

## 7X-5

## アレイ制御ディスクの構成方式に関する一考察

桜井 紀彦 小谷 尚也 木ノ内 康夫

NTT情報通信処理研究所

## 1. はじめに

ディスクサブシステムの構成方式として、アレイ制御ディスクサブシステムが注目されている。これは、ディスクユニットを複数個並列動作させ、論理的には、大容量磁気ディスク装置が存在している様に見える構成方式であり、エラー訂正用ディスク等冗長構成による高信頼度化、並列転送による高速転送の特徴がある反面、高頻度ランダムアクセス性能に関しては、単体のディスクに比べて低下するという問題もある。

本稿では、この関係を、論理的に意味があるデータの配置方式の違いによる信頼性と性能・コスト間のトレードオフの問題として捉え、ハードの独立動作単位に対するデータ配置方式を見直すことにより、並列動作と単体動作を必要に応じて動的に使い分けるアレイ制御方式を提案している。

## 2. 信頼性と性能・コストのトレードオフの関係について

アレイ制御ディスクサブシステムと単体の磁気ディスクサブシステムの構成方式に関する大きな相違は、HDA等ハードの独立動作単位(かつ障害単位;物理単位と呼ぶ)に対して、論理的に意味があるデータブロック(アクセス単位と呼ぶ)が如何に配置されているかにあり、この配置方式によって、信頼性・性能等に関する特徴が異なってくる。

アクセス単位の物理単位に対する配置方式としては、以下の方法が考えられる。

- (1) アクセス単位自身を横断的に複数の物理単位に配置する。(静的アレイ方式と呼ぶ)
- (2) アクセス単位は同一の物理単位内に配置し、複数の物理単位に横断して、連続したアクセス単位を配置する。
- (3) 連続したアクセス単位を同一の物理単位に配置する。(現状最も一般的な使い方;単体方式と呼ぶ)

各方式のイメージ図を図1に示す。ここでは、この配置方式による信頼性と性能・コストのトレードオフの関係について述べる。

## 2.1 性能・コストとの関係

データベース処理等では、(a)フルサーチをかける様な大容量かつシーケンシャルアクセスと、(b)インデックスを経由してアクセスされる様な小容量ランダムアクセスが混在し、(b)については、高頻度アクセスに耐え得る性能(アクセス耐力)を要求される。

静的アレイ方式は、並列動作による高速転送で(a)の性能向上は期待できるが(b)については、大容量のディスクとして動くため、サブネットにより容量が使いきれずに、不経済となる場合がある。

単体方式では、(b)については、物理単位の性能を發揮できるものの、(a)の負荷が性能に大きく影響する。

一方、アクセス単位毎に横断的に配置する方式は、(a)のアクセスに対しては、並列同期による高速転送が可能であり、(b)のアクセスに対しては、必要な物理単位にのみアクセスが可能で、ランダムアクセス性能も低下しない。この方式は同期を取るか否かを、転送データ量等で動的に変更可能なことより、性能面では、静的アレイ方式と単体方式の両方の長所を活かした制御が可能である(動的アレイ方式と呼ぶ)。ただし、各物理単位が独立動作している状態から並列転送に移る時には、アームの整列同期をとる必要があり、約5割程度シーク時間が延びる。(図2参照)

## 2.2 信頼性との関係

静的アレイ方式では、異なる障害単位から情報を読み出しており、障害発生時には複数の物理単位に跨ったエラー訂正機能による自動訂正等が期待できる。(eg ECC-DK搭載等)

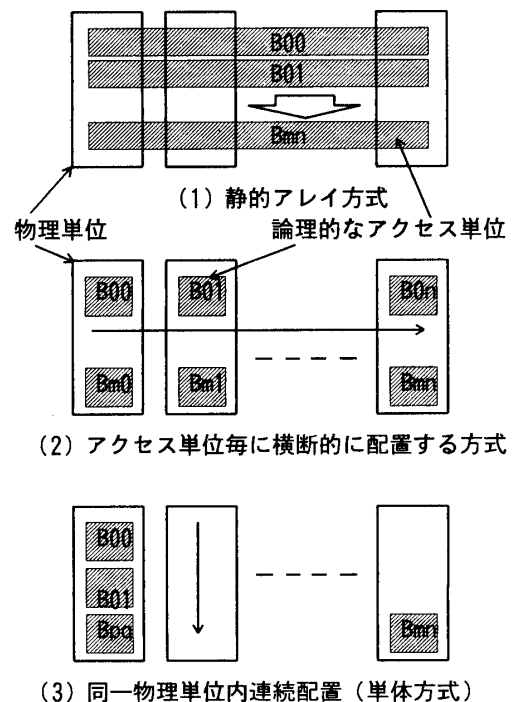


図1 物理単位に対するアクセス単位の配置の比較

単体方式では、アクセス単位が同一障害単位に格納されており、障害時にはこの情報に対するアクセスはできない。信頼性の向上のためには、二重化等 (egミラーDK) データ自身を冗長化する必要がある。このことは、単体方式の経済性を低下させる原因となる。

動的アレイ方式では、物理単位での独立動作で更新を行う場合には、複数物理単位に跨ったエラー訂正機能は活用できない。従って単体方式と同様データ自身の冗長化が必要である。ただし、read only のデータについては、並列転送時には、静的アレイ方式同様エラー訂正機能が活用できる

