

7K-1

RAID5 型ディスクアレイの  
二重化パリティスペアリング†福本 聡<sup>††</sup> 林 逸樹 中川 暉夫  
愛知工業大学 名古屋工業大学 愛知工業大学

## 1. まえがき

近年、ディスク装置の高速化、高信頼化を目的として、複数台のディスクによって並列冗長構成されたディスクアレイが注目されている。ディスクアレイは一般に RAID (Redundant Arrays of Inexpensive Disks) と呼ばれ、その主な記憶・構成方式として RAID1 から RAID5 までの5つが提案されている<sup>(1)</sup>。本研究では、トランザクション処理向けに最も実用化が進んでいる RAID5 の性能と障害回復処理時間について考察する。

RAID5 型ディスクアレイでは、データ更新時にパリティの再生成と更新を伴うため、ライトアクセスがリードモディファイライトとなり、性能低下の大きな要因となっている。このことを考慮し、性能を改善するための設計手法の一つとして、通常動作時にスペアディスクを活用する方法、スペアリングが提案されている<sup>(1),(2)</sup>。

本研究では、スペア領域を第二のパリティとしてアレイ内に分散し、二つのパリティの内容を一致させ、かつ、一時的に非同期、独立に更新することを可能とした方法、二重化パリティスペアリングについて検討する（図1参照）。

## 2. 性能シミュレーションモデル

はじめに、二重化パリティスペアリングによる RAID5 の性能改善について考察する。

二重化パリティスペアリングでは、パリティのアクセス競合の緩和による性能の改善が期待される（図1）。一方、この方法で全体としての障害回復機能を保つには、二つのパリティを相互に一致させる必要があり、そのための処理時間がオーバーヘッドとなる。

ここでは、モデル構築の柔軟性を考慮して、シミュレーションによる性能評価を行う。比較のため、二重化パリティスペアリングの他にスペアリングしない RAID5 および分散スペアリング<sup>(2),(3)</sup>の評価も考える。これらの各方法についてコマンド数 30000 のリード/ライトコマンドをランダムに発生させ、総実行時間からコマンドスループットを求め

る。コマンドキューイング容量は 10 と設定し、ライト比率は可変とする。

二重化パリティスペアリングにおいて、一時的に二つのパリティが独立に更新され、内容が不一致となったときのパリティ制御として以下の二つの規律を考える。

**[規律1]** 更新されていないパリティを含むディスクへの最優先コマンドとしてパリティ一致処理コマンドを発行する。

**[規律2]** 更新されていないパリティをただちに更新せず、後に続くコマンドによるディスクアクセスを優先して、その間隙を縫う形でパリティの更新を行う。

## 3. シミュレーション結果と考察

図2の規律1の場合から、パリティ一致処理優先の二重化パリティスペアリングでは、分散スペアリングによる性能の改善を上回ることはできないことがわかる。むしろ、ライト比率が高い範囲では、スペアリングしない場合よりも性能が低下している。この傾向は、ライト比率が高いほど著しいことから、ライトコマンド実行に伴うパリティ一致制御が性能低下の支配的要因になっているものと考えられる。

しかしながら、二重化パリティスペアリングによっても性能の改善は可能である。パリティ一致制御のオーバーヘッドを抑えることで、この方式についての性能が図2の規律2の場合に近づくからである。規律2の結果は分散スペアリングのそれをも上回っている。ただし P、これは最も望ましい場合、すなわち、パリティ一致処理が 100% 外部コマンドの実行に隠れる場合の結果である。

以上から、二重化パリティスペアリングを性能改善の方法として効果的に適用するためには、パリティのバッファリングやコマンドスケジューラの活用などによって、一致処理のオーバーヘッドをできるだけ減少させることが重要な課題となる。

## 4. 障害回復時間の評価

つぎに、二重化パリティスペアリングによる構成の障害回復処理時間について考察する。

二重化パリティスペアリングでは、故障ディスクに含まれていたパリティは、すべて他のいずれかの

†Dual Parity Sparing for RAID Level 5 Disk Arrays.

††Satoshi Fukumoto, Aichi Institute of Technology.

