

DBMS を適用した複数システムにおける Application Platform Energy Effectiveness (APEE) の適用方法の検討

五十嵐和人[†] 吉野 松樹[†] 大田原 実[†] 原 聖宣[†]

[†] 株式会社日立製作所 〒244-0817 神奈川県横浜市戸塚区吉田町 292 番地

E-mail: [†]{kazuto.igarashi.mk,matsuki.yoshino.pw,makoto.ootahara.ob,kiyonori.hara.zq}@hitachi.com

あらまし IT を活用した様々なサービスが登場し、我々の生活が便利で豊かになることに伴い、IT で消費する電力量は増加している。また、日本においては 2050 年までのカーボンニュートラル実現が宣言されており、実現に向けては IT のさらなる活用が必要であり、このためにも IT 分野自身の省エネルギー化が求められている。IT 分野におけるエネルギー効率指標として、データセンタの Power Usage Effectiveness (PUE) やサーバの Server Energy Effectiveness Metric (SEEM) などが国際標準規格となっており、これらを活用して省エネルギー化が進んでいる。加えて、ハードウェアだけでなく、DBMS などミドルソフトウェアまで含めたプラットフォームとしてのエネルギー効率も重要になっている。日本が 2018 年に提案した、ソフトウェアまで含めたエネルギー効率指標である Application Platform Energy Effectiveness (APEE) の規格開発が進められ、国際標準規格として発行間近となっている。データセンタの規模を問わず、リソース利用効率の改善のために、仮想化によるサーバの集約が広く行われている。集約後のサーバのリソースが十分であれば、集約後の性能が単独実行時の性能値から推定可能である。本研究では、DBMS を適用した単独サービスで測定した APEE から、DBMS を適用した複数サービスが混在した環境におけるエネルギー効率を推定する方法を提案する。また、予備実験として仮想化による集約環境を作成してエネルギー効率を計算してモデルを評価し、報告する。

キーワード エネルギー効率, ベンチマーク, 省エネルギー化, データセンター, IoT

1 はじめに

SDGs やカーボンニュートラルなどの全世界的な流れの中で、IT を活用することの重要性は増しており、一方では IT 自体の一層の省エネルギー化が求められる。IT のエネルギー効率向上のために、データセンタの単位では、データセンタ全体での IT 機器および付帯設備で使用する電力量に着目した PUE や、サーバのエネルギー効率 SEEM が国際標準規格となっている。しかし、ハードウェアのエネルギー効率の向上だけでは、限られたエネルギー供給の中で爆発的に増加するデータ処理の要求を満たすことは難しい。

ハードウェアの使い方の工夫による消費電力の削減は、従来から様々な研究が行われてきた [1] [2]。近年では、データベース内のデータの配置を制御することにより、ソフトウェアの大幅な性能低下を起こすことなくソフトウェア実行中のストレージの消費電力量を削減できることが示されている [3]。これは、同程度の性能をもつアプリケーションであっても、アプリケーション部分の工夫によってエネルギー効率が異なることを意味する。そのため、ハードウェアだけでなく、DBMS などミドルソフトウェアまで含めたプラットフォームとしてのエネルギー効率も重要になっている。

ソフトウェアまで含めたプラットフォームとしてのエネルギー効率指標が日本から提案されて、ISO/IEC 23544 Application Platform Energy Effectiveness (APEE) として規格開発が進

んでいる。2020 年 12 月現在、Draft International Standard (DIS) として公開されており、正式な International Standard (IS) としての発行が近い。この指標は、サーバやストレージ装置などの IT 機器、および OS とミドルウェアからなる基盤をアプリケーションプラットフォームと定義し、ベンチマークプログラムを選び、ベンチマークプログラムの実行期間での IT 機器の消費電力量を測定し、エネルギー効率を算出するものである。APEE の測定事例も公開され始めている [4]。

