

## エコモデルに基づいた大規模分散システムの研究

滝沢, 誠 / TAKIZAWA, Makoto

---

(雑誌名 / Journal or Publication Title)

科学研究費助成事業 研究成果報告書

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2019-05-16

令和元年5月16日現在

機関番号：32675

研究種目：基盤研究(B) (一般)

研究期間：2015～2018

課題番号：15H02695

研究課題名(和文) エコモデルに基づいた大規模分散システムの研究

研究課題名(英文) Scalable Distributed Systems based on Echo Model

研究代表者

滝沢 誠 (TAKIZAWA, Makoto)

法政大学・理工学部・教授

研究者番号：80188119

交付決定額(研究期間全体)：(直接経費) 13,900,000円

研究成果の概要(和文)：情報システムで消費される電力量を低減するために、応用プロセスを実行したときにコンピュータ全体で消費される電力量を考えるソフトウェアの観点にたったマクロレベル方法論に基づいて研究を行った。本研究課題では、プロセスをサーバ間で移行させることにより、サーバのクラスタ全体の消費電力量を低減させる方式を新たに提案している。特に、仮想マシンを用いて、仮想マシンをサーバ間で移行させるアルゴリズムを研究した。開発したエコ分散シミュレータを用いて評価を行い、仮想マシン移行アルゴリズムにより、クラスタ全体の消費電力を低減し、各プロセスの実行時間を短縮できることを示した。

研究成果の学術的意義や社会的意義

地球環境を守るために、情報システムでも消費電力を低減することが喫緊の課題となっている。これまでCPU等のハードウェア要素の省電力化はメーカー等で行われてきたが、コンピュータではプロセスが実行されることにより電力が消費される。このため、本研究課題では、あらたにソフトウェアの観点から、各種の応用プロセスを実行したときに消費される電力を低減することを目的としている点に新規性と有用性がある。IoTに代表されるような大規模化する情報システムで必須の課題となる省電力化のソフトウェア基盤となる研究である。

研究成果の概要(英文)：In this research, we newly proposed a macro-level approach to reduce the total energy consumption of a server to perform application processes. We newly discussed a migration approach where processes migrate to another server which is expected to consume smaller energy. Especially, processes migrate from a host server to a guest server by migrating virtual machines. By using virtual machines, processes on virtual machines can migrate without suspending the processes, i.e. live migration. As number of processes increases, virtual machines are newly created. A group of virtual machines migrate among servers. We also developed an eco simulator where the electric energy to be consumed by servers to perform application processes can be obtained. In the evaluation using the eco simulator, the total electric energy consumption [J] of servers and average execution time [sec] of processes can be reduced in the virtual machine algorithm proposed in this research compared with other algorithms.

研究分野：情報工学

キーワード：ネットワーク 情報システム 分散システム 消費電力モデル アルゴリズム エコ情報システム

### 1. 研究開始当初の背景

近年の情報システムは、コンピュータのみならずセンサ、携帯端末、家電、車載機器等の多種多様な情報機器をノードとして無線、有線等の種々のネットワークで相互接続された大規模なシステムとなり、「もののインターネット(IoT)」と呼ばれている。情報システムの形態には、クラウド・システムのような集中的な管理を行うモデルと各ノードが自律的かつ対等に動作を行う分散型とがある。P2P(Peer-to-Peer)モデルが分散型の例である。京都議定書以来、地球規模でのCO<sub>2</sub>削減に向けた環境問題が喫緊の課題となっている。社会の重要なインフラである情報システムでも、消費される電力の低減が重要な課題となってきた。本研究では、大規模な分散型システムを対象として電力消費を低減するモデルについての研究を行なう。

CPU、メモリ、二次記憶装置等のハードウェア構成要素、アーキテクチャについては、インテル等の企業が開発をすすめて、低消費電力のコンピュータ、情報機器が提供されてきている。サーバ等のコンピュータの動作は、ソフトウェアを含めて決定されるものであり、コンピュータ全体の消費電力は、個々のハードウェア要素の消費電力を合成することによって得ることは困難である。申請者の研究グループでは、これまでにサーバで種々のプロセスを実行させたときのコンピュータ全体の消費電力を実測することにより、形式的な電力消費モデルを求める研究を行ってきた。

