

データベース管理システムのエネルギー高効率化手法とその評価試験

早水 悠登[†] 合田 和生[†] 茂木 和彦^{††} 木村 耕治^{††} 喜連川 優[†]

[†] 東京大学 生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

^{††} 株式会社日立製作所 〒185-8601 東京都国分寺市恋ヶ窪 1-280

E-mail: [†]{haya,kgoda,kitsure}@tkl.iis.u-tokyo.ac.jp, ^{††}{kazuhiko.mogi,uv,kohji.kimura.zn}@hitachi.com

あらまし 所謂ビッグデータブームに牽引され、多くの応用分野に於いて、大規模データを戦略的に活用する機運が高まってきている。大規模データの管理・処理のために膨大な IT 資源がデータセンタに投入され、自ずとデータセンタの消費エネルギーを増大させている。当該傾向は今後も続くと考えられ、大規模データの管理・処理を担うデータセンタに於いて、エネルギー効率を向上することは重要な課題と言える。著者らは、関係データベース管理システムを対象として、そのエネルギー効率の向上を目指した新たなソフトウェア構成法に関する研究を進めてきた。従前の関係データベース管理システムが、利用可能な IT 資源の制約の下、実行性能を最大化する設計を追及してきたのに対して、本研究は消費エネルギーを基軸と定めた新たなデータベース管理システムの技術的な展開を目指すものである。本論文では、データベース管理システムに於けるエネルギー高効率化のための技術課題を述べるほか、著者らが問合せ実行の高速化のために考案したアウトオブオーダ型実行方式によるエネルギー高効率化効果を議論し、実験による考察を報告する。

キーワード データベース管理システム, 消費エネルギー, 問合せ最適化, アウトオブオーダ型実行

1 はじめに

所謂ビッグデータブームに牽引され、多くの応用分野に於いて、大規模データを戦略的に活用する機運が高まってきているのは周知のとおりである。大規模データの管理・処理のために膨大な IT 資源がデータセンタに投入されるようになってきており、自ずとデータセンタの消費エネルギーを増大させている [11]。当該傾向は今後も当分の間、続くと考えられ、大規模データの管理・処理を担うデータセンタに於いて、エネルギー効率を向上することは重要な技術課題に他ならない。

著者らは、関係データベース管理システムを対象として、そのエネルギー効率の向上を目指した新たなソフトウェア構成法に関する研究を進めてきた [4, 7, 10, 20–24]。従前の関係データベース管理システムが、利用可能な IT 資源の制約の下、実行性能を最大化する設計を追及してきたのに対して [1]、本研究は新たな基軸としてエネルギー効率を向上することを目指す。

本論文では、著者らのこれまでの研究の積み重ねをもとに、データベース管理システムに於けるエネルギー高効率化のための技術課題を紹介する。また、著者らが問合せ実行の高速化のために考案したアウトオブオーダ型実行方式 [5, 19] によるエネルギー高効率化効果を議論し、実験による考察を報告する。

本論文の構成は以下の通りである。2 節では、データベース管理システムに於けるエネルギー高効率化のための技術課題と構想を示す。3 節では、小規模実験環境に於いて著者らが行ったアウトオブオーダ型実行によるエネルギー効率向上効果の観察実験を示し、その効果を考察する。4 節では、関連研究を簡潔にまとめ、最後に 5 節で論文を纏める。

2 データベース管理システムに於けるエネルギー効率の向上のための技術課題

データベース管理システムのエネルギー効率の向上のための技術課題を、4 つの観点から整理する。また、各々の技術課題について、解決方法の試みを紹介する。

第一に、データベース管理システムのエネルギー効率を向上するためには、問合せ等の与えられた仕事に係る消費エネルギーを低減することが求められる。ソフトウェアはハードウェアの上で動作するものであり、この際にハードウェアは電力を消費する。よって、ソフトウェアの動作に必要な範囲で、ハードウェアの消費する電力を低減する制御を実現することが求められる。例えば、ハードウェアのうち、ソフトウェアが利用する回路や機構へのみ給電を行い、それ以外の回路や機構への給電を停止するといった制御が考えられる。