APEE は、プラットフォームをベアで使用情况の場合のエネルギー効率である。現在の IT においては、仮想化技術を活用して、様々なシステムを一つのサーバに集約することが、リソースの利用効率の観点から進んでいる。本研究では、DBMS を適用した単独サービスをベア環境で測定した APEE を使って、そのサービスを複数のサービスが混在する仮想化による集約環境に入れた際のエネルギー効率を推定するモデルを提案する。このモデルを活用することで、ベア環境でのエネルギー効率を用いて、実際の集約環境の省エネルギー化を進めることが可能となる。本研究は、このモデルを検証するための予備実験として、ベア環境で測定したエネルギー効率と集約後の環境で測定したエネルギー効率を比較する。

本論文の構成を以下に示す。2 節ではエネルギー効率推定のモデルを示し、3 節において実測結果とモデルの比較を行い、4 節で本論文をまとめる。

2 提案モデル

本研究では、複数のシステムが稼働している環境での、特定のシステムのエネルギー効率を推定するモデルを提案する。ISO/IEC 23544 Application Platform Energy Effectiveness で定義しているエネルギー効率指標 APEE は、特定の環境をベアで使った場合の、アプリケーションプラットフォームのエネルギー効率である。そのため、APEE の規格としては、他のシステムの実行の電力分を差し引くことは許していない[5]。筆者らは、ベア環境で測定したエネルギー効率 APEE の値から、システムを集約した環境における特定のシステムのエネルギー効率 e を推定するモデルとして、式 1 を提案する。

$$e = \alpha \text{APEE} \quad (1)$$

ここで、 α は集約することによるエネルギー効率の変化の割合を示す。このモデルが意味するところは、ベア環境におけるエネルギー効率の優劣は、仮想化の環境に移行しても同様の優劣を持つことである。

一般に、エネルギー効率は、成果と、その成果を得るために投入したエネルギーの比率である。そのため、集約後の環境における特定のシステムのエネルギーを求めるには、ホストマシンそのものの全体の電力量ではなく、特定のシステムが稼働する仮想マシンにおいて成果を得るために消費した電力量を求める必要がある。そこで筆者らは、仮想マシンの消費電力量を推定する方法として、[6] で提案されているサーバの消費電力量の推定手法を参考に、稼働中の負荷などの統計情報を基にして推定することを検討した。一般に、システムを構成する IT 機器としては、サーバに加えて、ストレージ装置やネットワーク装置などからなる。そのため、機器全体の消費電力量を推定するためには、個々の機器の推定値を積み上げる方法が考えられる。しかし、この方法では、機器ごとに統計情報を取得して推定値を計算する必要があり、容易に推定することが難しい。そこで、機器全体の電力量を容易に推定するために、サーバの電力と、ストレージおよびスイッチの電力の 2 つに分けて計算した。

サーバの消費電力は、統計情報として CPU 利用率を取得し、この CPU 利用率によって計算した。サーバにおいては、CPU 利用率が増加するに従って、消費電力が増加する。これは、CPU そのものの消費電力や、処理に伴うメモリアクセスなどによって、電力を消費するためである。また、ストレージ装置とスイッチの消費電力は、統計情報として I/O 転送量を取得し、この転送量によって計算した。これは、I/O の発生にともない、ディスクへのアクセスやデータ転送に伴いスイッチが稼働して電力を消費するためである。

集約後の環境においては、複数のシステムが稼働するため、IT 機器の電力を、それぞれのシステムに按分する必要がある。そこで本研究では、ベンチマークプログラムが動いておらず CPU 利用率が最も低い状態での消費電力をアイドル電力とし、サーバのアイドル電力を p_{idle} 、ストレージのアイドル電力を o_{idle} とする。また、最も負荷が高い状態での、サーバの消費電

力を p_{max} 、ストレージの消費電力を o_{max} とする。そして、これらの電力とサーバ及びストレージの統計情報から、システム当たりの電力を計算する。個々のシステムは、それぞれ一つの仮想マシン上で稼働させる。

