

データベース処理実行時における 省電力化のためのストレージ制御手法の提案

飯村 奈穂¹ 西川 記史² 中野 美由紀² 小口 正人¹

概要：近年のデジタル情報量の爆発的な増加により、ストレージの出荷容量台数も急増している。これによるストレージの管理運用コストは見過ごせないものとなっており、データの効率的管理に注目が集まっている。データセンタのエネルギー消費量は2050年には2010年度の日本の発電電力量の約3倍になると予測されており、社会全体での節電が求められる中でデータセンタの消費電力を削減することは急務である。また、データセンタの消費電力割合の中でストレージの消費電力は約13%であることから、ストレージの消費電力を削減することにより、データセンタ全体を一定量省電力化することが可能であると言える。そこで、本研究ではデータの効率的管理という点からクラウド上のデータベースの省電力化を考え、本論文では業界標準のデータベースベンチマークであるTPC-Hの実行時省電力化に向けて、TPC-H実行時のシステム性能と消費電力量の解析を行い、ディスクの省電力状態とI/O発行間隔を利用することによるTPC-H実行時の省電力化が可能であるということを示した。また、実行時のI/O発行間隔の制御を行うためのデータ配置についても検討を行い、本研究で提案した手法がTPC-Hの実行時省電力化に有効であることを示した。

A Proposal of Storage Control Method for Energy Saving on Runtime Database Processing

NAHO IIMURA¹ NORIFUMI NISHIKAWA² MIYUKI NAKANO² MASATO OGUCHI¹

1. はじめに

近年デジタル情報量は爆発的に急増しており、今後10年で約44倍になるとも言われている。これに伴いストレージの出荷容量も急増していることから、ストレージの管理運用コストは見過ごせないものとなっており、データの効率的管理に注目が集まっている。データセンタのエネルギー消費量は2050年には2010年度の日本の発電電力量の約3倍になると予測されており、社会全体での節電が求められる中でデータセンタの消費電力を削減することは急務になっている [1]。

また、データセンタの消費電力割合の中でストレージの消費電力比率は約13%であることから、ストレージの消費電力を削減することにより、データセンタ全体を一定量

省電力化することが可能であると言える。そこで、本研究ではデータの効率的管理という点からクラウド上のデータベースの省電力化を考え、アプリケーションの性能劣化を抑えつつ、ストレージの消費電力を削減することを研究目的とする。

ストレージはCPUインテンシブアプリケーションに比べて、データインテンシブアプリケーションを動作させた場合における、コストや性能、消費電力が重要である。これらのことからストレージの性質を考慮して、本研究における測定対象のアプリケーションは、データインテンシブアプリケーションとする。本研究では特に、業界標準のデータベースベンチマークであるTPC-H [2] に焦点を当てて測定を行う。

また、TPC-Hの省電力化に向けて、Break-Even Timeを詳細に算出し、TPC-H実行時のディスクI/O発行間隔を調査した。そして実行時消費電力量を見積と実測により調査し、これらの値を比較することにより、本研究で提案した

¹ お茶の水女子大学
2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610 JAPAN
² 東京大学生産技術研究所
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 JAPAN

見積りが妥当であるということと、ディスクの省電力状態と I/O 発行間隔を利用することによる TPC-H 実行時の省電力化が可能であるということを示した。また、実行時のディスク I/O 発行間隔の制御を行うためのデータ配置手法についても検討を行った。

2. 関連研究

これまでも多くのストレージ省電力手法が提案されてきた [3]-[7]。これらの研究では、ストレージレベルの入出力頻度に従ってディスクを停止する等の手法が提案されている。しかし、実際にストレージレベルでの入出力を予測することは必ずしも容易ではない。

省電力アプリケーション協調型のストレージ省電力手法もある [8]-[10]。特に [9] では、アプリケーション協調型のストレージ省電力システムの構築を目指し、データインテンシブアプリケーションの I/O 挙動特性を解析、評価して、ストレージ電力制御モデルの提案を行っている。これらの研究では、提案手法におけるストレージの消費電力削減を考慮した上で、アプリケーションの性能評価を行っている。

そこで、本研究では、データインテンシブアプリケーションの SLA(Service Level Agreement) に注目し、アプリケーションの性能劣化を最小限に抑えつつ、ストレージの消費電力を削減し、性能評価を行っていくことを研究方針とする。