### 2. 研究の目的

情報システムは、IoTといわれるように、サーバ、センサ等の多種多様な情報機器ノードが多様なネットワークにより相互接続されてきている。これまで研究してきた消費電力モデルは、一定の型、例えばCPUのみを利用する応用プロセスをサーバで実行したときのサーバ全体の消費電力を与えるモデルであった。これに対して、本研究では、種々のマルチコアCPUを備えたコンピュータ、センサ等の種々のノードと、CPUのみならず通信、ストレージも利用する種々のプロセスを対象とした電力消費モデルを構築することが目的である。CPU、通信、二次記憶装置等を複合的に利用する種々の応用プロセスを考え、これらを実行したときのノード全体の消費電力と各プロセスの実行時間を測定し、消費電力に最も影響を与えるパラメータを明確化する。これにより、種々のノードのマクロレベルな消費電力モデルを構築する。これまで研究し提案してきた単純電力消費(SPC: simple power consumption)モデルを基本として、コンピュータのCPUのコア数、スレッド数、通信速度、二次記憶装置のアクセス量を加味した**多レベル電力消費(MLPC: multi-level power consumption)モデル**を構築する。

情報システムでは、高性能化と高信頼化、さらに低消費電力化が必要となっている。クラウド等のサーバクラスでは、CPU、ストレージ等の計算資源を仮想化し、各サーバのアーキテクチャ、OS等の異種性と位置と独立に仮想計算資源を利用できるように仮想マシン(virtual machine)が応用に提供され広く利用されている。本研究課題では、仮想マシン技術を利用して、クライアントのアプリケーションから発行された応用プロセス(以降、プロセスとする)を、どの仮想マシンで実行するかをまず決定することが必要となる。信頼性を向上するためには、異なるサーバ上の複数の仮想マシンで実行する。このために、電力消費モデルを用いて、プロセスを、より消費電力量を少なく実行できるサーバを検出する。さらに、プロセスが実行されているサーバで、予想以上の電力量が消費されるときには、仮想マシンを他のサーバに移行する(migrate)ことにより、クラスタ全体の消費電力の低減をはかる。本研究課題では、仮想マシンの選択アルゴリズムと、仮想マシンの移行アルゴリズムを研究し、本研究課題で開発したエコ分散シミュレータにより評価を行う。

### 3. 研究の方法

分散型の大規模システムの消費電力低減方式を研究するためには、システムを構成する各種コンピュータが各種の応用プロセスを実行するためにどのように電力を消費するかを示す形式的な**電力消費モデル**と、各プロセスの実行時間を示す**プロセス実行モデル**が必要となる。まず、各コンピュータで種々の応用プロセスを実行させたときにコンピュータ全体で消費される電力[W]を電力測定器で測定する。このときのノード全体の消費電力を電力測定器により実測する。これらの実測データを解析することにより、ノードの消費電力を決定するパラメータを明確化する。これまでに申請者等は、コンピュータの消費電力モデルとしてSPC(Simple Power Consumption)モデルを提案している。ここでは、コンピュータでプロセスが一つでも実行されていれば最大の電力maxE[W]をコンピュータが消費し、一つのプロセスも実行されていないときには最小の電力minE[W]を消費するものである。SPCモデルでは、多コア、多スレッドのCPUを検討していなかったが、本研究では多コア・多スレッドのCPUのコンピュータを対象として、電力消費モデルを研究する。一つでもプロセスが実行されているコアとスレッドをアクティブ(active)であるとする。本研究では、消費電力をアクティブなコアとスレッドの関係を明らかにする。さらに、通信型、ストレージ型のプロセスについても、同様なアプローチにより、応用プロセスを実行させたときの消費電力の実測データを基に消費電力モデルを構築する。また、計算、通信、ストレージ型が混在した一般的プロセスが実行されたときの電力消費モデルと実

行時間の関係を示すプロセス処理モデルを明らかにする。

コンピュータで種々のプロセスを実行させたときの消費電力を 100 [msec]以下の周期で測定できる電力測定器を導入する。実験データをデータベース・システムに収集し解析するためのサーバを導入する。

