

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 10 日現在

機関番号：12601
研究種目：若手研究(B)
研究期間：2012～2013
課題番号：24700044
研究課題名(和文) ヒート・スプレッド指向プロセッサに関する研究

研究課題名(英文) A Study of Heat-Spreading-Aware Processor

研究代表者

三輪 忍 (Miwa, Shinobu)

東京大学・情報理工学(系)研究科・助教

研究者番号：90402940

交付決定額(研究期間全体)：(直接経費) 3,400,000円、(間接経費) 1,020,000円

研究成果の概要(和文)：近年のマイクロプロセッサの性能はチップ温度によって制限を受けることが多い。そこで本課題ではチップ温度を低減する方法を研究開発し、それによるプロセッサ性能の向上を目指した。提案手法では、近年の商用プロセッサにおいて広く採用されているターボ・モードに着目し、アクティビティ・マイグレーションと呼ばれる手法を用いることでターボ・モード中にプロセッサが利用可能な最大周波数を引き上げる。アクティビティ・マイグレーションを空間的に細粒度に行う手法とそれを適用したプロセッサの設計手法を開発し、シミュレータを用いて評価を行った結果、提案手法により2.8%のチップ面積の増加で最大14.5%の性能向上を達成した。

研究成果の概要(英文)：Performance of modern microprocessors is often constrained by the chip temperature. Therefore, we studied a method to reduce chip temperature for performance improvement of microprocessors. We focused on Turbo Mode, which is widely used in modern microprocessors. Our technique is to use Activity Migration for upgrading the maximum CPU frequency during Turbo Mode. We developed a method to apply Activity Migration to a microprocessor with spatial fine granularity and its design methodology. We evaluated our technique with the combination of some simulators. The result shows that our technique achieves 14.5% of performance improvement in exchange for 2.8% of area increase.

研究分野：情報学

科研費の分科・細目：計算基盤・計算機システム

キーワード：マイクロプロセッサ アーキテクチャ 熱 ターボ・モード アクティビティ・マイグレーション

1. 研究開始当初の背景

近年のプロセッサでは、チップから発生する熱が性能を制限する大きな要因となっている。冷却装置に求められる冷却能力を表す TDP (Thermal Design Power) は、この 20 年間で急激に上昇した。その結果、TDP は実用的な冷却能力の限界に達し、頭打ちとなっている (図 1)。出荷時のチップの動作周波数は TDP に基づいて定められる。実際、特殊な冷却装置を用いれば、近年のプロセッサは 5GHz を超える速度で動作させることも可能である。このように、冷却装置の能力向上率は、LSI の微細化によってもたらされるプロセッサの性能向上率に追いついていないのが現状である。

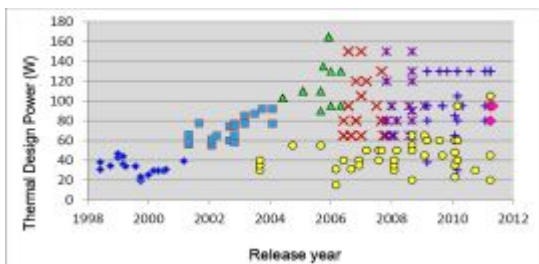


図 1 プロセッサの TDP の推移

発熱の問題は、チップの動作周波数だけでなく、稼働率も制限する。チップは、冷却装置の不具合等により、システムの稼働中に TDP を上回る熱を発生することもある。そのような場合でもシステムに深刻なダメージを与えるのを防ぐため、近年のシステムには温度を管理する機構が設けられている。チップ内に埋め込まれたセンサがチップの温度が所定の値を上回ったことを検知すると、ある程度の期間クロックを止めることで温度が下げられる。逆に言えば、所定の温度を下回る限りチップは連続稼働する。

チップにはホット・スポットとコールド・スポットが存在することが知られている (図 2)。前者の例は演算器、後者の例はキャッシュである。両者の温度差は、場合によっては、30 度にも達する。回路は、一部分でも発熱量が正常に動作する限界値を上回ると、全体として正常に動作することを保証できなくなる。そのため、チップが正常に動作する限界の発熱量は、ホット・スポットのそれによって決まる。なお、ホット・スポットはチップのごく一部の領域に限られる。

