

省電力を目的とした OS による細粒度パワーゲーティング制御方式の研究

砂田 徹也^{†1} 木村 一樹^{†1} 近藤 正章^{†2} 天野 英晴^{†3} 宇佐美 公良^{†4} 中村 宏^{†5} 並木 美太郎^{†1}^{†1} 東京農工大学 ^{†2} 電気通信大学 ^{†3} 慶應義塾大学 ^{†4} 芝浦工業大学 ^{†5} 東京大学

1 緒言

現代社会では省電力の気運が高まっており、LSI は高度情報化社会を支える基盤として広く利用されているが、その性能に対する消費電力について注目されている。特に近年の LSI プロセスの微細化によって消費電力の中でも支配的となっているリーク電力を削減する技術として、パワーゲーティング (PG: Power Gating) や Dual Vth, 基板バイアスなどが注目されている。この中で筆者らは PG 技術に着目する。PG 技術は、Intel Atom や ARM11 などにおいて実用化がなされているが、これらのプロセッサは、ハードウェアが CPU のアイドル期間を予測することで自律して PG 制御を行い、ソフトウェアによる制御は行われていない。そのため、既存の OS に変更を加えることなく PG による省電力効果を得られるが、ハードウェアによる制御のため、実行時の情報を基にした柔軟な制御を行うことは難しい。

本研究は、PG を細粒度に適用したプロセッサ Geyser に対して、PG を行う際に問題となる電力損益分岐点 (BEP: Break Even Point) を考慮した省電力機構を OS により実現し、細粒度 PG の省電力化を実現する。

2 Geyser アーキテクチャ

Geyser は、省電力を目的とした MIPS R3000 アーキテクチャをベースとしたプロセッサコアである。Geyser の省電力技術の特徴として、細粒度 PG を適用している。細粒度 PG は、PG の適用粒度を、時間的に毎クロックごと、空間的にプロセッサコア内の演算器 (ALU, SHIFT, MULT, DIV) ごとに PG を行うことを意味する。

2.1 細粒度 PG のインタフェース

細粒度 PG のインタフェースとして、特権レジスタ PGStatus レジスタを用意する。PGStatus は、図 1 として用意され、各演算器ごとに細粒度パワーゲーティングのスリープポリシーを制御することが可能となる。

A Study on Fine Grain Power Gating Control with Operating System for Conservation of Electric Power
Tetsuya SUNATA^{†1}, Kazuki KIMURA^{†1}, Masaaki KONDO^{†2}, Hideharu AMANO^{†3}, Kimiyoshi USAMI^{†4}, Hiroshi NAKAMURA^{†5}, Mitaro NAMIKI^{†1}

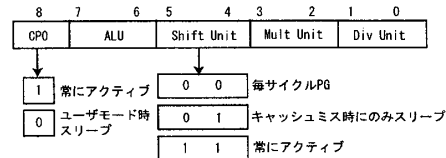
^{†1}Tokyo University of Agriculture and Technology^{†2}The University of Electro-Communications^{†3}Keio University^{†4}Shibaura Institute of Technology^{†5}The University of Tokyo

図 1: CP0 PGStatus レジスタ

2.2 細粒度 PG の電力的特徴: BEP

細粒度 PG により、その演算器にはアクティブとスリープの期間がうまれる。その状態を遷移する際に電力オーバーヘッドが発生するため、頻繁な PG を行うことは、逆に電力を増加させる要因となる。ここで「あるサイクル数スリープさせることができれば、電力オーバーヘッドを考慮しても省電力を実現できるサイクル数」として BEP を定義する。BEP は対象とする演算器や動作温度によって大きく異なる。また BEP は先行研究により既知である。

3 GeyserOS の概要と省電力機構

Geyser アーキテクチャで動作する OS として GeyserOS を設計、実装した。GeyserOS が提供する基本機能として、マルチタスク OS, 仮想メモリ管理, 省電力機構の三つがある。

3.1 GeyserOS による細粒度 PG 制御機構

図 2 に、GeyserOS における細粒度 PG 制御機構の概要を示す。本省電力機構は OS スケジューラによって実現され、個々のタスクの演算器使用特性に基づいて制御を行う。入力として、BEP を越えないスリープの割合である、BEP ミス率を新たに定義する。これは、あ

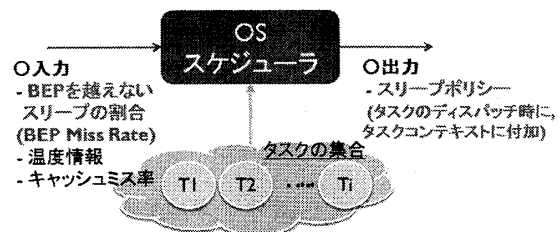


図 2: 細粒度 PG 制御機構の概要

る期間における総スリープサイクル数に対して、BEP を越えずに電力ロスを発生するスリープサイクルの割合である。

GeyserOS は、「BEP ミス率が越えるスリープの割合を削減する」という方針で制御する。BEP ミス率を削減する手段として、PGStatus レジスタを用いたスリープポリシーの変更を OS により行う。各スリープポリシーは、以下の図で示すように三つの閾値を定めることで変更する。また PG 無しから毎サイクル PG

