

論文 / 著書情報  
Article / Book Information

題目(和文)	高可用かつ高信頼なストレージシステムの省電力化に向けたデータ管理に関する研究
Title(English)	A Study on Data Management for Saving Power of Highly Available and Reliable Storage Systems
著者(和文)	引田諭之
Author(English)	Satoshi Hikida
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11231号, 授与年月日:2019年6月30日, 学位の種別:課程博士, 審査員:横田 治夫,佐伯 元司,権藤 克彦,吉瀬 謙二,金子 晴彦
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11231号, Conferred date:2019/6/30, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

## 論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	引田 諭之	
論文審査 審査員		氏名	職名	氏名	職名
	主査	横田 治夫	教授	金子 晴彦	准教授
	審査員	佐伯 元司	教授		
		榎藤 克彦	教授		
	吉瀬 謙二	准教授			

### 論文審査の要旨 (2000 字程度)

本論文は、“A Study on Data Management for Saving Power of Highly Available and Reliable Storage Systems (高可用かつ高信頼なストレージシステムの省電力化に向けたデータ管理に関する研究)”と題し、英文6章から成っている。

第1章“Introduction”では、ストレージシステムの動向と、高可用性と高信頼性を保った省電力化の重要性、それを実現するための課題を述べ、その解決策としての本論文におけるアプローチの概要を述べている。さらに、本論文の貢献、および本論文の構成に関して述べている。

第2章“Background”では、本論文の背景となる、HDDからなるストレージシステムの構成、HDDの動作とその電力特性、およびHDDの省電力化の際に重要となる回転停止間隔に対するBreak Even 時間の概念に関して述べている。

第3章“Related Work”では、ストレージシステムにおけるこれまでの省電力化の関連研究および信頼性を考慮した省電力ストレージシステムに関する関連研究について述べている。

第4章“RAPoSDA: Replica-Assisted Power Saving Disk Array”では、本論文の主体となるPrimary/Backup 構成の各 HDD の回転状況を考量して Read/Write 制御を行う省電力化ストレージシステムである RAPoSDA の提案内容と、その評価について述べている。

RAPoSDA では、高信頼化するために HDD 上のデータを Chained Declustering 化した Primary/Backup 構成にするとともに、Read/Write 性能と各 HDD の Break Even 時間以上の回転停止間隔を確保するため、HDD の前段に不揮発化したバッファを用意し、バッファも Primary/Backup 構成としている。その上で、バッファ閾値を超えた時点で回転させる HDD に Primary と Backup データをまとめて Write Back を行う Group Write の機構と、回転中 HDD から優先して Read する Selective Read の機構を実現している。これらの機構の導入により、従来以上に回転停止間隔を確保することができる。

提案する RAPoSDA の効果の評価のため、ディスクの回転確率やバッファへの書き込み可能率等のパラメータを用いた概算式を作成し、HDD の回転を止めない Normal と従来の省電力化手法である MAID (Massive Arrays of Inactive Disks) との比較を行い、省電力効果があることを確認している。さらに、概算式ではワークロードに対する評価ができないことから、RAPoSDA、Normal、MAID の構成のシミュレータを構築し、システムの規模を 32 台構成から 1024 台構成まで変え、Read/Write の比率を変えたワークロードでの評価を行っている。評価の結果、構成と Read/Write の比率を変化させても Normal と MAID に対する RAPoSDA の高い省電力性が達成できていることを示している。それとともにワークロードに対する応答性能の評価も行い、Read/Write 比率や規模を変化させた場合でも MAID に比べて高い応答性能が得られることを示している。さらに、実際の HDD を用いたプロトタイプを構築し、実環境への適用可能性の検証も行っている。20 台の HDD と実電力計測環境を用いたプロトタイプによる実験により、RAPoSDA は MAID と比較して 36%の省電力効果があることを実証している。

システムの信頼性に関する評価では、構成要素の平均故障間隔(MTTF)から RAPoSDA と MAID の構成の平均データ損失時間(MTTDL)を算出し、その差は無視できる程度であることを示している。また、システム構成コストに関する評価も行い、初期コストは、システム構成に依存するが、省電力化による電力費用の削減により 2 か月から 2 年半程度で償却できることを示している。

第5章“Energy Efficient Data Placement and Buffer Management”では、第4章で前提としていた構成が 2 重化による高信頼化であったのに対し、近年の傾向として 3 重化以上の多重化によるストレージの高信頼化の流れがあることから、冗長度を上げた場合にも十分な回転停止間隔を確保するためのバッファの制御に関する手法の提案と、その評価について述べている。

冗長度を上げるとバッファフラッシュ回数が増加するため、HDD のグループとバッファの対応選択、およびバッファフラッシュ時のデータ選択によりバッファフラッシュ回数を抑制する手法を提案している。HDD のグループとバッファの対応選択によるバッファフラッシュ回数抑制手法としては、バッファフラッシュ後のバッファ上のデータ量を偏らせる **Disk Group Aggregation** と、データ量を均等化する **Cache Striping** をアクセス動向によって切り替える手法を提案している。また、バッファフラッシュ時のデータ選択によりバッファフラッシュ回数を抑制する手法として、全ての回転中の HDD にバッファフラッシュする **With All Spins** と、次にバッファフラッシュされるまでの期間が最長と予測される HDD をスピニアップしてバッファフラッシュを行う **Spin-up-EE** の提案をしている。シミュレータを構築し、HDD 台数の構成と Read/Write の比率を変化させた評価を行い、それぞれの手法によりバッファフラッシュ回数の抑制ができていていることを示している。

第 6 章 “**Conclusion**” では、本論文で得られた成果をまとめ、将来の研究における課題について述べている。

以上を要するに、本論文では、今後益々重要性の高まる情報システムの省電力化の一環として、高い可用性と信頼性を確保したストレージシステムにおける省電力化のためのデータ管理に関し、独自の複数の手法を提案し、様々な面からの評価を行い、その効果を示したもので、その成果は工学上貢献するところが大きい。よって我々は、本論文が博士（工学）の学位論文として十分価値があるものと認める。

注意：「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。