3. 測定環境

本研究では、サーバ PC として、CPU が AMD Athlon 64 FX-74 3GHz (4 cores) × 2, 主記憶が 8GB, HDD が Seagate Barracuda 1TB × 6, OS が CentOS 5.6 64 ビット版, DBMS は Hitachi HiRDB Single Server Version 9 を使用する。また、電力計は YOKOGAWA WT1600 Digital Power Meter を使用する。

これらの測定環境は全て遠隔アクセスによる実験が可能である。電力計はサーバ PC の HDD に繋がれており、電力計は電力計操作用 PC で操作する。またサーバ PC, 電力計, 電力計操作用 PC は、ローカル PC と全てリモート接続されている。測定環境の簡単な模式図と、電力計の操作画面を図 1 に示す。

4. 基本性能測定

データベースベンチマークである TPC-H を動作させた際のディスクの消費電力と I/O トレースの解析を行う。測定対象のディスクは 2 台, HDD1 には TPC-H の LINEITEM 表を, HDD2 には LINEITEM 表の主キー索引, およびその他の表を配置してある。DB の規模を決めるスケールファクタ (SF) を 1, 2, 3 と変化させ, それぞれの SF ごとに DB とクエリを用意して測定を行う。本計測では, blktrace と btreord(I/O トレース取得ツール)[11] を用いて I/O トレー

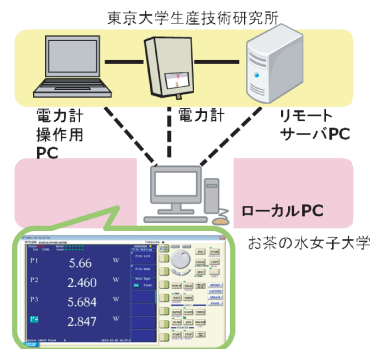


図 1 測定環境

スの取得・解析を行った。

図 2 には SF=3 の時の HDD1, 図 3 には SF=3 の時の HDD2 の測定結果を示す。SF=1, 2 の時もそれぞれ同様の傾向であった。上段は I/O トレースの取得結果を, 下段は実行時の消費電力をクエリごとに色分けしたものを表している。図 2 の I/O トレースと消費電力を比較すると, 消費電力の上下とディスクアクセスの頻度がほぼ一致しており, 応答時間が長いクエリはランダムアクセスを行っていることがわかる。これはクエリプランが複雑であり, ランダムアクセスの方がシーケンシャルアクセスよりも時間がかかるためである。また, 図 3 の経過時間 10,000 秒以降に注目すると, クエリ実行中の後半部分において HDD2 側のアドレス範囲が狭いのは, 前半部分は Order 表と Lineitem 表の主キー索引の両方にアクセスが行われているのに対し, 後半部分は Lineitem 表の主キー索引のみにアクセスが行われているためであると考えられる。従って, この結果は妥当な振舞いであると考えられる。

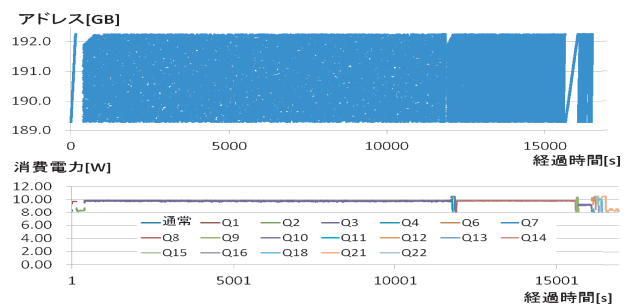


図 2 HDD1(SF=3)

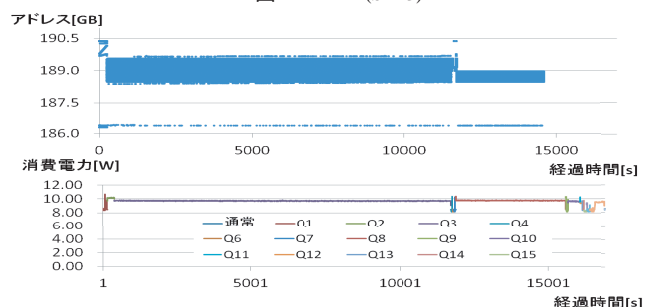


図 3 HDD2(SF=3)