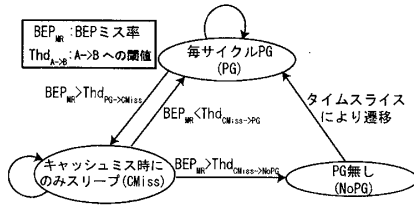


図 3: スリープポリシーの遷移条件

へは、一度のタイムスライスにより遷移することとする。本省電力機構は、パフォーマンスカウンタを用いて各タイムスライスごとに BEP ミス率を計算し、スリープポリシーを実行時に動的に変更することで、省電力を実現する。

4 実装と評価

Geyser アーキテクチャおよび GeyserOS の実行/評価環境として、Verilog による RTL シミュレーションを用いる。本シミュレーション環境は、Geyser コアに加え細粒度 PG によるスリープ期間とその頻度を観測するパフォーマンスカウンタを備え、BEP ミス率の計算に用いる。RTL シミュレータには、Cadence 社の NC-Verilog を用い、電力評価には、Synopsys 社の HSIM および PowerCompiler を用いる。ベンチマークは、QSORT および Matrix を使用し、省電力機構におけるスリープポリシー変更のための BEP ミス率の閾値は、10%~90%まで変動させる。また制御対象の演算器を今回は使用頻度の多さから ALU とする。

評価として、細粒度 PG における ALU の使用頻度を図 4 に示す。OS による制御を加えない場合は、ALU など使用頻度の高いもので 70~80% 利用されていることが分かる。また OS による制御を加えることで、BEP ミスを削減するためにスリープ期間の制御が行われ、制御前と比べて使用率が 5~10% 程度大きくなっていることが分かる。

この演算器使用頻度を用いて、最良の値が出た時の閾値の各演算器のリーク電力を求めた図を図 5 に示す。OS による制御の結果、制御前と比べて 5~最大 25% の平均リーク電力削減を実現した。これは、OS の制御により BEP ミスを起こす ALU の利用を削減できたため

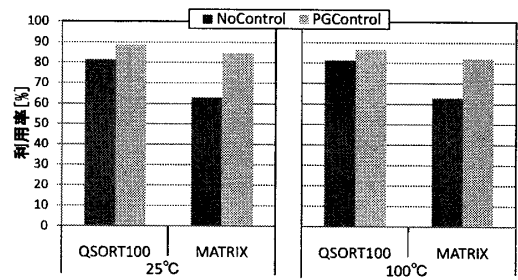


図 4: ALU の使用頻度

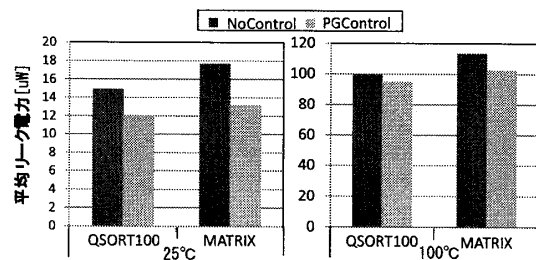


図 5: ベンチマークごとの平均リーク電力

であると考えられる。またこの際省電力機構により発生した OS オーバーヘッドは最大 5% 程度である。また温度が 25 °C から 100 °C へ上昇したことにより、平均リーク電力全体が増大し、さらに温度上昇により BEP が短くなるため BEP ミス率が減少することから、OS による制御の効果はあまり出ていない。

5 結言

本研究では、BEP ミス率を考慮した OS スケジューラによる細粒度 PG 制御機構について提案した。本機構により、少ない OS オーバーヘッドでリーク電力の削減を実現した。

今後の課題として、Geyser の実機による評価に加え、BEP が動作温度によって大きく変動することを考慮して、タスクの実行順序を制御することにより一つの演算器の温度上昇を抑えることで、省電力を実現する機構の設計、評価などが挙げられる。

謝辞 本研究は、科学技術振興機構「JST」の戦略的創造研究推進事業「CRSET」における研究領域「情報システムの超低電力化を目指した技術革新と統合化技術」の研究課題「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」によるものである。

参考文献

[1] Andreas Merkel et al.: Task Activity Vectors: A New Metric for Temperature-Aware Scheduling, Proceedings of the 3rd ACM SIGOPS/EuroSys European Conference on Computer Systems 2008, pp.1-12 (2008).