

ミッドレンジビジネスコンピュータ用 ディスクアレイのアーキテクチャ

3B-5

峯村 治実 中村 俊一郎 吉村 啓二 吉森 幹夫 早川 孝之 鹿島 理華

三菱電機情報システム研究所

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^[1]の開発を行った。本論文では、このディスクアレイのアーキテクチャ（全体構成、基本制御方式など）、および優位性について述べる。以下、および後続の論文^[2, 3, 4]で述べるように、本ディスクアレイは信頼性・性能・容量において高レベルを達成している。

2. 特長

本ディスクアレイの特長を以下に示す。*は、本ディスクアレイ独自の機能/仕様である。

(1) 高信頼性

- ・RAIDレベル5（パリティ）により、ディスク1台故障時も継続運転可能（縮退運転）
- ・スタンバイディスクへのオンライン（業務継続中の）自動/手動復元
- ・システム動作中にオンラインで故障ディスクの活線挿抜による交換が可能
- ・無停電電源（UPS）
- ・二重系

(2) 大容量

*1 制御装置（30cm四方のカード1枚）あたり49台のディスクを接続可能

(3) 高性能

- ・コマンドキューイング（1論理ユニットあたり240個までのコマンドを並列実行可能）による高スループット
- ・高度なF/Wによる最適化制御

*RAIDキャッシュ（ライト処理性能低下防止）^[5]

*専用LSI（ホストI/F, EOR演算）の使用

3. 全体構成

図1に本システムの全体構成（ディスク49台接続時）を示す。本ディスクアレイでは、SCSI IDが同じ7台のディスク（図1の点線で囲んだ部分）を

RAIDレベル5の1組のグループとして使用する。ホストからは、この7台のディスクを1つの論理ユニットとしてアクセスすることになる。ただし、7台のうち1台をスタンバイディスクとして使用する6+1台の構成定義も可能である。スタンバイディスクは制御装置あたり2台まで定義することができる。また、スタンバイディスクは複数の論理ユニットで共有して定義することも可能である。

4. 基本制御方式

4.1 データ転送方式

本ディスクアレイでは、ホストとディスクとの間で直接データを転送するのではなく、以下のように一度内部のバッファ（SRAM）を経由してデータ転送を行うようにしている。

(1) リード時のデータの流れ

- ① ディスク→SIOP→SRAM
- ② SRAM→ホストI/F→主記憶

(2) ライト時のデータの流れ

- ① 主記憶→ホストI/F→SRAM
- ② 排他的論理和（EOR）演算
- ③ SRAM→SIOP→ディスク

複数のリード/ライトを同時に実行する時、例えば(1)①と(2)①のように並列に実行できる処理があるためスループットが向上する。また、ディスク側の処理をホスト側と切り離すことにより、複数のディスクからの転送を同時に実行できるため、ディスクの回転待ちが減り、平均応答時間改善の効果もある。

4.2 制御用F/Wの構成

縮退、復元等、RAIDレベル5の処理は複雑なため、本ディスクアレイ制御用のF/Wは、通常のディスク制御装置のF/Wに比べて、かなり大規模（数十Kステップ）なものになっている。図2にF/Wの構成を示す。以下、F/Wの各構成要素について説明する。