5. ディスクの消費電力特性

ディスクの遷移状態の種類と、各状態における消費電力を調査し、それに基づいて、省電力可能性の1つの指標となる、Break-Even Time を算出する。

5.1 ディスクの遷移状態と消費電力

本研究で使用したディスクの遷移状態は Standby1, Standby2, Idle, Active の4種類である。Idle/Active 状態から Standby1 状態に移行することを Spindown, Standby1 状態から Idle/Active 状態に移行することを Spinup1, Standby2 状態から Idle/Active 状態に移行することを Spinup2 と呼ぶ [8]。

[8]では使用するディスクの状態を Standby, Idle, Active の3種類としているが、本研究で使用するディスクの遷移状態を詳細に調査したところ、Standby 状態時に消費電力が異なる2種類の期間がみられたため、本研究では2種類の状態を Standby1, Standby2 と区別している。

各状態におけるディスクの消費電力の測定を行った。測定対象のディスクは、4節の測定に使用したものと同様のディスク2台である。Standby1 時, Standby2 時, Idle 時, Active 時の最大消費電力と, Spindown, Spinup1, Spinup2 に必要なエネルギーを表3に示す。

表1 ディスクの遷移状態における消費電力とエネルギー量

Disk	Standby1(W)	Standby2(W)	Idle(W)	Active(W)
HDD1	1.81	1.21	8.42	10.5
HDD2	1.92	1.24	8.43	10.8
Disk	Spindown(J)	Spinup1(J)	Spinup2(J)	
HDD1	16.31	159.03	184.41	
HDD2	13.77	181.39	180.05	

5.2 Break-Even Time

ディスクの Spindown 及び Spinup により消費されるエネルギーと、ディスクを Standby 状態に移行し、その状態を維持することにより削減できるエネルギーが等しくなる Standby 状態の持続時間を Break-Even Time と呼ぶ。これは Spindown に必要なエネルギーを E_d , Spinup2 に必要なエネルギーを E_{u2} , Standby1 状態の消費電力を P_{s1} , Standby2 状態の消費電力を P_{s2} , Idle 状態の消費電力を P_i , Spindown と Spinup2 に必要な時間をそれぞれ T_d , T_{u2} , Standby1, Standby2 状態の持続時間をそれぞれ T_{s1} , T_{s2} とすると、Break-Even Time T_{be} は、

$$T_{be} = (E_d + E_{u2} - P_{s2} * T_d - P_{s2} * T_{u2} + T_{s1} * (P_{s1} - P_{s2})) / (P_i - P_{s2})$$

により求めることができる。

本研究では Standby を2種類の状態に区別するため、Break-Even Time の算出式は [12] を参考に作成した。

本研究で用いた HDD1, HDD2 では Break-Even Time はそれぞれ約 26 秒であった。これより Standby 状態を利用して省電力化を実行するためには、ディスクへの I/O 発行間隔が HDD1, HDD2 それぞれのディスクにおいて約 26 秒以上必要である。

図4は HDD1 において Idle 状態から Standby1 状態を経て、Standby2 状態に移行した後、再び Idle 状態に移行した時の消費電力の推移を示している。

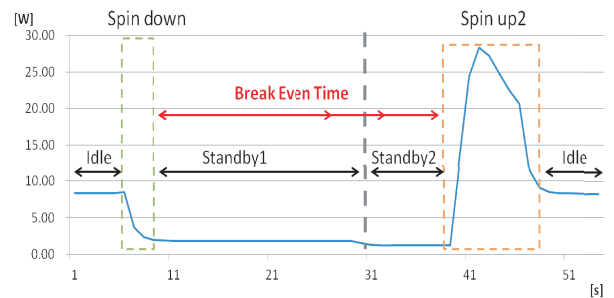


図4 ディスクの状態遷移における消費電力と Break-Even Time(HDD1)

6. 実行時省電力可能性

本節では、前節をふまえて TPC-H 実行時の I/O 発行間隔を調査する。また、ディスクの省電力状態を適用することにより、TPC-H の実行時消費電力をどの程度削減することができるのか、という点について見積りをもとに実測を行い、省電力状態を適用した場合と、そうでない場合の実行時消費電力量を比較し、評価を行う。

6.1 I/O 発行間隔

ディスクに I/O が行われてから、次の I/O が発行されるまでの時間を I/O 発行間隔と呼ぶ。本研究では、TPC-H 実行時のディスク I/O の利用状況を取得・解析し、I/O 発行間隔を調査する。測定環境は3節同様で、調査対象のディスクは前節と同様のディスク2台とする。測定期間は TPC-H クエリ実行中で、時間は 16,903 秒 (4 時間 41 分 43 秒) である。