データベース管理システムの場合、対象とするデータや負荷によって多様性があるものの、大規模データの管理と処理を目的とする点が特徴的であり、他のシステムと比較した場合に、自ずとストレージシステムが抱える資源が多く、当該資源の消費する電力が多く、支配的である傾向が強い。即ち、エネルギー効率の向上のためには、磁気ディスクドライブ、フラッシュメモリ等の記憶デバイスの電力制御が肝となるが、これらのデバイスはデータを格納することが主たる機能であることから、プロセッサ等の他のデバイスと異なり、給電を停止した場合に他のデバイスに機能を代替させることができないという技術的な困難を伴う。データベース管理システムに於いて当該問題を克服するために、著者は過去に能動的電力制御と称する手法を提案した [20, 22, 24]。例えば、データベース管理システムは問合せ

せを受け付けると、実行計画を生成し、当該計画に基づき問合せの実行を進める。この際、能動的電力制御は、実行計画を活用することにより、高い確度で将来の記憶デバイスへのアクセスの必要性を判断することが可能となり、その判断をもとに記憶デバイスの電力状態の制御を行う。同様のアプローチは、問合せ実行以外にも拡張可能であり、例えば遠隔レプリケーションに於いても有効性を確認するに至っている [4]。

記憶デバイスの消費電力が大きいとは言え、プロセッサによる消費電力も無視できず、同時に制御を行うことが望ましい。プロセッサに関しては、負荷の必要性に応じてスレッドを特定のコアに配置して動作させ、それ以外のコアの動作を停止し、消費電力を低減するアフィニティ制御が肝要である。また、近年のプロセッサでは更に、Dynamic Voltage and Frequency Scaling (DVFS) と呼ばれる周波数・電圧を多段階に制御可能とする機能が広く実装されており、データベース管理システムに於いても当該機能を活用することが有望と言える。過去に著者らは、トランザクション処理や解析系問合せを対象するアフィニティ制御ならびに DVFS 制御手法を提案し、実験により有効性を確認してきた [7,10]。

第二に、消費エネルギーの制御に適したデータ編成も重要な課題である。上述の通り、記憶デバイスは、データを格納することが主たる機能であることから、給電を停止した場合に他のデバイスに機能を代替させることができない。即ち、どのようにデータが編成され、記憶デバイスに格納されているかによって、消費エネルギー制御の可能性が大きく影響を受ける。著者らは、過去に上述の能動的電力制御の研究を進める過程で、データ編成の選択が与えるエネルギー高効率化への影響について考察を行ってきた [24]。

第三に、問合せ最適化の消費エネルギーへの拡張が挙げられよう。データベース管理システムは、受け付けた問合せをもとに、その実行性能を最大化する、即ち、与えられた問合せを処理するための実行時間を最小化する、或いは、単位時間に処理する問合せの数を最大化するといった実行計画を生成し、当該計画に基づいて問合せ処理を実行する。問合せ最適化に於いて、目的関数は常に実行性能の最大化であったが、消費エネルギーを考慮することにより、従前とは異なるクラスの最適化問題を設定することが可能となる。

例えば、問合せを実行する際に、一定の実行時間の制約内で完了するのであれば、消費エネルギーを最小化して欲しい、という要請は、データセンタに係る電力料金を抑制するという事業上の利得にも沿う。また、データセンタの設備を設計し維持管理する観点では、消費電力を一定範囲内に抑制することが強い制約であることから、問合せを一定の消費電力内で実行することにより、ソフトウェアによる電力のピークカットを実現することも有力な解法と言える。

著者はこのような洞察をもとに、消費エネルギーを考慮したコストベース問合せ最適化のための枠組みを提案してきた [21,23]。問合せの実行計画を構成するブロック（一連の演算パイプライン）の実行に要する消費電力の推定精度が鍵となるが、複数の異なるストレージ構成に於いて実験を行ったところ、15%以下

の誤差での推定が可能となることを確認するに至っている。

第四に、エネルギー効率の高い実行方式の探求が求められる。一般にプロセッサや記憶デバイス等、IT 機器を構成する部品は、無負荷状態であっても一定の電力を消費すると同時に、負荷に応じて電力が変化するため、当該特性を考慮して、ハードウェアに負荷を印加することにより、消費電力とスループットのバランスを確保することが鍵となる。

