

# データ配置の違いが分散ストレージにもたらす 省電力効果の測定と配置手法の検討

井元 啓貴<sup>1,a)</sup> 阿部 洋丈<sup>2,b)</sup> 町田 文雄<sup>2,c)</sup> 長谷部 浩二<sup>2,d)</sup> 加藤 和彦<sup>2,e)</sup>

**概要:** SNS やクラウドストレージの普及により、データセンターにおけるハードディスクドライブ (HDD) 使用量や、それに伴う電力消費量が増加している。本研究は、SNS やクラウドストレージのデータを格納する HDD を用いたストレージシステムの省電力化を目的とする。データセンターのストレージシステムに格納されるデータは、よくアクセスされるものと滅多にアクセスされないものがあり、多くのデータは時間経過と共にアクセスされなくなる。アクセス頻度の異なるデータを無作為に HDD に格納する方法では、ストレージシステム内の全ての HDD に低頻度でアクセスが発生することになるため、HDD をスピンドウンさせて低消費電力モードに切り替えることが難しくなる。そこで、最終アクセスからの経過時刻により、よくアクセスされるデータを集めた HDD とほとんどアクセスされないデータを集めた HDD に分ける手法を提案する。この手法によって後者の HDD のスピンドウン時間を伸ばし、省電力化を実現する。本発表では、実際の HDD を用いたストレージ群を用意し、フォロワー・フォロワー機能やタグ機能などを持った SNS モデルによるアクセスを行うことで、上記手法の実環境における省電力性能を測定した結果を示す。

## 1. 序論

近年では YouTube や Instagram などのように、ユーザが撮影、または作成したコンテンツをインターネット上に投稿し、投稿されたコンテンツに誰でもアクセスできるサービスが提供されている。特に YouTube においては、1 秒間に 1 時間分の動画がアップロード [1] され、1 日あたり 10 億時間分の動画が視聴 [2] されている。このようなサービスにおいては、ユーザが自分自身でコンテンツを削除しない限りデータはサーバーに残り続け、利用者がいつアクセスしても閲覧できる状態にしておく必要があるため、データセンターのストレージの利用量とそれに伴う消費電力は増加する一方である。

また、2020 年のはじめより流行し始めた新型コロナウイルスの影響で、多くの国や地域で外出自粛制限やテレワークの拡大などが余儀なくされたこともクラウドサービスの利用量増加に拍車をかけた。マイクロソフトの報告 [3] によると、

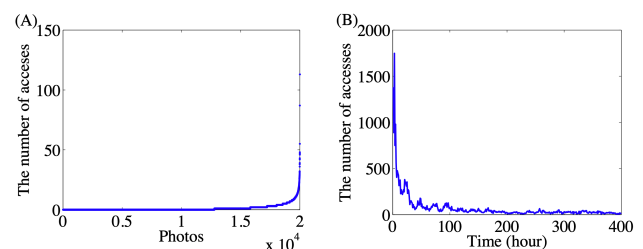


図 1 投稿データごとのアクセス回数 (A) と累計アクセス数の時間変化 (B)[5]

オンライン会議サービスやバーチャルデスクトップサービス、クラウドコンピューティングサービスである Microsoft Azure などの需要が急増したため、同社はデータセンター等のコンピューティングリソースの大幅な増強を行った。これらのようなクラウドサービスは便利なサービスではなく、業務上必須なシステムへ急激に変化している。

2014 年のデータセンターレポート [4] によると、米国でのデータセンターによる消費電力は、国全体での消費電力の 1.8% を占めると報告されており、2010 年-2014 年から約 4% 増加している。同レポートでは 2020 年にかけて同じく 4% の増加が予測されていたが、新型コロナウイルスの影響でその予測を上回る電力が消費されると予測される。新しい生活様式にはクラウドコンピューティングが欠かせなくなっており、消費電力への考慮もより一層重要となっている。

大越ら [5] は、写真共有コミュニティサイト「Flickr」に

<sup>1</sup> 筑波大学 システム情報工学研究科 コンピュータサイエンス専攻  
Department of Computer Science, University of Tsukuba

