

仮想計算機環境における省電力化を 目的としたサーバ資源制御

今田 貴之^{†1} 佐藤 三久^{†1} 木村 英明^{†1}

本稿では、仮想計算機 (VM) を運用するサーバ群において QoS (Quality of Service) 性能維持とエネルギー削減を両立させるために必要な要素の評価について述べる。近年ではデータセンタ等においてインターネットサービス等を提供する際に VM が利用されてきている。VM 運用環境においても VM とそれが動作しているサーバを適切に制御できれば更なる消費電力削減が可能である。そのためには VM 運用環境に特有な“負荷のマイグレーション”および“どのプロセッサコアに VM を割り当てるべきか”について QoS 性能および消費電力の観点から考慮する必要がある。本稿での評価により、マイグレーションを利用して VM を 1 台のサーバへ集約することで QoS 性能を維持しながら消費電力削減を行えることが分かった。また、VM のプロセッサコア割り当てと動作周波数を適切に制御することでさらなる消費電力削減を行えることが分かった。

Server Resource Management for Power Saving on Environment with Server Virtualization

TAKAYUKI IMADA,^{†1} MITSUHISA SATO^{†1}
and HIDEAKI KIMURA^{†1}

In this paper, we describe some promising evaluation results which contribute to strategies for energy saving and keeping QoS (Quality of Service) on virtualized servers. Virtual Machines (VMs) are often used for Internet services like cloud computing in data centers. It is possible to reduce energy consumption by managing VMs on servers and the servers themselves. New issues in such virtualized servers compared to conventional physical servers are “migration of load between specific two servers” and “which processor cores we should pin down VMs to” from the view point of QoS performance and energy consumption. We found that server consolidation using VM migration can realize power reduction with keeping QoS performance, and that allocation of VMs to multiple processor cores running at less frequency can bring additional power reduction on a server.

1. はじめに

近年、多数の計算機資源を有する計算機センタやデータセンタ施設での消費電力問題が多く言及されており、この問題を解決する効果的な手法が必要とされている。現状では、インターネットサービスやデータベース等に用いられる各種サーバは処理能力を優先する構成をとり、常に電源が投入されたままである。また、このようなサーバを冷却するため装置が当然ながら必要であり、装置自体が電力を消費することになることから熱源となるサーバの性能に対する消費電力効率を高めていかなければならない。

現在では、クラウドコンピューティングをはじめとして仮想計算機 (VM) を用いた計算資源の運用方法が注目を集めている。VM を用いる利点の一つとして、各所に散在している多数のサーバ上で動作しているサービスを少数のサーバ上の VM で運用できることがあげられる。満たすべき QoS 性能を維持することができれば、結果として既存よりも少ないサーバ台数と消費電力で運用可能であると考えられる。しかし、このような手法を適用したとしてもそのシステムは想定される最大要求を満たすように構成されなければならない。よって性能-電力効率の良いシステムを構築するためにはさらにソフトウェア側で省電力化を行う手法が必要であると考えられる。

本稿では、VM 上で一般的な HTTP の Web サービスを提供する複数のサーバ運用時における省電力化手法について考える。従来研究¹⁾において我々は分散型 Web サーバ上での省電力化手法について提案を行っているが、VM 運用環境下では異なる特性を持つと考えられる。そこで、本稿では最初に 1 台のサーバで VM 運用する場合の性能と消費電力特性について基本的な評価を行う。VM を使用するサーバの運用では

- VM のプロセッサコアへの割り当て方法
- VM のマイグレーション

という新たな事項について考慮する必要がある。本稿では上記 2 項目に関しても QoS 性能および消費電力の観点から評価を行い、複数のサーバ上で VM を運用しながらシステム全体の省電力化を行う方法について考える。

本稿の構成は以下ようになる。まず次章にて関連研究について述べる。3 章では、1 台

^{†1} 筑波大学大学院システム情報工学研究科

Graduate School of Systems and Information Engineering, University of Tsukuba

のサーバ上での VM 運用時の QoS 性能と消費電力特性について述べる．そして 4 章において負荷のマイグレーションによる QoS 性能と消費電力特性について提案する．5 章において VM のプロセッサコア割当てによる QoS 性能と消費電力特性について述べる．最後に 6 章でまとめと今後の展望について述べる．

2. 関連研究

最近では、VM 運用時のサーバにおけるエネルギー削減を行う手法が提案されている．Nathuji らは VirtualPower と呼ばれるハードウェアベースの動作周波数制御とソフトウェアベースのタイムスライス制御を使用してサーバを省電力化する手法を提案している²⁾、VirtualPower では VM から透過的に設定することが可能で動作周波数である VPM states を用いて消費電力の削減を行う．Verma らは消費電力とライブマイグレーションコストを考慮して VM を複数のサーバ上へ配置する手法を提案している³⁾．この手法では現在の VM 配置に対して全サーバの消費電力とマイグレーションコストに基づいて次に考えるべき VM 配置を算出する．Hu らは複数サーバに存在する低い負荷の VM をある特定のノードの集約して消費電力を削減する手法について提案している⁴⁾．この手法では Magnet と呼ばれる VM 負荷の具合を表す 3 階層の ring に基づき、ライブマイグレーションを使用して低い負荷の VM の集約、および過負荷なサーバの VM を軽い負荷のサーバへ分散させる．Kusic らは待ち行列理論に基づく VM 資源制御によって分散型サーバの消費電力削減手法を提案している⁵⁾．しかし、Kusic らの手法では VM のライブマイグレーションとプロセッサコアに対する VM の pin down 手法について考慮がなされていない．