測定の結果、図5,6に示すように、I/O 発行間隔が Break-Even Time 以上である回数が HDD1 では4回、HDD2 では10回であった。また、I/O 発行間隔が Break-Even Time 未満である回数は HDD1 では8回、HDD2 では11回であった。ここでは、I/O 発行間隔が Break-Even Time 未満のうち、最短1秒以上のものをカウントしている。Break-Even Time 以上の I/O 発行間隔のうち、最長は HDD1 では 322 秒、HDD2 では 157 秒であった。

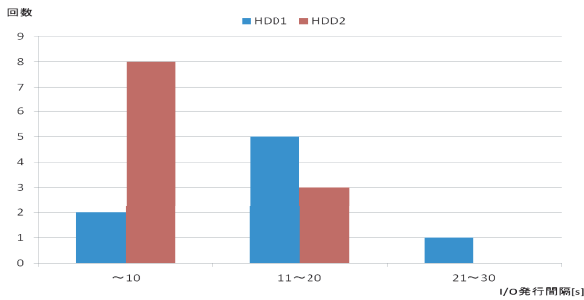


図 5 Break-Even Time 未満の I/O 発行間隔の回数



図 6 Break-Even Time 以上の I/O 発行間隔の回数

6.2 省電力状態適用時の実行時消費電力量

6.2.1 削減可能エネルギー見積

ディスクのスタンバイ状態を利用する場合を省電力状態適用あり、利用しない場合を省電力状態適用なしとする。さらに、省電力状態適用ありの場合は、ディスクをスタンバイ状態に移行するまでのタイムアウトを設定すると仮定して、実行時削減可能エネルギーの見積を行う。タイムアウトを設定することにより、タイムアウト時間より長くディスクへの I/O が発行されなかった場合に、ディスクを省電力状態へ移行することとする。すなわち、本研究では、【I/O 発行間隔 (s) - タイムアウト時間 (s)】を省電力状態適用期間とする。また、Spinup の契機は省電力状態 (Standby1, Standby2) 時にディスクに I/O が発行された時とする。

しかし、タイムアウトを設定する場合、Break-Even Time より短い I/O 発行間隔にも省電力状態を適用することになるため、省電力状態適用期間によって見積式を選択する必要がある。また、Break-Even Time より短い I/O 発行間隔に省電力状態を適用する場合、削減可能エネルギーがマイナスになるため、削減可能エネルギーの見積式ではタイムアウト時間も考慮する必要がある。提案する見積式では、Spindown に必要な時間、Standby1 状態の持続時間の合計時間と、省電力適用期間を比較することにより、見積式を選択する。5.2 節で使用した項目に加えて、I/O 発行間隔を T_i 、設定するタイムアウトを T_t とするとき、それぞれの I/O 発行間隔における削減エネルギー E_s は、

- 省電力適用期間 $< T_d + T_{s1}$ の場合

$$E_s = (T_i - T_t) * P_i - E_d - E_{u1} + P_{s1} * (T_i - T_t - T_d)$$

- 省電力適用期間 $\geq T_d + T_{s1}$ の場合

$$E_s = (T_i - T_t) * P_i - E_d - E_{u2} + P_{s1} * T_{s1} + P_{s2} * T_{s2}$$

により求めることができる。I/O 発行間隔ごとに算出した値の合計値を、削減可能エネルギーの見積値とする。

表 2 は、ディスクにタイムアウト時間を設定した場合の TPC-H 実行時の削減可能エネルギーを、見積式により算出した値を示している。タイムアウトが 20 秒以上の場合には、タイムアウトの増加に伴い、見積値が減少していたため、ここではタイムアウトが 20 秒までの見積値を載せている。表 2 より、削減可能エネルギーが最も大きいタイムアウト時間は、HDD1 では 15 秒 (3498.29J)、HDD2 では 10 秒 (2547.81J) であることがわかる。

表 2 実行時削減可能エネルギー (J)

Timeout(s)	5	10	15	20
HDD1	3224.64	3403.89	3498.29	3368.55
HDD2	2489.51	2547.81	2228.56	2028.73

クエリの遅延時間についても見積によって求めることができる。今回の見積では、Spinup の契機をディスクに I/O が発行された時としているため、ディスクの起動 1 回に必要な時間 (Spinup1, Spinup2 に必要な時間) の合計をクエリの遅延時間とすることができる。よって、各ディスクのクエリの遅延時間 T_{late} は、

$$T_{late} = T_{u1} * Spinup1 \text{ が行われた回数} \\ + T_{u2} * Spinup2 \text{ が行われた回数}$$

