

消費エネルギー予測に基づいた KVM 仮想化環境における省電力制御の研究

DOUANGCHAK SITHIXAY[†] 佐藤 未来子[†]
山田 浩史[†] 並木 美太郎[†]

1. はじめに

近年、クラウドの成長に伴い、計算機の台数増加による莫大な電力消費が問題となっており、仮想化環境における消費電力の削減に注目が集まっている。従来の計算機単体の省電力化の研究では、システムの演算性能を予測し、演算性能をできるだけ落とさずに、DVFS 手法を用いて周波数・電圧制御を行う研究¹⁾がある。しかし、周波数・電圧を落とすと、実行時間が延びてしまい、消費エネルギーを最も削減できるとは限らない。また、CPU、メモリ、ディスク I/O、ネットワーク I/O バウンドアプリケーションの混在など、最適な動作周波数が異なるアプリケーションが並列動作する場合などにおいては従来の手法では不十分である。

そこで本研究では、仮想化環境における省電力化を目的とし、仮想マシンモニター (VMM) レイヤで仮想マシン (VM) ごとの消費エネルギー予測に基づいた省電力化を行う。実装には、Linux カーネルを用いた VMM である KVM を用いた。本稿ではこれらの設計・評価の詳細について述べる。

2. 仮想化環境における省電力制御

本研究では、コアを n 個持つホモジニアスマルチコアプロセッサ上の仮想化環境に省電力制御を適用する。システム管理者は各 VM の演算性能とスループットを閾値として与え、本省電力制御がその両方を上回る演算性能とスループットを保ちつつ、各 VM の消費エネルギーが最小となる最適な周波数・電圧を求め、VM 単位およびコア単位で DVFS を行う。

本研究では、完全仮想化環境を提供する KVM の VMM レイヤにおいて、個々の VM の消費エネルギーを予測する。消費エネルギー予測に使われる VM の演算性能、消費電力、ネットワーク I/O、ディスク I/O の予測は定量的に多変量回帰分析し、モデル化を行う。

その予測モデルに対して、キャッシュミス率やプロセッサ全体のメモリアクセス頻度などのパフォーマンスカウンタの情報と仮想 NIC、仮想 I/O の情報を用いて、VM ごとに消費エネルギーを予測する。

本省電力制御は、上記予測のモデル化をする「学習モード」と VM の消費エネルギーが最小化となるように周波数・電圧の決定を行う「省電力モード」からなる (図 1)。学習モードでは、1 台の VM 上であらかじめ様々なベンチマークを各種の周波数で動作させ、統計的な学習を行う。その後、多変量回帰分析により予測に用いる回帰係数を算出する。省電力モードでは、求められた回帰係数を用いて、演算性能比、消費電力比、スループット比を予測し、その予測に基づいて複数の VM が動作している状態で、最適な周波数・電圧を求め、DVFS による省電力制御を行う。

本周波数決定機構を通じて、VMM レイヤで VM のタイムスライス単位で最適な周波数・電圧を求め、VCPU コンテキストスイッチ毎に DVFS を行う。具体的には、任意の VM_i に対して、演算性能比とスループット比から演算性能とスループット閾値の両方を上回る周波数の候補を求める。次に、周波数の候補の中から、最も消費電力比が小さい周波数を選択し、 VM_i の最適な周波数とする。最後に、 VM_i の全ての VCPU が割り当てる物理コアの周波数および電圧を変更する。

3. 実装

Linux カーネル自体を VMM とする仕組みである KVM を用いて実装を行った。本研究では、ターゲットとなるアーキテクチャは x86 とし、同じ x86 の中でも特に電源電圧・周波数制御は CPU やチップセット依存となるため、AMD の Phenom9500 をターゲットとした。カーネルモジュールである `kvm-kmod-3.4` と I/O などをエミュレートする `qemu-kvm-1.0` に機能を追加した。既存の KVM カーネルモジュールと `qemu-kvm` に 77 行のコードを変更し、新規部分は合計 1396 行となった。