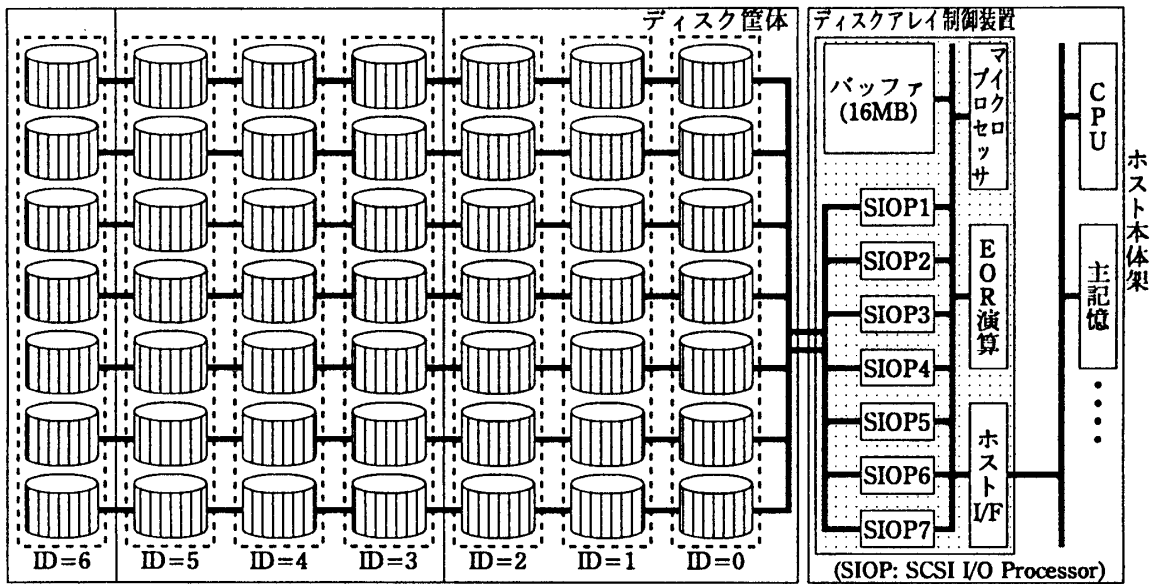


図1 全体構成

- ・ LIT (L-bus Interface Task) … コマンド、およびステータスの送受信を行う。
- ・ CST (Command Scheduler Task) … 論理ユニット単位の排他制御、コマンドスケジューリング等を行う。
- ・ CET (Command Execution Task) … ホストから指令された各I/O処理を実行する。
- ・ DST (Disk Scheduler Task) … ディスク装置毎にシーク距離の最適化（セクタアドレスでのソート）等を行う。
- ・ ドライバ … 専用LSI（ホストI/F、EOR演算）とSIOPのH/W制御を行う。

- [2] 小林剛 他, 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速化機構, 情報処理学会第50回全国大会
- [3] 山口智久 他, 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二重系制御方式, 情報処理学会第50回全国大会
- [4] 鈴木和雅 他, 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, 情報処理学会第50回全国大会
- [5] 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

5. おわりに

以上、我々が開発した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アーキテクチャについて述べた。このディスクアレイにより、ビジネスコンピュータ用の高信頼・大容量・高性能のストレージシステムの構築が可能となる。

最後に、本講演原稿作成にあたり多くの有益なご助言を頂いた静岡大学工学部情報知識工学科渡辺尚助教授に感謝いたします。

参考文献

[1] D.A.Patterson他, A Case for Redundant Arrays of Inexpensive Disks (RAID), Proc. of ACM SIGMOD, pp.109-116, 1988

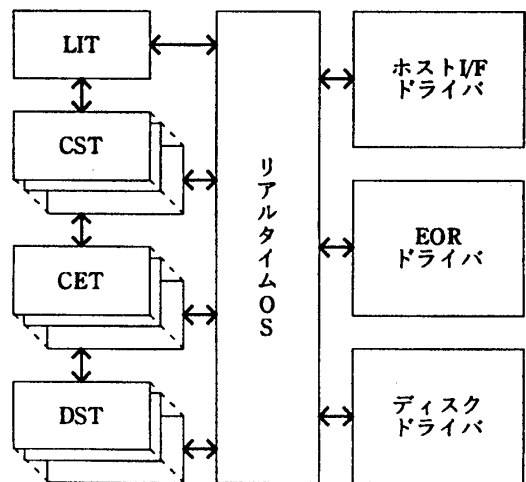


図2 制御用F/Wの構成