

キャッシュ付き磁気ディスク制御装置の マルチプロセッサ制御方式

7N-2

佐藤孝夫* 北嶋弘行* 神林公咲**

*(株)日立製作所システム開発研究所

**同小田原工場

1. はじめに

計算機システムに対する高速化のニーズに応えるため、磁気ディスクサブシステムでは、その制御装置内に半導体のキャッシュメモリを設けることにより、データアクセスの高速化を図っている。本報告で対象とする磁気ディスク制御装置は、キャッシュ機能として、従来のリード要求に対するキャッシュ機能に加え以下の機能を持つ。

・高速ライト機能

上位装置からのライトデータをキャッシュメモリとバッテリー電源によりバックアップされた半導体メモリの両方に記録する。これにより、ライトデータの信頼性を確保しつつキャッシュメモリによる入出力処理の高速化をライト処理に拡張することができる。

・自動2重書き機能

特に信頼性が要求されるデータを、2つの磁気ディスク装置に自動的に書き込む。

上記機能を持つ磁気ディスク制御装置のプロセッサ構成と、その制御方式を開発したので以下に報告する。

2. 磁気ディスクサブシステムにおけるデータ転送形態

図1に高速ライト機能、および自動2重書き機能を備えた磁気ディスクサブシステムにおけるデータ転送形態を示す。この図に示したデータ転送形態を大別すると以下のようになる。

(1) CACHE $\leftrightarrow$ HOST

キャッシュヒット時の上位装置とのデータ転送処理である。

(2) HOST $\leftrightarrow$ DKU

キャッシュミス時の上位装置とのスルー転送処理である。

(3) CACHE $\leftrightarrow$ DKU

キャッシュメモリと磁気ディスク装置間のデータ転送処理であり、具体的には以下を示す。

- ・高速ライト、自動2重書き機能におけるキャッシュメモリ上のデータの磁気ディスク装置への反映処理
- ・キャッシュヒット率向上のための先読み処理

ここで、(1)および(2)のデータ転送形態は、上位

装置からの入出力要求に同期して実行される処理であり、(3)のデータ転送形態は、上位装置からの入出力要求とは、非同期に実行される処理である。この同期/非同期処理の並列実行を効率良く行うことが磁気ディスク装置の性能上の課題である。

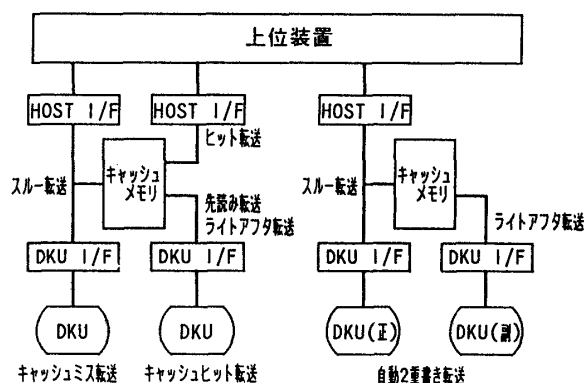


図1. データ転送形態

3. マルチプロセッサ構成方式とその制御方式

本報告の磁気ディスク制御装置は、大型計算機用であり、特に、高性能、高信頼なものでなければならない。以下に、これらの観点から検討し、採用したマルチプロセッサ制御方式と、その制御方式について報告する。

3.1 マルチプロセッサ構成方式

先に示したように、本報告の磁気ディスク制御装置は、上位装置からの入出力処理要求に同期した処理と、非同期な処理が混在する。そこで、上位装置とCACHE間のデータ転送を制御するHOST側プロセッサ(MP1)と磁気ディスク装置とCACHE間のデータ転送を制御するDKU側プロセッサ(MP2)をデータ転送路対応に独立に配置するマルチプロセッサ構成とした。上位装置と磁気ディスク装置間のデータ転送であるスルー処理は、MP1とMP2を物理的に結合することにより実行され、上位装置とキャッシュメモリ間のデータ転送であるヒット処理は、MP1のみで実行する。また、上位装置からの入出力要求とは非同期に実行される先読み処理、およびライトアプタ処理は、MP2で実行する。

