

ストレージ階層仮想化機能の実現方式検討

須藤 粹[†] 坂 弘明[†] 大平 良徳[†] 江口 賢哲[†] 山本 政行[†]
 (株)日立製作所 システム開発研究所[†]

1. はじめに

電子メールや電子文書など、さまざまな業務分野において情報のデジタル化が進み、企業のデータ量は急速に増え続けている。しかし、その一方で、企業のITコストは増加が見込めないという問題がある。そこで、増え続けるデータを格納するストレージの容量を増やすために、安価だが低速なドライブを追加すると、これまでより性能が低下する懸念がある。このため、コストパフォーマンスの高いストレージシステムが求められている。

この要求に応える一手段として、我々は、ストレージ装置が自動的に、データのアクセス頻度に応じて適切なストレージ階層(高価で高速なメディア/安価で低速なメディア)にデータを配置する「ストレージ階層仮想化機能」を開発した。

本研究では、ストレージ階層仮想化機能の実現方式を検討し、その効果をシミュレーションにより検証した。

2. 背景と目的

ある企業における、サーバからのストレージに対するデータアクセス頻度の分布を分析した結果、20%の領域に80%のI/Oが集中していることが報告されている(図1左)[1]。このように、ストレージ内のアクセス頻度には偏りがあるため、企業では、アクセス頻度に応じて適切なストレージ階層にデータを配置する「階層管理」を行っている[2]。

この階層管理は従来、ボリューム単位で行っていたが、サーバ仮想化の普及などによりボリューム内でアクセス頻度の異なるデータが混在するようになった(図1右)。

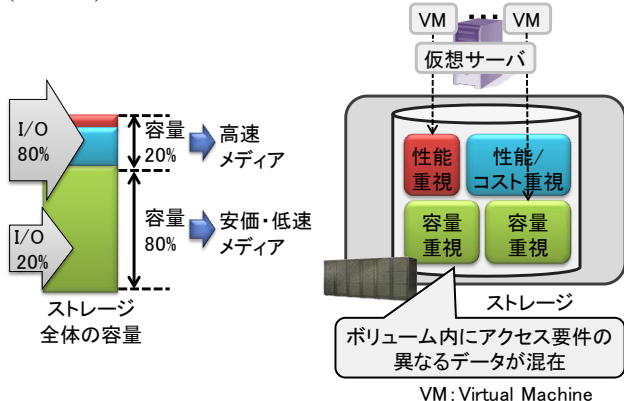


図1 ストレージおよびボリューム内のアクセス頻度の偏り

このため本研究では、ボリュームより小さな単位で階層管理を行うことによる、更なるコストパフォーマンス向上を目的とし、その実現方式の確立と効果検証方法の確立が課題となった。

3. ストレージ階層仮想化機能

ストレージ装置が動的にボリュームより小さな単位で階層管理を行う機能として、ストレージ階層仮想化機能を開発した。本機能は、容量の利用効率を向上させることを目的とした容量仮想化機能の拡張機能として提供する[3]。容量仮想化機能とは、サーバからの書き込みに応じて、ボリュームより小さい実領域(ページ)をストレージプールから仮想ボリュームに動的に割当てする機能である。

ストレージ階層仮想化機能の処理の流れを図2に示す。

- (1) サーバからの書き込みに応じて、ストレージプールからページを割当て、仮想ボリュームにどのページを割当てたかを示すアドレスをアドレス管理テーブルに記録する。
- (2) ページのアクセス頻度をモニタし、その結果をアクセス頻度管理テーブルに記録する。
- (3) (2)のモニタ結果に基づき、ページの適切な階層を判定する。
- (4) (3)の判定結果に基づき、ページを適切な階層に再配置する。

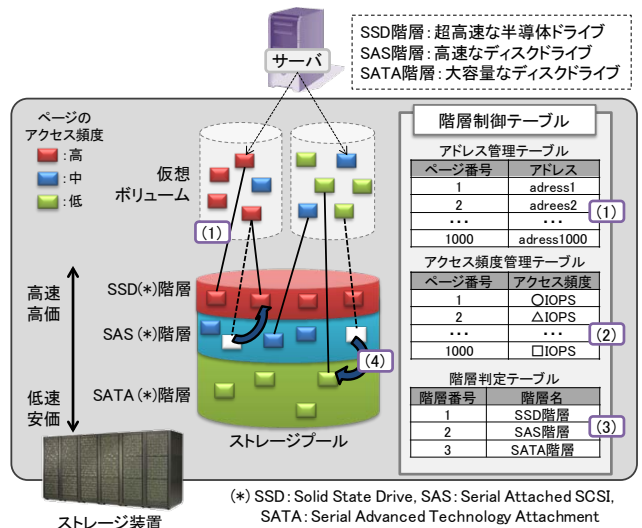


図2 ストレージ階層仮想化機能

本機能により、アクセス頻度の高いページを高速なメディアに、アクセス頻度の低いページを安価で低速なメディアに自動的に配置することができる。

A Study of Virtual Storage Tiering Function

Azusa Sudo[†], Hiroaki Akutsu[†], Yoshinori Ohira[†], Yoshiaki Eguchi[†], Masayuki Yamamoto[†]

[†] Systems Development Laboratory, Hitachi, Ltd.

