

ディスクの消費電力データストリームを用いたストレージの省電力化手法

小栗 寛生[†] 引田 諭之^{††} 渡辺 陽介[‡] 横田 治夫^{††}

[†] 東京工業大学 工学部情報工学科 ^{††} 東京工業大学 大学院情報理工学研究所

[‡] 東京工業大学 学術国際情報センター

1 はじめに

近年の爆発的なデータ量の増加はストレージシステムを大規模化させ、それに伴いハードディスクを含むストレージシステムの消費電力の増加が問題となっている。こうしたストレージシステムの消費電力を抑えるためにストレージシステムの省電力化手法 [1][2] が提案されている。

しかし既存の手法では、ディスクの動作や構成をディスクの消費電力の状況に応じて変更することは行っていない。そのため消費電力の状況によってはディスクの動作や構成を変更することで更なる省電力効果が期待できると考えられる。

そこで本論文では、複数のディスクの消費電力をリアルタイムに測定できる環境を構築し、実際に測定したディスクの消費電力の状況から最適なディスクの構成やディスクの動作を動的に決定する手法を提案する。

2 前提システム

2.1 ストレージシステム

本論文で前提とするストレージシステムは、複数のディスクを持ち、hdparm などを用いたディスクの回転状態の制御が可能なものとする (図 1 下部)。

こうした環境における既存の省電力化手法には、MAID [1] や RAPoSDA [2] などが挙げられる。MAID は、ディスクをオリジナルなデータを格納する **データディスク** と、最近アクセスされたデータをキャッシュとして配置し常時回転させている **キャッシュディスク** の 2 グループに分ける。そして、キャッシュディスクにアクセスを集中させることで、データディスクへのアクセスを減らしアクセスの少ないデータディスクをスピンドルダウンさせ省電力化を実現している。RAPoSDA は MAID を拡張した手法で、データをプライマリ・バックアップ構成にすることで信頼性を向上させ、データディスクへアクセスする際にディスクの回転状況を考慮することで消費電力の抑制を行っている。

しかし既存の手法は、ディスクの利用状況に応じて

A Power Saving Method Using Energy Consumption Data Stream from Disks

Hiroki OGURI[†] Satoshi HIKIDA^{††} Yousuke WATANABE[‡] Haruo YOKOTA^{††}

[†]Department of Computer Science, School of Engineering, Tokyo Institute of Technology

^{††}Graduate School of Information Science and Engineering, Tokyo Institute of Technology

[‡]Global Scientific Information and Computing Center, Tokyo Institute of Technology

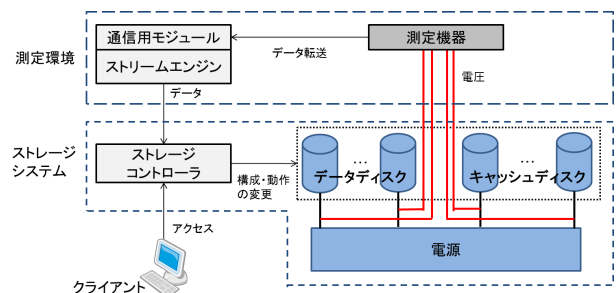


図 1: 前提システム

ディスクの動作や構成の変更を行うことは考慮しておらず、アクセス数が時間によって大幅に変化し、アクセス数に対するキャッシュディスク数の割合が大きい場合や、一度アクセスされると長い時間アクセスされないデータディスクがある場合などは余計な消費電力が発生すると考えられる。

2.2 測定環境

ディスクの消費電力を提案手法で利用するためには、複数のディスクが同時に測定可能で、測定したデータをリアルタイムに利用できる環境が必要となる。我々は、ディスクの消費電力を測定するための回路を作成し、測定機器により測定したデータをストリームエンジンで収集することで、リアルタイムに提案手法で利用できる環境を構築した (図 1 上部)。

2.2.1 消費電力測定

ディスクを接続した回路と測定機器を接続して 5V 線、12V 線の電力を測定できるようにした。そして通信用モジュールを用いて測定機器と通信し、ストリームエンジンを用いて消費電力の平均や合計を計算し、算出した結果をストレージコントローラに送信している。ストレージコントローラでは、受け取った測定結果を元に 3 節で述べるキャッシュディスク数やアイドル時間閾値の変更をする提案手法の処理を行う。

図 2 は実際に 5 台のディスクの消費電力を測定した様子を示している。縦軸はそれぞれのディスクにおける消費電力、横軸は時間を表している。

作成した回路は、ディスクとディスクの電源の間に抵抗を配置し、その抵抗間の電圧降下からディスクの電圧を測定できるようになっている。ディスクの電圧の測定には HIOKI メモリハイログャー 8423 を使用した。測定器によるディスクの電圧の測定間隔は 10ms で、測定器からストリームエンジンへの測定値の送信

