

3B-8

# ミッドレンジビジネスコンピュータ用ディスクアレイの 高信頼化機能及び性能

鈴木和雅† 峯村治実† 中村俊一郎† 桑田圭三‡ 小林清志‡ 木岐将之‡

†三菱電機情報システム研究所 ‡三菱電機東部コンピュータシステム

## 1 はじめに

ミッドレンジビジネスコンピュータ用ディスクアレイ (RAID) の開発を行なった。このディスクアレイでは RAID レベル 5 の技術を採用しており、たとえ冗長グループ内の 1 台のディスクが故障してもグループ内の他のディスクからデータを自動復旧することにより、システムの稼働をそのまま続行させることができる。また、スタンバイディスクへの復元、ディスクの活線挿抜も可能であり高い信頼性を有している。そして、データストライピング処理、旧データ・旧パリティのキャッシュング [1]、コマンドキューイング等により高性能化も図っている。本稿ではこのディスクアレイの高い信頼性及び性能について述べる。

## 2 高信頼化機能

本システムでは RAID 5 の技術を採用することにより、冗長グループ内の 1 台のディスクが故障してもそのまま業務を継続させることができる (図 1)。これは、故障したディスクの情報を、パリティを用いて他のディスクから計算させて行なうことによる (縮退運転)。また、スタンバイディスクを定義しておくことにより、グループ内のディスクが故障した際、故障したディスクの内容をスタンバイディスクに復元させることができる。設定により、故障を検出した際自動的に復元を開始させること (自動復元) も、故障した後ホストからのコマンドにより復元を開始させること (手動復元) もできる。また、故障したディスクを新しいディスクに交換した後、この新しいディスクに故障したディスクの内容を冗長グループ内の他のディスクのデータから回復させることもできる (回復)。これら復元・回復処理は、通常の業務

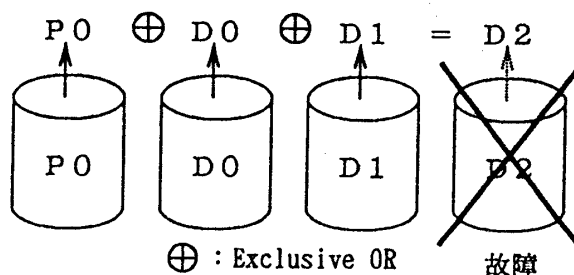


図1 故障ディスクのデータの再生

のバックグラウンドで行なうことができ、故障したディスクの交換も、業務を継続させたままで行なうことができる (活線挿抜機能)。

本システムの筐体内には、無停電電源 (UPS) が備えてあり、停電等で電源の供給が途絶えても処理を継続させることができる。この UPS は、フル充電時には約 5 分の電源の供給が可能であり、UPS からの電源供給が一定の時間経過した場合には、ホスト側でシャットダウンプロセスがスタートするようになっている。また、ディスクアレイ制御装置やホストの故障に備え、二重系システムの構築が可能となっている。

## 3 性能

RAID 5 では、通常データと同時にパリティも同時に書きこまなくてはならない。そのため、1 回のライト処理に対し、旧パリティ・旧データの読みだし、新パリティ・新データの書き込みの 4 回のディスクアクセスが必要となり、一般に RAID 5 のライト処理は遅くなる。そのため、本システムでは旧パリティ・旧データ用のキャッシュ (RAID キャッシュ) を設けている (図 2)。

The High Reliance Faculty and Performance of a Disk Array System for Mid-range Business Computers  
Kazumasa Suzuki†, Harumi Minemura†, Shunichiro Nakamura†, Keizo Kuwada†, Kiyoshi Kobayashi†,  
Masayuki Kiki†

†Computer and Information Systems Laboratory, Mitsubishi Electric Corporation

‡Mitsubishi Electric Computer Systems

ホストがあるデータをディスクから読みだし (図 2 (1))、そのデータの内容を変更した場合に、そのデータを再びディスクに書き込む可能性は非常に高い。そこで、読み込んだデータの内容を変更した場合には、ホスト側では変更する前のデータ (旧データ) を専用のキャッシュに保持し、それと同時にディスクアレイ制御装置に対して対応するパリティ (旧パリティ) の読みだしの命令を発行する (図 2 (2))。この旧パリティは、制御装置内の専用のキャッシュ領域に保持され、実際にホストからそのデータの書き込み命令が来た場合には、旧データ・旧パリティをディスクから読み込む代わりに、これらあらかじめキャッシングしておいたデータを用いる (図 2 (3))。これにより、RAID 5 のライト処理の遅さを補っている。