<sup>2</sup> 筑波大学 システム情報系  
Faculty of Engineering, Information and Systems, University of Tsukuba

a) imoto@oss.cs.tsukuba.ac.jp

b) habe@cs.tsukuba.ac.jp

c) machida@cs.tsukuba.ac.jp

d) hasebe@cs.tsukuba.ac.jp

e) kato@cs.tsukuba.ac.jp

において 20000 件の画像データのアクセスをトレースしたデータにより、分散ストレージへのアクセスの特徴について分析した。図 1(A) では、各画像データにおいて、投稿後 200 時間経過した時のアクセス回数を少ない順に並べたデータである。約 70% のデータにはアクセスされていない一方で、一番アクセスの多い画像データへのアクセス回数は 133 回であった。また図 1(B) では、全データにおけるアクセス回数の累計を 1 時間ごとに計測し、その時間変化を表している。アクセス回数がもっとも最多になったのは 3 時間経過後の 1748 回であり、17 時間経過後には 219 回にまで落ち込んでいる。この結果により、投稿ファイルにはアクセス頻度の高い人気なファイルとそうでない不人気ファイルが存在し、アクセス頻度と経過時間の間に強い負の相関があることがわかる。

本研究の目的は、大量のファイルが連続的にアップロードされる分散ストレージの省電力方法を提案することである。アクセス頻度の高い人気のファイルを集めた Working Disk と、アクセス頻度の低い不人気のファイルを集めた Archiving Disk の 2 つの分けることによる省電力効果と、Working Disk と Archiving Disk の間でファイルをマイグレーションする方法を検討する。本研究の 2 つ目の目的は実際に HDD を用いた分散ストレージを構築し実際の省電力効果を測定することにより、アクセスやマイグレーションコストを考慮した省電力効果を評価することである。

今回は分散ストレージにハードディスクドライブ (HDD) を利用した分散ストレージシステムを対象とする。近年ではソリッドステートドライブ (SSD) のような NAND 型フラッシュメモリが普及し、一般のコンピュータにも広く使われている。しかしながら HDD は、高容量化による導入コストの低下や、頻繁には使わないが長期的に保存しておきたいデータなどを格納する「ニアラインストレージ」としての期待により、これからもデータセンターを支えるストレージとして使用し続けられると予想される [6]。したがって HDD は、前述のような長期間の保管が必要でアクセス頻度の少ないものを格納するサービスの省電力効果の検証に適していると考えられる。

本研究では、SNS モデルを作成して分散ストレージへのアクセス過程を生成し、実際の HDD にアクセスすることで消費電力を測定した。SNS モデルによりアップロードされたファイルをそのままにしたストレージ群と、それぞれのファイルのアクセス回数でファイルを並び替えられたストレージ群では、ファイルを並べ替えたストレージ群で省電力効果が見られた。また、動的なファイルマイグレーションを行った消費電力測定では、小規模な実験ではマイグレーションコストにより消費電力が大きくなる結果となった。



図 2 HDD の 3 つの状態

## 2. 先行研究

最近では分散ストレージに対する様々な省電力手法が提案されている。代表的な省電力手法として、Massive Arrays of Idle Disks (MAID)[7] や Popular Data Concentration(PDC)[8] などがある。どちらも分散ストレージへのワークロードを偏らせることで省電力化を行う考え方に基づいている。

HDD には図 2 のように 3 つの状態があり、それぞれ、アクティブモード、アイドルモード、スタンバイモードである。データにアクセスされている時 HDD はアクティブモードとなり回転・データ転送・制御などに電力が用いられる。アクセスがない場合は回転したまま待機するアイドルモード、一定時間アクセスがない場合は主軸を停止させ、ヘッドを安全な場所に待機させるスタンバイモードとなる。スタンバイモードになると消費電力を大きく抑えることができるが、アクセスがあった場合、ヘッドを動かしシークし始めるのに大きな電力と時間が必要である。そのためスタンバイモードに移行することによる省電力のためには、長時間スタンバイモードに移行する必要がある。