3. 1 台のサーバ上で VM を動作させる場合の性能と消費電力

本章では、1 台のサーバ上で VM を動作させる場合の性能と消費電力について評価を行うために、SPECweb2005⁶⁾ を用いてベンチマークを行った結果について述べる．

3.1 評価システム

図 1 に評価環境の概要を示す．システムは VM ホスティング用サーバ (Core i7 940, 12GB DDR3-1333, kernel-xen-2.6.18-128.1.10.el5) 2 台およびロードバランササーバからなる分散型のシステムである、本評価ではサーバ 1 台のみを使用した．サーバ上では Xen 3.3.1⁷⁾ による VM 環境 (VM イメージは iSCSI ストレージ上に存在) を構築した．本評価では VM の構成として 仮想 CPU 数を 1, メモリ 1GB を割当てたものを VM の単位とした．なお、Xen における VM のスケジューラには標準の Credit スケジューラを用い

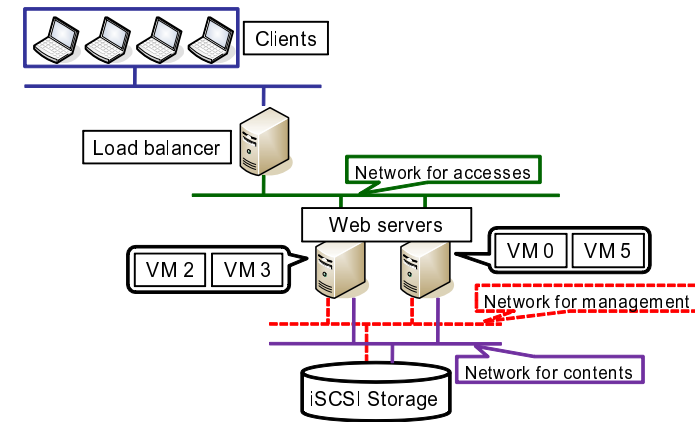


図 1 評価環境の概要

た．一方、ロードバランササーバ上では ipvs-1.2.0⁸⁾ による Layer-4 の負荷分散をソフトウェアで行う．評価環境では、クライアントから発行されたリクエストがロードバランササーバを通じてサーバ上の VM で動作している http サーバで処理される．その際、VM はバックエンドの iSCSI ストレージから静的データを取得し、PHP によって生成された動的データと併せてリクエスト応答がクライアントへ返される．なお、ネットワークはすべて Gigabit Ethernet を用いて構成されている．

消費電力測定には我々が開発した PowerWatch⁹⁾ を用いた．本評価ではマザーボードへ供給される CPU+12V DC, ATX3 系統 (+12V, +5V, +3.3V DC), HDD へ供給されている +12V, +5V DC, そして FAN へ供給されている +12V, +5V DC それぞれの電流値を測定し、これらの値より上記 Web サーバの消費電力を算出した．

3.2 評価方法

ベンチマークには SPECweb2005 (Non-SSL 構成) を用いた．SPECweb2005 は特性の異なる三つのベンチマーク (Banking, E-commerce, Support) により構成されており、それらの特性は

- Banking: 計算性能と I/O 性能がバランスされている
- E-commerce: 計算性能に比重が置かれている
- Support: I/O 性能に比重が置かれている

となっている．本評価では 1 台のサーバ上で 1 つもしくは 2 つの VM を使用してベンチ

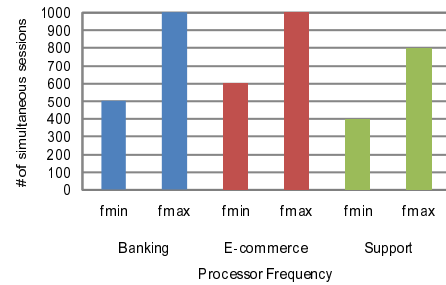


図 2 1 VM 動作時の最大要求同時セッション数

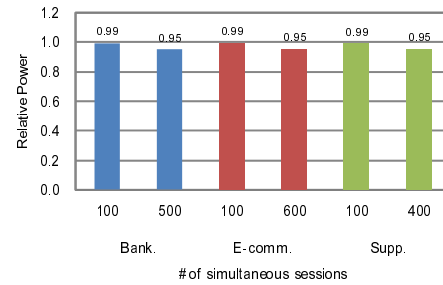


図 3 f_{min} 動作時の相対消費電力 (1 VM 動作時)