その他、本システムでは、データを複数のディスクに分割配置し複数のディスクへの同時アクセスが可能となるデータストライピング方式や、コマンドキューイング機能、従来ディスク制御装置ではチャンネルを介してホストと接続されていたのを、ホストの内部バスに直結するような形にしている等性能面での配慮がなされている。性能については、実行するジョブの内容にかなり依存するところが多いため、従来のディスク装置との単純な比較はできない。しかし、一般的にリード処理は得意であり、ライト処理は多少苦手である。また、ディスクへの I/O がシリアルに来るようなジョブでは、RAID の特性を生かすことができない。図 3 は、従来型のディスクシステムとの比較を行なったものである。図中のジョブは、それぞれ以下のようなものである。

ジョブ A … 主に 2KB から 10KB の write 比が約 38% の r/w を行なうバッチ処理

ジョブ B … 主に 2KB から 6KB の write 比が約 66% の r/w を行なうバッチ処理

ジョブ C … 2KB のランダムリードを繰り返す処理 (10 多重)

図中ジョブ C は、RAID 向きのジョブであり、ジョブの内容によっては、これほどの差がでる。

さらに、正規のシステムに対し、

- ・制御装置の内部クロックの高速化
- ・MPU を高速なものに変更 (68040→68060)
- ・MPU に対するバスモードの変更

等を実験的に行ない、性能の考察を行なった。これらの変更により、例えば 2KB のリードを繰り返

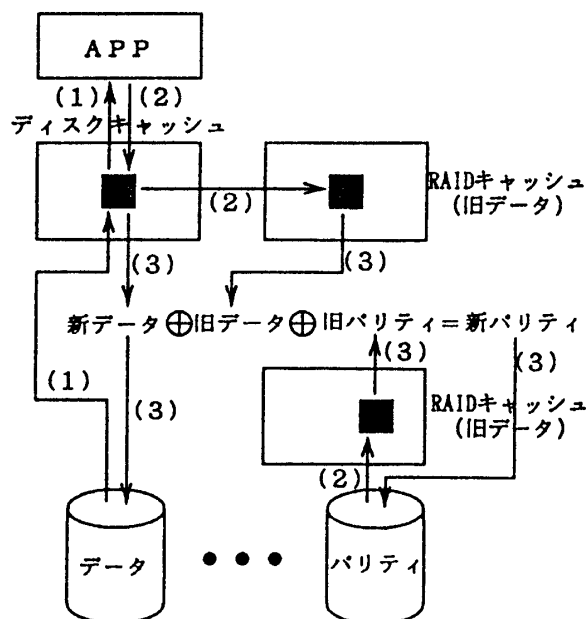


図 2 RAID キャッシュ

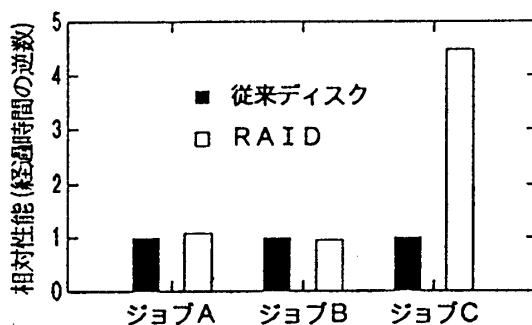


図 3 従来ディスクを基準にしたときの相対性能

す処理では約 1.5 倍以上の性能向上がみられ、1058 IOPS 以上のスループットが出ることで測定により確認された。この値は、ディスクキャッシュを用いないシステムとしては、最高水準のレベルを達成している [2]。

#### 4 おわりに

以上、本システムの高信頼化機能及び性能について簡単に述べた。今後性能面では、ポトルネットとなる要素をどのように解決していくかが課題となる。

#### 参考文献

[1] 早川孝之他：RAID レベル 5 のためのバッファ管理方式とその性能評価，電子情報通信学会技報，Vol.93, No.251, pp.51-57

[2] 百瀬晴男他：F6401A ディスクアレイ装置，FUJITSU.45,6, pp.534-540(11,1994)