

RAIDシステムにおけるパリティ生成方式

2F-6

廣藤 進、笹本 享一、米山 正（（株）東芝 府中工場）

1. はじめに

近年、計算機の高速度が進むにつれて、大量のデータを高速に処理し、安全に保管する必要性が今までにも増して要求されている。

データ転送の性能と安全性を向上させる方式として、RAID (Redundant Arrays of Inexpensive Disks) 方式のディスクサブシステムが実現されている。RAIDでは、ハードディスク故障時に、データを復元するために必要なパリティの生成方式がシステムの性能向上の一つの鍵となる。

本報告では、RAIDレベル3から5において使用されるRAIDパリティの生成方式と、ディスクドライブ故障時のデータ復元方式を紹介する。

2. ディスクサブシステムの構成

図1に今回開発したディスクサブシステムのブロック図を示す。ホストとの接続ポートを4つ持ち、HDD接続用ポートとして6本のSCSIバスを備えている。また、内部バスとしてPCIバスを用いており、大容量ディスクキャッシュが接続される。このキャッシュにおいて、RAIDパリティを生成させる。また、本システムは、コントローラ、電源の2重化を行い、3.5型ハードディスクを12台実装することができる。ホストからのアクセスは、SCSIを用いており、ホストからは、1台の大容量ディスクとして使用することができる。

装置の大きさは、高さ220mm、幅440mm、奥行き550mmである。外観を図2に示す

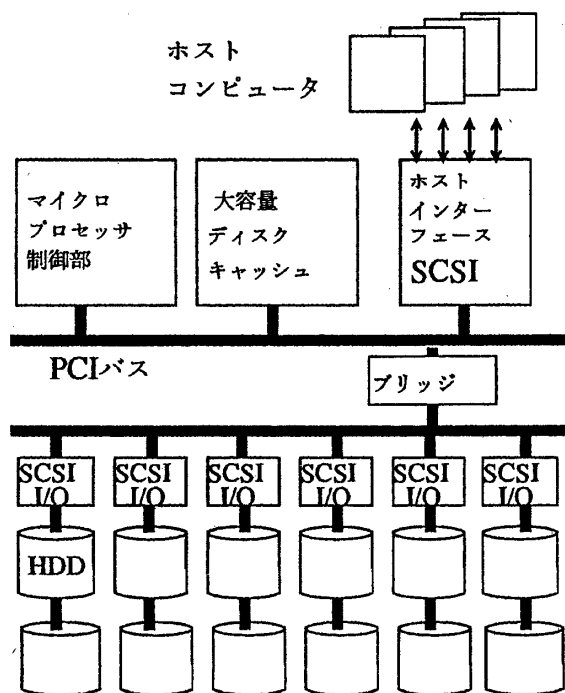


図1 システムブロック図

3. ディスクキャッシュ

今回開発したディスクキャッシュの構成を図3に示す。基本構成は、メモリ、PCIバスとのブリッジ、パリティ生成で使用するFIFOメモリ、キャッシュ全体の制御を行うコントローラ

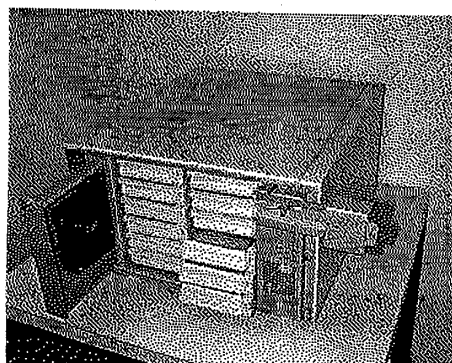


図2 装置外観

Parity Generating Method
for RAID System
S.Hirofujii, K.Sasamoto, T.Yoneyama
TOSHIBA CORPORATION

で構成されている。このディスクキャッシュ上のデータは、制御部の制御プログラムによって管理されており、RAIDパリティの生成、およびディスク故障時のデータの復元を行えるようになっている。

