

階層ストレージにおける性能向上のためのデータ配置適正化方式 Data Location Optimization Method for Performance Improvement for Tiered Storage

林 真一[†]
Shinichi Hayashi

薦田 憲久[‡]
Norihiro Komoda

1. はじめに

近年、IT (Information Technology) システムの高性能化に伴い、HDD (Hard Disk Drive) よりも高速なメディアである SSD (Solid State Drive) が普及してきている。SSD は HDD よりも高価であるため全ての HDD を SSD に置換するのではなく、I/O が集中する領域に高性能な SSD を割り当て、I/O が集中しない領域に大容量・低価格である HDD を割り当てる階層ストレージが発表されている [1]。

階層ストレージでは性能の異なるメディアを階層と定義し、I/O 数に応じて適切な階層にデータを自動的に移動する。この機能を本論文では動的階層制御と呼ぶ。移動すべき階層を決定するために I/O を計測する期間を I/O 数計測期間と呼ぶ。システム性能を向上させるには I/O 頻度が高い領域を高速階層に移動する必要がある。そのための 1 つの方法として I/O 数計測期間を短くし、I/O 頻度が高くなった領域を直ちに高速階層に移動することが考えられる。しかし、これによりデータを階層間で頻繁に移動しなければならず I/O 数が増加する。I/O 数計測期間を長くすることによりデータの階層間移動量を削減できるが、この場合はシステム性能が低下する。

本論文では、従来手法では単一である I/O 数計測期間を分割し、長短 2 種類の I/O 数計測期間を適用することにより階層間移動量を削減し、システム全体性能を向上する手法を提案する。

2. 従来方式とその課題

階層ストレージは、高速階層ボリュームおよび低速階層ボリュームを実体とする仮想的なボリュームをアプリケーションに提供する。このボリュームを仮想ボリュームと呼ぶ。高速階層は例えば SSD であり、低速階層は HDD である。仮想ボリュームをページと呼ぶ領域に分割し、ページ単位で高速階層ボリュームまたは低速階層ボリュームを割り当てる。従来の動的階層制御ではページごとにアプリケーションが発行した I/O 数を一定期間計測し、ストレージ内のページに順位を付け、順位が高いページから高速階層を割り当てる。このときに、データを異なる階層に移動する必要があり、これを階層間移動と呼ぶ。

ページへの I/O 数の順位付けを 1 つの仮想ボリューム内で行うのではなく、ストレージ内全ての仮想ボリューム内で行うことにより、ストレージを利用するシステム全体の性能向上を見込める。ここで、ストレージへの I/O 数に占める高速階層への I/O 数の割合を高速階層 I/O 率と呼ぶ。I/O 数が多いページを高速階層に移動し、高速階層 I/O 率を向上させることによりシステム全体性能を向上させる。

従来手法において高速階層 I/O 率を向上させるには、I/O 数計測期間を短くすることが考えられる。これにより、ア

プリケーションからの I/O が集中する領域が変化した場合に、I/O が集中していない領域に割り当てた高速階層を I/O が集中する領域により早く割り当てることができる。しかし、これにより階層間移動量は多くなる。

従来手法において階層間移動量を減少させるには I/O 数計測期間を長くすることが考えられる。しかし、この場合は I/O が集中しない領域に高速階層が割り当たったままとなり、高速階層 I/O 率は低下し、I/O 数計測期間が短い時と比較してシステム全体性能が低下する。

仮想ボリュームクラスタリング法 [2] は、従来の動的階層制御が適用されているストレージにおいて仮想ボリュームへの I/O 数のオーダーと目標レスポンスタイムが類似している仮想ボリュームをクラスタリングし、高速階層や低速階層の容量をクラスタごとに調整し、メディアコストを低減する方式である。これは各階層の容量を調整する方式であり、各ページに割り当てる階層を決定する方式ではない。

3. 提案方式

この課題を解決するために、長短 2 種類の I/O 数計測期間に基づいて高速階層を割り当てるページを決定する方式を提案する。

I/O 数計測期間を短くせずに高速階層 I/O 率を向上させるために、I/O が集中すると予測できる領域に高速階層をあらかじめ割り当てておくことを考える。長期間を通して I/O 数が多い領域は、短期間で I/O 数が少なくなったとしても再び I/O 数が多くなると考えられる。よって、短期間で I/O 数が多い領域に高速階層を割り当てるだけでなく、長期間を通して I/O 数が多い領域にも高速階層を割り当てておく。