により求めることができる。

タイムアウトを HDD1 では 15 秒、HDD2 では 10 秒に設定した際の、TPC-H 実行時のクエリ遅延時間の見積値は、HDD1 では 34.6 秒、HDD2 では 99 秒、全体の遅延時間は 133.6 秒であった。

6.2.2 省電力状態適用時の実行時消費電力量

見積式から得られた最適なタイムアウトを各ディスクに設定し、TPC-H 実行時の消費電力量を測定する。最適なタイムアウトとは、HDD1 では 15 秒、HDD2 では 10 秒を指す。測定に使用した TPC-H の SF は 3、測定期間は TPC-H の実行開始から終了までとする。測定値は 3 回の測定の平均値を使用する。

図 7 は省電力状態適用なし、省電力状態適用ありの実測値、省電力状態適用ありの見積値の TPC-H 実行時の消費電力量の比較を示している。省電力状態を適用しなかった場合の消費電力量は 164,794J、省電力状態を適用した場合の消費電力量の実測値は 162,135J、省電力状態を適用した場合の消費電力量の見積値は 161,296J であった。

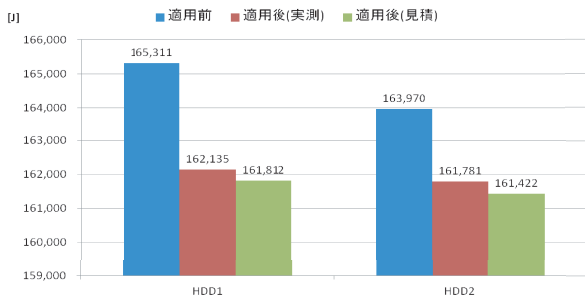


図7 省電力状態適用による TPC-H 実行時消費エネルギー比較

クエリ遅延時間の実測値は 125 秒であった。見積値と比較して誤差が生じているが、これは 2 台のディスクに異なるタイムアウトを設定したことにより、TPC-H 実行中に、片方のディスクの処理を待つ等の動作が生じ、Spinup のタイミングや時間に誤差が生じたためであると考えられる。

消費エネルギー削減率の見積値は、HDD1 では 2.1%、HDD2 では 1.6% であるのに対し、実測値は HDD1 では 1.9%、HDD2 では 1.3% であり、誤差は 0.2~0.3% に収まった。

TPC-H 実行時の I/O 発行間隔の長さや回数、ディスクの消費電力は毎回若干異なるため、見積値と実測値の誤差は許容範囲であると考えられる。これらのことから、ディスクの省電力状態と I/O 発行間隔を利用した TPC-H の省電力化は可能であるといえる。

6.2.3 削減可能エネルギー見積式の整合性

6.2.1 節で提案した、実行時削減可能エネルギーの見積式の整合性を示すために、6.2.2 節でディスクに設定したタイムアウト以外の値をタイムアウトとしてをディスクに設定し、同様の測定を行う。測定に使用した TPC-H の SF は 3、測定期間は TPC-H の実行開始から終了までとし、測定値は 3 回の測定の平均値とする。ディスクを省電力状態に移行するまでのタイムアウトを、HDD1 には 5 秒、10 秒、HDD2 には 5 秒、15 秒のタイムアウトをそれぞれ設定し、TPC-H 実行時の消費電力量を測定する。

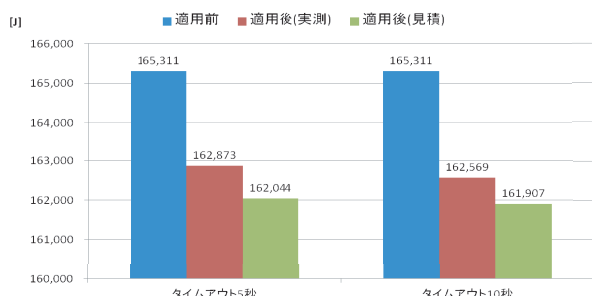


図8 タイムアウト設定時消費エネルギー比較 (HDD1)

図8は、HDD1 に 5 秒、10 秒のタイムアウトを設定した時の、省電力状態適用なし、省電力状態適用ありの場合の実行時消費エネルギーの見積値、実測値を表している。図