クラウド・コンピューティング・システムのような情報システムの特徴の一つは大規模化、高信頼化、高性能化に加えて、サーバで消費される電力量を低減する必要がある。本研究課題では、仮想マシン技術を利用して、以下を研究する。

(1) クライアントのアプリケーションから発行された応用プロセス(以降、プロセスとする)を、どの仮想マシンで実行するかをまず決定することが必要となる。信頼性を向上するためには、異なったサーバ上の複数の仮想マシンで実行する。このために、電力消費モデルを用いて、プロセスを、もっとも消費電力を少なく実行できるサーバを選択する。

(2) さらに、プロセスが実行されているサーバで、予想以上の電力が消費されるときには、仮想マシンを他のサーバに移行する(migrate)ことにより、クラスタ全体の消費電力の低減をはかる。

(3) 本研究課題では、仮想マシンの選択アルゴリズムと、仮想マシンの移行アルゴリズムを研究し、本研究課題で開発したエコ分散シミュレータにより評価を行う。

#### 4. 研究成果

本研究課題により、ソフトウェアの観点から情報システムの消費電力量を低減するためのモデルとアルゴリズムを明らかにできた。サーバ等のコンピュータの消費電力は、CPU等の個々のハードウェアを起動させるプロセスの動作によって決まる。この観点から、応用プロセスを動作させたときのコンピュータの消費電力を測定し、消費電力に最も影響を与える因子を明らかにすることにより、電量消費モデルを明らかにした。個々のハードウェア要素の低消費電力化に対して、応用プロセスを実行したときに消費されるコンピュータ全体の電量を考えるマクロレベル方法論である。

クラスタ内のサーバ  $s_t$  を考える。ここで、 $np_t$  を、サーバ  $s_t$  の CPU 数、 $cc_t$  を各 CPU 内のコア数、 $ct_t$  を各コア内のスレッド数とする。サーバ  $s_t$  の総スレッド数  $nt_t$  は、 $nc_t \cdot cc_t \cdot ct_t$  である。サーバ  $s_t$  で、 $n$  個のプロセスを実行するときに消費される電力[W]は以下となる。

$$NE_t(n) = \min E_t + nap_t(n) bE_t + nac_t(n) cE_t + nat_t(n) tE_t$$

ここで、 $nap_t(n)$ 、 $nac_t(n)$ 、 $nat_t(n)$  は、アクティブ(プロセスが実行されている)CPU、コア、スレッド数である。

次に、プロセス  $p_i$  を、サーバ  $s_t$  で実行することを考える。プロセス  $p_i$  を、他のプロセス無に単独で実行したときの実行時間を  $\min T_{ti}$  [tu(time unit)] とする。クラスタ内で、 $\min T_{ti}$  が最小となるサーバ  $s_t$  を最速サーバとし、これの  $\min T_{ti}$  を  $\min T_i$  とする。 $\min T_i$  は、プロセス  $p_i$  の計算総量を与えるものである。最速サーバ  $s_f$  の計算速度  $CRT_f$  を 1 と仮定する。他のサーバ  $s_t$  の計算速度  $CRT_t$  は、 $CRT_f \cdot \min T_i / \min T_{ti} = \min T_i / \min T_{ti}$  となる。 $n$  個のプロセスを実行しているときのサーバ  $s_t$  の計算速度は以下となる。

$$n \cdot \min T_i / \min T_{ti} \quad (\text{for } n < nt_t)$$

$$NSR_t(n) = \{$$

$$nt_t \cdot \min T_i / \min T_{ti}$$

このときの、サーバ  $s_t$  で実行中の各プロセス  $p_i$  の計算速度  $NPR_{ti}(n)$  は  $NSR_t(n) / n$  となる。

この電力消費モデルに基づいて、各種のプロセスを実行させたときのコンピュータの消費電力とプロセスの実行時間の見積もりを与えるエコ分散シミュレータの開発を行った。データベースを中心に、SQL により開発を行った。シミュレータの概要を以下に示す。

for each time

for each server  $s_t$ ,

if a process  $p_i$  starts,  $\{pl_i = \min T_i; \text{state}_i = \text{Active};\}$

