

Web サービスを対象とした実行時ストレージ省電力手法の検討

西川 記史[†], 中野 美由紀[†], 喜連川 優[†]
[†]東京大学 生産技術研究所

1. はじめに

デジタルデータの急増に伴い、データセンタの運用コストは増加の一途を辿っている。特にストレージの電力コストの成長率は他のコストを圧倒しており[1], データセンタの運用コスト削減のためにはストレージの電力コストを下げるのが最重要課題である。

我々は、これまで単一のアプリケーションの入出力挙動特性をストレージ省電力に活用する手法を研究してきた[2]。しかし、一般にデータセンタでは複数のアプリケーションが協調してサービスを提供する。

そこで、本論文では、アプリケーション間の関係をストレージ省電力に利用することを目的に、まずアプリケーションが協調してサービスを提供する Web サービスの代表的ベンチマーク TPC-W[3]の入出力挙動を解析する。さらに、解析結果に基づき複数のアプリケーション間の関係をストレージ省電力に利用する手法の提案、および TPC-W を用いた提案手法の初期評価結果を示す。

2. アプリケーション協調型大規模ストレージ省電力方式

我々が提案しているアプリケーション協調型大規模ストレージ省電力手法[2]について述べる。本手法の特長は、i) アプリケーション実行時のストレージ省電力、ii) アプリケーションレベルにおける入出力発行間隔の長さや read/write 入出力の頻度等のモニタリング結果に基づくアプリケーションレベルでの入出力挙動のパターン化、及び iii) アプリケーションレベルの入出力挙動のパターンに基づく、データ配置制御、キャッシュを利用した入出力発行制御である。図 1 に、我々が提案するアプリケーション協調型大規模ストレージ省電力を実現する実行時ストレージ省電力モデルを示す。

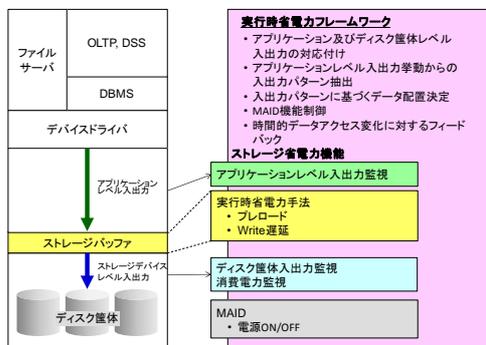


図 1. 実行時ストレージ省電力モデル

Study on Application Runtime Storage Energy Saving Method for Web Service Application.

Norifumi NISHIKAWA[†], Miyuki NAKANO[†], and Masaru, KITSUREGAWA[†]

[†] The University of Tokyo Institute of Industrial Science

アプリケーションは、バッファに対してアプリケーションレベルの入出力を行う。そして、バッファを用いた先読み、書き込み遅延など省電力に有効な手法をデータごとに適用、その後ストレージへの入出力を行う。

3. Web サービスの入出力挙動の解析

Web サービスの代表的ベンチマークである TPC-W を対象に、入出力トレースを採取し、入出力挙動の解析を行った。

3.1 計測環境および計測方法

TPC-W の入出力トレースの採取に用いたサーバおよびストレージの仕様を表 1 に示す。サーバとストレージ間は 2GB の Fibre Channel 1 本で接続されている。

表 1. サーバおよびストレージ仕様

サーバ	日立製作所 HA8000 ・CPU Intel Xeon X5670 2.93GHz (24 コア) ・メモリ 48 GB ・OS Red Hat Enterprise Linux 5.4 (64 bit)
ストレージ	日立製作所 Adaptive Modular Storage 2500 750GB, SATA 7200 rpm ×150 ・RAID 構成 13D+2P RAID 6 ・RAID グループ数 10 ・容量 11.25TB (RAID 構成前) ・キャッシュ 2GB

Web サーバには Apache tomcat を、DBMS には日立製作所の HiRDB Single Server Version 9 をそれぞれ用いた。また、クライアントプログラムとして TPC-W の Java 実装[4]を用いた。TPC-W の DB は Item 数 100 万件、容量は約 20GB、Web サーバに格納される画像ファイル数は 200 万件、約 488GB である。DB、画像ファイルをそれぞれ 5 RAID グループに格納した。我々は、TPC-W を 24 時間稼働し、その入出力トレースを採取した。クライアントのスレッドは数 10、Think Time は 0 秒である。

3.2 入出力挙動の解析結果

TPC-W の入出力挙動の解析結果を図 2 に示す。左側の図は DB の表・索引毎の入出力回数とその累積度数分布、右側の図は Web サーバが管理する画像ファイルの入出力回数とその累積度数分布である。X 軸はデータを入出力回数の降順でソートした際の順位である。

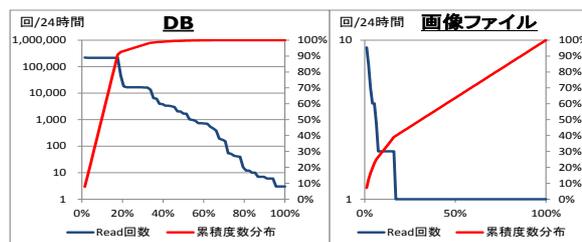


図 2 TPC-W 入出力挙動解析結果

図から分かるように、DBの少数の表・索引、および画像ファイルに入出力が集中していることが分かる。一方、画像ファイルに対する入出力頻度は低く、約85%は1度しか入出力が行われていない。