集約後の環境に含まれるシステムの集合を S とし、 $s_i \in S, 1 \leq i \leq n$ とする。 s_i におけるエネルギー効率測定のためのベンチマークプログラムの実行時間を d_i とし、 s_i における d_i の期間の CPU 利用率の平均を u_i 、I/O 転送量を b_i とする。 u_i は、 $0 \leq u_i \leq 1$ を満たす。 s_i におけるサーバの消費電力を x_i 、ストレージの消費電力を y_i とし、それぞれ式 2 および式 3 で定義する。

$$x_i = \frac{p_{idle}}{n} + (p_{max} - p_{idle}) \frac{u_i}{\sum_{i=1}^n u_i} \quad (2)$$

$$y_i = \frac{o_{idle}}{n} + (o_{max} - o_{idle}) \frac{b_i}{\sum_{i=1}^n b_i} \quad (3)$$

x_i と y_i の合計が IT 機器全体の電力であり、ベンチマークプログラムの実行時間 d_i をかけたものが、 s_i の電力量 c_i である (式 4)。

$$c_i = (x_i + y_i)d_i \quad (4)$$

消費電力量 c_i と、 s_i における応用成果量としてのベンチマークプログラムの実行回数 a_i を用いて、システム s_i のエネルギー効率 e_i を、式 5 のように定義する。

$$e_i = \frac{a_i}{c_i} \quad (5)$$

3 評価

本節では、エネルギー効率のモデル式 1 を評価する。まず、複数の測定パターンにおいて、ベア環境での APEE を測定する。そして、Linux 上に docker のコンテナによる仮想化での集約環境を作成し、式 5 を使ってエネルギー効率を計算する。最後に、ベア環境での APEE と仮想環境でのエネルギー効率を比較する。

APEE は汎用的なエネルギー効率を求めるではなく、特定の目的を持つシステムを想定し、そのシステムの基盤としてアプリケーションプラットフォームを使った場合のエネルギー効率である。そのため、APEE では、想定したシステムを模すベンチマークプログラムを選び、エネルギー効率の測定に使用する[5]。今回の評価では、アプリケーションプラットフォームにおけるミドルウェアの中で、DBMS に焦点をあてた測定を行った。DBMS のベンチマークプログラムである TPC-H および TPC-E は、測定対象の DBMS を変更することができるため、DBMS を変更することで、実装の異なるシステムを手軽に作ることができる。そこで、ベンチマークプログラムとして TPC-H および TPC-E のクエリを利用した。TPC-H で DBMS の異なる 2 パターン、および、TPC-E で同様に DBMS の異なる 2 パターン、合計 4 パターンのシステムを準備した。

エネルギー効率の計算に必要となる CPU 利用率の測定に

は、docker のリソース利用状況の取得機能を用いた。この機能は、一定間隔で、各コンテナの CPU 利用率やメモリ利用率を出力する機能である。これをベンチマーク実行期間中に 5 秒間隔で取得し、すべての CPU 利用率の相加平均を求めて u とした。今回の APEE の測定ではベンチマークプログラムとして TPC-H 2.18.0 [7] および TPC-E 1.14.0 [8] の表定義や問い合わせ文を使用した。それらのベンチマークプログラムは、Transaction Processing Performance Council (TPC) が公開している RDBMS 向けのベンチマークである。TPC-H は、卸売業者における経営戦略の意思決定を支援するシステムを模したものであり、商品や顧客、取引のデータなどをもとにして、意思決定のためのレポートの作成のクエリからなっている。TPC-E は、証券取引のシステムを模したものであり、証券の売買、レポートの作成などのクエリからなっている。

本測定におけるベンチマークプログラムの 1 回の実行は、TPC-H では、表定義、データロードおよび 22 個のクエリを処理することとした。TPC-E では、表定義、データロードおよび 12 個のトランザクションの処理であり、トランザクションについては TPC-E の既定する割合で合計 16,000 回処理することとした。

評価対象のハードウェアおよびソフトウェアの構成は、表 1 および表 2 の通りである¹。

表 1 ハードウェア構成

#	Item	Detail
1	Server	Hitachi BS500
2	CPU	2 x Intel Xeon Processor E7-8880v2
3	Memory	512GB
4	Storage	Hitachi VSP G600 with 12 x HT-F40SC-3R2FM (3.1TB)
5	FC switch	Hitachi HT-4990-SW6510V
6	Power Analyser	Yokogawa Test & Measurement Corporation WT1500e