$n$  = number of active processes on  $s_t$ ;

$c$  = computation rate of each process =  $NSR_t(n) / n$ ;

if ( $n > 0$ )  $EE_t = EE_t + NE_t(n)$ ;

for each active process  $p_i$  on  $s_t$ ,

$pl_i = pl_i - c$ ;

if  $pl_i \leq 0$ ,  $p_i$  terminates;

$EE_t$  は、サーバ  $s_t$  の総消費電力量[J]である。サーバのクラスタ全体の消費電力を低減するために、まず、クライアントから発行されたプロセスを実行するサーバを選択するアルゴリズムの研究を行った。特に、サーバ上の仮想マシンでプロセスを実行させるものである。電力消費モデルを用いて、プロセスを実行したときの消費電力量が最小のサーバ内の仮想マシンでプロセスを実行させる。クラスタ内の仮想マシンを実行されるプロセス数が増加するにともない、新たに仮想マシンを生成させることにより、各仮想マシンの負荷を低減する動的アルゴリズムを考案した。このためには、プロセスを実行させたときの実行時間と消費電力を見積もることが必要となる。各種のプロセスの実行時間の見積もりは困難であるので、以下の仮定を設ける。

[仮定] 実行中のプロセス  $p_i$  は、計算量  $t_t(n_t) \cdot n_t$  を実行せねばならない。 $n_t$  はサーバ  $s_t$  上で

実行されているプロセス数である。ここで、

$$\alpha_t(n_t) = \begin{cases} 0.5 & \text{for } n_t \leq nt_t. \\ 0.6 & \text{for } nt_t < n_t \leq 2 \cdot nt_t. \\ 0.8 & \text{for } 2 \cdot nt_t < n_t \leq 4 \cdot nt_t. \\ 1 & \text{for } n_t > 4 \cdot nt_t. \end{cases}$$

$n_t$  個のプロセスは実行中のサーバ  $s_t$  に、 $k$  個の新しいプロセスが実行されること考える。このときの実行時間  $SET_t$  と消費電力量  $SEE_t$  は以下で与えることができる。

$$SET_t(n_t, k) = (\alpha_t(n_t) \cdot n_t + k) / NSR_t(n_t + k).$$

$$SEE_t(n_t, k) = (\alpha_t(n_t) \cdot n_t + k) \cdot NE_t(n_t + k) / NSR_t(n_t + k).$$

プロセス  $p_i$  が、サーバ  $s_t$  に発行されたときに消費される電力量は、 $SEE_t(n_t, 1)$  となる。また実行時間は、 $SET_t(n_t, 1)$  となる。したがって、 $SEE_t(n_t, 1)$  が最小値なるサーバ  $s_t$  をクラスタから選ぶ。ついで、サーバ  $s_t$  上の仮想マシンの中で、実行しているプロセス数が最少の仮想マシンを選び、プロセス  $p_i$  を実行させる。

さらに、仮想マシンで実行されるプロセス数が増大するにともない、仮想マシンを新たに生成する動的なアルゴリズムも研究した。

次に、サーバ上の仮想マシンを、より消費電力の小さいサーバに移行させる仮想マシン移行アルゴリズムの研究を行った。仮想マシンを移行させたときと、移行させなかったときの消費電力を求め、より消費電力を低減できる方を選択するものである。さらに、複数の仮想マシンを移行させる仮想マシンのグループ移行アルゴリズムを考案した。

これらのアルゴリズムは、エコ分散シミュレータにより評価を行った。RR, SGEA アルゴリズムは、仮想マシンを移行させないものである。SGEA は、仮想マシンを動的に生成しないアルゴリズムで、DM は同的に生成するアルゴリズムである。SM(1)とDM(1)は、仮想マシンを移行するとき、1つの仮想マシンを移行するアルゴリズムである。SM(\*)とDM(\*)は、複数の仮想マシンのグループを移行させるアルゴリズムである。図2からわかるように、仮想マシンを、プロセス数の増加に伴い造作させる動的で、かつ複数の仮想マシンを移行させるアルゴリズムにより、クラスタ全体の消費電力を低減できることがわかる。