マークを実行し、要求同時セッション数をプロセッサ動作周波数を変化させてそのときの性能と消費電力を測定した。1つのVMでの評価では上記3種のベンチマークをそれぞれ実行し、2つのVMの評価では (Banking, E-commerce) および (Support, E-commerce) の組合せでそれぞれベンチマークを実行した。このとき、1つまたは2つのVMを1台のサーバ上で動作させ、それぞれのVMがプロセッサコアへ独立して pin down されるように設定した。また、プロセッサ動作周波数には最高周波数 f_{max} (=2.93GHz) および最低周波数 f_{min} (=1.6GHz) を使用した。

3.3 評価結果

3.3.1 1つのVM動作時

図2に1VM単体を1プロセッサコアに pin down させてそれぞれのベンチマークを実行したときの QoS 満足条件下における最大要求同時セッション数をプロセッサの最低周波数 f_{min} および最高周波数 f_{max} の場合で示す。ここで、QoS の定義は SPECweb2005 ベンチマークで定義されているものの一つである 95% TIME_GOOD 指標である。図2より、QoS 満足条件の観点からそれぞれのベンチマークにおいてある要求同時セッション数までは f_{min} でも QoS 条件を満足できることが分かる。図3に QoS 満足条件下における要求同時セッション数でプロセッサを f_{min} で動作させた場合の相対消費電力を示す (f_{max} における消費電力を1とする)。図3より、すべてのベンチマークにおいて1-5%の消費電力削減が可能であることが分かる。つまり、要求同時セッション数が少ない (100 や 500) 場合にはプロセッサ DVFS を用いることで同じ性能を達成しながら消費電力削減が可能であることが分かる。

表 1 2 VM 動作時のワークロードセットと要求同時セッション数

Workload	VM0		VM1	
	S_{bank}	S_{ecom}	S_{supp}	S_{ecom}
B-E1	100	100		
B-E2	500	600		
B-E3	1000	1000		
Workload	VM0		VM1	
	S_{supp}	S_{ecom}	S_{supp}	S_{ecom}
S-E1	100	100		
S-E2	400	600		
S-E3	800	1000		

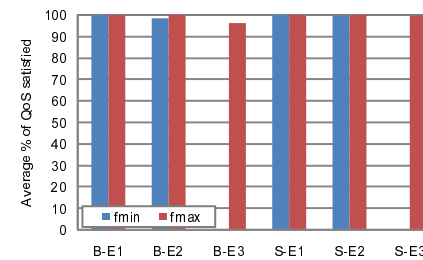


図 4 2 VM 動作時の QoS 性能

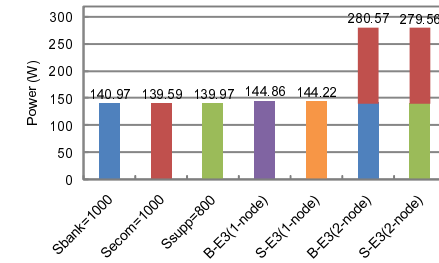


図 5 サーバ台数と VM 数の変化による平均消費電力

3.3.2 2つのVM動作時

3.3.1 項での結果に基づき、それぞれのベンチマークにおける QoS 満足下での最大要求同時セッション数までの負荷を与えて2つのVM (VM0, VM1) を同時動作させて評価を行った。表1に行ったベンチマークセットと要求同時セッション数を示す。ここで、 S_{Bank} , S_{ecom} , S_{supp} はそれぞれ Banking, E-commerce, Support ベンチマークの要求同時セッション数を示す。プロセッサ動作周波数 f_{min} の場合はベンチマークセット B-E1, B-E2, S-E1 および S-E2 を実行し、 f_{max} の場合はすべてのベンチマークセットを実行した。

図4に2VM動作時の平均 QoS 性能 (TIME_GOOD を満足する ページ/download の割合) を示す。図4より、 f_{min} で動作する場合にはほぼ性能低下がないことが分かる。一方、 f_{max} で動作させた場合には B-E3 ベンチマークセットで明らかな性能低下が起こった。B-E3 ベンチマークセットは比較的 CPU 負荷の高い特性を持つ。I/O 負荷が高い S-E3 ベンチマークセットではそのような性能低下が起こらなかったことから、この性能低下はプロセッサ資源不足によるものであると考えられる。また、図5に

- 1台のサーバ (f_{max}) で各ベンチマークの最大要求同時セッション数を処理するときの平均消費電力 $P(1\text{-node}, 1\text{-VM})$
- B-E3 および S-E3 ベンチマークセット (2VM 使用) を1台のサーバ (f_{max}) で処理するときの平均消費電力 $P(1\text{-node}, 2\text{-VM})$

● B-E3 および S-E3 ベンチマークセット (2VM 使用) を独立した 2 台のサーバ (f_{max}) 上で処理するときの平均消費電力 $P(2\text{-node}, 2\text{-VM})$

