

令和 2 年 7 月 2 日現在

機関番号：12102

研究種目：基盤研究(B) (一般)

研究期間：2017～2019

課題番号：17H01718

研究課題名(和文)高精度なアクセスパターン予測に基づく省電力大容量ストレージシステムに関する研究

研究課題名(英文) Research on Power-Aware Large-Scale Storage Systems Based on Highly Accurate Access Pattern Prediction

研究代表者

加藤 和彦 (Kato, Kazuhiko)

筑波大学・システム情報系・教授

研究者番号：90224493

交付決定額(研究期間全体)：(直接経費) 14,600,000円

研究成果の概要(和文)：本研究の遂行による成果は以下の通りである。まず、アクセスパターンの予測手法に関して、アクセス頻度の低いファイル群を、現在と将来のアクセス頻度の高低による4つのグループに分類する手法を得ることができた。また、代表者らによる先行研究を基にしたストレージシステムの省電力化のためのデータ管理手法を確立した。さらに、これらの成果を基にしたストレージシステムの性能評価を行うために、今日のSNSをはじめとする大規模サービスのアクセスログを模倣するシミュレータを開発し、実装実験と併せて性能評価を行うことで、本研究で得られた省電力化手法の有用性を示すことができた。

研究成果の学術的意義や社会的意義

これまで数多くの省電力ストレージが提案されており、特に近年では大規模なサービスの基盤システムでの利用を想定した研究が数多く行われている。そうした中で本研究は、アクセスパターン予測を取り入れることにより、従来の省電力化の基本的なアイデアをさらに発展させたものである。また、本研究を通じて開発されたシミュレータは、近年入手が困難となっている実サービスのアクセスログを、いくつかのパラメータを設定することで容易に模倣することができるものとなっている。これにより、大規模システムの開発のための新しい方法論を提供するものとなっている。

研究成果の概要(英文)：The results of this study are as follows. First, for the access pattern prediction method, we have obtained a method of classifying rarely accessed files into four groups depending on the current and future access frequencies. We have also established a data management method for power saving of storage systems based on our previous study. Furthermore, in order to evaluate the performance of the storage system based on these results, we have developed a simulator that mimics the access log of SNS and other large-scale services. Also, we have conducted performance evaluation together with implementation experiments. As a result, we have demonstrated the usefulness of our method.

研究分野：ソフトウェア

キーワード：ストレージシステム 省電力 アクセスパターン予測 SNS

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

近年、データ量の急増により、データセンターにおけるストレージの占める電力消費量は著しく増加しており、その省電力化が重要な課題となっていた。これまでに提案されてきたストレージの省電力化手法の多くは、一部のディスクにアクセスを集約することにより、残りのディスクへのアクセスを回避して省電力モードに移行させるというものであった。最近では、ビッグデータ処理向けストレージシステムのための省電力化手法も数多く提案されている。しかしながら、大量のデータが継続的にアップロードされる動画共有サービスやビッグデータ処理などの基盤システムに応用するためには、いくつかの課題が残されていた。その一つは、大量のストリームデータを効率良く格納しながら、急増し続けるデータの最適な配置を少ないデータの移動で維持し続けなければならないという問題である。また、人気度の低いデータを集めたディスクに人気度の高いデータが一つでも混入すると、そのディスクは頻繁にスピニングし省電力性や応答性を著しく悪化させる。そのため、データの人気度をいかに高精度で予測するかということも重要な課題であった。

こうした課題に応えるため、研究代表者らは Popular Data Concentration (PDC) と呼ばれる手法のアイデアを基にした省電力ストレージを提案した。具体的には、ディスクアレイの一端にクライアントからアップロードされたデータを順次格納し、ディスクが一杯になる度に新しい空のディスクを追加することを繰り返す。その上で、PDCと同様、定期的にデータを最新の人気度順に再配置するというものである。その評価実験の結果は十分な有用性を示唆するものであったが、実用化のためにはアクセス予測の精度向上が必要であった。また、より効率的なデータの配置方法や高速なルックアップ機構の実現、さらには耐障害性や拡張性の強化といった課題が残されていた。

### 2. 研究の目的

本研究は、データセンタ規模のストレージシステムを対象とした省電力化手法の開発を目的とした。特にここでは、大量のデータが継続的にアップロードされる環境を想定する。この目的を達成するために、本研究では、研究代表者らによるこれまでの研究成果をもとに、以下の3つの課題の解決を目指した。

