T2R2 東京工業大学リサーチリポジトリ Tokyo Tech Research Repository

論文 / 著書情報 Article / Book Information

論題(和文)	ストレージ省電力化手法RAPoSDAのデータ多重化時における制御方式 が与える影響
Title(English)	
著者(和文)	│ │ 引田諭之, Hieu Hanh Le, 横田治夫 │
Authors(English)	Satoshi Hikida, Hieu Hanh Le, Haruo Yokota
出典(和文)	7 7 7
Citation(English)	, , ,
 発行日 / Pub. date	2013,

ストレージ省電力化手法 RAPoSDA のデータ多重化時における 制御方式が与える影響

引田 諭之[†] HieuHanh Le[†] 横田 治夫[†]

† 東京工業大学大学院情報理工学研究科計算工学専攻 〒 152-8552 東京都目黒区大岡山 2-12-1 E-mail: †{hikida,hanhlh}@de.cs.titech.ac.jp, ††yokota@cs.titech.ac.jp

あらまし IT 機器の発達やクラウドサービスの普及等で大量のデータを格納する大規模ストレージが要求されるよう になり,ストレージ自身の消費電力量の増加を抑制することは重要な課題となっている.我々はこれまでにストレー ジ省電力化手法である RAPoSDA を提案し,その有効性を検証してきた.RAPoSDA では可用性確保のためにデー タをプライマリ・バックアップ構成で二重化しているが,可用性を向上させるためには更なる多重化の検討も必要で ある.本研究ではデータを三重化以上として構成する際に,データ配置方法やアクセスタイミング等を様々に変化さ せ,それらの違いが省電力効果や応答性能に与える影響について評価を行う.

キーワード ストレージ,省電力,データ多重化

Satoshi HIKIDA[†], Hieu HANH LE[†], and Haruo YOKOTA[†]

 † Department of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology 2–12–1 Ookayama, Meguro-ku, Tokyo 152–8552, Japan E-mail: †{hikida,hanhlh}@de.cs.titech.ac.jp, ††yokota@cs.titech.ac.jp

1. はじめに

近年における情報量の爆発的増加に伴い,それら大量の情報 を格納するためにストレージも益々大規模化している.米調査 会社の IDC によれば,世界中のデータ量は2年毎に倍増して おり,2015年には全世界で蓄積されるデータ量はおよそ7.9ゼ タバイトに達すると予想されている[7].これら大量のデータ を格納するための大規模なストレージシステムは,データセン ターにおける IT 機器の中でも近年もっとも消費電力量が増加 していると報告されている[10].そのため,ストレージシステ ムの消費電力量を削減することは重要な課題である.

我々はこれまでにストレージシステムの省電力化手法で ある RAPoSDA(Replica Assisted Power Saving Disk Array) [5], [16], [17] を提案してきた.RAPoSDA では,信頼性および 可用性を確保するためにデータをプライマリ・バックアップ構 成で二重化し,個々のディスクドライブの回転状況を考慮した ディスクへのアクセス制御によりストレージシステムの省電力 化を実現している.

ストレージシステムの信頼性と可用性を高めるには,データの 複製数を増やすことが有効な手段の一つであり,RAPoSDA は 二つの複製データを持つことで信頼性と可用性を確保している. しかし,現実に運用されている大規模ストレージシステムでは 二つよりも多く複製データを持つものもある.例えば Goolge 社 が自社で独自開発した Google File System(GFS) [4] や,GFS に触発されてオープンソースとして開発された Apache Hadoop プロジェクトの Hadoop Distributed File System(HDFS) [1] においては,初期設定でデータは三つに複製され,それぞれ 別々のノードに格納される.

データの複製数を増やすことはストレージシステムの信頼性 および可用性を高めるが,ストレージシステムの省電力化の観 点からは問題もある.例えば,複製数が増加するとディスクへ の書き込みも必然的に増加し,省電力化のために停止していた ディスクへのアクセスも増えてしまい,ディスクのスピンアッ プに起因する省電力効果の低下や応答性能の劣化等が起こって しまう.