図3に性能の観点からみた各方式の特性の比較を、図4に信頼性の観点からみた各方式の比較を示す。

これらの図からわかる様に、動的アレイ方式は、性能面では、静的アレイ方式の大容量転送時の高速性と単体方式の小容量転送時の高ランダムアクセス耐力の両方の長所を持った特性を持つことがわかる。一方信頼性の面では静的アレイ方式には劣るものの、READ ONLY 時にはハード冗長構成による自動訂正を期待できる点では、改善がある。

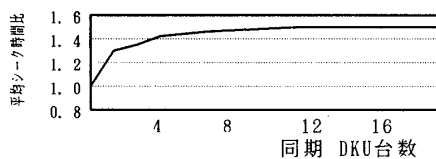


図2 整列同期時間

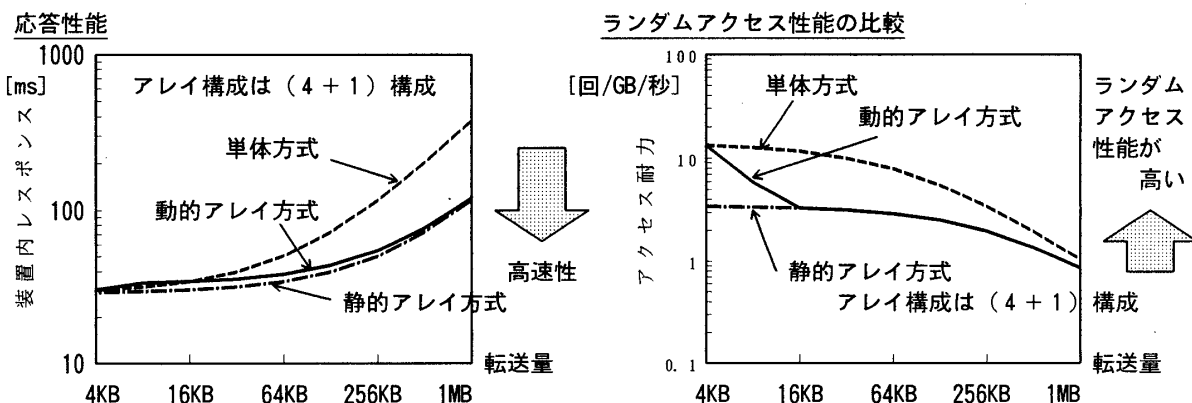


図3 性能の観点からの各方式の比較

### 3. 各方式のコスト性能比の評価

各方式について、信頼性に向上によるコスト増分を考慮し、同一性能を満足する条件でのコスト比較を行った結果を図5に示す。評価においては、4KB (アクセス単位) 転送と 256KB転送の割合を 3:1として混在させたケースを考え、スループット条件 (横軸アクセス頻度) およびレスポンスタイム条件を満足する範囲でコストが最小となる場合のコストを比較した。(\*) 信頼性条件については、単体方式および動的アレイ方式について、データの二重化を行った場合のコストについて検討した。図5より以下のことが言える。

- (1) データの二重化コスト増分を考慮しても、動的アレイ方式が有利な領域がある。(数回/GB/秒以上) これは静的アレイ方式では、小容量ランダムアクセスに対するスループットネックで、ファイル収容率の極端に低下、また、単体方式では、大容量転送の負荷が大きくパスネックが生じるためである。
- (2) 信頼性条件を考慮すると、低アクセス頻度域では、静的アレイ方式が有利となる。

### 4. まとめ

アレイ制御ディスクの構成方式について、独立に動作できる単位に対するデータ配置方式を見直すことにより、並列動作と単体動作を必要に応じて動的に使い分ける構成方式を提案し、コスト性能比の向上が可能であることを示した。

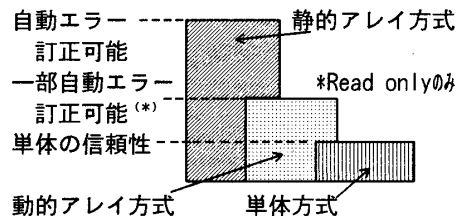


図4 信頼性レベルの比較

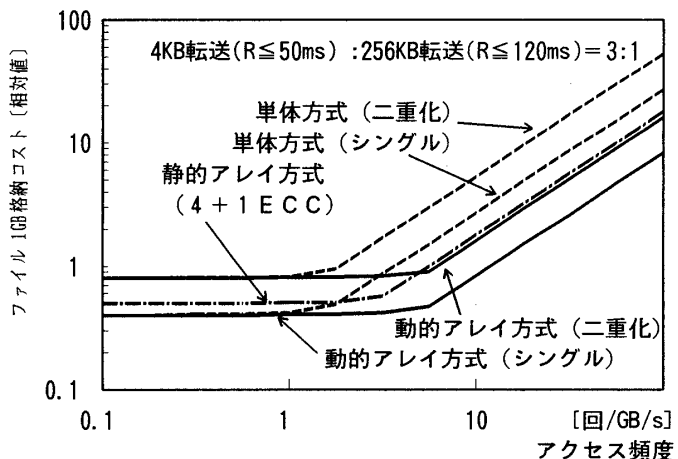


図5 信頼性を考慮したコスト比較

[参考文献]

(\*)1) 木ノ内・山口・桜井:オンラインファイル系記憶装置のコスト評価に関する一考察 :情報処理学会論文誌Vol127 no. 6(1986)