上記を実現するために、図 1 に示すように、高速階層の容量を 2 つに分け、長短 2 つの I/O 数計測期間を 2 つに分けた高速階層にそれぞれ適用する。具体的には、過去 24 時間で I/O 数が多かったページに高速階層の一部の領域を割り当てる。このとき割り当てた領域を長期 I/O 数計測期

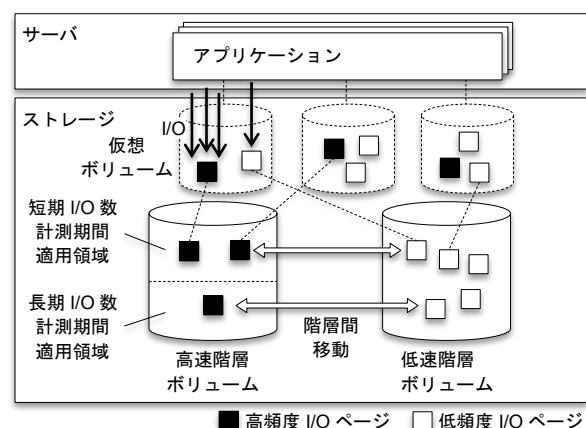


図 1 提案方式の概要

[†] 株式会社 日立製作所, Hitachi, Ltd.

[‡] 大阪大学, Osaka University

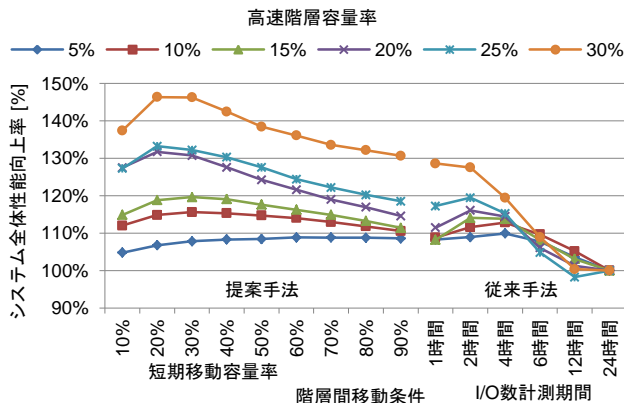


図 2 システム全体性能

間適用領域と呼ぶ。この処理を 24 時間おきに行う。次に過去 1 時間で I/O が多かったページのうち長期 I/O 数計測期間適用領域が割り当てられていないページに残りの高速階層の領域を割り当てる。この処理を 1 時間おきに行う。

4. 評価実験

従来方式と提案方式を比較するために、実稼働環境で取得した I/O ログデータ[3]に基づいてストレージへの I/O およびページの階層間移動のシミュレーションを行った。以下、シミュレーション条件について述べる。

従来方式および提案方式のどちらの場合においても、割り当て済み容量に対する高速階層の容量の割合を変化させ、システム全体性能および階層間移動量を計算した。この割合を高速階層容量率と呼ぶ。従来手法ではさらに、I/O 数計測期間を 1, 2, 4, 6, 12, 24 時間と変化した。

提案手法では長短 2 種類の I/O 数計測期間を指定する必要がある。I/O ログデータは 24 時間を 1 つの周期としていることから、長期 I/O 数計測期間を 24 時間とした。従来方式において、I/O 数計測期間は一般的に 1 から 24 時間であるため、短期 I/O 数計測期間を最も短い 1 時間とした。ここで、短期 I/O 数計測期間を適用する高速階層の容量の割合を短期移動容量率と呼ぶ。さらに、提案手法のシミュレーションでは短期移動容量率を 10% 単位で変化させた。先行研究[2]を参考にし、ページサイズを 10MB、高速階層 I/O 性能と低速階層の I/O 性能の比を 180:1 とした。仮想ボリュームに割り当てられたページ数は 8,679 であった。I/O ログデータに基づいてシミュレートしたストレージへの I/O は、10% のページに 50% の I/O が集中し、20% のページに 71% の I/O が集中し、30% のページに 82% の I/O が集中する特性であった。高速階層 I/O 率(r_{High} [%])、高速階層 I/O 性能(P_{High} [IOPS])、低速階層 I/O 性能(P_{Low} [IOPS])、階層間移動に伴う I/O 数($P_{Migration}$ [IOPS])に基づき、システム全体性能(P_{Total} [IOPS])を 1 時間ごとに下式(1)に基づき算出し、1 か月分のシミュレーションを実施し、それぞれ平均値を求めた。