を示す。図 5 より, $P(1\text{-node}, 1\text{-VM})$ に対する $P(1\text{-node}, 2\text{-VM})$ の消費電力増加は高々約 5W 程度であり, ベンチマークセットの種類によらないことが分かる。一方, $P(2\text{-node}, 2\text{-VM})$ は $P(1\text{-node}, 2\text{-VM})$ と比較して約 2 倍の消費電力となっており, 単純にサーバを 1 台増加させただけの消費電力増加にほぼ等しい。つまり出来る限り 1 つのサーバに VM を集約させていくアプローチが消費電力の観点から非常に重要であることが分かる。

3.4 考 察

3.3.1 項における評価結果より, DVFS を利用するメリットは VM 運用環境下においても消費電力削減に重要な要素であることが分かる。3.3.2 項での 2VM 動作時においても同様の評価を行ったが, プロセッサ動作周波数を f_{min} に設定することで最大約 8% の消費電力削減を示した。DVFS 制御にはその時間的ペナルティがほぼ無いことから, 省電力化を行うためには積極的に利用すべきであると考えられる。

一方, 3.3.2 項での評価により 複数の VM を 1 台のサーバへ集約させていく方が省電力化には非常に重要な要素であると考えられる。これを実現するには負荷のマイグレーションを行う必要がある。このとき, 複数の VM を 1 台のサーバで運用するためには評価結果で見られたような性能オーバーヘッドによる QoS 性能の低下を十分に考慮する必要がある。

4. マイグレーション時の性能と消費電力

VM による計算資源運用での大きなメリットの一つとして考えられるのは VM のマイグレーション機能である。前章での評価からも VM を集約して動作サーバ台数を削減させることは消費電力削減において非常に重要である。また, サーバ台数の削減までには至らなくとも VM の配置状況を変更することでシステム全体の消費電力を下げられる可能性がある。しかし, マイグレーションの適用に対する影響をサーバの負荷状況に応じて考慮しなければ要求された QoS 性能を維持できない可能性が生じる。また, マイグレーションの適用がどの程度消費電力に影響するか考慮することが省電力化を行うためには必要である。本章ではマイグレーション時に考えられる状況として,

- VM 移動先のサーバが無負荷の場合
- VM 移動先のサーバに高負荷の場合

を考え, この 2 つの状況の違いに対する QoS 性能と消費電力特性について評価を行う。

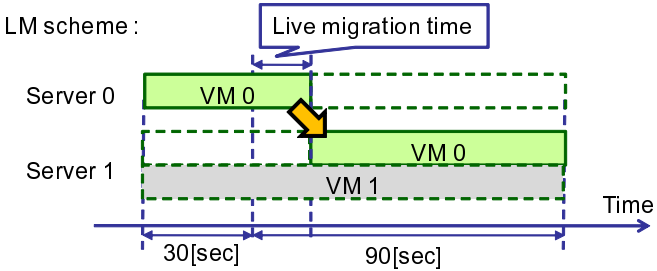


図 6 LM 手法の実行手順

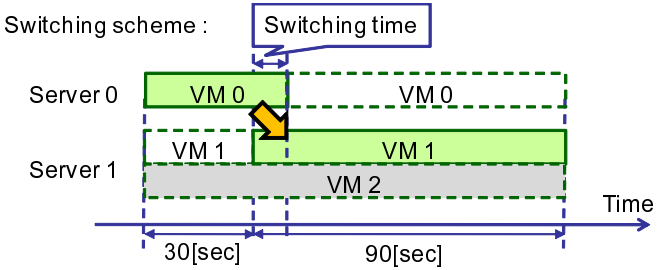


図 7 Switching 手法の実行手順

4.1 評価方法

評価システムは 3 章と同様の環境でサーバを 2 台用いた。この評価ではマイグレーション先サーバの負荷状況の影響を調べるため, ベンチマークの組合せを

- A: 無負荷サーバへの Banking VM のマイグレーション
 - B: E-commerce 実行中サーバへの Banking VM のマイグレーション
- として評価実験を行った。

さらに, 本評価ではマイグレーションの比較対象として

- Xen で提供されているライブマイグレーション手法 (LM)
- 2 つの VM を用いたロードバランサによるリクエスト転送先 VM の切り替え (Switching)

を設定した。図 6 に LM 手法の概要を示す。LM 手法は Xen でのライブマイグレーションを libvirt¹⁰⁾ によりパラメータ (“xenmigr://Server1”) で操作したものである。よって図中 VM0 のメモリイメージを転送する時間がマイグレーション時間となる。なお, ライ

表 2 LM および Switching 手法での VM 構成

LM	VM 0	VM 1	Switching	VM 0	VM 1	VM 2
A	Banking	None	A	Banking	Banking	None
B	Banking	E-commerce	B	Banking	Banking	E-commerce