PDC の考え方に基づき Hasebe ら [9] は分散ストレージを Working Disk と Archiving Disk に分類し、継続的にアップロードされたファイルを定期的にアクセス頻度により再配置することで、アクセス頻度の少ないファイルを集めたディスクをスタンバイモードにする方法を提案した。しかしながらスタンバイモードに格納されているファイルへのアクセスが頻繁に行われるとスピンアップに大きな電力が必要となる。

Horleang ら [10] は分散ストレージを Working Disk と Archiving Disk に分ける考えに基づき、定期的な再配置ではなくアクセスに基づいた動的なマイグレーションによる省電力効果を検証した。

本研究では Horleang らの結果をもとに、マイグレーション方法を検討し、実環境での電力測定を実施した。このことにより、実際の HDD を利用した分散ストレージの省電力効果の測定によりマイグレーションによるコストを調査した。

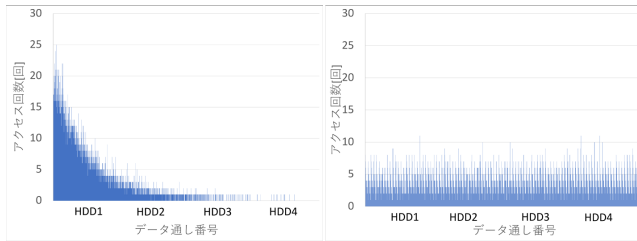


図 3 実験 (i) における各ファイルへのアクセス回数  
図 4 実験 (ii) における各ファイルへのアクセス回数

### 3. 予備実験

この節では予備実験として、以下の2つの場合について分散ストレージの消費電力を測定した。

- (i) アクセス回数の多い順番にファイルが格納されている場合
- (ii) アクセス回数にかかわらずランダムにファイルが格納されている場合

これは、ファイルによりアクセス頻度が大きく異なることから、アクセス頻度順に並べた場合に最も省電力効果が高いという仮定で予備実験を行った。ランダムなアクセス過程を作り出すために、ポアソン分布を用いてそれぞれのファイルについてアクセスのされやすさを設定した。実験後にそれぞれのデータにおいてアクセスされた回数は図3、図4のようになった。

測定の結果を図5、図6に示す。予備実験 (i) では19.7Wh、予備実験 (ii) では22.5Whで、アクセス回数の多い順番にファイルが格納されている場合の方が省電力効果が高い結果となった。

### 4. アクセス過程のモデル

分散ストレージへのアクセス過程を生成するために、SNSモデルを作成した。このSNSモデルでは100人のユーザが利用することを想定し、それぞれのユーザが他のユーザを3人ずつフォローする。他のユーザをフォローすることにより、フォローしたユーザがファイルを投稿した場合にフォロワーの閲覧リストに入り、フォロワーが閲覧することができる。また、ユーザが投稿するファイルにはそれぞれ3つのタグがついており、作成したファイルと同じタグが付けられた最新のファイルにアクセスすることにより、過去へのランダムなアクセスを促している。

私たちが提案するシステムでは、より実際のアクセスに近い過程を作り出すためにバラバシ・アルバートモデルとユールサイモン過程を利用したSNSモデルを作成する。バラバシ・アルバートモデルは現実のネットワークにおいて成長と優先的選択が共存するという考え方に基づいたスケールフリーネットワークを生み出す最小のモデルである [11]。そのモデルを用いたアルゴリズムでは、ノードを