この結果は、画像データを格納したRAIDグループは、従来のアクセス頻度に基づく省電力方式では高い省電力効果を得ることが困難であることを示している。

4. アプリケーション間の入出力挙動の関係を用いた省電力の可能性

TPC-Wが稼働するストレージの省電力を行うために、TPC-WがWebサーバとDBMSの2種類のアプリケーションから構成されていることに着目する。図1で示した実行時ストレージ省電力モデルをWebサーバとDBMSそれぞれに適用しモニタリングを行うことで、これらのアプリケーション間の入出力を関連付ける。

TPC-Wでは、画像ファイルはDBMSが管理するItem表によって管理される。このため、画像ファイルに対する入出力は、Item表に対する入出力と関連を持つと考えられる。

4.1 Item表に対する入出力を利用した画像ファイルの先読み方法

提案手法の動作を以下に示す。

- Step 1. Item表のブロックとそれが管理する画像ファイル名の対応関係を取得
- Step 2. Item表のブロック毎の入出力を一定時間モニタ
- Step 3. 監視時間内にブロックにアクセスがあると、当該ブロックと対応関係のある画像ファイル群を、モニタリング期間の終了時にバッファに先読み

4.2 シミュレーションによる評価

省電力制御なし、MAIDのみ、提案手法、の3種類の方式について、それらの消費電力と秒当りWebアクセス数(Web Interaction Per Second; WIPS)をシミュレーションにより求め、提案手法の効果を確認した

(1) 評価パラメタ

シミュレーションに用いたパラメタを表2に示す。シミュレーションには、3章で取得した入出力トレースを使用した。なお、4.1節Step1のItem表と画像ファイルの対応関係は与えられるものとした。

表2. シミュレーションパラメタ

DB用RAIDグループ数	5台
画像ファイル用RAIDグループ数	5台
モニタリング期間	600秒
RAIDグループスピンドウンタイムアウト時間	60秒
RAIDグループ消費電力 Active/Idle時	280W
電源OFF時	0W
RAIDグループSpin upエネルギー	14,717J
RAIDグループSpin up待ち時間	60秒
計測期間	3,600秒
Item表ブロック長	1MB
スレッド数	10個
Think Time	0秒

(2) 消費電力とWIPS値の計算方法

RAIDグループの消費電力とWIPS値は以下の計算式により求めた。WIPS値の計算では、Spin up 1回につき、1スレッドが起動待ち時間だけ停止するものとした。

$$\text{消費電力} = \text{RAIDグループ電源ON時間} \times \text{Active/Idle時電力} + \text{Spin up回数} \times \text{Spin up時エネルギー}$$

$$\text{WIPS} = \text{省電力制御なし時のWIPS} \times (\text{計測期間} - (\text{Spin up合計回数} \times \text{spin up待ち時間} \div \text{スレッド数}))$$

(3) 評価結果

図3, 4に省電力制御なし、MAIDのみを適用した場合、および提案方式のそれぞれについて、RAIDグループの消費エネルギーと秒当りWebアクセス数(WIPS値)を示す。なお、省電力なしのWIPS値は実測値である。

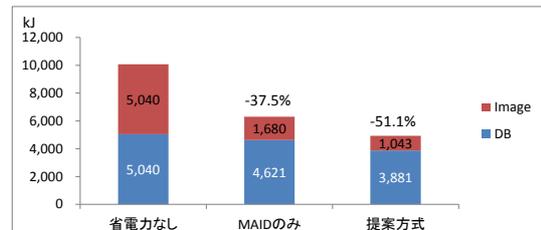


図3. RAIDグループ消費電力

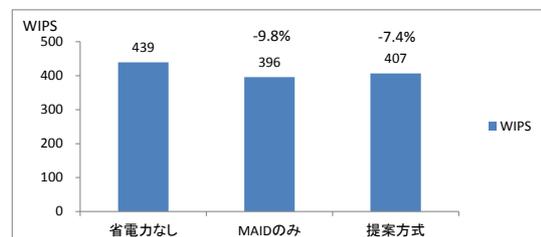


図4. WIPS値

図3から分かるように、MAIDのみを用いた場合のRAIDグループの消費電力の削減率は約37.5%、提案方式を用いた場合は51.1%であった。また、図4よりWIPS値の低下率はMAIDのみの場合は9.8%、提案手法は7.4%であった。

これらの結果より、DBMSが管理するItem表とWebサーバが管理する画像ファイル間の入出力の関係性を利用することで、ストレージ省電力効果を高める可能性があることを示された。

5. まとめ

本論文では、実行時ストレージ省電力モデルを適用することで複数のアプリケーション間の入出力挙動の関連を利用することが可能なストレージ省電力手法を提案し、シミュレーションによりその有効性を調査した。評価の結果、WIPS値低下率を7%程度に抑えつつ、ストレージの消費電力を50%以上削減できる可能性を確認した。

今後は、提案手法を実装・ストレージ上で稼働させ、その有効性を明らかにする。

参考文献

- [1] S W Worth. Green Storage - The Big Picture. In Storage Networking World Spring 2010 Conference, 2010.
- [2] 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), 2012.
- [3] TPC BENCHMARKTMW (Web Commerce) Specification Version 1.8, http://www.tpc.org/tpcw/spec/tpcw_V1.8.pdf
- [4] H.W.Cain, R.Rajwar, M.Marden, M.H.Lipasti, An Architectural Evaluation of Java TPC-W, 7th Intl. Conf. on High-Performance Computer Architecture, pp.229-240, 2001