[†] 東京農工大学
Tokyo University of Agriculture and Technology

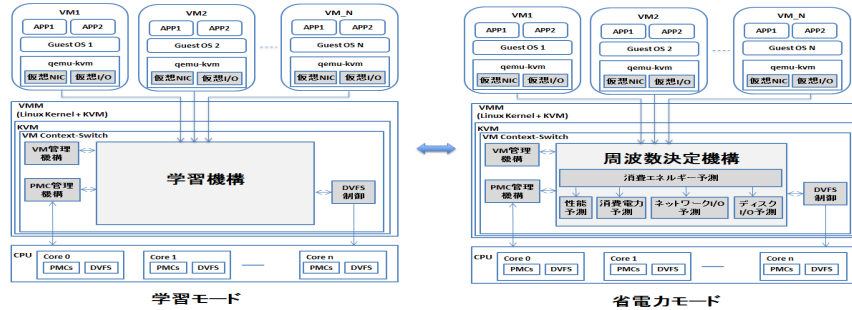


図 1 省電力制御の全体構成

4. 評価

本手法について、SPEC2006、ネットワーク、ディスクのベンチマークを用いて評価を行った。評価には、Phenom9500 を搭載した PC を評価マシンとして利用した。閾値 (= 演算性能閾値 = スループット閾値) は 70% と 40% に設定した。VM は三台があり、三台に異なるベンチマークを表 1 の組合せで実行する。評価結果を図 2 と 3 に示す。評価に対して、VM の演算性能閾値とスループット閾値の両方を上回る性能で終了させることができた。

図 2 は SPEC2006 のベンチマークのみの組合せ (scenario1,2,3) の評価結果であり、図 3 は SPEC2006、ネットワークベンチマークとディスクベンチマークの組合せ (scenario4,5) の評価結果である。横軸は scenario 番号、縦軸は scenario に各評価項目を、全コアで最高周波数に設定した場合を 1 とした時の相対値を表している。

表 1 ベンチマークの組合せ

scenario	ベンチマーク		
	VM1	VM2	VM3
1	444.namd	444.namd	-
2	444.namd	470.lbm	-
3	470.lbm	470.lbm	-
4	444.namd	httperf_500rate	bonnie++
5	httperf_200rate	bonnie++	470.lbm

VM の閾値が 40% の時に、図 2 により、メモリバウンドとメモリバウンドの組合せである scenario3 が最大で 44.3% であり、図 3 により、scenario5 が最大で 39.7% の CPU の消費エネルギーを削減した。本手法は、VM のタイムスライス単で周波数を決定するため、VM が複数台あっても各 VM の消費エネルギー予測ができ、本手法が有効であると確認した。

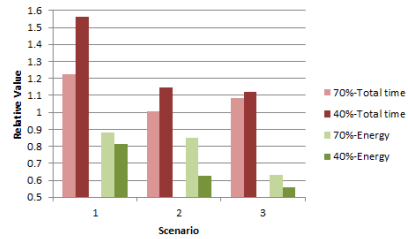


図 2 VM 複数台の評価 1

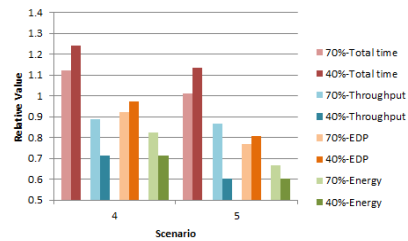


図 3 VM 複数台の評価 2

5. おわりに

仮想化環境における省電力化を目的とし、消費エネルギー予測に基づいた省電力化を行う VMM の設計、実装と評価について述べた。今後の課題として、VM 上で複数のベンチマークが動作時のリソース競合による消費エネルギーの予測精度の分析が挙げられる。

参考文献

- 1) W.Wu, M.Martonosi, D.W.Clark, V.J.Reddi, D.Connors, Y.Wu, J.Lee, D.Brooks: A Dynamic Compilation Framework for Controlling Microprocessor Energy and Performance, In Proc. of the 38th annual IEEE/ ACM International Symposium on Microarchitecture, pp. 271-282, 2001.