

データインテンシブアプリケーションの 実行時省電力に関する考察

飯村 奈穂[†] 西川 記史[‡]
[†]お茶の水女子大学

中野 美由紀[‡] 小口 正人[†]
[‡]東京大学生産技術研究所

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台数も急増しており、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ストレージにおけるデータの効率的な管理に注目が集まっている [1]。

本研究ではデータの効率的な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、本稿ではデータベースベンチマークであるTPC-Hの省電力化に向けて、実行時のシステム性能と消費電力量の解析を行い、TPC-Hの省電力化が可能であることを示した。

2. 基礎性能測定

2.1 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。

2.2 ディスクアクセス時の性能測定

ディスクに対してシーケンシャルおよびランダムアクセスを行った際のディスクの消費電力とスループットを1秒毎に測定する。アクセスの方法はRead処理, Write処理の2種類で測定を行う。測定対象のディスクは1台とし、アクセス単位(バッファサイズ)を4, 8, 16KBと変化させる。データの読み書き量はシーケンシャルの場合は4, 8, 16GB(アクセス単位*1M回), ランダムアクセスの場合は4, 8, 16MB(アクセス単位*1K回)とする。I/Oの方法としてはDirect I/Oを用いる。スループットは/proc/diskstatsの内容を1秒毎に読み込んで測定する。

表1, 2の消費電力とスループットはそれぞれ最大値を示している。表1より、シーケンシャルアクセスの場合、バッファサイズが大きいほど消費電力は大きく、スループットも向上した。これはバッファサイズが大きくなったことで一度にI/Oを行うデータ量が増加したためである。表2より、ランダムアクセスの場合もスループットは同様である。ランダムアクセス時の方が同じ条件のシーケンシャルアクセスより消費電力が大きいのはI/Oを行う際のディ

スクアドレスが毎回異なるためである。これらのことから測定結果は妥当であると考えられる。

表 1: シーケンシャルアクセス時

R/W	BufferSize(KB)	消費電力 (W)	throughput(MB/s)
Read	4	9.4	33.65
	8	10.7	77.67
	16	10.7	77.77
Write	4	9.2	30.9
	8	9.7	52.94
	16	10.1	72.99

表 2: ランダムアクセス時

R/W	BufferSize(KB)	消費電力 (W)	throughput(MB/s)
Read	4	11.97	0.28
	8	11.92	0.55
	16	11.83	1.06
Write	4	9.69	0.45
	8	9.65	0.92
	16	9.62	1.79

2.3 TPC-H 実行時のディスクの性能測定

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LINEITEM表を, HDD2にはその他の表と索引を配置してある。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。

図1, 2にはSF=3の時のHDD1, 2の測定結果を載せる。SF=1, 2の時も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。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り、この結果は妥当な振舞であると考えられる。

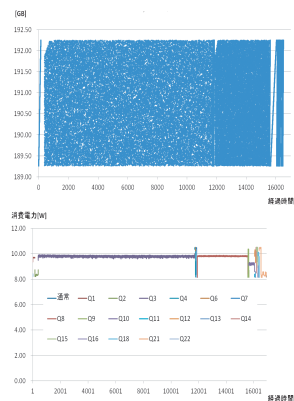


図 1: HDD1(SF=3)

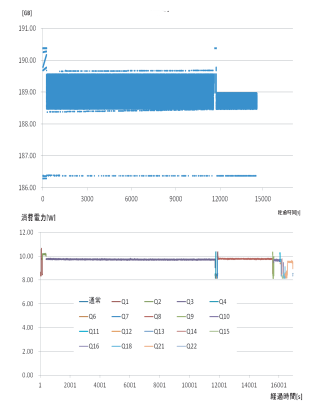


図 2: HDD2(SF=3)