一つの試みとして、著者らはアウトオブオーダ型実行方式 [5,19] を、消費エネルギーの制御に活用することを試みている。アウトオブオーダ型実行方式は、問合せ実行を高速化するものであり、多数の非同期入出力とスレッドの同時駆動によって、ストレージの帯域とマルチコアプロセッサの演算能力を有効活用するものである。問合せ実行の高速化は、多くの場合、消費エネルギーの削減に繋がるものの、必ず最適化点は一致せず、また、消費電力の制約に応じた実行並列度の調整等、新たな技術課題も少なくない。3 節では、当該アプローチを検討するにあたり、著者らが実施した小規模実験を示し、エネルギー効率向上効果を議論すると共に、技術課題を考察する。

3 小規模実験環境に於けるアウトオブオーダ型実行によるエネルギー効率向上効果の観察

アウトオブオーダ型実行によるデータベース管理システムのエネルギー消費を分析するために、小規模実験環境を構築し、測定試験を実施した。

実験環境は、主にサーバ装置、ストレージ装置、電力計から構成される。図 1 に実験環境の概観を示す。サーバ装置は、Dell PowerEdge 2900 であり、2 基の Intel Xeon (3.2GHz) プロセッサ（合計 2 コア）と、2048MB の DRAM メモリを搭載し、RedHat Linux を稼働させた。また、2 ポートのファイバチャネル HBA を搭載し、後述のストレージ装置と接続した。ストレージ装置は、4 台の外部磁気ディスク JBOD (JMR Fortra 2G6) から構成され、各々の JBOD は 6 台の Seagate 146GB HDD (ST3146807FC) を搭載し、合計 24 台の磁気ディスクを稼働させた。また、サーバ装置と磁気ディスク装置の各々への給電回路に電力計（Yokogawa WT320）を接続し、各々への電