MP1とMP2の台数は、必ずしも同数である必要がなく、キャッシュメモリのヒット率が高い場合は、MP1の台数を増やすことにより、スループットを向上することができる。

3. 2マルチプロセッサ制御方式

次に、マルチプロセッサ制御方式を示す。

(1) マスタレス型協調制御方式

上記マルチプロセッサの制御方式として以下の2つの方式が考えられた。

(a) マスタ/スレーブ型制御 (図2)

マスタプロセッサMPOを設け、MP1、MP2の実行の制御は、全てこのMPOが制御する方式

(b) マスタレス型協調制御 (図3)

全てのプロセッサから共通にアクセス可能な制御メモリを設け、この制御メモリの内容に基づいて各プロセッサが自律的に処理を実行する方式

(a)の場合、MPOの信頼性がボトルネックとなる。このため、MPOを複数台設けることが考えられるが、制御上複雑なものになる。これに対し、(b)の場合、制御メモリの信頼性が問題となるが、メモリを2重に設けることにより比較的容易に信頼性を高めることができる。また、性能の観点より両者を比較すると、(a)の場合、MPOが単一資源であるのに対し、(b)の場合、制御メモリを分割管理することにより、複数のプロセッサからの同時アクセスが可能となり、性能面で拡張性が高い。以上から、(b)のマスタレス型協調制御方式を採用した。

(2) 制御メモリによるジョブ管理

マスタレス型協調制御方式のマルチプロセッサのソフトウェア制御方式として、上位装置からのデータ転送要求処理や、キャッシュメモリへの先読み処理といった制御装置内で処理される全ての仕事をジョブとして制御メモリ上で管理する。各プロセッサは、自らが

実行可能なジョブを制御メモリから取り出して実行する方式とした[1]。

(3) 制御メモリによる資源一元管理

ジョブを実行するために必要な資源の管理情報を制御メモリ上に配置し、各資源単位にハードウェアによるロック機構を制御メモリに設けた。これにより、高速なプロセッサ間の排他制御を可能とした。

(4) 仮想デバイス方式による資源管理

HOST側プロセッサとDKU側プロセッサの独立動作を可能とするため、1台の磁気ディスク装置資源を論理デバイスと物理デバイスに分割し、HOST側プロセッサが確保するデバイス資源を論理デバイスとし、DKU側プロセッサが確保するデバイス資源を物理デバイスとした。

以上により、MP1、MP2は、各々独立した動作が可能となり、磁気ディスクサブシステムの性能向上が可能となる。具体的には、上位装置とキャッシュメモリ間のデータの転送処理と並行し、キャッシュメモリと磁気ディスク装置間のデータの先読み処理、あるいは反映処理が実行でき、データ転送路の有効利用が図れ、高スループットが期待できる。また、MP1と、MP2の接続を柔軟な構造とすることにより、物理的なデータ転送路の組み合わせ数を増やすことができ、性能および信頼性が向上する[2]。

4. おわりに

キャッシュメモリ付き磁気ディスク制御装置のプロセッサ構成方式と、その制御方式を示した。

5. 参考文献

- [1] 横畑他：磁気ディスク制御装置用マルチプロセッサ対応実時間OS，情報処理学会第41回全国大会
- [2] 猪股他：キャッシュ付き磁気ディスク制御装置におけるマルチプロセッサ制御方式の性能評価，情報処理学会第41回全国大会

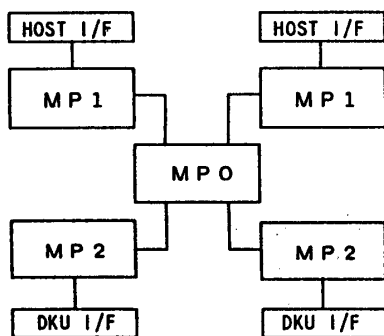


図2. マスタ/スレーブ型

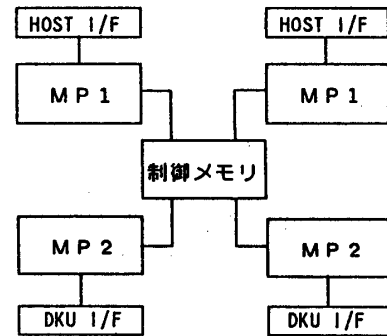


図3. マスタレス型