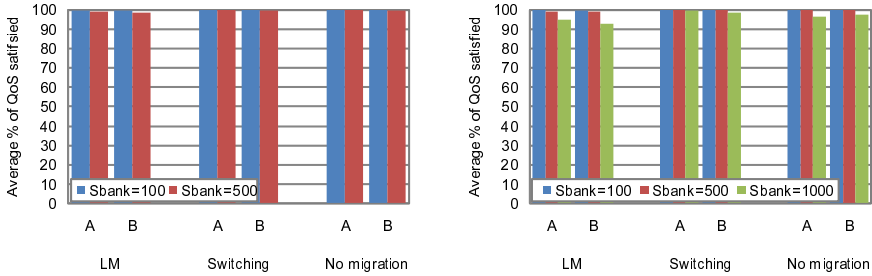
ブマイグレーションは図 1 における管理用ネットワークを用いて行われる。また、図 7 に Switching 手法の概要を示す。この手法は分散型 Web サーバ上での省電力化手法で用いられる方法をそのまま VM 運用環境へ適用したものである。対象サービスの VM を予め 2 台のサーバ上で動作させておき、ロードバランスによるリクエスト転送先の切替えにより負荷のマイグレーションを行う。Switching を発行したとき 図中 VM1 がサービスを開始するが、VM0 即座にサービスを止めずに処理中のリクエストを返さなければならない(新たなリクエストを受け付けずに処理中リクエストの返信だけを行う)。つまり 2 つの VM がオーバーラップしてリクエストを処理する時間帯が存在し、この時間が Switching に要した時間となる。

表 2 に LM および Switching 手法 でのベンチマーク組合せ) 毎の VM 構成を示す。LM 手法では VM0 と VM1 が独立したサーバ上で動作する。一方、Switching 手法では VM0 が 1 つのサーバ上、VM1 と VM2 が別のサーバ上で動作する。なお、LM および Switching 両手法において E-commerce については単独 VM 時での最大性能でベンチマークを実行し、マイグレーション先サーバに負荷を与えておくようにする(以下では、Banking および Support 実行 VM を Foreground VM、E-commerce 実行 VM を Background VM と呼ぶ)。このとき、それぞれの VM が独立したプロセッサコアへ割り当てられるように設定した。また、すべての評価においてベンチマーク実行時間を 120[sec] とし、ベンチマーク開始 30[sec] 後にマイグレーション開始させた。

4.2 評価結果

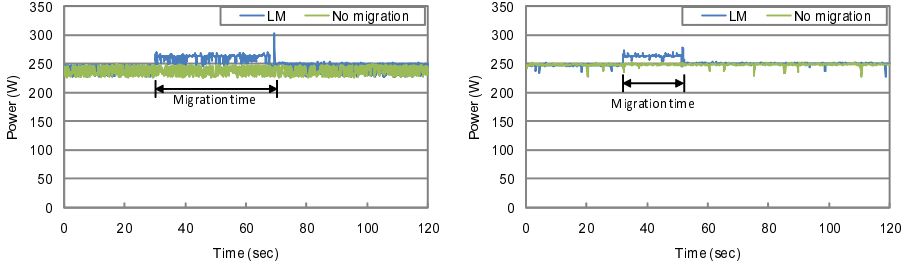
4.2.1 QoS 性能

図 8 に 4 種ベンチマークセットで Foreground VM への負荷を変化させたときの QoS 満足率を示す。ここで、No migration はマイグレーション非実行時の値である。図 8 よりプロセッサ動作周波数と Foreground VM の負荷状況によらず、LM 手法と比較して Switching 手法の QoS 満足率が同等もしくは若干高いことが分かる。特に LM 手法では Foreground VM への要求同時セッション数が 1000 場合、すべてのベンチマークセット実行時に QoS 満足率が 95% を下回った。原因の一つとして、ライブマイグレーション時のメモリ転送による Hypervisor 性能の低下が考えられる。また、LM 手法ではマイグレーション時間に



(a) f_{mix} 動作時 (b) f_{max} 動作時

図 8 マイグレーション環境下での QoS 性能



(a) S_{bank} = 500 (b) S_{bank} = 1000

図 9 LM 手法適用時の消費電力プロファイル (ベンチマークセット A)

最低でも約 20[sec] を要した。一方、Switching 手法では同条件においても QoS 満足率は 95% を維持できており、満足率の値は No migration 手法の場合とほぼ同じであることから 1 台のサーバ上で Foreground VM と Background VM を同時実行させても良好な QoS 満足率を達成していることが分かる。さらに Switching 手法では Foreground VM への負荷が最大の場合において No migration 手法よりも QoS 満足率が高くなった。その主な原因として Switching 手法ではマイグレーション時に別のサーバの VM を一時的にオーバーラップして使用するためであると考えられる。Switching 手法では Switching に費やした時間は最大でも約 3.2[sec] であった。

4.2.2 平均消費電力

2種のベンチマークセットを実行した場合、No migration 手法と比較して Switching 手法は平均消費電力の差が無かったのに対し、LM 手法では約 1-4%程度その値が増加した。特に要求同時セッション数が 500 の場合にその増加値が最大となった。図 9 にベンチマークセット A ($S_{bank} = 500, 1000$) に対してプロセッサ動作周波数 f_{max} で LM 手法を適用したときの消費電力プロファイルを示す。図 9 より $S_{bank} = 500$ の場合の方が約 2 倍マイグレーションに時間を費やしていることが分かる。これは Xen で採用しているライブマイグレーションアルゴリズム¹¹⁾ による影響であると考えられる。このアルゴリズムでは差分となるメモリーイメージの転送の条件として頻繁に更新されるメモリー領域は差分転送を行わない。そのため Foreground VM への負荷が $S_{bank} = 1000$ の場合はメモリーイメージ転送量が抑制される。これにより上記のような平均消費電力の差になって現れたと考えられる。