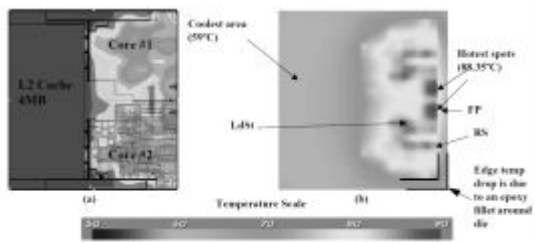


図 2 チップの温度分布 [Black, MICRO'06]

このような温度差が生じるのは、回路のア

クティビティを局所化することでプロセッサは高性能化してきたためである。レイテンシを短くするためには、短時間の内に使用されるモジュール同士はレイアウト上で集中的に配置した方がよい。また、データを分散するとその制御のために余分なレイテンシや回路が必要となることから、データは集中的に管理した方がよい。

発熱の問題は、基本的には、デバイスや冷却装置の問題と考えられてきた。そのため、これまでは主にこれらのレベルで対策が行われてきた。しかし、前述のように、解決には至っていない。

レイアウト・レベルやアーキテクチャ・レベルにおける対策もあるが、現在のところ、両者は独立に行われている。レイアウト・レベルでは、モジュールの並びを工夫することで発熱量を空間的に分散させるものの、モジュール内のアーキテクチャを変えることはない。一方、アーキテクチャ・レベルでは、個々のモジュールの消費電力 (発熱量) を減らすことが重視される。プロセッサ全体を考慮することがあっても、全体の稼働率を調整することで発熱量を時間的に分散させる程度であり、チップのどこにアクティビティが集中しているかはほとんど意識されない。

2. 研究の目的

本課題では、チップの発熱量が空間的にも時間的にも分散されたプロセッサ、すなわち、ヒート・スプレッド指向プロセッサを開発する。ホット/コールド・スポットの区別なく、各エリアが一様な温度を持つプロセッサを開発する (図 3)。これを、チップのレイアウト、および、アーキテクチャを統合的に見直すことで実現する。

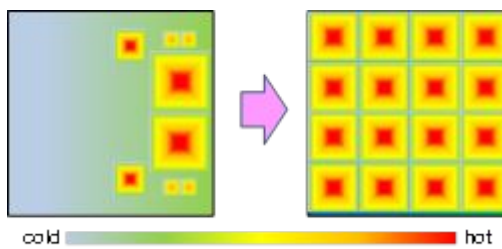


図 3 チップ温度のイメージ (右: 従来, 左: 提案)

前述のように、現在のチップにはホット/コールド・スポットが存在し、ホット・スポットの発熱量によってチップ全体の性能が決まっている。しかし、ホット・スポットの発熱量がシステムが正常に動作する限界に達しても、コールド・スポットのそれはまだ限界に達していない。理論上は、コールド・スポットの一部のトランジスタが動作しても、ホット・スポットの発熱量が限界を超えることはない。すなわち、現在の回路は、発熱量が限界に達した際のアクティビティにまだ余裕がある。

本課題では、この潜在的なアクティビティを利用することで、LSIの微細化の恩恵を真に享受できるプロセッサを実現する。本研究期間内では、回路のアクティビティがチップ内に分散するレイアウト、および、アーキテクチャを考える。アクティビティが回路全体に分散すれば、同じ処理を行う場合であっても、チップのピーク温度を抑えることができる。その結果、温度管理機構によってシステムの活動が低下する機会が減る、同一TDPでより高い動作周波数を使用できるなどのメリットが得られる。

アクティビティの分散化は各種のレイテンシの増加を招くが、上述のメリットにより、結果として、システム全体の性能が向上する可能性がある。本研究期間内では、提案するチップと従来のチップとを比較し、同一TDPを仮定した場合、前者の性能が後者のそれを上回ることをシミュレーションによって示す。

このようにして、本課題では、冷却能力の限界によってもたらされる、現在のプロセッサの性能限界を打破することを目指す。

3. 研究の方法

本課題では、チップの発熱量を均一にするという方針のもと、プロセッサの設計を抜本的に見直す。プロセッサ・アーキテクチャ、および、回路のレイアウトを統合的に見直し、回路のアクティビティが分散化されたアーキテクチャ、および、レイアウトを考える。