RAPoSDA においても,信頼性および可用性をより向上させ るためにデータの複製数を増やすという多重化は重要であるが, その際は上記問題にも対応する必要がある.そこで,本研究で は RAPoSDA においてデータを三重化以上に対応させる際に, それまでの省電力効果を出来る限り維持するようなデータ配置 方法やディスクアクセスのタイミングを制御する方法について 検討し,それぞれの方法が省電力効果や応答性能に与える影響 について考察を行うこととする.

なお,本論文の構成は以下の通りである.2.節で我々が提

案している RAPoSDA について概要を説明する.次に3.節で は RAPoSDA を多重化させる際のデータ配置方法およびアク セスタイミングの制御方法について述べ、それぞれの方法が RAPoSDA に与える影響について考察を行う.そして 5.節で 関連研究を紹介し,最後に6.節でまとめと今後の課題について 述べる.

2. RAPoSDA の概要

2.1 RAPoSDA の構成

RAPoSDA [5] はキャッシュメモリとディスクドライブで構成 され,ディスクドライブは更に少数のキャッシュディスクと多 数のデータディスクに分けられる.キャッシュメモリは個別の 電源系統に接続された複数のメモリからなり,無停電電源装置 (UPS)等で断電対策されているものとする.ただし,磁気抵 抗メモリ (MRAM) や強誘電体メモリ (FeRAM) 等の書き換 え可能回数が大きな不揮発性メモリ(NVRAM)が実用化され, これらをキャッシュメモリに採用するならば,このような個別 のメモリに対する断電対策は不要となる.

キャッシュディスクは常に回転しており,クライアントから の読み出し要求に対するキャッシュとして働く.一方データディ スクは一定時間アクセスがなければ回転を停止し, それにより ストレージシステム全体での省電力化を実現している.

また,キャッシュメモリとデータディスクはプライマリ・バッ クアップ構成によりデータを二重化して保持している.RA-PoSDA は一つのキャッシュメモリと複数のデータディスクを 紐付けてこれを論理的なグループとして扱う. あるグループに 所属するデータディスク群をディスクグループ (Disk Group: DG)と呼ぶ.

2.2 RAPoSDA におけるデータ配置

これまで報告してきた RAPoSDA では, データはキャッシュ メモリとデータディスクのそれぞれで二重化されている.キャッ シュメモリでは,プライマリデータは予め格納先ディスクを 割り当てられており, そのディスクが所属する DG に紐付く キャッシュメモリのプライマリデータ格納領域に書き込まれ る.バックアップデータはそれとは別のキャッシュメモリの バックアップデータ格納領域に書き込まれる.一方,データ ディスクでは Chained declustering [6] を採用しており, i 番目 ディスクのプライマリデータに対応するバックアップデータは, $(i+1) \mod N_{DD}$ 番目のディスクに格納される.ここで, N_{DD} はデータディスクの総数である.

3. RAPoSDA のデータ多重化

本節以降で用いる記号とその説明を表1に示す.本研究では 議論を容易にするため, RAPoSDA においてデータを三重化さ せた場合 ($N_R = 3$) におけるデータ配置方法およびアクセス 制御方法について議論するが,これらは三重化以上に一般化さ せる場合でも適用可能である.

3.1 データ配置

3.1.1 データディスクにおける三重化時のデータ配置

データディスクにおける三重化時のデータ配置は, Chained

表1 記号の説明

Table 1 Symbols notation

記号	説明
N _{CM}	キャッシュメモリの総数
N _{DD}	データディスクの総数
N_{DG}	ディスクグループに所属するデータディスク数
N_R	データの複製数
i	データディスク番号($0 \leq i \leq N_{DD} - 1$)
j	キャッシュメモリ番号($0 \leq j \leq N_{CM} - 1$)
	海制レベル (… 手化口の海制を辛吐すて)



(b) CS のデータ配置 図 1 DGA と CS のデータ配置例($N_{CM} = 3$, $N_{DG} = 2$, $N_{DD} = 6$, $N_{R} = 3$)

 D_{2}

Fig. 1 A example of data layout with DGA and CS ($N_{CM} = 3$, $N_{DG}=2$, $N_{DD}=6$, $N_R=3)$

declustering を拡張し. i 番目ディスクのプライマリデータ P_i に対する r 番目の複製データ (r = 0 はプライマリとする) R_i^r は, $i + r \mod N_{DD}$ 番目のディスクに格納されるように配置 する.

