

パーソナルコンピュータ用 ディスクアレイ (MD-2) の動作評価

9 L-1

杉本 欽一 瀬川 哲 辻澤 隆彦
NEC 機能エレクトロニクス研究所

1 はじめに

パーソナルコンピュータをはじめ小型コンピュータシステムの処理速度の高速化には目覚ましいものがあるが、それに比較して、ファイル装置の入出力性能は不十分であるといわざるを得ない。そのため、ファイル装置の効率的使用に際してはハードウェア上のアクセスの仕様やソフト的なアクセス速度の向上策が必要不可欠となっている。本報告では、小型コンピュータシステムの入出力の特徴をのべ、その転送性能の向上のためのハードウェア機構として、筆者等が開発したディスクアレイサブシステムを例に、その制御方式の評価結果をのべる。

2 ファイルシステムの特徴

現在の小型コンピュータシステムで使用されているディスクファイル装置の管理は、BIOSあるいはファームウェアと呼ばれるハードウェアに依存した制御ルーチンと、それを使用しハードウェアの管理を行うオペレーティングシステムにより行われている。このような場合に、ハードウェアの性能はそのアクセスを管理するハードウェア及びファームウェアのハードウェア上の転送性能と、オペレーティングシステムのバス使用効率でシステムとしてのパフォーマンスが決まる。

小型のコンピュータシステムにおいては、ディスクに対し同期入出力か非同期入出力かが、アプリケーションレベルでの性能を左右するケースが増加している。DOSのようなシングルタスクのOSでは、同期I/Oが一般的である。よって通常データの処理の際にアプリケーションレベルで同期を取る必要なく並行処理が可能であってもディスクアクセスの期間中待つことになり、処理速度が遅く見えることになる。一方UNIXなどの非同期I/Oを実現しているOSにおいては、ディスクに対する処理を行う期間でもバスを使用する必要がない期間に、ファイルアクセスを管理するドライバの処理を休止し他のプロセスの処理時間に回すことが可能となる。それによりトータルの処理時間を、みかけ上高速化することが可能となる。更に近年一般化されつつあるI/O専用プロセッサを使用した入出力システムの場合はさらにドライバのハードウェア制御処理の大部分を周辺装置側に任せるこ

とが可能となるため、特に非同期I/Oのシステムにおいてはその性能の改善が顕著にあらわれる。

このように現在では小型コンピュータにおいてもCPUの処理速度の向上にあわせ周辺装置のI/O処理に関して機能分担がすすみ、OSレベルでのバスなどのハードウェア資産の有効利用が重要となっている。

3 MD-2の構成

図1のような構成をとっているが、このうち今回開発を行った部分は、NESAバスインタフェースボードおよび分配統合ユニット部である。図1から判るようにCPUをのせたインテリジェントインタフェースとなっており、I/O専用プロセッサとしての機能を実現している。

本システムのディスクバッファ領域は、ローカルCPUであるV53のメモリ空間にマップされ、ホストCPUからはNESAバスインタフェースのI/O空間にマッピングされた

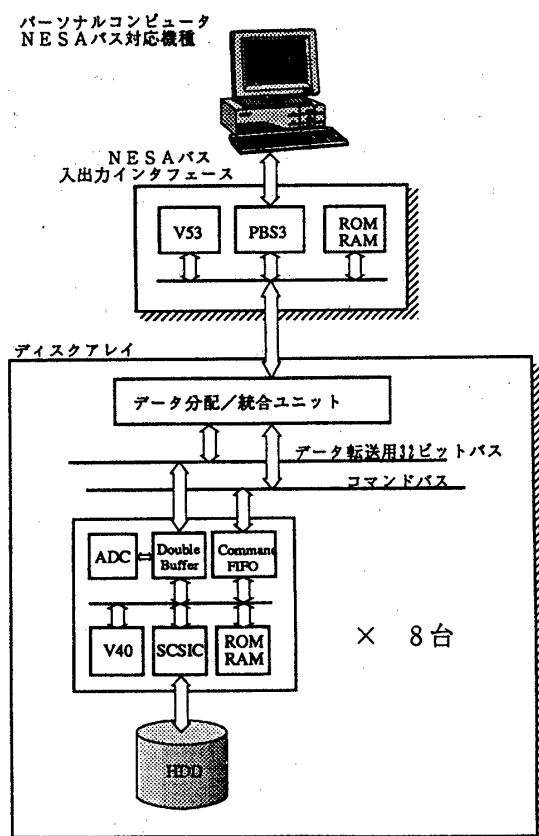


図1 MD-2のハードウェア構成

An Evaluation of A Disk-Array Subsystem(MD-2) for Personal Computers.