ディスクによって二重化されているため、それらの再生成は省略できる。ただしこれは、一致処理前に失われたパリティをコントローラ側のバッファで保障することを前提としている。図3は、ディスク数6台のときの二重化パリティスペアリングの回復処理の例を示している。この場合、処理時間が再生成データ量に単純に比例するものとして見積もれば、回復処理時間はスペアリングしない場合と較べて2/3、分散スペアリングと較べて4/5に短縮できる。一般に、 $N$ 台構成のRAID5では、スペアリングしない場合の $(N-2)/N$ 、分散スペアリングの場合の $(N-2)/(N-1)$ の回復処理時間となる。これらから、二重化パリティスペアリングは回復処理時間の点では極めて有効なスペアリング方法であることがわかる。

5. むすび

本研究では、RAID5型ディスクアレイのスペアリング方法について議論した。スベア領域を第二のパリティとしてアレイ内に分散し、二つのパリティの内容を一致させ、かつ、一時的に非同期、独立に更新することを可能とした方法について、性能と障害回復処理の観点から評価した。

文献

- (1) Patterson, D. A., Gibson, G. and Katz, R. H. :“A Case for Redundant Arrays of Inexpensive Disks (RAID)”, *Proc. of ACM SIGMOD*, pp. 109-116 (June 1988).
- (2) 喜達川優 :“最近の2次記憶装置:ディスクアレイ”, *情報処理学会誌*, Vol. 34, No. 5, pp. 642-651 (1993).

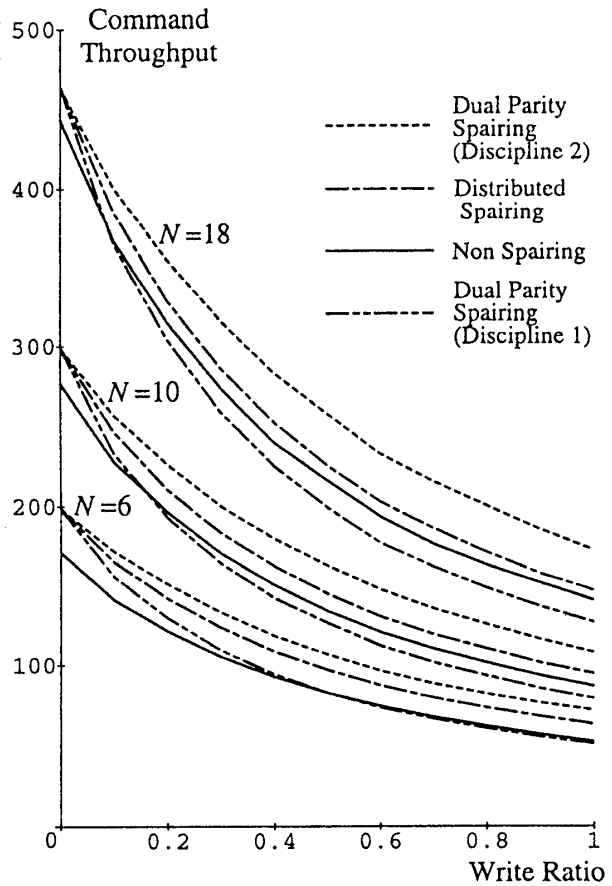


図2: コマンドスループットのシミュレーション結果。

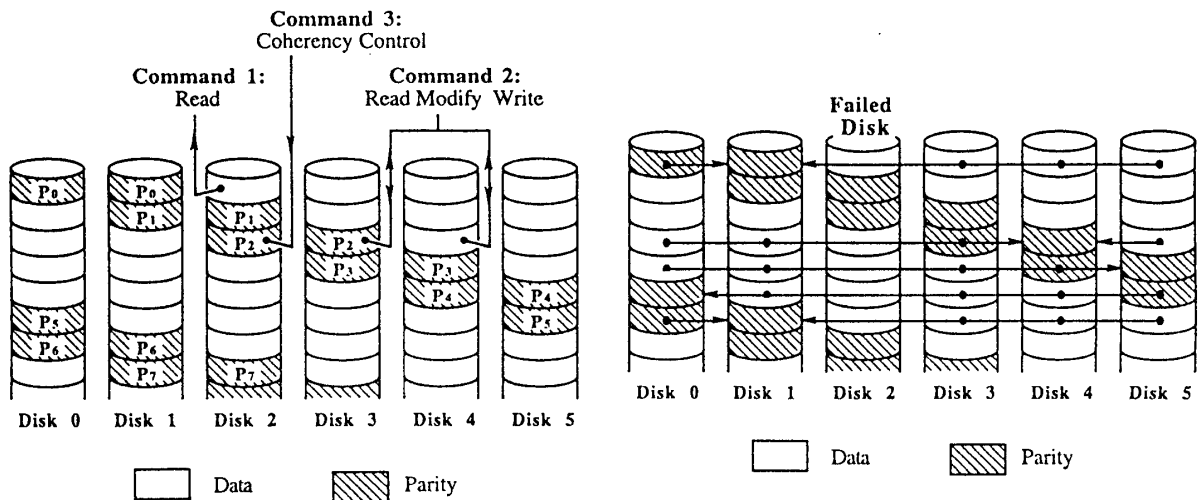


図1: 二重化パリティスペアリングの構成。

図3: 二重化パリティスペアリングの回復処理時のデータの流れ。