

# 複数ディスクからなるストレージシステムの省電力化手法 における電力削減効果の比較

引田 諭之<sup>†</sup> LE HIEU HANH<sup>†</sup> 横田 治夫<sup>†</sup>

<sup>†</sup> 東京工業大学大学院情報理工学研究科計算工学専攻

## 1 はじめに

ストレージシステムの省電力化への取り組みは近年における重要な課題の一つである。我々はこれまでに、プライマリ・バックアップ構成を有効活用した省電力化手法として RAPoSDA[3] を提案し、概算式による消費電力量の見積もりによってその効果を確認してきたが、アクセスが時間的に変化するようなワークロードを用いた検証が出来ていなかった。今回我々はシミュレーション実験によりワークロード下における RAPoSDA の省電力効果や性能について検証する。なお実験は関連手法である MAID[1] についても実施し、両手法の省電力効果を比較することによって RAPoSDA の優位性を示す。

## 2 評価対象の省電力化手法

### 2.1 RAPoSDA

RAPoSDA[3] は、多数のディスクドライブを組み合わせたストレージシステムの省電力化を対象としている。データセンター等で実際に運用されるときは信頼性の確保が重要であるため、データは冗長化されて複数のディスクに保存されることが多い(図1)。

RAPoSDA ではデータの信頼性を確保するために、キャッシュメモリとディスクドライブの双方でプライマリ・バックアップ構成をとるようにし、データ配置方法やディスクへの書き込みのタイミングを工夫することで省電力化を実現している。ディスクドライブにおけるデータの配置方法は Chained Declustering[2] を用いている。

### 2.2 MAID

MAID[1] はキャッシュディスクを用いたストレージの省電力化手法である。データアクセスの局所性に着目し、頻繁にアクセスされるデータをキャッシュディスクと呼ばれる常時回転中で小数のディスクドライブに集約し、その他の大多数のディスクドライブへのアクセスを抑制する。アクセスのないアイドル時間がある閾値を超えたディスクドライブはスピンドウンさせてスタンバイ状態にする。

書き込み要求も読み出し要求も最初にキャッシュディスクで受け付けているため、キャッシュディスクの容

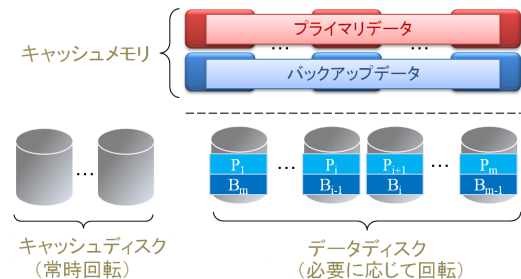


図1: RAPoSDAの全体構成図

量やデータ転送能力が全体性能のボトルネックとなるおそれがある。

また、MAID では大幅な省電力化に焦点をあてているが、ストレージの重要な要素である信頼性については考慮されていない。

## 3 シミュレーション実験

評価実験は、今回新規に作成したストレージのシミュレーションプログラムによって行う。このプログラムではシミュレーション環境上に構築したストレージサブシステムに対し時間的に変化するワークロードを与え、待ち行列をベースとしたモデルにより応答時間をシミュレートする。また各ディスクの状態とその期間を記録することで消費電力量も観測できる。

シミュレーション実験では RAPoSDA および MAID について電力削減率と性能に関して検証を行う。電力削減率については省電力化手法を用いない同数のディスクドライブで構成されるストレージ(以下 Normal と記す)に対する消費電力量の比率で表すこととする。

### 3.1 シミュレーション構成

シミュレーションで用いるハードディスクドライブのモデルは Hitachi Global Storage Technologies の Hitachi Deskstar 7K2000 に基づいたパラメータを使用する。シミュレーションは表1に示すパラメータを持つ人工的なワークロードを用い、データディスクの台数が 100, 500, 1000 台のそれぞれについて 12 時間分のワークロードを与える。RAPoSDA と MAID で用いるキャッシュディスクの台数はデータディスク数の 1% としている。

### 3.2 実験結果

人工的ワークロードを用いてシミュレーションした MAID および RAPoSDA の電力削減率の結果を図2に示す。図2より、RAPoSDA と MAID においては、全てのケースの全ての時間帯において省電力効果が得ら

An Comparison of Power Reduction for Storage Systems consists of Multiple Disk Drives

Satoshi HIKIDA<sup>†</sup> Le Hieu HANH<sup>†</sup> Haruo YOKOTA<sup>†</sup>

<sup>†</sup>Dept. of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology

<sup>†</sup>hikida@de.cs.titech.ac.jp hanhlh@de.cs.titech.ac.jp yokota@cs.titech.ac.jp

表 1: 人工的ワークロードの諸元

workload parameter	value
時間	約 12 時間
read:write	7:3
格納ファイル数	100,000 (1MB/file)
格納ファイルサイズ	200GB (Primary × Backup)
リクエスト数	$\lambda \times 3600 \times 12$
アクセス分布	Zipf 分布
アクセス到着分布	Poisson 到着
Zipf 係数 $s$	1.2
平均到着間隔	25 (request/sec)