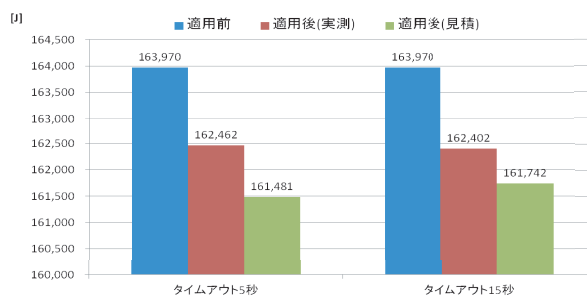


図9 タイムアウト設定時消費エネルギー比較 (HDD2)

表3 消費エネルギー削減率 (%) (HDD1)

Timeout(s)	5	10
見積値	1.9	2.0
実測値	1.5	1.7

表4 消費エネルギー削減率 (%) (HDD2)

Timeout(s)	5	15
見積値	1.5	1.4
実測値	0.9	1.0

9 は同様に HDD2 に 5 秒、15 秒のタイムアウトを設定した時の測定値を表している。2 つの図より、タイムアウトを 6.2.2 節で設定した値以外のものを設定した場合も、TPC-H 実行時の消費エネルギーを削減することが可能であることがわかる。

HDD1、HDD2 においてそれぞれタイムアウトを設定した場合の実行時消費エネルギー削減率の見積値と実測値の比較を表 3、4 に示す。見積値と実測値の間における誤差はそれぞれ約 0.3~0.8% 程度に収まっている。

この誤差は、ディスクにタイムアウトを設定したことによって起動オーバーヘッド等の待ち時間が生じ、I/O 発行間隔の回数と長さによずれが生じたためであると考えられる。本研究では、TPC-H の省電力化を目標としており、見積式の整合性はある程度保たれていれば良いものとする。従って、本研究で提案した、TPC-H 実行時削減可能エネルギーの見積式は妥当であるといえる。

7. データ配置制御

これまで、我々は TPC-H 実行時のディスクに対する I/O 発行間隔と Break-Even Time を利用することによる TPC-H の実行時省電力化が可能であることを示した。しかし、これはアプリケーションに手を加えずにディスクの単純な振り回しを利用しているナイーブな手法である。

本節では、TPC-H 実行時の I/O 発行間隔を制御するという目的のもと、TPC-H の各表と索引のバッファに対する I/O の状況を調査し、データの配置に手を加えることでさらなる実行時省電力化を目指す。

7.1 TPC-H 実行時の入出力状況

各バッファの I/O 状況の取得には pdbufsls(DBMS に付属する DB バッファ統計情報取得ツール)[13] を使用する。DB は raw デバイス上に直接配置しており、調査対象のバッファは 23 個、調査期間は TPC-H クエリ実行開始から終了までとする。また、SF=1, 2, 3 の 3 種類において測定を行い、データは 1 秒毎に取得する。今回は I/O 状況を把握するために、得られるデータの項目の中でも実 READ 回数に注目し、解析を行う。通常の DBMS はバッファに DB の一部が載った状態 (Hot 状態) で使用されるため、本測定は Hot 状態で行うものとする。DB バッファのサイズは、すべての SF において表のデータが格納されたバッファの合計が約 0.58GB、索引のデータが格納されたバッファの合計が約 0.21GB である。DB のサイズは SF によって異なるため、表 5 に示す。また、本測定は 4 節と同一の条件で行う。

解析の結果、実 READ 回数が 0 であったバッファは SF=1 のとき 12 個、SF=2 のとき 10 個、SF=3 のとき 9 個であった。また、SF=1, SF=2, SF=3 のすべての場合において実 READ 回数が 0 であったバッファは 9 個であった。そのうち、表のデータが格納されているバッファは 2 個、索引のデータが格納されているバッファは 7 個であった。図 10 は、SF=3 において実 READ 数に変化がみられた、つまり TPC-H 実行中における秒当り実 READ 回数の推移を表している。LINEITEM テーブルのデータが格納されているバッファは BUF_L2, BUF_L2_I1 であり、BUF_L2 は表データを、BUF_L2_I1 は索引データを格納している。主にこれらのバッファにアクセスが集中しているのは、LINEITEM 表のサイズが約 1GB であるのに対し、LINEITEM 用 DB バッファのサイズが約 390MB と小さいためであると考えられる。また、表のデータが格納されているバッファはシーケンシャルアクセスが行われているのに対し、索引はランダムアクセスが行われていることもこの図から把握することができる。