$$P_{Total} = \frac{1}{\frac{r_{High}}{P_{High}} + \frac{1-r_{High}}{P_{Low}}} - P_{Migration} \quad (1)$$

5. 実験結果と考察

図 2 に提案手法従来手法における I/O 数計測期間を変化させた場合のシステム全体性能を示す。縦軸は従来手法に

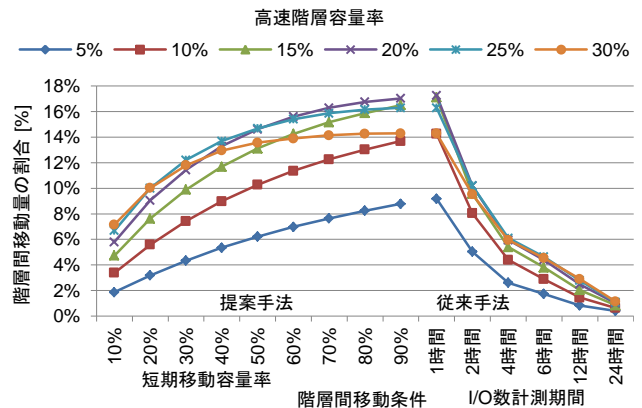


図 3 階層間移動量

おける I/O 数計測時間が 24 時間の場合のシステム全体性能を 100% としたときのシステム全体性能向上率を示す。横軸は階層間移動条件であり、従来手法の場合は I/O 数計測時間を、提案手法の場合は短期移動容量率を示す。

図 3 に従来手法および提案手法における階層間移動のための I/O 数の変化を示す。縦軸は、アプリケーションからストレージへの通常 I/O 数に対する階層間移動のための I/O 数の割合を示す。横軸は、従来手法の場合は I/O 数計測時間を、提案手法の場合は短期移動容量率を示す。

提案手法では、高速階層容量率が 10% 以上の場合それぞれにおいて、短期移動容量率が 20% または 30% のときにシステム全体性能が最大となり、従来手法のシステム全体性能より高くなった。提案手法のシステム全体性能は従来手法と比較して、高速階層容量率が 10% の場合は 2.5%、同 20% の場合は 13.5%、同 30% の場合は 13.8% 向上した。このとき、高速階層容量率が 20% および 30% の場合はにおいて階層間移動量は増加していない。

高速階層容量率が 30% の時に、提案手法の短期移動容量率が 10% の時の階層間移動量は、従来手法の I/O 数計測期間が 1 時間の時と比較して、システム全体性能が低下せず 50.0% 削減された。

6. おわりに

階層ストレージにおける高速階層 I/O 率を向上するために、長短 2 種類の I/O 数計測期間に基づいて高速階層を割り当てるページを決定する方式を提案した。シミュレーションにより有効性を確認した結果、階層間データ移動量を増加させずに I/O 性能が最大で 13.8% 向上した。また、長期 I/O 数計測期間適用領域の割合を変化させることにより、I/O 性能を低下させずに階層間データ移動量を 50.0% 削減できた。

参考文献

- [1] 鋤柄 力, 熊谷 奈緒子, “クラウドコンピューティングを支えるストレージソリューション”, 日立評論, Vol. 93, No. 7, p.p. 44-47 (2011)
- [2] 江丸 裕教, 高井 昌彰, “仮想ボリュームクラスタリング法による動的階層制御ストレージの性能管理”, 情報処理学会論文誌, Vol. 52, No. 7, p.p. 2234-2244 (2011).
- [3] Hewlett-Packard Development Company, “Cello99 Traces”, <https://tesla.hpl.hp.com/opensource/cello99/>