

3B-3

ストライプの動的再編成を伴う RAID5 型ディスクアレイにおける  
ホットブロックとコールドブロックの分離法とその性能評価茂木 和彦 喜連川 優  
東京大学 生産技術研究所

## 1 はじめに

2次記憶装置の高性能化と高信頼化を目的とした RAID5 型ディスクアレイの開発が進められている。RAID5 型では、高信頼化のためにパリティを記録しており、データの更新時にはデータブロックに加えてパリティブロックも更新する必要があるため、性能低下が生じ問題となっている。

この性能低下を抑えるため、ストライプの動的再編成を行う記録管理方式として仮想ストライピングによる記憶管理方式と、LFS[1]に基づく記憶管理方式について検討を加えて来た[2, 3]。これらの記憶管理方式では、新たに書き込まれるデータはその場所を変え、新たに書き込まれるデータのみでパリティの計算を行うことによりパリティ更新の負荷を減らしている。更に、これらの方式では、書き込みを連続した空き領域に対して行うことにより性能の向上を図っている。そのため空き領域を作成するための処理(ガベージコレクション; GC)を行う必要があり、このオーバーヘッドの削減が性能向上の大きな鍵になる。

実際の負荷を考えると、アクセスのローカリティが存在すると考えられ、従来の高性能化手法においてはこの利用が考えられて来た。ストライプの動的再編成を行う記録管理方式においても、ローカリティの利用により GC の効率化を行い、性能の向上が図る事が可能であると考えられる。本稿では、ホットブロックとコールドブロックの分離することによりアクセスローカリティを利用可能とし、それによりシステムの性能を向上させる手法について述べる。そして、この手法を適用したときの性能についての評価を行う。

## 2 アクセスローカリティの利用

実際の負荷においては、すべてのデータに対して均等にアクセス要求が存在する訳ではない。磁気ディスク装置においては、ホットな(アクセス頻度が高い)データを1ヶ所にまとめて配置する事が可能であれば、その部分のみがアクセスされる可能性は非常に高く、デー

タの読み出しまでに必要なヘッドの移動時間を削減する事が出来る。このように、アクセスローカリティを利用する事により、性能の向上を図る事が出来る。

それに加えて、ストライプの動的再編成を行う記録管理方式においては、必須の動作である GC の効率化のためにアクセスローカリティを利用する事が可能である。ホットなデータは更新される可能性が高いため、現在記録されている場所を移動する可能性が高い。そのため、ホットなデータが1ヶ所にまとめて書き込まれると、しばらくするとその場所に集中してデータブロックが作成される事になる。GC 処理における1つの空きブロックの作成に必要なコストはデータブロックが多い程小さくなる。つまり、ホットブロックを1ヶ所にまとめ、多数のホットブロックが別の場所に移った後に GC を実行する事により、GC のコストを大きく削減する事が出来る。

そこで、上記の効果を得るために、ディスクをホットブロックの領域とコールドブロックの領域に物理的に2分割することを考える。分離された領域ではそれぞれ独立に管理を行い、書き込みや GC もそれぞれ独立に行うことにする。

## 3 ホットブロックとコールドブロックの分離法

実際には、どのようにしてホットブロックとコールドブロックの分離を行うかが大きな問題となる。90%のアクセスが10%のブロックに集中する(以下 90-10 と略記)ようなローカリティを考える時、書き込みアクセスのそのほとんどはホットデータの書き込みである。そこで、[1]で述べられている cost-benefit と同様に、書き込みアクセスはすべてホットであると見なす。しかし、これを続けていくとホット領域内に書き込みのための空き領域を作成することが不可能となる。そこで、GC 実行時にコールドであると推定されたブロックをコールド領域に移す動作を行い、ホットブロックとコールドブロックの分離を行う。書き戻しに際し、基本的には書き込みアクセスによりコールド側からホット側に移されたブロック数と同じ数だけコールド側に移し替える。このとき、コールドブロックの推定は以下のように行う。アクセス頻度が低いコールドブロックは最後にアクセスが行われてからの経過時間が長い

A separation method of hot blocks and cold blocks and its performance evaluation on RAID5 disk arrays with dynamic stripe reorganizations

Kazuhiko Mogi and Masaru Kitsuregawa  
Institute of Industrial Science, University of Tokyo  
7-22-1, Roppongi, Minato, Tokyo 106, Japan