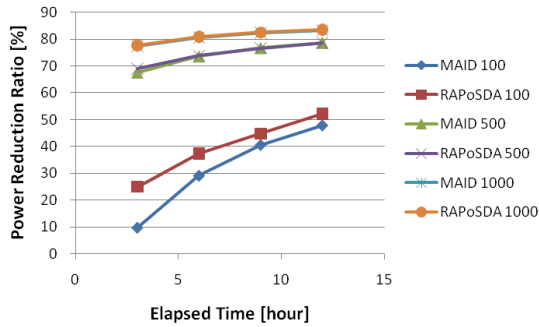


図 2: Normal に対する MAID および RAPoSDA における電力削減率

れていることが確認できる。また、使用するディスク台数が多くなるほど電力削減率は高くなる。ディスク台数が同じケースでは、MAID よりも RAPoSDA の方が高い電力削減率を得られている。ただし運用時間の経過に伴い、両手法間の削減率の差は縮小している。これは、両手法で同数のキャッシュディスクを用いており、時間の経過によってキャッシュディスクにアクセスが集中するためである。但し RAPoSDA はキャッシュメモリも併用しているため、MAID よりもキャッシュディスクへのアクセス頻度が抑制されて電力削減率に多少の差が出ている。

図 3, 4 はそれぞれ平均応答時間およびスループットをグラフに表したものである。平均応答時間は Normal が一番良い結果であるが、これはディスクを回転停止させないためであり、その代わりに消費電力量も最も大きい。MAID と RAPoSDA を比較した場合、データディスク数が 100 台では MAID は極端に平均応答時間が遅くなっている。RAPoSDA でも同数のキャッシュディスクを用いているが、キャッシュメモリの効果もあり 100 台の場合から 1000 台の場合にかけて安定した平均応答時間を維持している。ただし、ディスク数が 500 台と 100 台においては、MAID の平均応答時間が RAPoSDA よりも短い。

スループットは単位時間（秒）当たりに応答したリクエスト数によって求めている。グラフより、全体としては MAID, RAPoSDA とともに Normal と同程度のスループットを確保していることが確認できる。

#### 4 まとめおよび今後の課題

今回我々は新規に構築したシミュレーションプログラムを用いて、ストレージの省電力化手法である RAPoSDA および MAID についてその省電力効果と性能

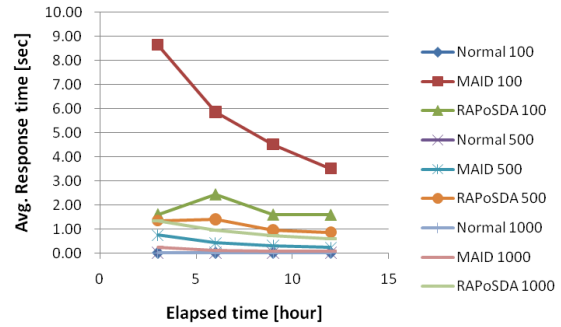


図 3: 平均応答時間

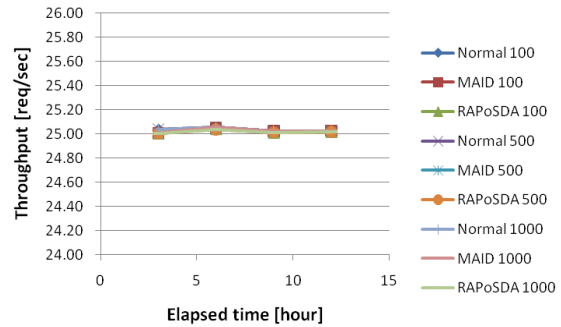


図 4: スループット

に関して検証を行った。省電力効果においては、MAID に対し RAPoSDA の優位性を示すことができた。性能に関しては更なる改善が必要である。今後は実環境での運用を想定して、外部公開されているファイルサーバー等のトレースをベースにしたワークロードによるシミュレーション実験を行う予定である。

#### 謝辞

本研究の一部は、日本学術振興会科学研究費補助金基盤研究 (A)(# 22240005) および文部科学省科学研究費補助金特定領域研究 (# 21013017) の助成により行われた。

#### 参考文献

- [1] Dennis Colarelli and Dirk Grunwald. Massive arrays of idle disks for storage archives. In *Supercomputing '02: Proceedings of the 2002 ACM/IEEE conference on Supercomputing*, pp. 1-11, Los Alamitos, CA, USA, 2002. IEEE Computer Society Press.
- [2] Hui-I Hsiao and David J. DeWitt. Chained declustering: A new availability strategy for multiprocessor database machines. In *Proceedings of the Sixth International Conference on Data Engineering*, pp. 456-465, Washington, DC, USA, 1990. IEEE Computer Society.
- [3] 引田諭之, 横田治夫. プライマリ・バックアップ構成を有効利用したストレージシステムの省電力化手法の提案. In *DEIM Forum 2010 E6-4*, 2010.