

VMの負荷を考慮したDBテーブル配置によるストレージ省電力 Storage Power Saving by DB Table Relocation Based on VM Loads

谷貝 俊輔[†] 山口 実靖[†]
Shunsuke Yagai Saneyasu Yamaguchi

1. はじめに

近年、情報技術が普及しデータセンター等において多数のサーバ計算機が稼働するようになった。これに伴い、サーバの消費電力の増加が問題となっている[1]。この問題に対する解決策の一つとして、アプリケーションの動作情報を用いてディスク上のデータレイアウトを変更し HDD の消費電力削減する方法がある[1][2]。

本研究では上記手法を仮想化環境に適用し、仮想計算機(VM)ごとのアクセス量を考慮したデータ再配置を行なう。具体的には Xen を用いた仮想計算機上に MySQL を立ち上げ TPC-C 実行時、各テーブルのアクセス量を観察する。そしてアクセス量を考慮してテーブルの再配置を行い、HDD アクセス間隔の拡大の程度とトランザクション性能の評価を行なう。最後に観察結果に基づくデータ配置手法の提案を行い、評価によりその有効性を検証する。

2. 応用情報を用いたストレージ省電力

西川らは、データ(テーブル)へのアクセス頻度を考慮しディスクへのデータ配置を制御することにより、ディスクに省電力機能を適用できるだけの I/O 発行間隔を生成する手法を提案している[1][2]。本手法ではアクセス数が多いデータを Hot データ、アクセス数が少ないデータを Cold データと呼び、図 1 の様にこの Cold データをひとつの HDD に集中させることで特定の HDD のアクセス間隔を拡大させ省電力化をはかっている。

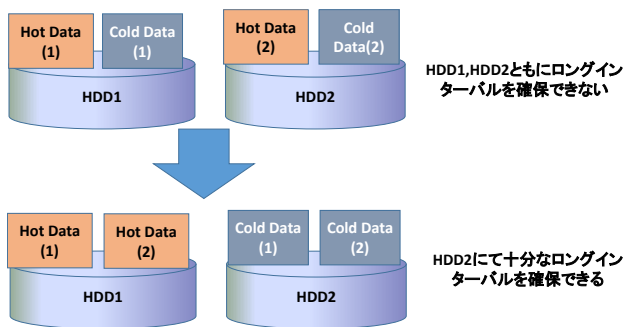


図 1 応用情報を用いたストレージ省電力

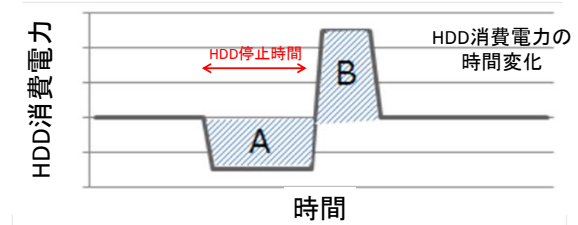


図 2 停止と再起動時の電力の変化

図 2 にストレージの停止と再起動時の電力の変化について示す。ストレージ停止により削減できる電力量とストレージ再稼働により失われる電力量が等しくなる(A=Bとなる)ストレージ停止時間をブレイクオープンタイムと呼び、それより長くなる(A>B)ストレージ停止時間をロングインターバルと呼ぶ。当該手法では配置制御によりロングインターバルを作り出すことで、省電力化を実現している。

ブレイクオープンタイムは HDD の実装に依存するが、文献[2]の例では 25 秒である。

3. 仮想環境における応用情報を用いたストレージ省電力

本章で、仮想化環境においてアプリケーションの動作情報を用いてデータ配置を制御し、特定の HDD におけるアクセス間隔を拡大する手法を提案する。アプリケーションとしては TPC-C を想定する。

本手法では、TPC-C 実行時の DB テーブルファイルへのアクセス要求数をカーネル内で監視し、各ファイルのアクセス頻度を調査する。そしてアクセス要求の少ないテーブルファイルを、特定の HDD 上に集中して配置する。これにより、要求の少ない HDD のアクセス間隔が大きくなると予想される。また別の HDD 群へのアクセス頻度は均等になるようアクセス頻度が高いファイルは別々の HDD に分散配置し、性能の劣化を抑える。

4. 性能評価

4.1 評価方法

仮想化システム Xen を用いて 1 台の物理計算機上に 2 台の VM(VM1, VM2)を起動させ、両 VM 上に MySQL を立ち上げた。そして VM ごとにサイズが異なる TPC-C の表を作成した。VM 1 には SF(スケールファクター)が 4 でサイズ 0.38[GB]の DB を、VM2 には SF が 64 でサイズが 5.8[GB]の DB を作成した。表の配置方法としては下記の 2 通りを用意し比較した。

HDD は 4 台(HDD1 から HDD4)使用し、HDD1 にはゲスト OS のシステムファイルを格納し、ほかの 3 台の HDD(HDD2, HDD3, HDD4)には MySQL のテーブルファイルを格納した。

[†] 工学院大学大学院工学研究科電気電子工学専攻
Electrical Engineering and Electronics, Kogakuin University
Graduate School

一つ目の配置方法では、HDD間でテーブルサイズの合計が均等になるよう配置する。この方法はアクセス頻度を考慮しておらず、本稿ではこれを標準的な配置方法として考える。この配置手法を“サイズ均等”と呼ぶ。