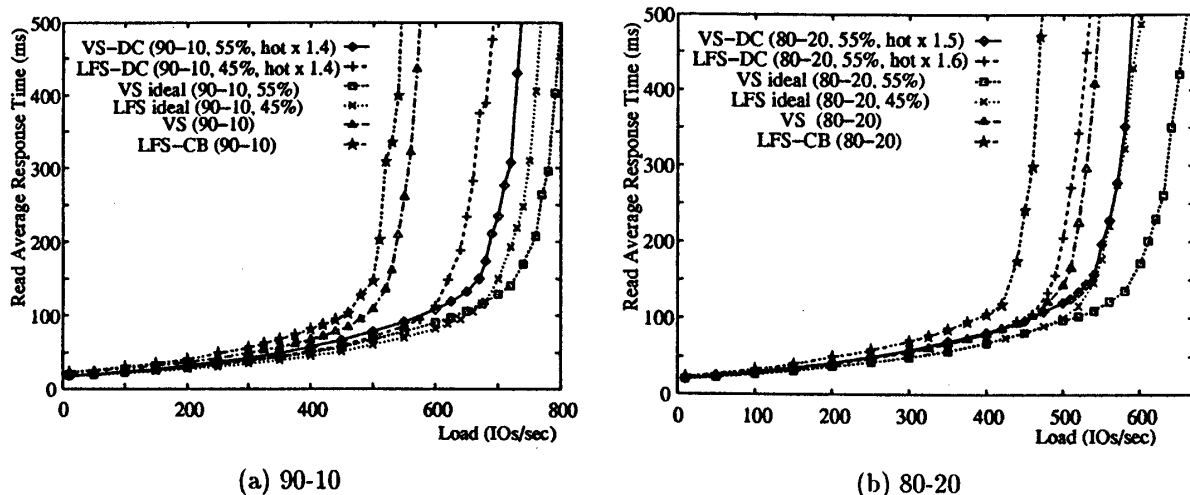


図 1: ホットブロックとコールドブロックの分離とその性能

と考えられる。そこで、すべてのホット側が存在するブロックに関しては、最後にアクセスが行われた時刻を記録しておく。(ただし、GC に関するアクセスは含めない。) この値から経過時間を求め、それが大きなものから順に移し替える分だけコールドであるとみなす。この推定を行う時、ホットブロックが誤ってコールドブロックであると推定される可能性が存在する。そのため、一部のコールドブロックはホット領域に存在するようにし、ホット領域内の経過時間が大きなコールドブロックの数を増やすことにより誤推定が行われないようにする。

#### 4 性能評価

上述の方法の有効性を調べるためにシミュレーションを行った。8D+P 構成に於いて、ディスク使用率が 85%、R:W = 1:1 の静的負荷を与えた場合の読み込みの平均レスポンスタイムを測定した。90-10 のローカリティが存在した時の結果を図 1(a) に、80-20 のローカリティが存在した時の結果を図 1(b) に示す。あらかじめわかっているブロックの属性を用いて理想的に分離を行った時と、分離を行わなかった時の性能も比較のために同図中に示した。空き領域の分配の割合は 5% 刻に、また、ホット領域中のコールドブロックの割合は総ホットブロック数の 10% 刻で計測し、最良であると思われるものを図に載せた。

ホットブロックとコールドブロックの分離の効果は大きく、90-10 のローカリティでは大幅な性能の向上が見られ、80-20 のさほど高くないローカリティの場合でも性能の向上が見られる。ただし、上述の分離法では理想的に分離が行われた時ほどの性能の向上は見られない。これは、ホット領域の GC 実行時にコールドブロックの移動処理を必要とすること、また、ホット領

域側に一定のコールドブロックが存在することによる GC 効率の低下によるものと考えられる。

#### 5 まとめ

パリティストライプの動的再編成を行なう記憶管理方式である、仮想ストライピングによる記憶管理方式と LFS に基づく記憶管理方式について、アクセスローカリティ存在時にホットブロックとコールドブロックの分離による性能の向上法を検討した。書き込まれるブロックはすべてホットブロックであるとみなして書き込み、GC 時にコールドブロックの分離を行うことにより、アクセスローカリティを利用することができ性能を向上させる事ができる。特に、90-10 のようなアクセスローカリティが存在する時には大幅に性能を向上させる事が出来る。

#### 参考文献

- [1] Mendel Rosenblum and John Ousterhout. The Design and Implementation of a Log-structured File System. In *Proc. of ACM SOSP*, pp. 1-15, October 1991.
- [2] Kazuhiko Mogi and Masaru Kitsuregawa. Dynamic Parity Stripe Reorganizations for RAID5 Disk Arrays. In *Proc. of PDIS*, pp. 17-26, September 1994.
- [3] 茂木和彦, 喜連川優. 仮想ストライピングを用いた RAID5 型ディスクアレイに於ける更新処理方式の性能解析. 第 48 回全国大会 3B-7. 情報処理学会, 1994 年 3 月.