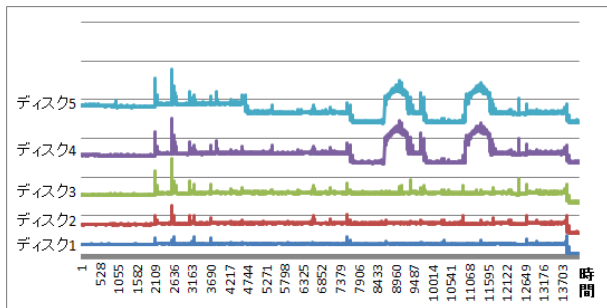


図 2: ディスク 5 台の測定状況

間隔は 1s, ストリームエンジンからストレージシステムへの送信間隔も 1s としている。

2.2.2 測定値のストリーム処理

測定したデータを提案手法で利用するためには, 測定した各ディスクの電圧データから消費電力の計算, 計算した消費電力の平均や合計をリアルタイムに算出する仕組みが必要となる. そこでリアルタイムなデータを処理するためにストリームエンジン [3] を用いて各ディスクの平均や電力量を得られるようにしている。

3 消費電力の測定結果を用いた省電力化

測定した消費電力を用いたキャッシュディスク数の動的変更とアイドル時間閾値の動的変更の 2 つの手法について述べる。

3.1 キャッシュディスク数の動的変更手法

消費電力を抑えるためには利用されないディスクをスピンドウンさせることが効果的であり, 既存の手法においてデータの利用状況やキャッシュディスク数によっては, 一時的に必要なないキャッシュディスクが存在すると考えられる. そこで以下のような条件を設け, キャッシュディスクをスピンドウン, スピンアップさせることでキャッシュディスク数を変更する。

キャッシュディスク数を減らす条件は, ディスクへの単位時間当たりのアクセス数が, 事前に設定した単位時間当たりのアクセス数の閾値よりも少ない場合とする. ただし, キャッシュディスク数が 1 台の時はそのディスクはスピンドウンさせない。

次に, キャッシュディスク数を増やす条件は, キャッシュディスクをスピンドウンさせたことによりデータディスクへ頻繁にアクセスするようになった場合に行う. つまり以下の式 1 が成り立つ場合とする。

$$W_{data,n} - W_{data,n-1} > \bar{W}_{cache,n} \quad (1)$$

ただし, $W_{data,n}$ は n 台のキャッシュディスクをスピンドウンさせた時のデータディスク全体の電力合計値, $W_{data,n-1}$ は $n-1$ 台のキャッシュディスクをスピンドウンさせた時のデータディスク全体の電力合計値, $\bar{W}_{cache,n}$ は現在のキャッシュディスクの電力平均とする。

スピンドウンしたキャッシュディスクの扱いについては, スピンドウンしているキャッシュディスクはアクセスを受け付けないようにする. そして, キャッシュディスクのスピンドウンは提案手法でのみ行い, データの一貫性を保つためにスピンドウンした際には保持していたキャッシュデータは全て削除する。

3.2 アイドル時間閾値の動的変更手法

本論文では, 書き込み/読み込みの終了時刻から, 同じディスクに対する次の書き込み/読み込み時刻までの時間をそのディスクのアイドル時間とする. 既存の手法では, データディスクのアイドル時間が事前に設定した閾値を超えた場合に, 該当するデータディスクをスピンドウンさせている. この閾値となる時間をアイドル時間閾値と呼ぶことにする。

既存手法では, アイドル時間閾値を稼働中に変更することは考慮していない. スピンアップ, スピンドウンすることが多いデータディスクがある場合は, アイドル時間閾値を大きくすることでスピンドウン, スピンドウンの回数を減らし, ほとんどアクセスされないデータディスクがある場合は, アイドル時間閾値を小さくし早めにスピンドウンさせることで, 更なる省電力効果が期待できる。

そこで, n 番目のデータディスクにおけるアイドル時間閾値 $T_{i,n}$ を事前に設定したパラメータ α を用いて, 次の式 2 が成立する場合は $T_{i,n} - \alpha$, 式 3 が成立する場合は $T_{i,n} + \alpha$ に変更する。

$$W_{i,n} * (T_{s,n} + \alpha) > W_{s,n} * (T_{s,n} + \alpha) + J_{spin,n} \quad (2)$$

$$W_{i,n} * (T_{s,n} - \alpha) < W_{s,n} * (T_{s,n} - \alpha) + J_{spin,n} \quad (3)$$

ただし, $W_{i,n}$, $W_{s,n}$ はデータディスクのアイドル時, スタンバイ時の平均電力を表し, $J_{spin,n}$ はデータディスクのスピンドウン, スピンドウンの合計電力量とする. また, $T_{s,n}$ はデータディスクをスピンドウンさせた時刻から次のアクセスによってスピンドウンが開始するまでの時間とし, これらを測定より求める。

4 まとめと今後の課題

本論文では, 実際に測定したディスクの消費電力を利用した省電力化手法の提案を行い, その手法の概要を説明した. そして, ディスクの消費電力を用いた提案手法によりディスクの構成や動作を動的に変更でき, 既存の手法と比べより効果的な省電力化が行えると考えられる。

今後の課題としては, 構築した実際の環境で提案手法を測定し, 提案手法がストレージの消費電力に対してどのような影響を与えるか評価する必要がある。

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。

参考文献

- [1] Dennis Colarelli and Dirk Grunwald. Massive Arrays of Idle Disks For Storage Archives. In Proceedings of the 2002 ACM/IEEE conference on Supercomputing, Supercomputing '02, pp. 1-11, Los Alamitos, CA, USA, 2002. IEEE Computer Society Press.
- [2] 引田諭之, HieuHanh 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. DEIM Forum 2012 D6-2, 2012.
- [3] "StreamSpinner." [Online]. Available: <http://www.streamspinner.org/>