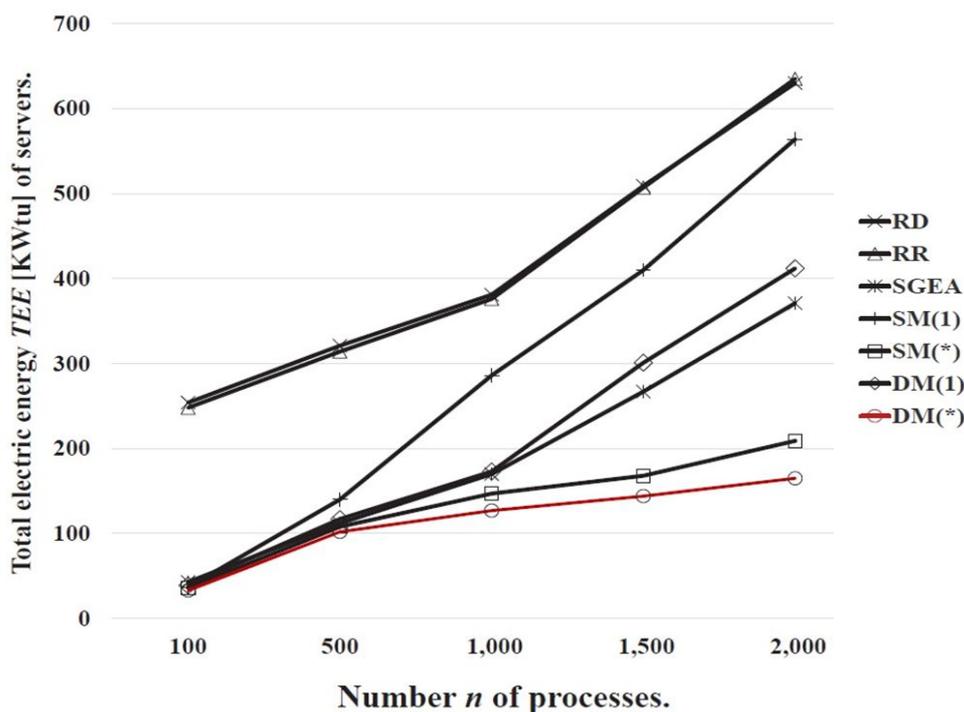


図2 消費電力量

5 . 主な発表論文等

[雑誌論文](計 15 件)

1. S. Sakamoto, K. Ozera, A. Barolli, M. Ikeda, L. Barolli, M. Takizawa: Implementation of an Intelligent Hybrid Simulation Systems for WMNs Based on Particle Swarm Optimization and Simulated Annealing: Performance Evaluation for Different Replacement Methods, *Soft Computing*, 23(9), 2019, pp.3029-3035 (DOI:10.1007/s00500-017-2948-1).
2. S. Nakamura, T. Enokido, M. Takizawa: A Flexible Read-Write Abortion Protocol with Role Safety Concept to Prevent Illegal Information Flow, *Journal of Ambient Intelligence and Humanized Computing*, 9(5), 2018, pp.1415-1425 (DOI: 10.1007/s12652-017-0541-1).
3. S. Nakamura, M. Sugino, M. Takizawa: Algorithms for Energy-efficient Broadcasting Messages in Wireless Networks, *Journal of High Speed Networks*, 24(1), 2018, pp.1-15 (DOI: 10.3233/JHS-170577).
4. S. Nakamura, L. Ogiela, T. Enokido, M. Takizawa: An information flow control model in a topic-based publish/subscribe system, *Journal of High Speed Networks*, 24(3), 2018, pp.243-257 (DOI: 10.3233/JHS-180593).
5. D. Duolikun, H. Kataoka, T. Enokido, M. Takizawa: Simple Algorithms for Selecting an Energy-Efficient Server in a Cluster of Servers, *International Journal of Communication Networks and Distributed Systems*, 21(1), pp.1-25, 2018 (DOI: 10.1504/IJCNS.2018.093383).
6. R. Oma, S. Nakamura, D. Duolikun, T. Enokido, M. Takizawa: An Energy-efficient Model for Fog Computing in the Internet of Things (IoT), *Internet of Things*, 1-2, 2018, pp.14-26 (DOI: 10.1016/j.iot.2018.08.003).
7. U. Ogiela, M. Takizawa, L. Ogiela: Cognitive Approaches for Medicine in Cloud Computing, online published at *Journal of Medical Systems*, 42(4), pp.70:1-70:5, 2018 (DOI: 10.1007/s10916-018-0918-5).