3.1.2 キャッシュメモリにおける三重化時のデータ配置

キャッシュメモリ上のデータ配置方法を工夫することでデー タディスクへのアクセスタイミングを制御し,省電力効果の 維持を図る.キャッシュメモリ上のデータ配置でも Chained declustering を応用するが,ここでは以下のようにプライマリ データに対応するキャッシュメモリの割り当て方法が異なる二 つのアプローチを検討する.

- (a) Disk Group Aggregation (DGA)
- (b) Cache Striping (CS)

図 1(a),図 1(b) はそれぞれ DGA および CS のデータ配置 例を示している.図中の D_i はi番目のデータディスク, CM_j

は j 番目のキャッシュメモリを表しており, P_i はデータディス ク D_i に対するプライマリデータ領域で, R_i^r はプライマリデー タ P_i に対する r 番目の複製データ領域である.また, 図中の 実線矢印は次のプライマリデータの位置を示しており, 破線矢 印は r 番目の複製データの位置を示している.

(a)の DGA では,同じ DG に所属するプライマリデータは 同一のキャッシュメモリのプライマリ領域に書き込まれる.複製 データはキャッシュメモリのプライマリ領域の単位で Chained declustering により配置先のキャッシュメモリが決定される.す なわち, i 番目ディスクの複製 r (r = 0 はプライマリとする) のデータが格納されるキャッシュメモリの番号 j とその複製領 域r は,次の式によって求まる.

$$j_r = \begin{cases} \lfloor i_0/N_{DG} \rfloor \mod N_{CM} \ (r=0) \\ j_{r-1} + 1 \mod N_{CM} \ (otherwise) \end{cases}$$
(1)

ここで, i_0 はr番目の複製データに対するプライマリデータが格納されているディスクの番号であり, N_{CM} はキャッシュメモリの総数である.

(b)の CS では,ディスクグループのデータ毎にまとめることはせず,キャッシュメモリへの割り当てはストライピングによって決定される.すなわち,*i*番目ディスクの複製r(r=0はプライマリとする)のデータが格納されるキャッシュメモリの番号*j*とその複製領域rは,次の式によって求まる.

$$j_r = \begin{cases} i_0 \mod N_{CM} \left(r = 0 \right) \\ j_{r-1} + 1 \mod N_{CM} \left(otherwise \right) \end{cases}$$
(2)

3.1.3 キャッシュオーバーフローの違いによる影響の考察 図 2(a),図 2(b)では、キャッシュメモリ上のデータがオー バーフローし、P0 およびその複製データ(R¹₀,R²₀)に対応す るディスクをスピンアップさせてバッファデータを書き込んだ 後の状態を表している.図中の矩形枠はその領域のデータが契 機となってディスクのスピンアップが発生していることを表し、 円形で囲まれている部分は、スピンアップのタイミングで同時 に書き込まれたデータ領域を表している.

図 2(b) より CS では全てのキャッシュメモリのすべてのバッファ領域で空き容量が確保出来ているのに対し,図 2(a)の DGA では CM_0 の r = 2 領域および CM_2 の r = 0,1 領域のデータ はそのまま残ってしまうため,これらの領域が直ぐにでもオーバーフローしディスクアクセスが頻発してしまう可能性が高い.そのため DGA では CS に比べ不要なスピンアップも増えてし まう可能性が高く,省電力効果が減少してしまうことが予想される.

3.2 アクセスタイミングの制御

データディスクに対するアクセスタイミングを制御するため に,キャッシュメモリのバッファ容量閾値を複製領域毎に異な る値に設定する方法について検討する.