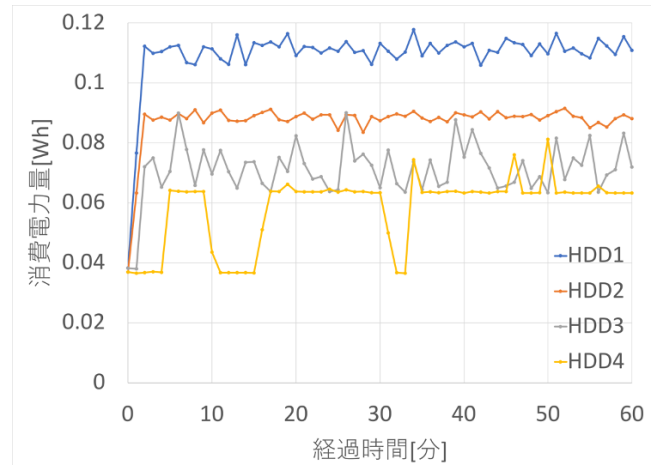


図 5 予備実験 (i) における消費電力量

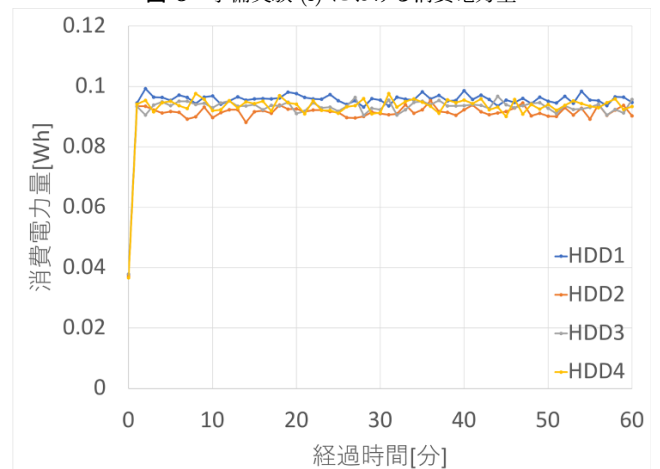


図 6 予備実験 (ii) における消費電力量

ネットワークに加え、そのノードが元から存在するノードと結びつくとき、その時点で次数が高いノードに結びつきやすくなる。このようなバラバシ・アルバートモデルを利用することにより、提案するシステムのSNSモデルにおけるフォロワー・フォロワー関係を構築する。一方、ユールサイモン過程は新たにファイルを作成した際のタグの作成過程をモデル化する。新しいタグを作成する際に一定の確率で新しいタグか既存のタグかを選択する。既存のタグが選択された場合は、これまでのタグの存在の割合でタグが付けられる。それによりタグが付けられた時点でより多くのファイルに属されているタグが優先的に付けられるのである。

構築したSNSモデルでは、1日に3回のアクションを行うことを想定して、365×3回のアクションを1タームとして実験を行った。Algorithm1に構築したSNSモデルのファイル作成、アクセス方法を示す。

上記のSNSモデルで1タームのファイル生成、アクセスを行ったところ、それぞれのファイルのアクセス回数は図7のようになった。これは第1節での分散ストレージへのアクセスの特徴と合致しており、適切なSNSモデルを

### Algorithm 1 SNS モデルにおけるアクセス過程の作成

```
1: for step = 1 to 365 × 3 do
2:   select action users = sample(Users,0.5)
3:   for user in action users do
4:     ispost = rand < 0.3
5:     if (ispost) then
6:       post file and new 3 tags
7:       read latest file that has above tags
8:       send message to follower
9:     end if
10:    read follow users' file
11:  end for
12: end for
```

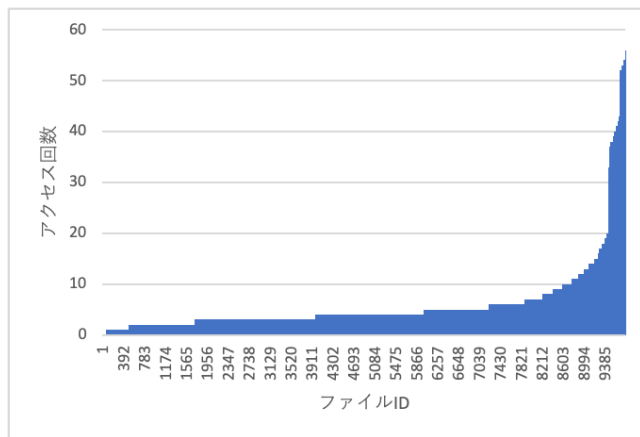


図 7 SNS モデルにおける、それぞれのファイルのアクセス回数

作成できたと考えられる。

この SNS モデルを用いることによって、次節の実験を行った。

## 5. 実験方法

実験ではクライアント PC とサーバ PC を用意し、サーバ PC に外付け HDD を 4 台取り付け、クライアント PC からアクセスした。外付け HDD にはワットチェッカーを取り付け、HDD で使われた電力を測定した。なお、アクセスするファイルは初めから HDD に格納しているためアップロードの負荷は考慮していない。このような環境を作り以下の 4 つの実験を行った。