設計を考える上で基本となるのは、多くのアクティビティを必要とする処理を細かい粒度に分割し、その回路を空間的に分散して配置することである。特に、演算およびオペランドの読み書きは継続的に複数の命令によって行われる、アクティビティが高い処理である。したがって、演算器とレジスタは分散させた方がよい。分散させる際の粒度は、意味のある範囲でなるべく細かい粒度を考えている。粒度を細かくするのは、その方がより多くの潜在的なアクティビティを利用できると考えるからである。

設計する上でのポイントは、データが複製されることを積極的に許す点にあると考えている。現在のプロセッサは、レジスタ・ファイルに代表されるように、データは集中的に管理されており、アクティビティはそのモジュールに集中している。開発するプロセッサでは、データの複製を許し、それらを分散配置し、アクセスを分散させることでアクティビティを分散させる。

実際にチップを制作するのは、検証までに膨大な時間と費用を要し、開発効率が悪い。そこで、本課題では複数のシミュレータを利用して開発を進める。性能はサイクル・アキュレートなアーキテクチャ・シミュレータを用いて評価する。温度シミュレータはHotSpotを用いる予定である。温度シミュレーションには、各モジュールの配置/サイズ/

消費電力が必要となるが、これらの情報も何らかのシミュレーションを行うことで得るものとする。

4. 研究成果

本課題では、チップ内のアクティビティを分散化させる手法として提案された、アクティビティ・マイグレーションに着目した。アクティビティ・マイグレーションは、図4および図5に示すように、プロセッサ内にホットなモジュールの複製をいくつか用意しておき、それらを交互に使用することでモジュールの温度上昇を抑える手法である。

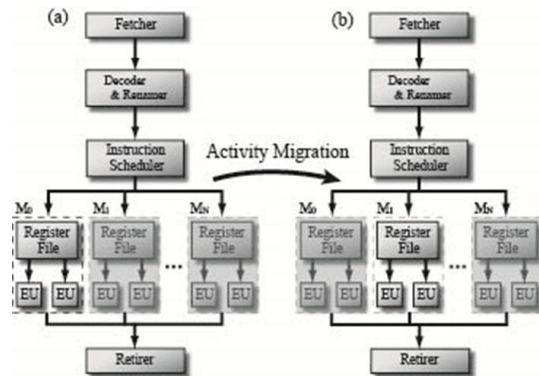


図4 アクティビティ・マイグレーション

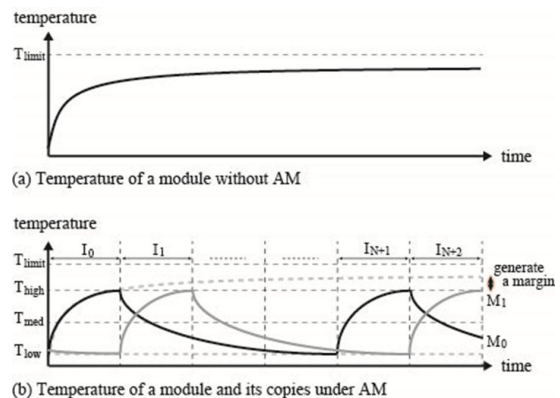


図5 アクティビティ・マイグレーションによるピーク温度の低減効果

ただし、従来のアクティビティ・マイグレーション研究には、以下の3つの問題点があった。

1. マイグレーションの対象とするモジュールの空間粒度が非常に粗く（例えば、実行ユニットとレジスタ・ファイルをひとまとまりとしたもの）、複製の追加による面積の増分に見合うだけの性能向上が必ずしも得られるわけではない
2. アクティビティ・マイグレーションの効果はマイグレーションの対象とするモジュールや追加の複製数に強く依存するが、それらは経験的に決定するしかない
3. 現実的な動作環境（プロセッサ構成やアプリケーションなど）において効果がどのぐらいのピーク温度の削減ができ、ま