図 3 はデータを三重化しているときの例である.この例の場合,これまでに報告していた RAPoSDA ではキャッシュメモリ



図 2 DGA と CS のオーバーフロー後の状態 ($N_{CM} = 3$, $N_{DG} = 2$, $N_{DD} = 6$, $N_R = 3$)

Fig. 2 The situation ater cache memory overflows of DGA and CS ($N_{CM}=3$, $N_{DG}=2$, $N_{DD}=6$, $N_R=3$)



図 3 複製領域毎にバッファ容量閾値に差を持たせた様子 ($N_{DG} = 2$, $N_R = 3$ の場合)

Fig. 3 A setting of differential buffer thresholds on each replication region(In the case of $N_{DG} = 2, N_R = 3$)

はプライマリデータ用領域 (r = 0),二重化目の複製データ用 領域 (r = 1),三重化目の複製データ用領域 (r = 2)に均等に 分けて使用し,それぞれのバッファ容量閾値 Th_{CM} は等しいと していた.ここで,ある割合 α ($0 \le \alpha \le 1$)に対し,プライマ リデータ用領域のバッファ容量閾値 Th_{CM}^0 を, $Th_{CM} \times (1-\alpha)$ とし,二重化目の複製データ用領域のバッファ容量閾値 Th_{CM}^1 を Th_{CM} ,三重化目の複製データ用領域のバッファ容量閾値 Th_{CM}^2 を $Th_{CM} \times (1 + \alpha)$ となるように設定し,プライマリ データ用領域のバッファ容量閾値が最も小さく,三重化目の複





図 4 バッファ容量閾値に差異を設けた場合におけるキャッシュメモリ のオーバーフロー時の様子 ($N_{CM} = 3$, $N_{DG} = 2$, $N_{DD} = 6$, $N_{R} = 3$)

Fig. 4 The situation when cache memory overflows with differential buffer threshold. ($N_{CM}=3$, $N_{DG}=2$, $N_{DD}=6$, $N_R=3)$

製データ用領域のバッファ容量閾値が最も大きくなるように各 領域のバッファ容量に差を持たせるようにする.

3.2.1 バッファ容量閾値に差を持たせることによる影響

図4はバッファ容量閾値に差を設けた場合におけるキャッシュ メモリのオーバーフロー時の状態を示している.このように複 製領域毎に異なるバッファ容量閾値を割り当てた場合,図4(a) にあるように,最初にプライマリデータ領域のデータがオー バーフローすることになる.この時,そのデータに対する二重 化目と三重化目の複製データが含まれる領域ではバッファ容量 に余裕があるためオーバーフローせず,アクセス対象ディスク は一台のみで済む.これらの複製データは同じキャッシュメモ リ中のプライマリデータ(P₂, P₃ または P₄, P₅)がオーバーフ ローしたタイミングで書き込まれるため,複製数の増加に伴う ディスクアクセス回数の増加を抑制することが出来る.

ただし,この方法の注意点としては,プライマリデータ用領 域のバッファ容量は通常よりも α の割合だけ小さくなってし まうため,その分キャッシュメモリのオーバーフローは早く起 こってしまう.更に,図4(b)の例では, CM_1 のr = 0,1の領 域や, CM_2 の全領域(r = 1,2,3)ではバッファ容量閾値近く までデータが蓄積されているため,最初のディスクアクセス後 から短い間隔で再度オーバーフローが発生する可能性が高く, 不要なスピンアップにつながる恐れがある.

表 2 各ストレージシステムの構成

Table 2 Configuration of eache storage systems

項目	Normal		RAPoSDA	
キャッシュメモリ数	-		24 GB $(8$ GB \times 3 $)$	
キャッシュディスク数	-		6	9
データディスク数	60	90	60	90
バッファ容量係数	-		0.2	
複製数	3		3	

表 3 シミュレーションで用いる HDD のパラメータ

Table 3 Prameters of Hard Disk Drive Used in Simulation

parameter	value
容量(TB)	2
プラッター数	5
ディスク回転数(RPM)	7200
ディスクキャッシュサイズ (MB)	32
データ転送速度(MB/s)	134
Active 時消費電力(Watt)	11.1
Idle 時消費電力 (Watt)	7.5
Standby 時消費電力 (Watt)	0.8
Spin-down 時消費エネルギー (Joule)	35.0
Spin-up 時消費エネルギー (Joule)	450.0
Spin-down 時間 (sec)	0.7
Spin-up 時間 (sec)	15.0

4.評価

本節では前節で述べたキャッシュメモリ上のデータ配置方法 とバッファ容量閾値に差を持たせる方法を RAPoSDA に適用 した場合について,消費電力量や応答性能について比較評価を 行う.