初めに、1 タームのファイル生成をした場合のアクセスによる消費電力を測定する。SNS モデルを利用して作成されたファイルを、予備実験をもとに次の 2 つの方法で分散ストレージに格納し、ファイルのアクセスによる消費電力量を測定した。

実験 1-(I) ファイルを作成された順番に格納する場合

実験 1-(II) 1-(I) 終了後のアクセスログを用いてアクセス回数の多かった順にファイルを格納する

次に、一定期間経過後のファイルアクセスによる省電力効果を見るために次の 2 つの実験をした。1 ターム目に作

成されたファイルのみを対象として、2 ターム目 (実験 2) と 5 ターム目 (実験 3) のアクセスにおける消費電力量を測定した。実験 1 と同様に以下の方法で分散ストレージに格納したファイルにアクセスした。

実験 2-(I) 1 ターム目でファイルが作成された順番で格納する

実験 2-(II) 2 ターム目におけるアクセスログを用いてアクセス回数の多かった順にファイルを格納する

実験 3-(I) 1 ターム目でファイルが作成された順番で格納する

実験 3-(II) 5 ターム目におけるアクセスログを用いてアクセス回数の多かった順にファイルを格納する

最後に、分散ストレージに格納されているファイルをマイグレーションすることによる省電力効果を調べる。SNS モデルにより作成されたファイルをディスク間でマイグレーションするアルゴリズムを Algorithm2 に示す。実験環境の外付け HDD のうち、1 台を Working Disk、3 台を Archiving Disk とした。アクセスが過去 90step 起こらなかった場合にファイルは Archiving Disk に、また、Archiving Disk のファイルに 1 度でもアクセスがあった場合は Working Disk にマイグレーションする。ここで、マイグレーションのコストは、移行先のディスクにアクセスすることで発生させた。

実験 4 では、SNS モデルで 1 ターム目に作成されたファイルのみを対象に、2 ターム目と 5 ターム目にファイルマイグレーションを行った際の消費電力を測定した。

実験 4-(I) ファイルマイグレーションを行った場合の 2 ターム目における消費電力量

実験 4-(II) ファイルマイグレーションを行った場合の 5 ターム目における消費電力量

## 6. 実験結果

### 6.1 実験 1

実験 1 の結果は図 8, 図 9 のようになった。1 ターム目では、ファイルが作成された順番で分散ストレージに格納していくと、後に利用するストレージの分がスタンバイモードで待機しているため消費電力量は 1-(I) の方が少なかった。

### 6.2 実験 2,3

実験 2 の結果は図 10 から図 13 のようになった。実験 2-(I) による消費電力量の合計は 44.25Wh、実験 2-(II) に

**Algorithm 2** SNS モデルにおけるマイグレーション方法

```

1: for step = 1 to 365 × 3 do
2:   select users = sample(Users,0.5)
3:   for user in select users do
4:     ispost = rand < 0.3
5:     if (ispost) then
6:       post file and new 3 tags
7:       read latest file that has above tags
8:       send message to follower
9:     end if
10:    read follow users' file
11:    if (file == file in archive disk) then
12:      file into working disk
13:    end if
14:  end for
15:  for file in all files do
16:    if (last access step < 90) then
17:      file into archive disk
18:    end if
19:  end for
20: end for