たそれによってどの程度の性能向上が達成できるかは確認されていない

そこで本課題では、まず3番目の課題に対し、サイクル・アキュレートなプロセッサ・シミュレータと温度シミュレータ、および、プロセッサの電力/面積シミュレータを用いて、現実的な環境におけるアクティビティ・マイグレーションの効果を評価した。具体的には、サイクル・アキュレートなプロセッサ・シミュレータとしてOnikiri2を使用し、各モジュールの利用率(アクティビティ)のトレースを取得する。同時に、プロセッサの電力/面積シミュレータであるMcPATを使用し、各モジュールのピーク電力と面積を計算する。これら2つのシミュレータから得られた結果から各モジュールの電力トレースを生成し、それを用いてHotSpotによる温度シミュレーションを行う。この方法で、先行研究で提案された、レジスタ・ファイルと実行ユニットを対象とするアクティビティ・マイグレーションの効果を、デスクトップ相当のプロセッサを想定して詳細に評価した結果、十分なピーク温度低減効果が見られた(学会発表)。

次いで、1番目の課題に対し、従来よりも細粒度なアクティビティ・マイグレーションの検討を行った。本課題では、(設計のし易さという意味で)考え得る限り最小のホットなモジュールであるALUに着目し、ALU単位でのアクティビティ・マイグレーションを行う手法を提案した。評価の結果、ALU単位でもアクティビティ・マイグレーションには十分な温度低減効果があることがわかった(学会発表)。

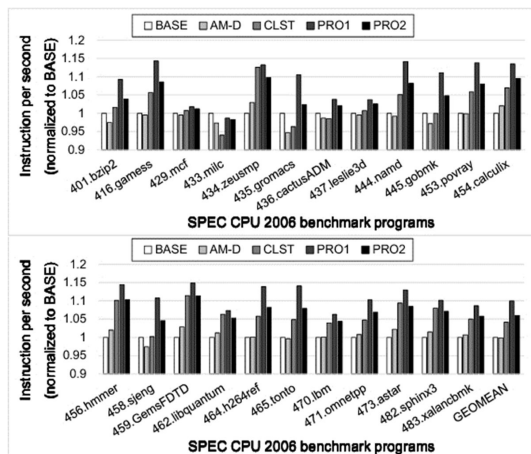


図6 提案手法による性能向上の効果

Processors	面積(mm ²)	相対値
BASE	30.481	1
AM-D	36.400	1.194
CLST	39.413	1.293
PRO1	31.327	1.028
PRO2	31.956	1.048

図7 提案手法の面積オーバーヘッド

上記の結果から、ALUなどの個々のユニットを単位とするアクティビティ・マイグレーション(細粒度アクティビティ・マイグレーション)を提案し、また、それを従来のプロセッサ設計のルーチンの中でシステムチックに行う方法を提案した。提案する設計手法では、ユニットの温度上昇モデルを用いて、マイグレーションの対象となるユニットと複製数を決定する。評価の結果、図6および図7に示すように、提案手法によって設計されたプロセッサは従来と比べて2.8%の面積増加で最大14.5%の性能向上を達成できることを確認した(雑誌論文 および学会発表)。

5. 主な発表論文等
(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計1件)

Shinobu Miwa, Takara Inoue, and Hiroshi Nakamura, Area-Efficient Microarchitecture for Reinforcement of Turbo Mode, IEICE TRANSACTIONS ON Information and Systems, Vol.E97-D, No.5, 2014 (to appear)

〔学会発表〕(計3件)

三輪忍, 井上聖等, 中村宏, ターボ・モード強化のための面積効率に優れたマイクロ・プロセッサとその設計手法, 情報処理学会研究報告 2014-ARC-208, No.12, 2014, pp.1-10, (2014/1/23, 東京)

井上聖等, 三輪忍, 中田尚, 中村宏, ALUローテーションによるスーパスカラ・プロセッサの性能向上, 情報処理学会研究報告 2013-ARC-204, No.10, 2013, pp.1-10, (2013/3/27, 白浜)

井上聖等, 三輪忍, 中村宏, レジスタ・ファイルと実行ユニットにおけるアクティビティ・マイグレーション, 情報処理学会研究報告 2012-ARC-201 (SWoPP2012), No.11, 2012, pp.1-9, (2012/8/1, 鳥取)

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

なし

6. 研究組織

(1)研究代表者

三輪忍 (MIWA, Shinobu)

東京大学・大学院情報理工学系研究科・助教
研究者番号: 90402940