図 1 実験装置の概観

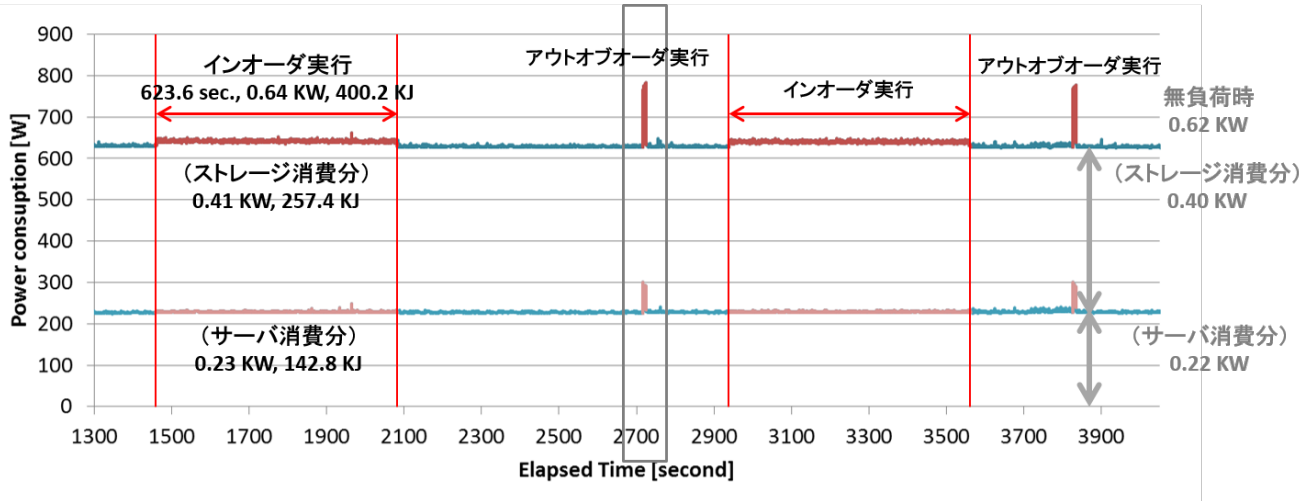


図 2 問合せ実行中の消費電力

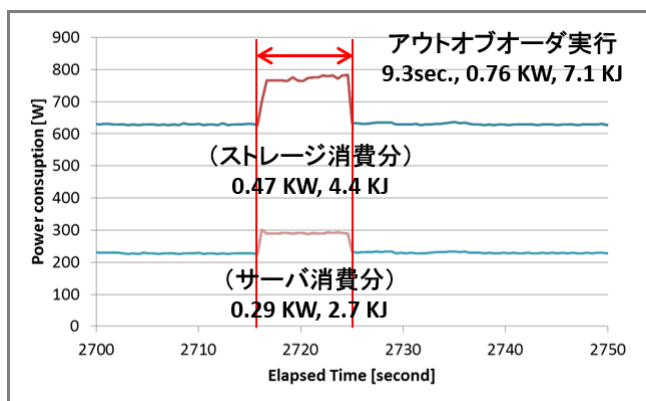


図 3 問合せ実行中の消費電力（アウトオブオーダ型実行方式の試行部分の拡大）

力供給を測定できるように構成した。24 ドライブの磁気ディスクからパリティなしストライピングによってデータベースボリュームを編成し、TPC-H ベンチマーク [15] の標準データセット（スケールファクタ：100）を生成して、組み込んだ。データベース管理システムは [5] で用いた試作データベースエンジンを用い、これに対して、Query 3 を簡略化した問合せを、インオーダ型実行方式とアウトオブオーダ型実行方式のそれぞれで実行し、それぞれの場合について実行時間を測定すると共に、サーバ装置とストレージ装置の各々について経過時間毎の消費電力を測定した。

図 2 に、問合せ実行中のデータベース管理システムの消費電力を示す。ここでは、同じ問合せをインオーダ型実行方式とアウトオブオーダ型実行方式で、それぞれ 2 回実行した。グラフは、サーバ装置の消費電力と、ストレージ装置の消費電力を積み上げて描画している。

まず、インオーダ型実行方式によると問合せ実行は 623.6 秒であり、この期間、平均的にサーバ装置は 0.23KW、ストレージ装置は 0.41KW の電力を消費していた。問合せを実行していない際の平均的な消費電力は、サーバ装置が 0.22KW、ストレージ装置は 0.40KW であることから、問合せ実行によって若

表 1 試験の結果の纏め

	実行時間	消費電力	消費エネルギー
無負荷状態	N/A	0.62KW	N/A
問合せ実行（インオーダ型実行方式）	623.6 秒	0.64KW	400.2KJ
問合せ実行（アウトオブオーダ型実行方式）	9.3 秒	0.71KW	7.1KJ
改善の率	67.1 倍高速化	11%上昇	98.2%削減

干の消費電力の情報はあるものの、著しいものではないことが判る。問合せ実行に要したエネルギーについては、サーバ装置で 142.8KJ、ストレージ装置で 257.4KJ、合計すると 400.2KJ であった。

次に、アウトオブオーダ型実行方式による問合せ実行に着目する。図 2 では視覚的に確認しづらいため、該当部分を拡大した図面を図 3 に示す。アウトオブオーダ型実行方式によると問合せの実行時間は 9.3 秒であり、平均的な消費電力はサーバ装置で 0.29KW、ストレージ装置で 0.47KW であった。これは問合せを実行していない際の消費電力と比較すると、有意な上昇が生じていると言え、このことはグラフ上はスパイク状の折れ線によっても視覚的に確認される。当該現象は、アウトオブオーダ型実行方式によって、演算ならびに入出力のスループットが著しく向上したことにより、とりわけプロセッサならびに磁気ディスクドライブの消費電力が上昇したことが伺える。問合せ実行に要したエネルギーについては、サーバ装置で 2.7KJ、ストレージ装置で 4.4KJ、合計すると 7.1KJ であった。なお更に実施した試行に於いても、同様の結果が得られた。

試験の結果を比較して表 1 に纏める。インオーダ型実行方式に対して、アウトオブオーダ型実行方式によると、問合せの実行時間は 67.1 倍に高速化した。この際、実行時の平均消費電力は 11%上昇したものの、消費エネルギーは 98.2%削減されており、即ち、問合せあたりのエネルギー効率は 56.4 倍に向上したと言える。