```

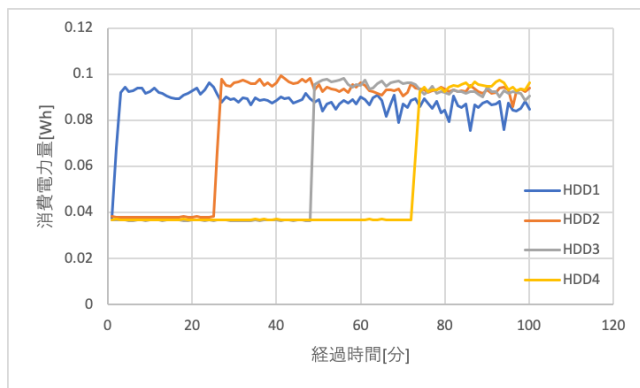


図 8 実験 1-(I) における HDD の消費電力量

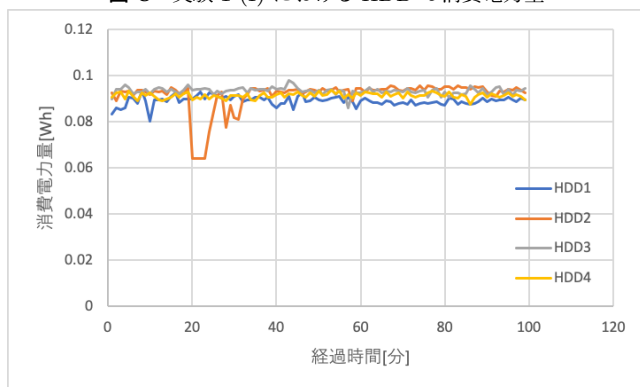


図 9 実験 1-(II) における HDD の消費電力量

よる消費電力量の合計は 44.40Wh、また、実験 3-(I) による消費電力量の合計は 6.21Wh、実験 3-(II) による消費電力量の合計は 8.33Wh である。実験 2 では (I),(II) ともに消費電力量は大きく変わらなかったが、実験 3 では (II) における消費電力量が圧倒的に小さかった。これは、実験 2-(II) では一番アクセスが少ないディスクでもわずかなアクセスでアイドルモードに移行していたのに対し、実験 3 ではワークロードがアクセスの多い 2 つの HDD に抑えら