A Study about Runtime Disk Energy Saving of Data Intensive Applications

[†] Naho Imura, [‡] Norifumi Nishikawa, [‡] Miyuki Nakano, [†] Masato Oguchi

Ochanomizu University ([†]), Institute of Industrial Science, The University of Tokyo ([‡])

3. ディスクの消費電力特性

3.1 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 Standby, Idle, Active の 3 種類である。Idle/Active 状態から Standby 状態に移行することを Spindown, Standby 状態から Idle/Active 状態に移行することを Spinup と呼ぶ [3]。

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、先程の測定 2.2 と同様のディスク 2 台である。Standby 時, Idle 時, Active 時の最大消費電力と, Spindown, Spinup に必要なエネルギーを表 3 に示す。

表 3: 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	Standby(W)	Idle(W)	Active(W)	Spindown(J)	Spinup(J)
HDD1	1.2	8.5	10.5	54.3	216.5
HDD2	1.24	8.2	10.8	53.2	172.7

3.2 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。これはディスクに必要なエネルギー (Spindown と Spinup に必要なエネルギー) を E , Standby 状態の消費電力を P_s , Idle 状態の消費電力を P_i , Spindown と Spinup に必要な時間を T_o とすると、Break-Even Time T_{be} は、

$$T_{be} = (E - P_s * T_o) / (P_i - P_s)$$

により求めることができる [5]。

HDD1, HDD2 では Break-Even Time はそれぞれ約 30 秒, 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 では 30 秒以上, HDD2 では 26 秒以上必要である。図 3 は HDD1 において Idle 状態から Standby 状態に移行した後、再び Idle 状態に移行した時の消費電力を示している。

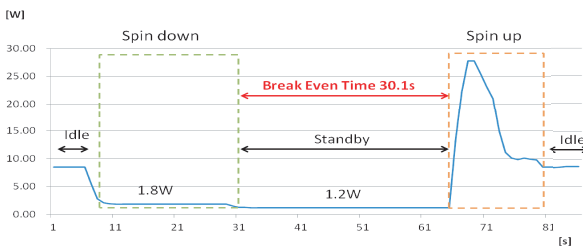


図 3: ディスクの状態遷移における消費電力と Break-Even Time(HDD1)

4. 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 発行間隔を調査し、Break-Even Time を考慮した際に、実行時の消費電力をどの程度削減することができるかを評価する。

4.1 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。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。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。

4.2 削減可能エネルギー

I/O 発行間隔が Break-Even Time 以上である期間に省電力状態を適用した場合、どの程度消費電力を削減することができるかを評価する。省電力状態とは、ディスクをスタンバイ状態に移行することを指す。図 4 は省電力状態適用なしの場合と Break-Even Time を考慮して省電力状態を適用した場合の TPC-H 実行時の消費電力量を表している。省電力状態を適用した場合、HDD1 では 2.5%, HDD2 では 2.1% 消費電力を削減できることがわかる。このことから、TPC-H 実行時のディスク省電力化は可能であると考えられる。

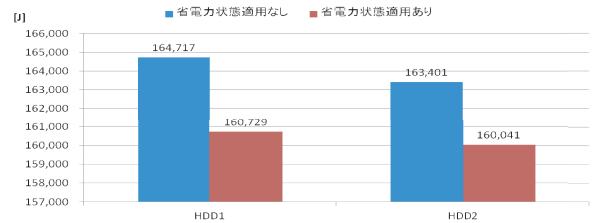


図 4: 省電力状態適用時の TPC-H 実行時消費エネルギー比較

5. まとめと今後の課題

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、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、実行時のシステム性能と消費電力量の解析を行った。また、I/O 発行間隔を利用することによる、TPC-H の省電力化が可能であることを示した。今後は実行時のディスク I/O 発行間隔制御など、更なる省電力化に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>
- [2] 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
- [3] 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
- [4] 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
- [5] Y.H. Lu, G.D.Micheli: Comparing System-Level Power Management Policies, IEEE Design & Test of Computers, 2010