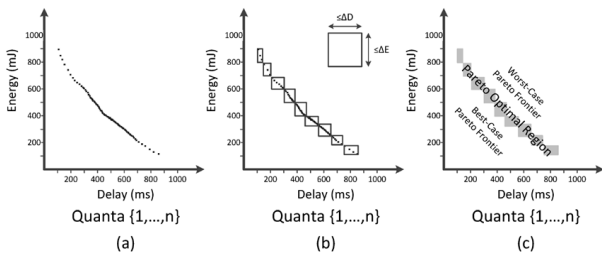


図 2: パレート最適曲線

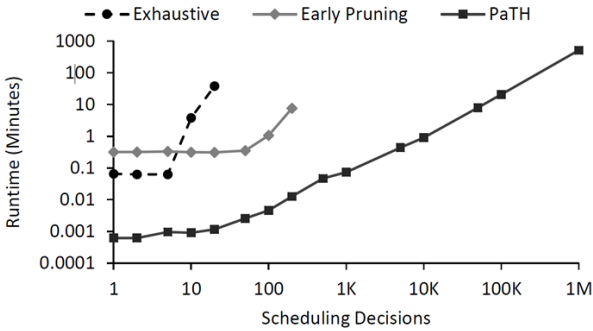


図 3: PaTH 計算時間

## 5 PaTH

本来、パレート最適曲線を求めるには全探索が必要となる。その計算量を削減するために、パレート最適領域を設定する。

図 2(a) はある Quantum 群におけるパレート最適スケジュールである。高消費エネルギー、低遅延である点から所定の誤差範囲に収まる点を長方形領域でグループ化する (b)。この領域をパレート最適領域と呼ぶ。領域内のすべての点を領域の右上 (この点に対して領域内のすべての点がパレート最適)、左下 (領域内のすべての点に対してパレート最適) のそれぞれに近似する。こうして作成した領域を同様の処理を施した次の Quantum 群と畳み込みを行うことで、パレート最適曲線に近似していく (c)。このとき、真のパレート最適曲線は、各領域の右上点と左下点の間に存在する。

ここで、表 1 の B → L から始まる結果は、L → B の場合よりも低効率である。このような場合が見つければ B → L で始まるスケジュールを予め (a) よりも前に枝刈りすることで計算量を削減できる。図 3 は PaTH のスケジュール決定までに要した時間である。全探索の 10 分間に 300 回以下に対し、PaTH は約 60000 回決定した。

## 6 HMs と DVFS の比較

PaTH を用いて DVFS と HMs がどれほどエネルギーの節約に貢献しているかを調査した。DVFS は理想的な DVFS と切り替えに制限があるデュアルレール DVFS (DR) を仮定し、HMs は判断間隔が 1K である Composite core [2] と、判断間隔を 10M に設定した ARM big.LITTLE を仮定した。PaTH の計算に用いる性能モデルは gem5 で計測したものを使用した。

図 4 は PaTH によって求められたパレート最適曲線である。HMs は DVFS よりも高いエネルギー節約を与えることがわかった。また、性能制約 60% まででは HMs と DVFS を組み合わせて用いても HMs 単体に比べて 3~4% ほどしか向上しないことがわかった。これは DVFS が

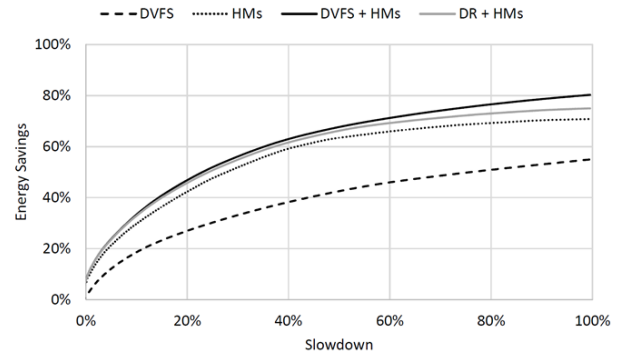


図 4: 削減技術間の比較

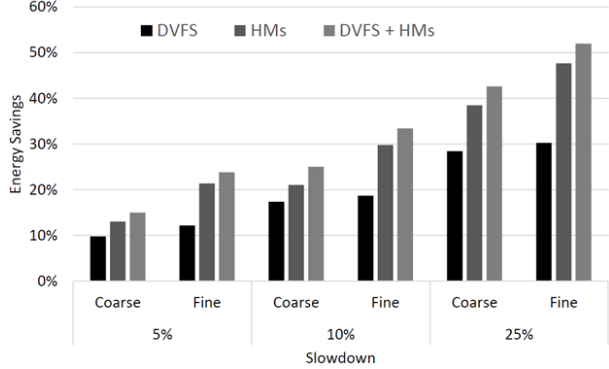


図 5: Quantum の大きさの違いによる比較

メモリバウンドな処理が多い時に有効な技術であることにに対し、HMs の削減は、それ以外の処理において性能を低くできる多くの他の区間にも対応することが原因である。

図 5 は判断間隔の長さによる比較である。同じ性能制約下において、間隔が長い場合に対して間隔を短くした際の処理は DVFS は 2%、HMs は 7~10% の向上が見られる。また、間隔を短くした HMs は広い間隔のどの方式よりもエネルギー削減を行える事がわかった。それぞれの間隔で DVFS での向上の差が小さい原因は、L2 キャッシュミスなど DVFS を行うべき区間を長い間隔、短い間隔どちらでも認識できるのに対し、HMs は間隔が短い場合のみ認識して切り替えを行う区間があるためである。

## 7 まとめ

パレート最適なスケジュールを現実的な時間で計算できるツール PaTH を開発した。PaTH を用いて DVFS と HMs、ならびにそれらの組み合わせによる性能、エネルギーのトレードオフを計算し、広く HMs が DVFS より高いエネルギー効率を与えることを示した。

## 参考文献

- [1] H.-P. Le, *et al.*, “A sub-ns response fully integrated battery-connected switched-capacitor voltage regulator delivering 0.19 w/mm<sup>2</sup> at 73% efficiency,” in Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International. IEEE, 2013, pp. 372-373.
- [2] A. Lukefahr, *et al.*, “Composite cores: Pushing heterogeneity into a core,” in Proc. of the 45th Annual International Symposium on Microarchitecture, 2012, pp. 317-328.