4.3 考 察

前項での評価結果より、LM 手法および Switching 手法ともに若干の消費電力増加でサーバ間にまたがるマイグレーションを行えることが分かる。しかしながら両手法はその特性が異なるため、適用すべき場面/状況が異なると考えられる。表 3 に両手法の主な特性を示す。マイグレーション時間に関しては LM 手法は VM に割り当てたメモリーサイズに依存し、通常は数十秒のオーダーであると考えられる。一方、Switching 手法はマイグレーション時間はそのときコネクションが張られているファイルサイズに依存する。取得ファイルサイズが数~数十 KB のコネクションが独立で多数になる場合は特に問題ないが、ストリーミング配信のような 1 つのコネクションで連続してデータを長時間取得、あるいはサイズが数百 MB のファイルの取得が主な用途ではコネクションが切れないためにマイグレーション時間が数分のオーダーになる可能性がある。このような環境下ではマイグレーション中でもコネクションを貼り続けられる点で LM 手法を適用するべきであると考えられる。また、必要な VM 数は LM 手法において必要最低限の VM 数を持っていけばよいために余分な資源は必要ない。一方で Switching 手法は余分な VM を立ち上げておくアプローチであるからメモリー等の資源を余分に使用したまま運用しなければならない。ただし、VM を低負荷もしくはアイドルで動作させておくことによる消費電力の上昇分はほぼ無いと考えられる。

QoS 性能に関しては LM 手法では性能低下が見られたために無視できないと考えられる。マイグレーションを行うかどうかの判断には性能低下によって生じる損失とマイグレーションの有無によるエネルギーコストを考慮する必要がある。

表 3 LM および Switching 手法の特性

	マイグレーション時間	1 コネクション時間	VM の数	QoS 性能低下
LM	VM のメモリーサイズに依存	長くてもよい	最低限の VM 数で運用可	考慮が必要
Switching	ファイルサイズに依存	短い方がよい	冗長に必要	考慮が必要

5. プロセッサコアへの VM 割り当てによる消費電力変化

VM をサーバ上運用していく場合には VM をどのようにサーバ内のプロセッサ上で動作させるかといった柔軟な運用方法が考えられる。たとえば QoS 性能制約から特定の VM を特定のプロセッサコアに pin down させることなどがあげられる。しかし、そのような場合性能制約という観点からすれば負荷が低い場合には pin down 無しに QoS 性能を維持できる可能性が十分にあり得る。このとき、pin down の制約から解放されてプロセッサ動作周波数を低くできればより低い消費電力での運用が可能であると考えられる。そして前章での評価のような負荷のサーバ間マイグレーションによって VM を集約するときにも同様の運用方法が可能であると考えられる。また、文献⁵⁾ ではワークロードが要求する計算資源をプロセッサ周波数ベースで評価して省電力化を行っているが、サーバノード内でプロセッサ周波数をどのように VM へ割り当てるかについて言及がなされていない。そこで、本章ではサーバノードが持つ最大性能を要求しない状況下での VM のプロセッサコアへの pin down と消費電力変化について述べる。

5.1 評価方法

本評価では 3 章で使用したシステムを用いて 1 台のサーバ上で VM を動作させてベンチマークを実行した。表 4 に実験環境を示す。図 2 での使用プロセッサコア数 1、動作周波数 f_{max} での動作時に得られる最大性能に対して、割り当てるプロセッサコア数 N と動作周波数 f を変化させて比較を行う。 $N \times f \geq f_{max}$ を満足するようにそれらの値を変化させた。また、ベンチマークの実行時間を 3 章での評価と同じく 120[sec] に設定した。以下ではベンチマーク環境を (Banking, 2 cores, 1.6GHz) のように

(ベンチマークセット名, 使用プロセッサコア数, 動作周波数)

の組合せを用いて表現する。

5.2 評価結果

図 10 に 1VM 動作時のそれぞれのベンチマークにおける QoS 満足率と相対消費電力値を示す。ここで、相対消費電力はプロセッサコア数 1、動作周波数 f_{max} での数値を 1 としている。図 10 より、VM に割り当てるプロセッサコア数を増やして動作周波数を下げ

表 4 評価ベンチマークセットと (N, f) の組合せ

Benchmark set	# of sessions	# of VMs	(N, f)	
Banking	1000	1	(1, 2.93GHz)	(2, 1.6GHz)
E-commerce	1000	1	(1, 2.93GHz)	(2, 1.6GHz)
Support	800	1	(1, 2.93GHz)	(2, 1.6GHz)
Bank. - E-comm.	1000(Bank.)/1000(E-comm.)	2	(1, 2.93GHz)	(3, 2GHz)
Bank. - Supp.	1000(Bank.)/800(Supp.)	2	(1, 2.93GHz)	(3, 2GHz)