ディスクのデータとディスクキャッシュのメモリ上のデータ配置の関係を図4に示す。通常、ディスクキャッシュのコントローラは、PCIバスから受け取るアドレスに対応した、メモリ上のデータをそのまま出力する。しかし、PCIバスからパリティ生成の要求を受け取ると、コントローラは、以下の手順でパリティデータを生成し、出力する。

- (1) HDD1のデータをFIFOメモリに記憶する
- (2) HDD2のデータとFIFOメモリに記憶されたデータとの排他的論理和をとって結果をFIFOメモリに記憶する。
- (3) HDD3のデータとFIFOメモリに記憶されたデータとの排他的論理和をとってFIFOメモリに記憶する。
- (4) HDD4のデータとFIFOメモリに記憶されたデータとの排他的論理和をとって、結果をPCIバスへ出力する。

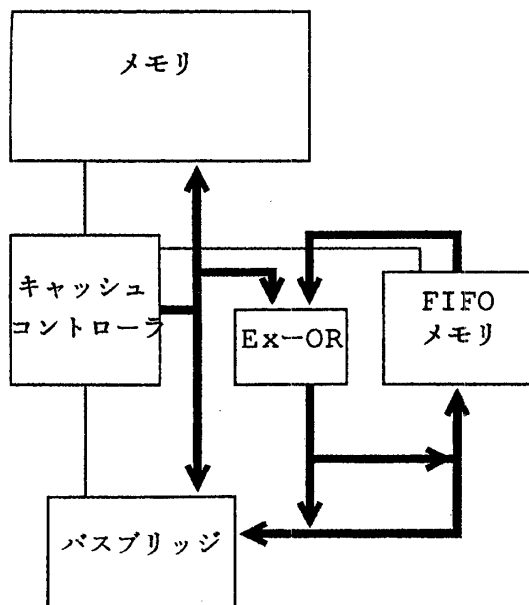


図3 キャッシュメモリ ブロック図

また、故障ディスク発生の際には、故障したディスクのキャッシュメモリの領域にパリティデータを格納し、パリティ生成と同じ操作を行う事によって、故障したディスクのデータを復元することができる。

以上のようにRAIDパリティの生成機能を、ディスクキャッシュの機能の一部として取り込むことによって、ハードウェアの構成を簡潔にし、かつ制御も容易となった。

4. まとめ

今回紹介したディスクキャッシュの構成は、排他的論理和によってパリティを生成するRAIDシステム全てにわたって有効である。

今後、パリティ生成のためにキャッシュを余計にアクセスしなければならないオーバーヘッドをどのようにクリアするかが、課題となる。

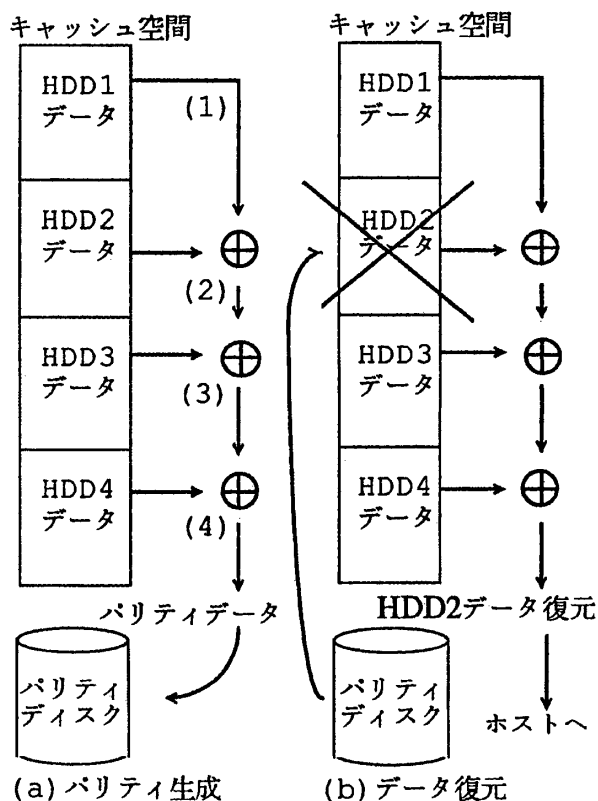


図4 パリティ生成とデータ復元