SF	1	2	3
表	1.38	2.75	4.13
索引	0.29	0.57	0.86

本手法ではバッファの利用状況を 1 秒毎に取得しているため、1 秒以内の短時間に急激に発生した入出力は確認することができない。そこで、TPC-H 実行時におけるクエリ単位での READ 回数を確認し、これを pdbufsls により取得した実 READ 回数と比較することにより、本手法で取得した実 READ 回数が妥当であることを示す。

クエリ単位での入出力を確認するために、pdobsls(DBMS に付属する SQL オブジェクト用バッファの統計情報表示ツール)[14] を用いる。SF=1,2,3 において TPC-H 実行時に

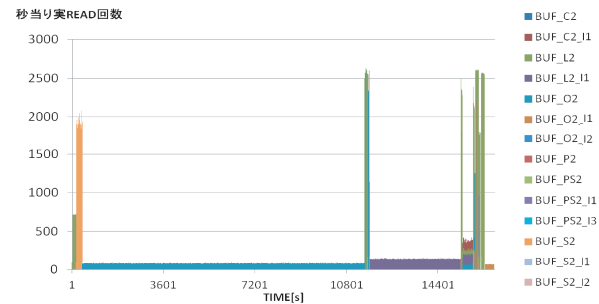


図 10 TPC-H 実行時における各バッファの秒当り実 READ 回数 (SF=3)

におけるクエリ単位での READ 数を計測し、pdbufsls を用いて取得したものと比較すると、誤差は 0.1~4% 程度に収まった。以上のことより、本手法により計測した各バッファの実 READ 回数は妥当であると考えられる。

7.2 データ配置の変更

7.1 節の結果を踏まえて、測定で使用したバッファを、TPC-H 実行開始から終了まで SF=1~SF=3 のすべての場合において「実 READ 回数が 0 であったデータ」、「実 READ 回数に変化が見られたデータ」の 2 種類に分類し、前者を HDD1 に、後者を HDD2 に配置する。ここで使用する HDD1, HDD2 はこれまでの測定で使用したものと同様である。データ配置のみを変更し、ディスクの省電力状態を適用した際の TPC-H 実行時における各 HDD の消費電力量とクエリの実行時間の測定を行い、データ配置変更前と変更後の各項目を比較する。測定はこれまでと同様、SF=1, 2, 3 の 3 種類について行う。また、ディスクがスタンバイ状態に移行するまでのタイムアウトを HDD1, HDD2 の 2 台とも 5 秒に設定する。測定期間は TPC-H 実行開始から終了までとする。消費電力量に関してはディスク 2 台の合計を比較対象とする。また、タイムアウトを 5 秒に設定したのはデータ配置を変更したことによって I/O 発行間隔が伸びたため、できるだけ早く HDD をスタンバイ状態に移行した方がより多く消費電力を削減できると考えたからである。

図 11 は TPC-H 実行時の消費電力量の比較を、図 12 は実行時間の比較を表している。図 11 から、データ配置の変更を行うことにより実行時の消費電力をどの SF においても削減していることがわかる。また、消費電力の削減率は SF=1 で 39.2%、SF=2 で 38.7%、SF=3 で 38.1% であった。一方、クエリの実行時間の遅延率については SF=1 で 6.6%、SF=2 で 9.3%、SF=3 で 11.3% の増加に留まった。

この結果はデータ配置を変更したことによるものだと考えられる。以上のことから本手法が TPC-H 実行時省電力化において有効であるということが示された。実行時の性能劣化は多少見られたが、こちらの検討に関しては今後の課題としたい。