4 関連研究

データベースシステムの省エネルギー化に関する議論は永らく限定的であり、ようやく最近になって徐々に本格化してきているのが実状である [2]。文献 [6] は、ストレージシステムの規模と走査アルゴリズムの相違に掛かる実験結果を示し、データベースシステムのソフトウェアの改変によってエネルギー効率を向上するためのアーキテクチャを議論している。文献 [9] は、上述のスケーリングに加えて、複数の問合せ処理の間で共通する演算処理を集約する技法を提案している。文献 [12, 16] は、データベースシステムに於ける消費エネルギー特性を解明する実験を報告している。文献 [14] は、整列アルゴリズムを対象としてエネルギー効率を評価するベンチマークを提案している。文献 [13] は問合せ最適化に消費エネルギーを組み込む構想を報告している。文献 [18] は、データベースシステムに於ける性能と消費電力のトレードオフ関係をコストモデルと著者らが実施した実験を用いて議論している。文献 [3] は、PostgreSQL を取り上げ、プロセッサのスケーリング機能を用いたエネルギー制御を取り入れるための課題を議論している。文献 [8] は、Energy Response Time Profile (ERP) なる指標を用いることにより、既存の実装を用いて簡便に消費エネルギーを取り入れた問合せ最適化を実現するフレームワークを提案している。文献 [17] は、データベース管理者が性能と消費エネルギーのトレードオフを調整するためのインターフェースを提案している。

5 おわりに

本論文では、データベース管理システムに於けるエネルギー高効率化のための技術課題を述べた。また、著者らがアウトオブオーダ型実行方式によるエネルギー高効率化効果を見極めるための実験を紹介した。当該実験によると、小規模実験環境に於いては、インオーダ型実行方式に対してアウトオブオーダ型実行方式によって、問合せの実行時間は 67.1 倍に高速化し、平均消費電力は 11% 上昇したものの、消費エネルギーは 56.4 分の 1 に低減することができた。即ち、アウトオブオーダ型実行方式によるエネルギー高効率化の潜在性が確認された。

本実験は予備的な実験に留まっており、例えば、消費エネルギーとしては問合せの開始から終了までの積算しているものの、問合せが実行されていない場合に、磁気ディスクドライブやプロセッサコアへの給電を停止すると言った、電力制御は行われておらず、これらに伴うオーバーヘッドは考慮されていない。通常、これらの制御は一定の時間と消費エネルギー上のオーバーヘッドを伴うため、これを考慮した制御方式が必要となるであろう。また、フラッシュメモリを用いたストレージ装置や、より大規模なサーバ装置・ストレージ装置を用いた検証も欠かさない。現在、更に規模や確度を高めた実験を実施しており、今後、これを議論してゆきたい。

謝 辞

本研究の一部は、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) 委託業務「エネルギー・環境新技術先端プログラム/革新的な省エネルギー型データベース問合せコンパイラの研究開発」及び「高度な IoT 社会を実現する横断的技術開発/先進 IoT サービスを実現する革新的超省エネルギー型ビッグデータ基盤の研究開発」による助成を受けている。