Kinichi Sugimoto, Satoshi Segawa, Takahiko Tsujisawa
Functional Devices Research Laboratories, NEC Corporation

インタフェースゲートアレイ (PBS3) を介してアクセスされる。ディスクの並列運転に伴うディスクとバッファとの間の転送速度は、6MB/Sec 以上 (バースト速度) と見込まれる。またバッファメモリとCPUとの間の転送はPBS3に内蔵するクロック同期式16ビット×16段FIFOを介して、NE SAバス上のメモリ空間にDMA転送される。

本システムでは、インタフェース上のローカルCPUにより、制御方式の変更が可能となっているため、OSのハードウェア制御が同期I/Oあるいは非同期I/Oのアクセスいずれにも対応可能である。

4 動作評価

4.1 バスインタフェースのDMA転送効率

NE SAバスDMAの連続転送 (ダイヤモンド転送) 時の転送速度に関して転送速度の評価を行った。ダイヤモンド転送回数に関しては、4, 8, 12, 16回のそれぞれの場合に関して計測した。その結果2.4MB/Sec~3.2MB/Secの転送速度が得られた。またこのNE SAバスの16ビットDMAの理論上の最大転送速度 (4MB/Sec) よりバスの利用効率を割り出すと、図2のように最大で80%を達成している。

このような高いバス使用効率を実現しており、ハードウェアの制御ルーチンはほとんどオプションボード上のCPUで行うため、非同期I/Oと組み合わせたハードウェア資源の高効率使用が可能であることがわかる。

4.2 ダブルバッファ制御

今回のディスクアレイはダブルバッファ機構を介して、アクセスするような形態を探っている。これにより、転送途中にデータの不連続箇所が存在しシーク動作及びアクセス動作を伴うような転送においても、その影響を最小限に押さえることが可能である。

このようなダブルバッファ制御では、ディスクのアクセス時間 T_a 、ディスク (SCSI) のデータ転送速度 V_d 、ホストインタフェースの連続読みだし速度 V_h の間に、次の関係が満足されるときに、最大性能を実現可能である。

$$T_a + L_t/V_d < L_t/V_h \quad (V_h < F_d)$$

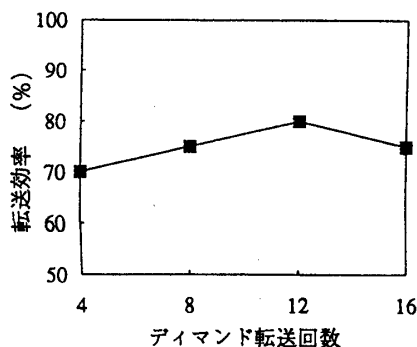


図2 DMA転送効率

ここで L_t は転送ブロック長さである。

表1 ディスクアレイ各部の転送速度

処理	転送速度
V53のDMA転送速度	5.0 MB/Sec (Max)
SCSIディスクの転送速度)	1.2 MB/Sec × 8 台
ディスクの常用転送速度(V_d)	6.0MB/Sec
NE SAバスインタフェースの実効データ転送速度(V_h)	3.2 MB/Sec (Max)

ここでアクセス時間を仮に $T_a=10\text{ms}$ と仮定すれば表1より

$$L_t > 69 \text{ kB}$$

となり、約70 kB以上のバッファ容量が必要となる計算となる。本ディスクアレイではこのデータを参考にダブルバッファ機構による連続読みだしを実現し、ディスクのシーク動作の影響を最小限に押さえることを可能としている。

この結果、ディスクに対するアクセスの長さを長くすることが可能なキャッシュ機構などと組み合わせることにより、同期I/O時における回転待ちなどの無駄な処理待ち時間も最少にするアクセスが可能であることがわかる。

5 終わりに

今回PCなどの小型コンピュータシステム用のディスクアレイ及びインタフェースを開発し評価を行った。その結果次のことが明らかとなった。

- ・ I/O専用プロセッサを持つディスクインタフェース及びディスクアレイを開発した。
- ・ 高いバス使用効率が達成可能であることが確認できた。
- ・ 同期I/Oのシステムにおいても、ダブルバッファ機構及びキャッシュを使用したデータ転送でオーバーヘッドを押さえることが可能である。
- ・ 非同期I/Oシステムにおいて特に高い性能が期待される。

参考文献

- [1]B.Lawrence "No More Data Loss: The BYTE Lab Tests Six Disk-Array Subsystems", BYTE August 1992.
- [2]D.A.Patterson et al., "Introduction to Redundant Arrays of Inexpensive Disks (RAID)", COMPCON '89 Spring, pp.109-117, 1989
- [3]杉本、菊地、辻澤, "マルチディスク装置 (MD-1) の性能評価", 電子情報通信学会春季全国大会 講演論文集, 6-139, 1991.
- [4]杉本、菊地、辻澤, "パーソナルコンピュータ用高速データ入出力方式", 情処学会春季全国大会 公演論文集, 6-69, 1991.