4. 効果の検証

ストレージ階層仮想化機能の効果を評価するために、従来行っていたボリューム単位とページ単位とで階層管理する場合のコストパフォーマンスを比較する。

4.1. 検証方法

本検証では、記憶媒体のコストが一定となるように、プールの構成を固定して、各単位で階層管理する場合のプールの最大スループット性能を算出する。算出の結果、前記性能が高い方が、コストパフォーマンスが高いと言える。

以下に、プールの最大スループット性能の算出方法を述べる。

(1) 各階層に対する I/O 比率の算出

まず、Cello99(*1)のアクセス分布に基づき、各単位でボリュームの領域を分割した場合に、各階層の I/O 比率(どの階層にどのくらいの I/O 数を集中させられるか)をシミュレーションし、算出する。

(2) プールの最大スループット性能算出

次に、(1)で求めた各階層の I/O 比率と下記計算式を用いて、プールの最大スループット性能を算出する。ここで、階層 i の I/O 比率を R_i (固定値)、階層 i が処理できる最大性能を P_i とし、 T を階層の集合とすると、 T 階層のプールの最大スループット性能は下記式で求められる。

$$\text{プールの最大スループット性能} = \min_{i \in T} \frac{P_i}{R_i} \dots (A)$$

4.2. 検証結果

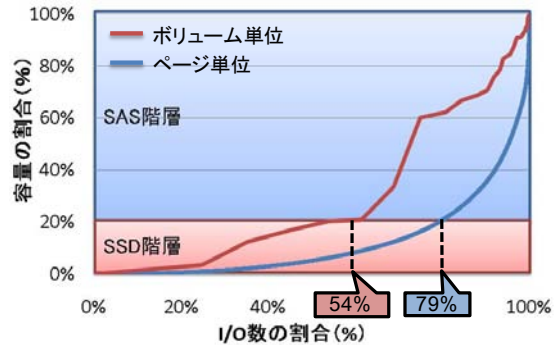
容量仮想化機能に関する研究によれば、ページサイズの上限值は、「128/3 MB」と定義している[1]。このため本検証では、ページサイズを 128/3 MB とした。一方、比較対象であるボリュームサイズおよびボリューム数は、Cello99 に基づき検証した。

また、本検証では簡単のため、SSD 階層と SAS 階層の 2 階層での性能を算出した。各階層の容量の割合は、図 1 左のストレージ全体のアクセス頻度の偏りより、SSD 階層を 20%、SAS 階層を 80% とした。

(1) 各階層に対する I/O 比率の算出結果

Cello99 の I/O アクセス分布に基づき、ボリューム単位およびページ単位の各階層への I/O 比率を算出した結果、ページ単位の階層管理では、20% の SSD 階層に 79% の I/O を集中させることが分かった。

一方、ボリューム単位の階層管理は、20% の SSD 領域に 54% の I/O を集中させるに留まることが分かった(図 3)。



階層管理単位	ボリューム単位	ページ単位
SSD階層のI/O比率	54%	79%
SAS階層のI/O比率	46%	21%

図 3 各階層に対する I/O 比率の測定結果

(2) プールの最大スループット性能算出結果

(1)で求めた I/O 比率を用いて管理単位毎に最大のスループット(相対値)を求めた。(A)式を用いると 2 階層でのプールの最大スループット性能は、下記式で求められる。

$$2\text{階層プールの最大スループット性能} = \min \left(\frac{P_1}{R_1}, \frac{P_2}{R_2} \right) \dots (B)$$

ここで、一般的に SSD の方が SAS より処理性能が非常に高いので、 $P_1 \gg P_2$ となることから、プールの最大スループット性能は、 P_2/R_2 となる。

これにより、各単位におけるプールの最大スループット性能を求めた結果、ページ単位での階層管理は、ボリューム単位よりも 2.2 倍スループット性能が高いことが分かった。

以上から、本機能によりコストパフォーマンスの向上が可能といえる。

5. まとめ

本研究では、ストレージ装置が自動的にページのアクセス頻度に応じて、適切なストレージ階層に配置する、ストレージ階層仮想化機能を開発した。また、本機能の効果を検証した結果、本機能はボリューム単位の階層管理よりも 2.2 倍性能向上できることが分かった。これにより、本機能がコストパフォーマンス向上可能であることを確認した。

今後は、ストレージ階層仮想化機能を導入および運用するに当たって必要となる、階層構成の決定方法を検討する予定である。

参考文献

- [1] L. Cherkasova, M. Gupta, "Characterizing locality, evolution, and life span of accesses in enterprise media server workloads", NOSSDAV'02, May 2002
- [2] 神吉琢磨, 山内敦広, "最適な運用を実現する日立ストレージ管理ソフトウェア", 日立評論, Vol.90 No.03 250-251, March 2008
- [3] 江口賢哲, "大規模ストレージシステムにおける動的容量割り当て(Dynamic Provisioning)機能の研究", FIT2008, C-019, September 2008

(*1) Hewlett-Packard社Hewlett-Packard Laboratoriesが提供するI/Oワークロードのトレースデータ (下記URL参照) <http://tesla.hpl.hp.com/opensource/>