ることで、すべてのベンチマークで約 10%のエネルギー削減を達成していることが分かる。しかし、QoS 性能に関して僅かにその値が低下した。これは Xen の Dom0 が動作するプロセッサコアも動作周波数が下がること、および VM へ割り当てるプロセッサ数が増えてスケジューリングへの負荷が増えることで Dom0 での負荷が相対的に増加したためであると考えられる。その結果 (Banking, 2 cores, 1.6GHz) 環境では元々の QoS 満足率が約 95% であったためにその値を下回ったが、その他のベンチマークでは QoS 満足率 95% 以上を維持できている。

図 11 に 2VM 並行動作時のそれぞれのベンチマークセットにおける平均 QoS 満足率と相対エネルギー値を示す。図 10 と同様、相対エネルギー値はプロセッサコア数 1、動作周波数 f_{max} での数値を 1 としている。なお、Xen の Credit スケジューラではプロセッサ 3 コアでの動作時に両方のベンチマークセットにおいて極端な性能低下が見られた。これは文献¹²⁾にて述べられているように pre-emptive なスケジューリング手法であることおよびそのスケジューリング間隔が長い (30[msec], 固定値) ため、リアルタイム性が損なわれたものと考えられる。そこで、Xen の SEDF(Simple Earliest Deadline First) スケジューラを利用してスケジューリング間隔を短くすることで (Banking, E-commerce:2[msec], Support:1[msec]) プロセッサ 3 コアを使用して 2VM を動作させた。しかし Banking - Support ベンチマークセットでは必要最低限の動作周波数 2GHz で 95% TIME_GOOD QoS を満足できなかったため、2.26GHz に設定してベンチマークを行った。図 11 より、プロセッサ 3 コアに 2VM を割り当てるとき、適切な割り当てを行うことができれば性能を低下させずにより低いプロセッサ周波数で QoS 性能を維持できることが分かる。その結果として Banking - Support ベンチマークセットでは約 9%, Banking - E-commerce ベンチマークセットでは約 5%のエネルギー削減を達成出来ていることが分かる。

5.3 考 察

前項での評価により、同じ要求性能を達成するのに使用プロセッサコアを増やし、動作周波数を低く設定することが省電力化にとって有効であることが分かる。これまでに提案され

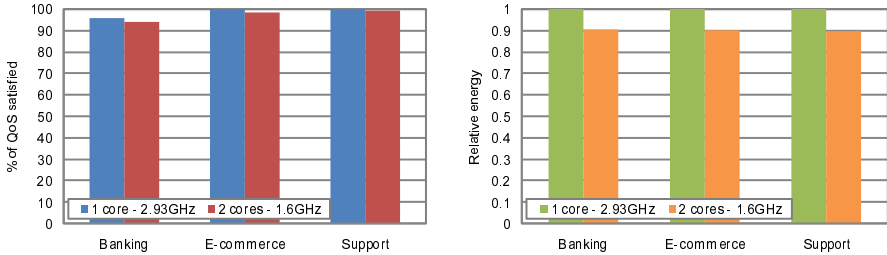


図 10 1 VM 動作時の QoS 満足率と相対エネルギー

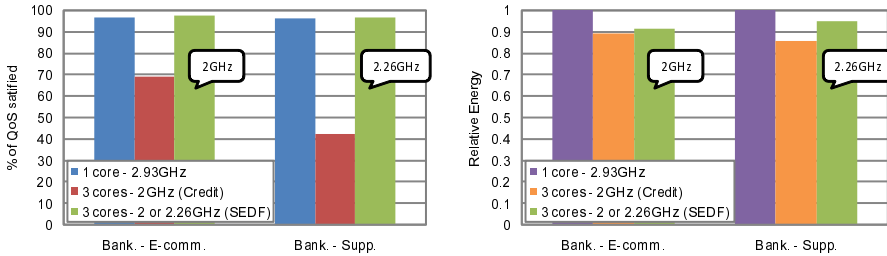


図 11 2 VM 並行実行時の QoS 満足率と相対エネルギー

てきている VM を使用しない分散型サーバ向け省電力化手法に対し、このような消費電力特性を考慮させることで VM 運用時にも省電力化が可能となる。これには文献²⁾にて提案されている VPM States を応用することで実現できる。VPM States はプロセッサで設定可能な動作周波数に対して VM プロセッサにスケジューリングされるタイムスライスを制御することで仮想的な動作周波数を実現する仕組みである。図 12 にどのようなアプローチによって VM と DVFS を制御するかについて、余剰プロセッサコア数 $N = 3$, 利用可能なプロセッサ周波数 f の数が 4 ($f(3) < f(2) < f(1) < f(0)$) の場合を例にして示す。このときプロセッサ全体で利用可能な資源 (プロセッサ周波数) の最大値は $N \times f(0)$ と表すことができる。その範囲内にて等間隔でいくつかの仮想プロセッサ周波数 VP を設定する。それぞれの VP の値に対し、 $N * f(i+1) < VP \leq N * f(i)$ を満たす $f(i)$ に対してタイムスライスを設定すれば電力最適な VP を作り出すことが可能となる。このようにして得られる VP を実際のプロセッサ周波数として考えれば、そのときに処理可能な最大のリクエストを基にして省電力化手法に適用することが可能となる。