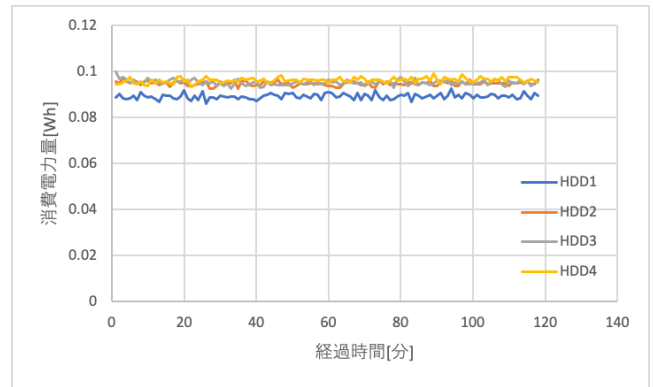


図 10 実験 2-(I) における HDD の消費電力量

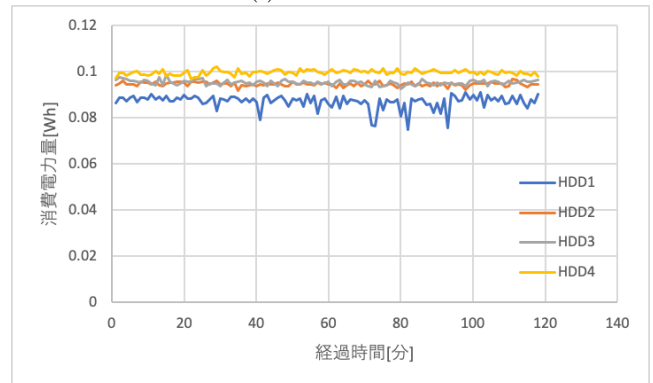


図 11 実験 2-(II) における HDD の消費電力量

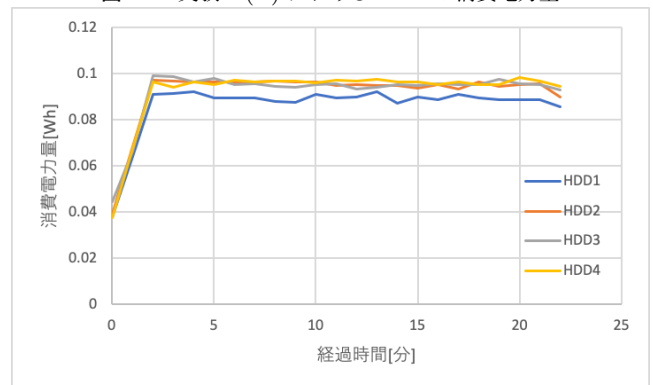


図 12 実験 3-(I) における HDD の消費電力量

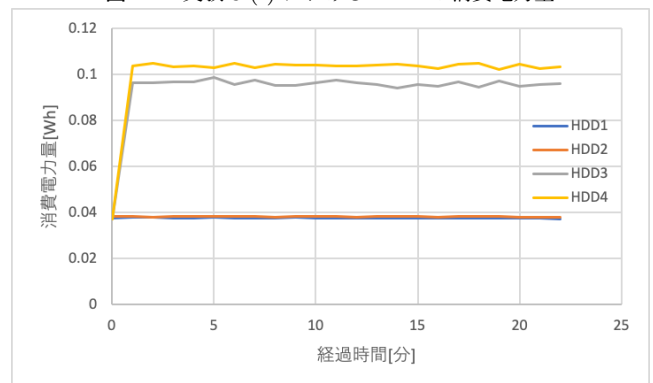


図 13 実験 3-(II) における HDD の消費電力量

れ残りの 2 つの HDD にアクセスが起こらなかったため、ディスクが常にスタンバイモードに移行していたからと考えられる。これまでの結果より、ファイルの投稿から時間

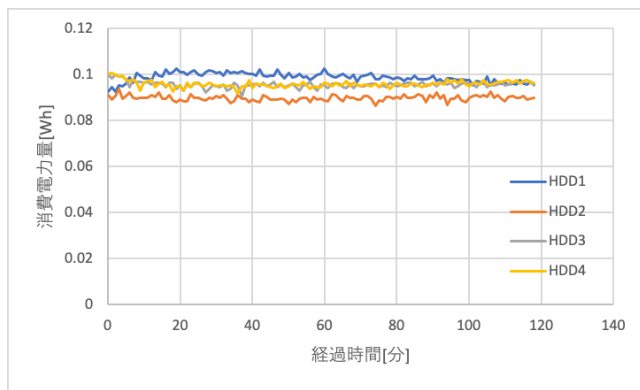


図 14 実験 4(I)における HDD の消費電力量

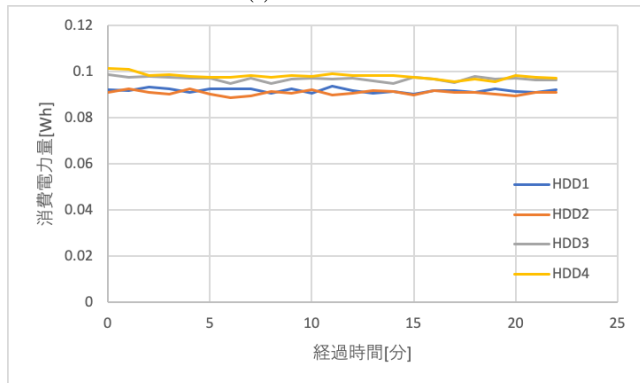


図 15 実験 4(II)における HDD の消費電力量

が経つことによってファイルのアクセス頻度による並べ替えが消費電力削減に貢献していることがわかった。

### 6.3 実験 4

実験 4 の結果は図 14, 図 15 のようになった。HDD の消費電力の合計はそれぞれ 45.18Wh, 8.67Wh となり、実験 2, 実験 3 の結果よりも増加する結果となった。これは、マイグレーションを行うことによるコスト、また、Archiving Disk へのわずかなアクセスによりディスクをスタンバイモードに移行できなかったためであると考えられる。