表 2 ソフトウェア構成

#	Item	Detail
1	OS	Red Hat Enterprise Linux 7.7
2	DBMS 1	Hitachi Advanced Data Binder 04-03 (HADB) ²
3	DBMS 2	PostgreSQL 12
4	DBMS 3	MariaDB 5.5.56

仮想化によるサーバの集約では、集約後に性能の劣化が少なくなるようにするために、集約後のホストマシンに十分なり

1: Intel Xeon は、米国およびその他の国における Intel Corporation の商標です。Linux は、Linus Torvalds 氏の日本およびその他の国における登録商標または商標です。Red Hat は、米国およびその他の国で Red Hat, Inc. の登録商標もしくは商標です。TPC-E, TPC-H は米国 Transaction Processing Performance Council の商標です。

2: 内閣府の最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的サービスの実証・評価」(中心研究者: 喜連川 東大教授/国立情報学研究所所長)の成果を利用。

ソースを確保することが多い。そのため今回の測定では、ベア環境と集約環境において同等のリソースを利用できるように、docker の機能によって上限を設定した。ベンチマークプログラムおよび DBMS の組み合わせを 4 パターン作成した。それぞれのパターンで APEE を求めるために、ベア環境でベンチマークプログラムを実行して消費電力量を計測した。各パターンのエネルギー効率 APEE を表 3 に、またサーバ及びストレージ装置のアイドル電力と最大電力を表 4 にまとめる。

表 3 ベア環境で測定した APEE 値の一覧

Benchmark	DBMS	APEE
TPC-H	HADB	0.575
TPC-E	PostgreSQL	0.223
TPC-H	PostgreSQL	0.192
TPC-E	MariaDB	0.0315

表 4 サーバ及びストレージ装置のアイドル電力と最大電力

変数名	値
p_{idle}	633.67
p_{max}	1009.7
o_{idle}	2877.47
o_{max}	3023.41

次に、モデルの検証のために、4 つの集約環境を構築した。それぞれ、DBMS とベンチマークが同じ構成のシステムの集約環境の v_1 と v_2 、および DBMS とベンチマークが異なる構成のシステムの集約環境の v_3 と v_4 である。これらの集約環境において、ベンチマークを実行して統計情報を取得し、式 5 でエネルギー効率 e_i を計算した。測定及び計算結果を表 5 にまとめる。

表 5 仮想環境におけるベンチマーク実行時の測定結果一覧

v_i	s_i	a_i	u_i	b_i	e_i
v_1	s_1	3	1790	4310	0.843
	s_2	3	1759	4320	0.861
v_2	s_1	3	406	827	0.352
	s_2	3	409	835	0.332
v_3	s_1	17	301	1492	1.11
	s_2	6	1644	24460	0.333
v_4	s_1	12	396	4825	0.350
	s_2	2	169	14650	0.0567

表 5 をもとに、ベア環境で測定した APEE と集約後のエネルギー効率 e_i を比較した。それぞれのシステムにおいて、式 1 に APEE の値と e_i をあてはめて、係数 α を算出した。結果を、表 6 にまとめる。今回の予備実験では、 α はおよそ 1.4 から 1.9 の範囲となった。以降で、 v_1 から v_4 の結果を考察する。

v_1 と v_2 では、それぞれ s_1 と s_2 で同じ DBMS およびベンチマークプログラムを実行した。結果、それぞれの s_1 と s_2 の α は、ほぼ同じ値となった。このことから、リソースの競合がない環境において、本モデルが有効であると考ええる。

v_3 と v_4 では、それぞれ s_1 と s_2 で異なる DBMS およびベ

表 6 APEE と EE の比較および係数 α の一覧

v_i	s_i	Benchmark	DBMS	APEE	e_i	α
v_1	s_1	TPC-H	HADB	0.575	0.843	1.46
	s_2	TPC-H	HADB	0.575	0.861	1.48
v_2	s_1	TPC-H	PostgreSQL	0.192	0.352	1.49
	s_2	TPC-H	PostgreSQL	0.192	0.332	1.52
v_3	s_1	TPC-H	HADB	0.575	1.11	1.94
	s_2	TPC-E	PostgreSQL	0.223	0.333	1.50
v_4	s_1	TPC-H	PostgreSQL	0.192	0.350	1.83
	s_2	TPC-E	MariaDB	0.0315	0.0567	1.80