二つ目の配置方法はアクセス頻度を考慮した配置方法であり、アクセス頻度が閾値以下のテーブルをHDD4に配置する。図3がアクセス頻度の観察結果である。図より、テーブルのアクセス頻度は、DBのサイズ(SF)によって変化することがわかる。本配置方法ではアクセス頻度が0.1[回/sec]以下のファイルをHDD4に配置する。残りのファイルはHDD2とHDD3に配置するが、図3により得られるアクセス頻度が同等になる様に配置した。この配置手法を“提案配置”と呼ぶ。両配置方法の詳細は表1の通りである。

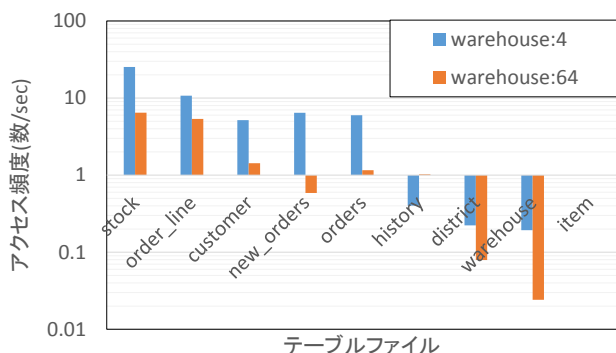


図3 各ファイルのアクセス頻度

表1 ファイル配置の詳細

		HDD2	HDD3	HDD4
サイズ均等	VM1	stock, district, warehouse	order_line, item, history	customer, new_orders, orders
	VM2	stock, district, warehouse	order_line, item, history	customer, new_orders, orders
提案手法	VM1	stock	order_line, new_orders, orders, history, district, warehouse	item
	VM2	stock, order_line, customer	new_orders, orders, history	item, warehouse, district

4.2 評価結果

図4, 5に両配置方法におけるトランザクション性能と最大HDDアクセス間隔を示す。図6に提案配置におけるHDD4のアクセス間隔の発生頻度分布を示す。

図4, 5より、サイズ均等配置と提案配置の性能はほぼ同等であることがわかる。わずかに提案配置が良い理由は、サイズ均等配置ではHDDのアクセス頻度が均等化されていないためであると予想される。図5より、提案配置により100秒を超えるロングインターバルを確保することができ、図6からロングインターバルが複数回発生していることを確認できる。総実験時間が1200秒であり、30秒以上のロングインターバルの合計が428秒であった。以上より、提案配置により性能劣化なくロングインターバルの確保が可能であることが確認された。

5. まとめ

本研究では、応用情報を利用したHDDのアクセス間隔の拡大手法に着目し、それをVMごとのデータサイズが異なる仮想化環境に適用し性能評価を行った。評価の結果、本手法による大幅なHDDアクセス間隔の拡大が確認された。

今後は3台以上の仮想計算機上での実装、メモリの変動による性能の変化、アクセス間隔のさらなる拡大について考察していく予定である。

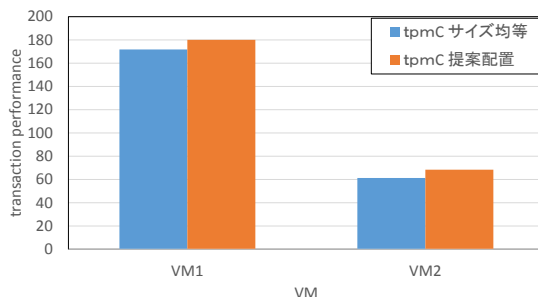


図4 トランザクション性能

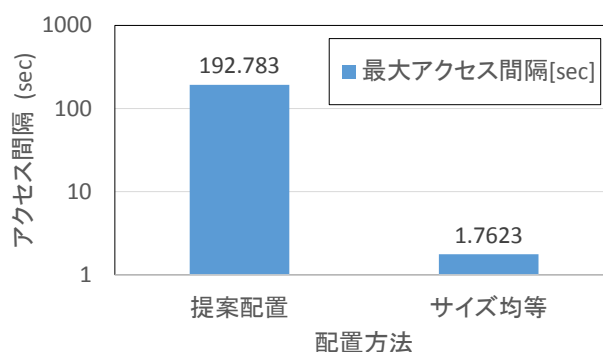


図5 HDD4の最大アクセス間隔

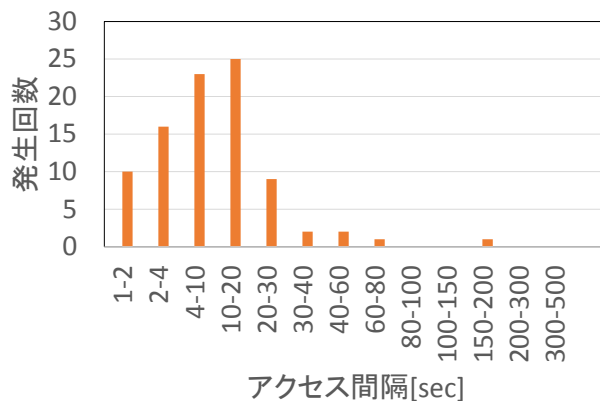


図6 HDD4のアクセス間隔頻度分布

謝辞

本研究はJSPS科研費24300034, 25280022, 26730040の助成を受けたものである。

参考文献

- [1] Norifumi Nisikawa, Miyuki Nakano and Masaru Kitsuregawa, "Energy Efficient Storage Management Cooperated with Large Data Intensive Applications," 28th IEEE International Conference on Data Engineering (IEEE ICDE 2012)
- [2] 西川 記史, 中野 美由紀, 喜連川 優 "アプリケーション処理のI/O挙動特性を利用したディスクの実行時省電力手法とその評価:オンライントランザクション処理における省電力効果" 電子情報通信学会論文誌, J95-D, 3, 1-13(2012.03)