他 9件

[学会発表](計 42件)

1. R. Oma, S. Nakamura, D. Duolikun, T. Enokido, M. Takizawa: Energy-Efficient Recovery Algorithm in the Fault-Tolerant Tree-Based Fog Computing (FTBFC) Model, *Proc. of the 33rd International Conference on Advanced Information Networking and Applications (AINA-2019)*, Matsue, Japan, Mar. 2019, pp.132-143.
2. M. Cuka, D. Elmazi, M. Ikeda, K. Matsuo, L. Barolli, M. Takizawa: Selection of IoT Devices in Opportunistic Networks: A Fuzzy-Based Approach Considering IoT Device's Selfish Behaviour, *Proc. of the 33rd International Conference on Advanced Information Networking and Applications*, Matsue, Japan, Mar. 2019, pp.251-264.
3. Barolli, S.Sakamoto, L. Barolli, M. Takizawa: A Hybrid Simulation System Based on Particle Swarm Optimization and Distributed Genetic Algorithm for WMNs: Performance Evaluation Considering Normal and Uniform Distribution of Mesh Clients, *Proc. of the 21st International Conference on Network-Based Information System*, Bratislava, Slovakia, Sept. 2018, pp.42-55.
4. D. Duolikun, S. Nakamura, T. Enokido, M. Takizawa: An Energy-Efficient Dynamic Live Migration of Multiple Virtual Machines, *Proc. of the 21st International Conference on Network-Based Information System*, Bratislava, Slovakia, Sept. 2018, pp.87-98.
5. S. Nakamura, T. Enokido, M. Takizawa: Evaluation of a Protocol to Prevent Malicious Information Flow in P2PPS Systems, *Proc. of the 12th International Conference on Complex, Intelligent and Software Intensive Systems*, Matsue, Japan, Jul. 2018, pp.102-114.D.
6. Duolikun, R. Watanabe, T. Enokido, M. Takizawa: An Eco IDMMV Migration Algorithm of Dynamic Virtual Machines in a Server Cluster, *Proc. of the 12th International Conference on Complex, Intelligent and Software Intensive Systems*, Matsue, Japan, Jul. 2018, pp.115-126.
7. R. Oma, Q. Cuiqin, S. Nakamura, T. Enokido, M. Takizawa: Replication Models of Fog and Device Nodes in IoT, *Proc. of IEEE the 32nd International Conference on Advanced Information Networking and Applications*, Krakow, Poland, May 2018, pp.301-306.
8. T. Enokido, D. Duolikun, M. Takizawa: An Energy-Efficient Process Replication Algorithm Based on the Active Time of Cores, *Proc. of IEEE the 32nd International Conference on Advanced Information Networking and Applications*, Krakow, Poland, May 2018, pp.165-172.
9. D. Duolikun, R. Watanabe, T. Enokido, M. Takizawa: An Eco Migration Algorithm of Virtual Machines in a Server Cluster, *Proc. of IEEE the 32nd International Conference on Advanced Information Networking and Applications*, Krakow, Poland, May 2018, pp.189-196.

他 33件

[図書](計 0件)

〔産業財産権〕

出願状況（計 0 件）

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年：  
国内外の別：

取得状況（計 0 件）

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年：  
国内外の別：

〔その他〕

ホームページ等

## 6 . 研究組織

### (1) 研究分担者

研究分担者氏名：榎戸 智也

ローマ字氏名：(ENOKIDO, tomoya)

期間名：立正大学

部局名：経営学部

職名：教授

研究者番号（8桁）: 10360158

分担者氏名：Barolli Leonard

ローマ字氏名：(BAROLLI, leonard)

期間名：福岡工業大学

部局名：情報工学部

職名：教授

研究者番号（8桁）: 40312722

### (2) 研究協力者

研究協力者氏名：

ローマ字氏名：

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。