ベンチマークプログラムを実行した。それぞれの α の値は、 v_3 では s_1 が 1.94、 s_2 が 1.50 と離れており、一方の v_4 では s_1 が 1.83、 s_2 が 1.80 でほぼ同じ値となり、それぞれで傾向が異なった。 α の傾向が異なる要因を、次のように推測する。 s_1 および s_2 で DBMS やベンチマークプログラムが異なることで、それぞれの CPU やメモリ、ストレージといったリソースの使い方が異なる。これは、本論文で使用したリソースの統計情報に基づく電力の算出方法において、CPU 利用率と I/O 転送量だけの案分に改善の余地があることを示しており、CPU や I/O の傾向が異なる環境における改善課題である。

4 ま と め

本論文では、単独実行によって求める APEE から、仮想化による集約環境でのエネルギー効率を計算するモデルを提案した。そして、DBMS とベンチマークが同じ構成のシステムの集約環境と、DBMS とベンチマークが異なる構成のシステムの集約環境において、ベンチマークプログラムを実行してリソースの統計情報を取得し、それぞれのシステムのエネルギー効率を計算した。

前者の、同じ構成のシステムの集約環境においては、APEE と集約後のエネルギー効率の比率は、ほぼ同じ値となった。このことから、リソースの競合がない環境において、本モデルが有効であると考えられる。

後者の、異なる構成のシステムの集約環境においては、APEE と集約後のエネルギー効率の比率は、それぞれの集約環境によって異なった。この要因として、DBMS やベンチマークプログラムが異なることで CPU やメモリ、ストレージなどのリソースの使い方が異なるために、本論文で使用したリソースの統計情報に基づく消費電力量の計算方法に改善の余地があるためと推測する。

今後は、測定パターンや測定回数を増やして、提案モデルの改善や有効性の検証を継続して実施することが課題である。また、仮想環境における仮想マシン単位での消費電力量の測定方法の研究がさらに発展することを期待する。

5 謝 辞

この成果は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の助成事業 (JPNP16007) の結果得られた

ものである。

文 献

- [1] Eduardo Pinheiro and Ricardo Bianchini. Energy conservation techniques for disk array-based servers. In *Proceedings of the 18th Annual International Conference on Supercomputing*, ICS '04, p. 68–78, New York, NY, USA, 2004. Association for Computing Machinery.
- [2] Ekow J. Otoo, Doron Rotem, and Shih-Chiang Tsao. Dynamic data reorganization for energy savings in disk storage systems. In Michael Gertz and Bertram Ludäscher, editors, *Scientific and Statistical Database Management, 22nd International Conference, SSDBM 2010, Heidelberg, Germany, June 30 - July 2, 2010. Proceedings*, Vol. 6187 of *Lecture Notes in Computer Science*, pp. 322–341. Springer, 2010.
- [3] N. Nishikawa, M. Nakano, and M. Kitsuregawa. Application sensitive energy management framework for storage systems. *IEEE Transactions on Knowledge & Data Engineering*, Vol. 27, No. 09, pp. 2335–2348, sep 2015.
- [4] Kazuto Igarashi, Makoto Ootahara, and Kiyonori Hara. アプリケーションを含めた IT サービスのエネルギー効率指標の提案. *DEIM Forum 2018*, 2018.
- [5] ISO. ISO/IEC DIS 23544 Information Technology Data Centres Application Platform Energy Effectiveness (APEE). 2020.
- [6] Ingolf Waßmann, Daniel Versick, and Djamshid Tavangarian. Energy consumption estimation of virtual machines. 03 2013.
- [7] Transaction Processing Performance Council. *TPC Benchmark H Standard Specification Revision 2.18.0*.
- [8] Transaction Processing Performance Council. *TPC Benchmark E Standard Specification Version 1.14.0*.