第一の課題は、データの持つメタ情報などを最大限に活用して、将来のアクセスパターンを高い精度で予測する方法の実現である。PDCやそれを基にした既存の手法の多くは、人気度という尺度を使って抽象化することで将来のアクセスパターンを表現していた。しかし、実際には、同じ人気度であっても将来のアクセスパターンが同じとは限らない。たとえば、特定の個人を写した写真と、一般的な風景を写した写真があるとすると、もし、それぞれの人気度が仮に同じだったとしても、その後のアクセスパターンが異なることは十分有り得る。特定の個人の場合は、その個人が注目を浴びることでアクセスパターンが変動する可能性があるが、一般的な風景写真ではそのような可能性は低い。そこで本研究では、格納されるデータ毎に付与されたメタデータや、データそのもののから抽出した意味情報などを利用することで、将来のアクセスパターンをより精密に予測可能とすることを目指した。

第二の課題は、データ量の増加に対する高い拡張性を備えた、高速かつ効率の良いデータの管理方法の実現である。先に述べたように、本研究が目指すストレージシステムでは、新しくアップロードされるデータを順次格納しながら、時々刻々と変化するアクセス予測に応じてデータを定期的に再配置しなければならない。そのためには、個々のデータに関するアクセスパターンをはじめとする膨大な属性情報を効率的に検索・更新する必要がある。また、外部からのアクセス要求を高速に処理するためのルックアップ機構を備える必要もある。こうした要求に応えるために、本研究では高並列順序付きインデックスの利用による解決を目指した。

第三の課題は、以上の研究で得られた手法の評価である。本研究課題で対象とするストレージシステムは、SNSやeコマースサイトに代表されるような数万から数百万ユーザを対象とするサービス基盤での利用を想定していた。一般に、こうした大規模なシステムの評価を行うためには、類似のサービスを提供する実際のシステムのログを利用する必要がある。しかしながら、実際のシステムのログは一般に公開されていなかったり、公開されていたとしてもセキュリティなどの問題から大きな困難が伴う。そのため本研究では、ネットワーク科学と呼ばれる分野の知見を利用しながら、実世界のSNSを模倣する仮想的なログを生成することが可能なシミュレータを開発し、現システムに近い評価手法の確立を目指した。

### 3. 研究の方法

先に述べた本研究の第一の課題であるアクセスパターンの予測手法については、以下の方針で研究を進めた。従来のアクセス予測は、人気度を予測するという問題に抽象化することで近似的に実現されていた。人気度は、データがアップロードされた日時、および、その日時からの累積アクセス回数という2つのパラメータを、過去の膨大なアクセスパターンに当てはめることで求められる。この手法によって同じ人気度と判定されたデータは、今後はほぼ同じパターンでアクセスされるとシステム内では見なされる。しかし、実際のアクセス数は、急峻なピークを示すもの、周期的な増減を示すもの、急速に増加した後に緩やかに減衰するものなど、さまざまな変化を見せる。意図しないディスクのスピニングを防いでシステムの省電力性能を向上させるためにはアクセス予測の精密化は不可欠である。

そこで本研究では、それらの情報を活用することで、将来のアクセスパターンそのものを予測する手法の実現を目指した。そのために、データに付与されているメタ情報、より詳細なアクセス実績データに加え、必要に応じてコンテンツそのものから抽出した情報(例えば、その写真に何が写っているか)を加えた情報を受け取り、将来予想されるアクセスパターンを予測する予測器の実現を目指した。

また第二の課題であるデータの管理方法については、研究代表者らがこれまでの先行研究において得た省電力ストレージを改良するという方針によって研究を進めた。先に研究代表者らが提案したシステムは、論理的に2次元の格子状に配置された数百から数千のディスクドライブからなる。ファイルがアップロードされると常に左上のディスクに順次格納され、ディスクの容量が上限に達すると右隣に移動する。この動作を続けて1つの行の全てのディスクの容量が上限に達すると、その行におけるアクセス頻度の低いデータが行の右端のディスクに集められ、次の行の左端に挿入される。この動作を繰り返すことで、大量のデータが継続的にアップロードされる環境においても、データの移動を抑えながらアクセス頻度順にデータを配置することができる。しかしながら、この方法では、アクセス頻度の低いデータを集めたディスク内で一部のファイルのアクセス頻度が上昇しても、そうしたファイルをアクセス頻度の高いディスクに移動できないといった問題が残されていた。本研究では、こうした欠点を改善しながら、アクセス頻度予測の結果に合わせたデータ再配置手法の考案を目指した。