評価には我々が構築したシミュレーションプログラム[18]を 用いる.シミュレーションの動作は,あらかじめ用意しておい たワークロードファイルを読み込み,ファイル中の各レコー ドからクライアントリクエストを生成してコントローラーに 要求を出す.コントローラーではリクエストの種類(read or write),リクエストの到着時刻やサイズ等を解析し,各種デバ イスへリクエストを転送する.コントローラー部分と使用する デバイスの構成はストレージシステムの特性に合わせて任意に 構成することが出来る.今回の実験では RAPoSDA と,省電 力化を行わない構成のシステム(以降 Normal と呼ぶ)を用意 した.Normal は各手法を適用した RAPoSDA の省電力効果を 評価するための基準として用いる.なお,各ストレージシステ ムの構成は表2に示す通りである.

4.1 実験パラメータ

シミュレーションに用いるディスクドライブのモデルは Hitachi Global Storage Technologies の Hitachi Deskstar 7K2000 [13] を用いており,詳細なパラメータは表3に示す通りである.

また,シミュレーションに用いるワークロードは表4に示す パラメータを持つ人工的に生成したワークロードを使用する.

表 4 人工的ワークロードの諸元

Table 4 Parameters of Synthetic Workload

workload parameter	value		
時間	約5時間		
read 比率	30%		
格納ファイル数	10,000,000 (32KB/file)		
格納ファイルサイズ	960 GB (3 replicas)		
リクエスト数	$\lambda \times 3600 \times$ 時間		
アクセス分布	Zipf 分布		
アクセス到着分布	Poisson 到着		
Zipf 係数 <i>s</i>	1.2		
平均到着率 (λ)	30 (request/sec)		



図 5 電力削減率





図 6 平均応答時間 Fig. 6 Average response time

4.2 実験結果

図 5 から図 8 はシミュレーション実験結果のグラフである. 各図中で CS, DGA とあるのはキャッシュメモリ上のデータ配 置方法がそれぞれ Cache Striping であるか Disk Group Aggregation であるかを示している.また,括弧中の Fix と Var は,それぞれキャッシュメモリのバッファ容量閾値をすべての 複製領域で固定値とするか,n重化目の複製領域毎に閾値を変 更する方法を採用しているかを示している.

図 5 は Normal の消費電力量に対する RAPoSDA の電力削 減率を表している.図より, CS と DGA の両方で,データディ スク数が 60 台の場合にバッファ容量閾値を固定とした方が電 力削減率は高くなっているが, 90 台では逆に低くなっているこ とが分かる.また,同一データディスク数における CS および







図 8 スピンアップ回数 Fig. 8 A number of spinups

DGA における省電力効果の差は1%程度であり,両手法におけ る省電力効果については顕著な差は認められなかった.

図 6 は,平均応答時間のグラフである.RAPoSDA の平均 応答時間をみると,データディスク数が 60 台では CS よりも DGA の方が平均応答時間は小さく,90 台では両手法でほぼ等 しくなっている.また,データディスク数が 60 台の場合にお ける CS を除き, Var は Fix よりも平均応答時間が小さくなっ ている.

図7,8はRAPoSDAにおけるキャッシュメモリのバッファ オーバーフローの回数およびデータディスクのスピンアップ回 数を表している.図7をみると、データディスク数が90台の 場合ではCSはDGAよりもバッファオーバーフロー回数を抑 制出来ているが,60台では逆にDGAよりもバッファオーバー フロー回数は多くなってしまっている.