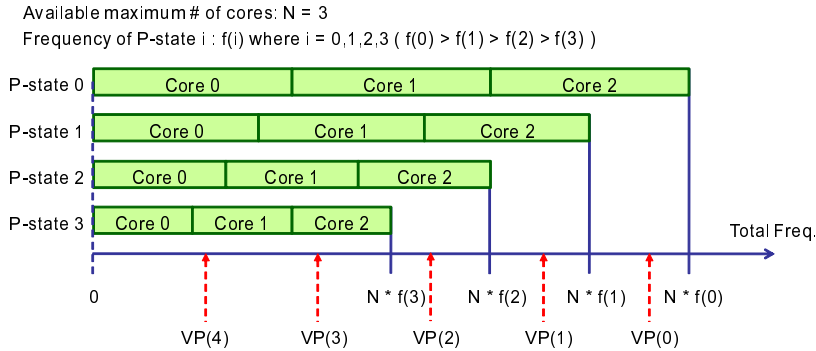


図 12 より電力最適な仮想 P-state

6. おわりに

本稿では仮想計算機 (VM) が動作するサーバシステムでの省電力化を目的として、1 台のサーバ上で VM を動作させるときの QoS 性能および消費電力効果について評価を行い、さらにマイグレーション時における QoS 性能と消費電力特性の評価および単一サーバ上における VM のプロセッサコア割当てによる QoS 性能と消費電力の特性について評価を行った。

1 台のサーバ上で VM を動作させると、DVFS を用いることで QoS 性能を低下させることなく消費電力を削減出来ることが分かった。そして 1 台のサーバに複数の VM を集約させることで大幅な消費電力削減が出来ることが分かった。マイグレーション時における QoS 性能と消費電力特性の評価では、ライブマイグレーション手法および従来の分散型 Web サーバ用の手法を適用することにより若干の性能低下と消費エネルギー増加でサーバ間の負荷マイグレーションを行えることが分かった。また、単一サーバ上における VM のプロセッサコア割当てによる QoS 性能と消費電力評価では使用プロセッサコア数を多くかつ動作周波数を低くすることで最大約 10%少ないエネルギーで同程度の性能を達成できることが分かった。

今後は、本稿での知見と考察、そして我々が以前に提案した Web サーバ向けの省電力化手法¹⁾を基にして VM を用いる分散型 Web サーバシステムでの省電力化について具体的なアルゴリズムの考案および実装を行う予定である。

謝辞 本研究の一部は、NEDO「グリーンネットワーク・システム技術研究開発プロジェクト (グリーン IT プロジェクト) / エネルギー利用最適化データセンタ基盤技術の研究開発 / データセンタのモデル設計と総合評価」による。

参考文献

- 1) 今田貴之ほか: 分散型 Web サーバでの負荷変動を考慮した省電力化のためのノード状態制御, 情報処理学会論文誌コンピューティングシステム, Vol.2, No.2 (To appear).
- 2) Nathuji, R. and Schwan, K.: VirtualPower: Coordinated Power Management in Virtualized Enterprise Systems, in *Proceedings of the 21st ACM SIGOPS Symposium on Operating Systems Principles(SOSP'07)* (2007).
- 3) Verma, A. et al.: pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems, in *Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware* (2008).
- 4) Hu, L. et al.: Magnet: A Novel Scheduling Policy for Power Reduction in Cluster with Virtual Machines, in *Proceedings of IEEE International Conference on Cluster Computing(Cluster 2008)* (2008).
- 5) Kusic, D. et al.: Power and Performance Management of Virtualized Computing Environments Via Lookahead Control, in *Proceedings of 2008 International Conference on Autonomic Computing(ICAC'08)* (2008).
- 6) Standard Performance Evaluation Corporation: SPECWeb2005. <http://www.spec.org/web2005/>.
- 7) Paul, B. et al.: Xen and the Art of Virtualization, in *Proceedings of the 19th ACM symposium on Operating systems principles(SOSP'03)* (2003).
- 8) LVS: Linux Virtual Server. <http://www.linuxvirtualserver.org/>.
- 9) Hotta, Y. et al.: Profile-based optimization of power performance by using dynamic voltage scaling on a PC cluster, in *Proceedings of High Performance Power-Aware Computing Workshop(HPPAC) in the 20th IEEE IPDPS'06* (2006).
- 10) libvirt: The virtualization API. <http://libvirt.org/>.
- 11) Clark, C. et al.: Live Migration of Virtual Machines, in *Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation(NSDI'05)* (2005).
- 12) Cherkasova, L. et al.: Comparison of the Three CPU Schedulers in Xen, *ACM SIGMETRICS Performance Evaluation Review*, Vol.35, No.2 (2007).