### 6.4 実験結果のまとめ

本実験では、SNS モデルの 1 ターム目, 2 ターム目, 5 ターム目においての分散ストレージでの配置を考慮する場合としない場合の実環境での消費電力を測定した。また、2 ターム目, 5 ターム目において、アクセスによるファイルマイグレーションを行った場合の実環境での消費電力を測定した。この実験で、2 ターム目においてはディスクの並べ替えによる省電力効果が見られなかったが、5 ターム目においてはディスクの並べ替えによる大きな省電力効果が見られた。したがって、アップロードから時間がたったファイルほど、ファイルの格納順による省電力効果が高いと考えられる。一方、アクセスによる動的なマイグレーションを行った実験では、2 ターム目, 5 ターム目ともに消費電力量が通常よりも大きくなった。省電力効果が見られなかつ

た原因の一つとして実験システムの規模が小さかった考えられる。ディスク数が少ないほど Archiving Disk へのアクセスの可能性が高まり Archiving Disk をスタンバイモードに移行できないためである。この実験で、実環境でのマイグレーションによる省電力効果を検証するためには、規模やスピンダウン時間を考慮した実験が必要であると分かった。

## 7. 結論・今後の課題

これまでの PDC に基づいた研究では、分散ストレージにアップロードされたファイルをアクセス頻度によるファイルの再配置によって省電力化する方法が提案されていた。しかし、実環境における HDD4 台の分散ストレージではマイグレーションによる省電力効果は見られなかった。実験環境の規模が省電力効果に影響していると考えられる。今後は大規模な分散ストレージを用いて、マイグレーションによる省電力効果を検証したい。

さらに、SSD と HDD を組み合わせた分散ストレージの省電力化も課題である。これは、1 つの HDD に過剰なワークロードを与えてしまうと処理速度が大幅に低下してしまう懸念があるからである。SSD はシークタイムが必要ないため処理時間を抑えることができる。また SSD は HDD より消費する電力が少ないため、多くのアクセスが起こる場合は HDD より消費電力を抑えられる可能性がある。これらの点を踏まえて、今後はユーザビリティを考慮した分散ストレージの省電力化方法を検討する。

### 参考文献

- [1] One Hour Per Second: YouTube(online), 入手先 <<http://www.onehourpersecond.com>> (2020.7.2).
- [2] YotTube Official Blog(online), 入手先 <<https://youtube.googleblog.com/2017/02/you-know-whats-cool-billion-hours.html>> (2020.7.2).
- [3] Microsoft news :Innovation Stories(online), 入手先 <<https://news.microsoft.com/innovation-stories/azure-covid-19/>> (2020.7.2).
- [4] Shehabi, Arman, et al. United states data center energy usage report. No. LBNL-1005775. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2016.
- [5] Okoshi, Jumpei, Koji Hasebe, and Kazuhiko Kato. "Power-aware autonomous distributed storage systems for internet hosting service platforms." International Conference on Cloud Computing. Springer, Cham, 2012.
- [6] 日経ビジネス:SSD 全盛時代でも HDD は死なず (online), 入手先 <<https://business.nikkei.com/atcl/report/15/226265/022800225/>> (2020.7.2).
- [7] Colarelli, Dennis, and Dirk Grunwald. "Massive arrays of idle disks for storage archives." SC'02: Proceedings of the 2002 ACM/IEEE Conference on Supercomputing. IEEE, 2002.
- [8] Pinheiro, Eduardo, and Ricardo Bianchini. "Energy conservation techniques for disk array-based servers." Proceedings of the 18th annual international conference on Supercomputing. 2004.

- [9] Hasebe Koji, Jumpei Okoshi, and Kazuhiko Kato. "Power-saving in storage systems for cloud data sharing services with data access prediction." IEICE TRANSACTIONS on Information and Systems 98.10 (2015): 1744-1754.
- [10] Horleang Choeng, Koji Hasebe, Hirotake Abe and Kazuhiko Kato. "Power-Saving Method in Storage Systems for File Sharing Services." 研究報告 システムソフトウェアとオペレーティングシステム, Vol.145, No.5, pp.1-6 (2019).
- [11] Albert-Laszlo Barabasi 著, 池田裕一, 井上寛康, 谷澤俊弘 監訳 (2019). ネットワーク科学-ひと・もの・ことの関係性をデータから解き明かす新しいアプローチ- 共立出版 (2019)