また,図8より,データディスク数が60台ではCSはDGA よりもスピンアップ回数が若干抑制出来ているが,こちらも90 台では逆の傾向を示している.FixとVarを比べてみると,す べての組み合わせにおいてVarの方がスピンアップ回数が多く なっている.

今回のシミュレーション結果ではキャッシュメモリ上のデー タ配置方法である CS と DGA を比較すると,データディスク 数が 60 台では,省電力効果,応答性能の両方で DGA が若干 優位である結果となった.但し,データディスク数が 90 台の場 合では逆に CS の方が若干優位であるという結果であり,メモ リ上のデータ配置方法とバッファ容量閾値を複製領域毎に差を 持たせる方法の組み合わせの違いには顕著な差はみられなかった.これは今回用いたワークロードの特性による影響が大きいことが原因だと考えられる.CSのデータ配置では write アクセスの分布が一様であるほど有効に働くと予想されるが,今回用いたワークロードではアクセスの局所性が高く,CSの効果を正確に検証することが困難となってしまった.

5. 関連研究

複製データの数を増加させることでストレージシステムの信頼性,可用性および応答性能を向上させる研究としてはShifted declustering(SD)[15]がある.これはChained declustering 手法を拡張しデータを三重化以上に多重化させ,著者らが定義した理想的なデータ配置の条件を満たすようにデータを各ディスクに割り当てることを提案している.しかしSDではストレージの省電力化については何も考慮されていない.

Willis ら [11] は Chained declustering 手法を活用したスト レージの省電力化に関する研究を行っている.彼らは Chained declustering で構成されたストレージの複製データを活用して 負荷分散と省電力化の両方を実現する手法を提案している.た だし, Chained declustering ではデータは二重化されるだけで あり,多重化までは考慮されていない.

MAID [3] や PDC [12] は,頻繁にアクセスするデータを特定 のディスクに集約し,アクセスの少ないデータを格納している 他のディスクをスピンダウンさせることでストレージシステム 全体の省電力化を実現する.MAID は少数のディスクをキャッ シュ用のディスクとし常に回転状態を維持させ,その他のディ スクは一定時間以上アイドル状態が続いた場合にスピンダウ ンさせて消費電力を削減する.PDC ではアクセス頻度の高い データと低いデータを区別し,一定時間毎にアクセス頻度の高 いデータ群と低いデータ群が別々のディスクに格納されるよう にデータを再配置する.これにより閾値時間以上アイドル状態 の続くディスクをスピンダウンさせて省電力化を実現する.こ れらの手法は比較的単純な仕組みで優れた省電力効果を得られ るが,信頼性を確保することは考慮されていない.

GreenHDFS [8] は,計算機ノードの利用率等を基に,ホット ゾーンとコールドゾーンに各ノードを分け,コールドゾーンの ノードを停止させることによって HDFS クラスターの消費電 力を削減している.

Rabbit [2], Sierra [14], FREP [9] は Power proportinality を実現するための手法である.これらは,システムに対する負 荷に応じて稼働するノード群の数を増減させることで,電力と 性能の比率を一定に保つことを目的としている.もっとも低い 負荷の時には最低限プライマリデータのみを格納するノード群 を稼働させ(低ギア),負荷が高くなった場合はそれに応じて 複製データを持つノード群を徐々に稼働させていく(高ギア). これらは大規模なデータセットを対象としており,プライマリ データ量自体が多い場合にはそれを格納するための最低限の ノードは常に起動させなくてはならないという制約がある.

6. まとめおよび今後の課題

本研究では我々が提案しているストレージ省電力化手法であ る RAPoSDA について,データを三重化させる際に省電力化を 考慮したデータ配置やバッファ容量閾値に差を持たせることに よるアクセスタイミングの制御を検討し,それらが RAPoSDA の消費電力量や応答性能に与える影響についてシミュレーショ ン実験による評価を行った.