第三の課題である評価手法については、上記のストレージシステムの動作と、クライアント側のデータのアクセスパターンの生成を模したシミュレータの開発を目指し研究を進めた。特に、クライアント側のアクセスパターンの生成については、動画や写真などのメディアを共有するためのSNSを想定し、ユーザ間のフォロー関係などのパラメータを設定するだけで、様々な環境におけるアクセスログが容易に生成できるものを目指した。また、シミュレーションによる評価に限らず、実システムによる実験を併せて行い、シミュレーションによる評価の妥当性の確認も行うことを計画した。

#### 4. 研究成果

第一の課題のアクセスパターンの予測に関する成果は以下の通りである。最初に、電力性能の向上のために有用性の高いアクセスパターンの予測は、ほとんどアクセスされない大多数のメディアの集合を更に細かく分類することであるという仮説を立てた。そこから、投稿時から一定時間経過後までのアクセスパターンの情報を用いて、ほとんどアクセスされないメディア群を、以下の手順で4つのクラスに分類する手法を得ることができた。具体的には、全メディアに対して2クラス分類を行い、HotクラスとColdクラスを得る。次に、Coldクラスに対して更に2クラス分類を行い、True Coldと呼ばれる将来的にほぼアクセスされないメディアのクラスと、False Coldと呼ばれる、ほぼアクセスされないように見えても長時間経過後に突発的なアクセスが発生し、電力消費の増加を招く恐れのあるクラスの2つに分類する。最後に、False Coldを更に2クラスに分類し、Keep Viewと呼ばれる突発的なアクセスの傾向が継続するクラスと、Fall Viewと呼ばれる突発的なアクセスが一定時間経過後に収まるクラスに分類するというものである。

第二の課題のデータ管理手法については、ディスクアレイを論理的に複数の階層に分け、上位層には人気のあるファイルを格納するための常時稼働しているWorking Diskと呼ばれるディスクを配備し、下位層には人気のないファイルを格納するためのArchiving Diskと呼ばれるディスクを配備するシステム構成を提案した。この階層構造は、アクセス予測を基にした定期的なデータの再配置により維持されるというものである。これにより、代表者等によって先に提案されたシステムの改良を行った。

第三の課題の評価手法に関する成果としては、特に、SNSにおけるユーザのフォロー関係の偏りやアップロードの頻度といった種々のパラメータをシミュレータの中で設定することにより、それに対応した環境から生成されるアクセスパターンを得ることができるシミュレータを開発した。これにより、近年入手が困難となってきた大規模システムのアクセスパターンを自由に生成することができるようになり、ストレージシステムの動作を模倣したシミュレータと組み合わせることで性能評価が容易に行えるようになった。さらに、上記の管理手法を模したシミュレータと組み合わせることによる性能評価を行い、ファイルの再配置を行わないシステムと比べて7~26%の省電力効果のあることを示すことができた。

以上の成果は、本研究課題の期間中にいくつかの学会で発表を行った。また最終的な成果を複数の論文にまとめ、査読付きの国際会議に投稿中である。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 Horleang Choeng, Koji Hasebe, Hirotake Abe, Kazuhiko Kato
2. 発表標題 Power-Saving Method in Storage Systems for File Sharing Services
3. 学会等名 情報処理学会 システムソフトウェアとオペレーティング・システム (OS) 研究会
4. 発表年 2019年

1. 発表者名 宮本大介, 阿部洋丈, 加藤和彦
2. 発表標題 アクセス予測を用いたストレージシステムの省電力化及び信頼性に関する研究
3. 学会等名 第17回ディベンドブルシステムワークショップ
4. 発表年 2019年

1. 発表者名 井元啓貴, 町田文雄, 阿部洋丈, 加藤和彦
2. 発表標題 アクセス頻度推定に基づくデータ配置方法によるストレージシステム省電力効果の測定
3. 学会等名 第17回ディベンドブルシステムワークショップ
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	長谷部 浩二  (Hasebe Koji)  (80470045)	筑波大学・システム情報系・准教授    (12102)	

## 6. 研究組織（つづき）

	氏名 (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	阿部 洋丈  (Abe Hirotake)  (00456716)	筑波大学・システム情報系・准教授    (12102)	
研究 分 担 者	岡 瑞起  (Oka Miyuki)  (10512105)	筑波大学・システム情報系・准教授    (12102)	