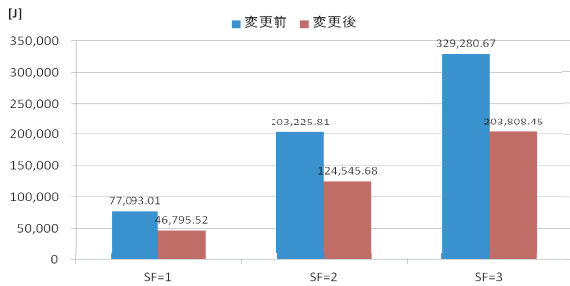


図 11 データ配置変更前と変更後における TPC-H 実行時の消費電力量の比較

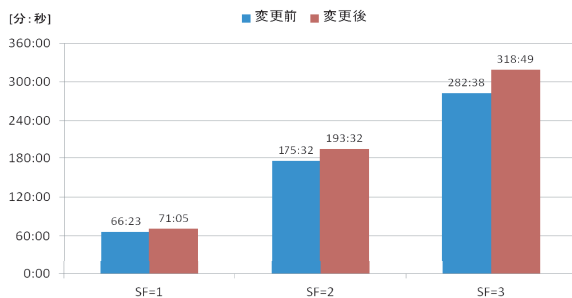


図 12 データ配置変更前と変更後における TPC-H 実行時の実行時間の比較

8. まとめと今後の課題

本研究では、TPC-H の省電力化に向けて、Break-Even Time を詳細に算出し、TPC-H 実行時のディスク I/O 発行間隔を調査した。そして実行時消費電力量を見積と実測により調査し、これらの値を比較することにより、本研究で提案した見積りが妥当であることと、ディスクの省電力状態と I/O 発行間隔を利用することによる TPC-H 実行時の省電力化が可能であることを示した。また、実行時のディスク I/O 発行間隔の制御を行うためのデータ配置手法についても検討を行い、I/O の頻度に応じてデータを 2 種類の HDD に分けて配置することで、実行時の消費電力量を約 40% 削減することに成功し、本手法が実行時省電力化に有用であることを示した。

今後の課題としては、性能劣化と消費電力の削減率のバランスを考慮し、研究方針である、SLA に注目した実行時省電力化に向けて取り組んでいきたい。

参考文献

[1] GIPC Survey and Estimation Committee Report FY2009 (Summary), <http://www.greenit-pc.jp/activity/reporting/100707/index.html>, 2009

[2] TPC-H: <http://www.tpc.org/tpch/default.asp>

[3] Jorge Guerra, Himabindu Pucha, Joseph Glider, Wendy Belluomini, Raju Rangaswami. Cost Effective Storage using Extent Based Dynamic Tiering. In 9th USENIX Conference on File and Storage Technologies, pp.1-14, 2011.

[4] Dushyanth Narayanan, Austin Donnelly, Antony Rowstron. Write Off-Loading: Practical Power Management for Enterprise Storage. In 6th USENIX Conference on File and Storage

Technologies, pp.253-267, 2008.

[5] Athanasios E Papatheasiou and Michael L Scott. Energy Efficient Prefetching and Caching. In Proceedings of the annual conference on USENIX Annual Technical Conference, 2004.

[6] Akshat Verma, Ricardo Koller, Luis Useche, and Raju Rangaswami. SRCMap: Energy Proportional Storage using Dynamic Consolidation. In Proceedings of the 8th USENIX Conference on File and Storage Technologies, 2010.

[7] Charles Weddle, Mathew Oldham, Jin Qian, An-I Andy Wang, Peter Reiher, and Geo Kuenning. PARAD: A Gear-Shifting Power-Aware RAID. In 5th USENIX Conference on File and Storage Technologies, Vol. 3, pp. 245–260, October 2007.

[8] 西川記史, 中野美由紀, 喜連川優: アプリケーション協調型大規模ストレージ省電力システムの開発と評価, DEIM Forum 2012, D6-1, 2012 年 3 月

[9] 西川記史, 中野美由紀, 喜連川優: アプリケーション処理の I/O 挙動特性を利用したディスクの実行時省電力手法とその評価: オンライントランザクション処理における省電力効果, 電子情報通信学会論文誌, Vol.J95-D, No.3, pp.447-459, 2012 年 3 月

[10] Norifumi Nishikawa, Miyuki Nakano, and Masaru Kitsuregawa. Energy Efficient Storage Management Cooperated with Large Data Intensive Applications, 28th IEEE International Conference on Data Engineering (IEEE ICDE 2012), pp.126-137, 2012.04

[11] Alan D. Brunelle: btreord and btreplay User Guide, <http://www.cse.unsw.edu.au/aaronc/iosched/doc/btreplay.html>, 2007

[12] Y.H. Lu, G.D.Micheli: Comparing System-Level Power Management Policies, IEEE Design & Test of Computers, 2010

[13] pdbufsl: <http://www.hitachi.co.jp/Prod/comp/soft1/manual/pc/d635540/W3550027.HTM>

[14] pdobils: <http://133.145.224.19/Prod/comp/soft1/manual/pc/d635540/W3550089.HTM>