今回のシミュレーション結果ではキャッシュメモリ上のデー タ配置方法およびメモリバッファ容量閾値に差を持たせる方法 の組み合わせの違いによる顕著な差はみられなかった.これは 今回用いたアクセスの局所性が強いワークロードの特性による 影響が大きいと考えられる.今後,書き込みが一様に発生する ようなロングテールな特性を持つ大規模なワークロードを用い て更なる検証を行う予定である.

また,その他としては実機環境での評価を行い,シミュレー ションプログラムとの誤差を検証し,シミュレーションプログ ラムの更なる精度向上に取り組むことも今後の課題である.

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)の助成により行われた.

文 献

- [1] Apache Hadoop Project. http://hadoop.apache.org/.
- [2] Hrishikesh Amur, James Cipar, Varun Gupta, Gregory R. Ganger, Michael A. Kozuch, and Karsten Schwan. Robust and flexible power-proportional storage. In *Proceedings of the 1st ACM symposium on Cloud computing*, SoCC '10, pp. 217–228, New York, NY, USA, 2010. ACM.
- [3] 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.
- Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung. The Google file system. SIGOPS Oper. Syst. Rev., Vol. 37, No. 5, pp. 29–43, 2003.
- [5] Satoshi Hikida, Hieu Hanh Le, and Haruo Yokota. A power saving storage method that considers individual disk rotation. In *The 17th International Conference on Database Systems for Advanced Applications (DASFAA)*, Vol. 7239/2010, pp. 138–149, April 2012.
- [6] 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.
- [7] IDC. Extracting value from chaos, 2011. http:// idcdocserv.com/1142.
- [8] Rini T. Kaushik and Milind Bhandarkar. GreenHDFS: towards an energy-conserving, storage-efficient, hybrid Hadoop compute cluster. In *Proceedings of the 2010 international conference on Power aware computing and systems*, HotPower'10, pp. 1–9. USENIX Association, 2010.
- [9] Jinoh Kim and Doron Rotem. Energy proportionality for disk storage using replication. In *Proceedings of the 14th International Conference on Extending Database Technology*, EDBT/ICDT '11, pp. 81–92, New York, NY, USA, 2011. ACM.

- [10] Jonathan Koomey. Growth in data center electricity use 2005 to 2010, 2011. http://www.analyticspress.com/ datacenters.html.
- [11] Willis Lang, Jignesh M. Patel, and Jeffrey F. Naughton. On energy management, load balancing and replication. SIG-MOD Rec., Vol. 38, No. 4, pp. 35–42, June 2010.
- [12] Eduardo Pinheiro and Ricardo Bianchini. Energy conservation techniques for disk array-based servers. In Proceedings of the 18th annual international conference on Supercomputing, ICS '04, pp. 68–78, New York, NY, USA, 2004. ACM.
- [13] Hitachi Global Storage Technologies. Hard disk drive specification, hitachi deskstar 7k2000. http://www.hgst.com/ tech/techlib.nsf/products/Ultrastar_7K4000.
- [14] Eno Thereska, Austin Donnelly, and Dushyanth Narayanan. Sierra: practical power-proportionality for data center storage. In *Proceedings of the sixth conference on Computer* systems, EuroSys '11, pp. 169–182. ACM, 2011.
- [15] Huijun Zhu, Peng Gu, and Jun Wang. Shifted declustering: a placement-ideal layout scheme for multi-way replication storage architecture. In *Proceedings of the 22nd annual international conference on Supercomputing*, ICS '08, pp. 134–144, New York, NY, USA, 2008. ACM.
- [16] 引田諭之, Hieu Hanh LE, 横田治夫. ストレージ省電力化手法の 電力削減におけるシステム構成の影響. In DEIM Forum 2012 D6-2, 2012.
- [17] 引田諭之,横田治夫. 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および評価. In DEIM Forum 2011 D10-1, 2011.
- [18] 引田諭之, LE Hieu Hanh, Koh Kai Hung, 横田治夫. ストレー ジシステムにおける省電力効果検証のためのシミュレータ. 情 報処理学会研究報告. データベース・システム研究会報告, Vol. 2010, No. 19, pp. 1–8, 2010-07-28.