文 献

- [1] Ramez Elmasri and Shamkant B. Navathe. *Fundamentals of Database Systems*. Pearson, 7th edition, 2016.
- [2] R. Agrawal et al. The Claremont Report on Database Research. *Comm. ACM*, Vol. 52, No. 6, pp. 56–65, 2009.
- [3] Tobias Flach. Optimizing Query Execution to Improve the Energy Efficiency of Database Management Systems. Master's thesis, Hasso Plattner Institut, 2010.
- [4] K. Goda and M. Kitsuregawa. Power-aware Remote Replication for Enterprise-level Disaster Recovery Systems. In *Proc. USENIX Annu. Tech. Conf.*
- [5] Kazuo Goda, Yuto Hayamizu, Hiroyuki Yamada, and Masaru Kitsuregawa. Out-of-order Execution of Database Queries. *PVLDB*, Vol. 13, No. 12, pp. 3489–3501, 2020.
- [6] Stavros Harizopoulos, Justin Meza, Mehul A. Shah, and Parthasarathy Ranganathan. Energy Efficiency: The New Holy Grail of Data Management Systems Research. In *Proc. Conf. on Innovative Database Syst. Research*, 2009.
- [7] Yuto Hayamizu, Kazuo Goda, Miyuki Nakano, and Masaru Kitsuregawa. Application-aware Power Saving for Online Transaction Processing using Dynamic Voltage and Frequency Scaling in a Multicore Environment. In *Proc. Int'l Conf on Arch. of Comput. Syst.*, pp. 50–61, 2011.
- [8] Willis Lang, Ramakrishnan Kandhan, and Jignesh M. Patel. Rethinking Query Processing for Energy Efficiency: Slowing Down to Win the Race. *IEEE Data Eng. Bull.*, Vol. 34, No. 1, pp. 12–23, 2012.
- [9] Willis Lang and Jignesh M. Patel. Towards Eco-friendly Database Management Systems. In *Proc. Conf. on Innovative Database Syst. Research*, 2009.
- [10] Boming Luo, Yuto Hayamizu, Kazuo Goda, and Masaru Kitsuregawa. Modeling Query Energy Cost in Analytical Database Systems with Processor Speed Scaling. In *Proceedings of the 29th International Conference on Database and Expert Systems Applications (DEXA2018)*, pp. 310–317, 2018.
- [11] Natural Resources Defense Council. Data Center Efficiency Assessment: Scaling Up Energy Efficiency Across the Data Center Industry: Evaluating Key Drivers and Barriers. <https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IP.pdf>, 2014.
- [12] Meikel Poess and Raghunatha Othayoth Nambiar. Energy Cost, The Key Challenge of Today's Data Centers: a Power Consumption Analysis. *Proc. Very Large Data Bases*, Vol. 1, No. 2, pp. 1229–1240, 2008.
- [13] Rafael Alonso and Sumit Ganguly. Energy Efficient Query Optimization. Technical Report, Matsushita Info Tech Lab., 1992.
- [14] Suzanne Rivoire, Mehul A. Shah, Parthasarathy Ranganathan, and Christos Kozyrakis. JouleSort: A Balanced Energy-Efficiency Benchmark. In *Proc. ACM SIGMOD Conf.*, pp. 367–376, 2007.
- [15] Transaction Processing Performance Council. TPC-H, an ad-doc, decision support benchmark. <http://www.tpc.org/>

tpch/.

- [16] Dimitris Tsirogiannis, Stavros Harizopoulos, and Mehul A. Shah. Analyzing the Energy Efficiency of a Database Server. In *Proc. ACM SIGMOD Conf.*, pp. 231–242, 2010.
- [17] Zichen Xu, Yi-Cheng Tu, and Xiaorui Wang. PET: Reducing Database Energy Cost via Query Optimization. *Proc. Very Large Data Bases*.
- [18] Zichen Xy, Yi-Cheng Tu, and Xiaorui Wang. Exploring Power-Performance Tradeoffs in Database Systems. In *Proc. Int'l Conf. on Data Eng.*, pp. 485–496, 2010.
- [19] 喜連川優, 合田和生. アウトオブオーダーデータベースエンジン OoODE の構想と初期実験. 日本データベース学会論文誌, Vol. 8, No. 1, pp. 131–136, 2009.
- [20] 合田和生, Wenyu Qu, 喜連川優. 複数問合せ処理を意識したディスクストレージ省電力化に関する一考察. 信学会データ工学ワークショップ (DEWS), pp. D5–2, 2008.
- [21] 合田和生, 早水悠登, 喜連川優. ストレージシステムの消費エネルギーを考慮したコストベース型のデータベース問合せ最適化手法の提案. In *The 1st Cross-disciplinary workshop on computing Systems, Infrastructures, and programming (xSIG 2017)*, pp. 3A–1, 2017.
- [22] 上野裕也, 合田和生, 喜連川優. データベースシステムの問い合わせ実行計画を利用したディスクアレイ省電力化に関する一考察. 日本データベース学会 Letters, Vol. 6, No. 1, pp. 85–88, 2007.
- [23] 早水悠登, 合田和生, 喜連川優. ストレージ消費電力特性に基づく関係データベース演算子の省電力指向コストモデル. 電子情報通信学会第9回データ工学と情報マネジメントに関するフォーラム/第15回日本データベース学会年次大会 (DEIM2017), pp. G1–3, 2017.
- [24] 平井遥, 星野喬, 合田和生, 喜連川優. データベースシステムにおけるプロアクティブなディスクアレイ省電力化手法に関する一考察. 信学会データ工学ワークショップ (DEWS), pp. D5–2, 2008.