

高速並列処理ワークステーション (TOP-1)

— 入出力システム —

7N-7

中田武男 若林真一

日本アイ・ビー・エム株式会社 東京基礎研究所

1. はじめに

TOP-1は、小規模並列処理のハードウェアおよびソフトウェアの研究を目的とした高性能マルチプロセッサ・ワークステーションである^{[1][2]}。本稿ではTOP-1の入出力システムについて説明する。また、入出力処理を考慮して設計されたキャッシュの機能についても述べる。

2. 入出力システム

2.1. 概要

TOP-1は本体であるハードウェアユニットとIBMパーソナルシステム/55モデル5570(以下IBM5570)から成る(図1)。TOP-1の入出力システムは、TOP-1ハードウェアユニット内に実装されるハードディスクドライブ装置(以下HDD, 300MB/台)、ハードディスク制御カード(以下HDCカード)、I/Oインタフェースカード(以下I/O-I/Fカード)、及びIBM5570から構成される。システム中の1枚のプロセッサカードがHDCカードに接続され(このプロセッサをHDCプロセッサと呼ぶ)、HDCカードを介して最大4台のHDDを制御する。HDD以外のすべての入出力装置はIBM5570に接続され、IBM5570によって制御される。IBM5570はI/O-I/Fカードと同軸フラットケーブルによりTOP-1ハードウェアユニットに接続される。HDCプロセッサとIBM5570はキャッシュを介してTOP-1の共有バスにアクセスし、共有バス上の他のプロセッサと協調して入出力処理を行う。

2.2. キャッシュの入出力処理向き機能

TOP-1のキャッシュコントローラの動作モードの1つに、ディスクとのデータ転送を考慮したHDCオプションがある^[3]。HDCオプションでは、共有メモリにデータを書き込む際キャッシュ内に残ったデータを無効(インバリデート)にする。他のプロセッサの要求に基づいてHDCプロセッサが通常のキャッシュ動作モードのもとにディスクからデータを読み込んだ場合、読み込んだデータはキャッシュ内に共有(シェアード)データとして残る。すなわち、HDCプロセッサは、自分自身では使用しないデータを、本来そのデータを必要とする他のプロセッサと共有することになる。この場合、データの書き替えの度にキャッシュコントローラよりバスにアクセスがでるため、バスの使用率が増し、システム効率が低下する。一方、H

DCオプションを使用した場合は、ディスクから読み込む際にHDCプロセッサのキャッシュに残ったデータを無効にするので、データの書き替えによる無駄なバスアクセスを避けることができる。

3. ハードディスク・サブシステム

3.1. ハードウェア構成

ハードディスク・サブシステムは3つの部分、プロセッサ・カード、ハードディスクコントロールカード(HDCカード)、そしてハードディスクドライブ装置(HDD)から成る(図2)。HDCカードはTOP-1のシステムバスには接続されず、1枚のプロセッサカードのローカルバスに接続される。このプロセッサカード(HDCプロセッサ)は他の9枚のプロセッサカードと同一のものである。HDCカードは、4MBのローカルメモリ、64KB×4のディスクバッファメモリ、そして4つのディスクインタフェースを持つ。HDDはTOP-1の本体内部に最大4台まで実装することができる。

3.2. 機能

4MBのローカルメモリは、HDC固有のプログラムの常駐領域、およびディスクキャッシュ用のメモリとして使用される。ディスクキャッシュとして用いる場合、制御は全てソフトウェアで行う。

64KBのディスクバッファメモリは、CPUおよびSCSI-LSIのどちらからも、同時にランダムアクセス可能なデュアルポートメモリとなっている。このため、大量にデータ転送を行う場合、バッファメモリへの転送とSCSIバスへの転送はオーバーラップして行える。

ディスクI/FにはSCSIを採用し、最大4台のHDDを接続可能である。詳細は次節で述べる。

3.3. ディスクインタフェース

ディスクI/Fを単に性能の面からだけ見た場合は、SCSIよりもESDI等の方がHDDを細かく制御でき効率もよい。さらに、SCSIはコマンド実行の際のオーバーヘッドタイムがやや大きい等の欠点もある。一方、ディスクI/FとしてSCSIは次に示す利点を持つ。

(1) インテリジェントHDD

SCSI仕様のHDDは、ドライブ側にプロセッサを内蔵しているため、大変インテリジェントな機能を有する。ディスクフォーマットや簡単なエラーハンドリングなどは

High-Performance Multiprocessor Workstation (TOP-1) -- I/O System --

Takeo Nakada and Shin'ichi WAKABAYASHI

IBM Research, Tokyo Research Laboratory

HDD内部で処理を行う。そのため、従来は詳細なコントロールプログラムを記述しなければならなかったOSのデバイスドライバが、かなり簡略化でき、OSの開発効率がよくなる。また、同期式データ転送を使えば4MB/秒の転送能力がある。また、標準I/Fであるからドライブ装置を置き換える場合にも、ほとんどソフトウェアの変更なしに行える。

(2) マルチホストサポート

SCSIバスには、ホストシステムと周辺装置を合わせて8台まで接続することが可能である。そのため、開発中のシステム内のHDDをシステムに組みこんだまま既存の他のシステムに接続させることができ、OSのデバッグを行う際、正常に動作するOSでディスクをインシャライズした後、すぐに動作の確認が容易にできる。

以上の利点を考慮して、TOP-1ではSCSIを採用している。

4. I/Oサブシステム

4.1. ハードウェア構成

TOP-1とIBM5570とは、それぞれの本体内部に実装されたインタフェースカード間を同軸のフラットケーブルで結ぶことによって接続される。HDD以外の入出力装置(ディスプレイ、キーボード、プリンタ、通信ポート、ネットワークアダプタ、フロッピーディスク装置等)は、すべてIBM5570側で制御される。

4.2. バスインタフェース

IBM5570とTOP-1とは、バスのアーキテクチャが全く異なる。IBM5570側のMicro Channel⁴⁾は、32ビットの非同期バスであり、一方、TOP-1側は、64ビット×2のデュアル同期バスであるため、両者を接続するには、データとバス信号をそれぞれのバスの仕様に合わせて変換する必要がある。そこで、両者のインタフェースはTOP-1内部に実装するI/O-I/FカードとIBM5570内に実装するMicro Channel-I/Fカードによって行なわれ、2枚のカード間は同軸フラットケーブルにより接続される(図3)。

他のプロセッサとの間でのキャッシュを介したメモリデータの一貫性を保つため、I/O-I/Fカードにもプロセッサカードと同一のキャッシュコントローラとキャッシュメモリが実装され、これを介してIBM5570はTOP-1システムバスにアクセスする。Micro Channel-I/Fカードは32ビットデータ幅をサポートし、バス信号の変換とバッファリングを行う。

この2枚のインタフェースカードにより、IBM5570はTOP-1の共有メモリにアクセスしたり、メッセージバッファリングの機能を使うことができ、他のプロセッサと協調して入出力処理を行う。

5. おわりに

TOP-1の入出力システムについて、主にハードウェア

を説明した。入出力処理のソフトウェアは別の機会に報告する。

【参考文献】

- [1]鈴木:"高速並列処理ワークステーション(TOP-1)一開発方針",第37回情報処理学会全国大会論文集(1988).
- [2]大庭,小原,清水,中田,森脇,若林:"同上アーキテクチャー",同上.
- [3]大庭,清水:"同上一スヌープ・キャッシュ",同上.
- [4]"パーソナルシステム/55 マイロチャネル 技術解説書",日本アイ・ビー・エム株式会社.

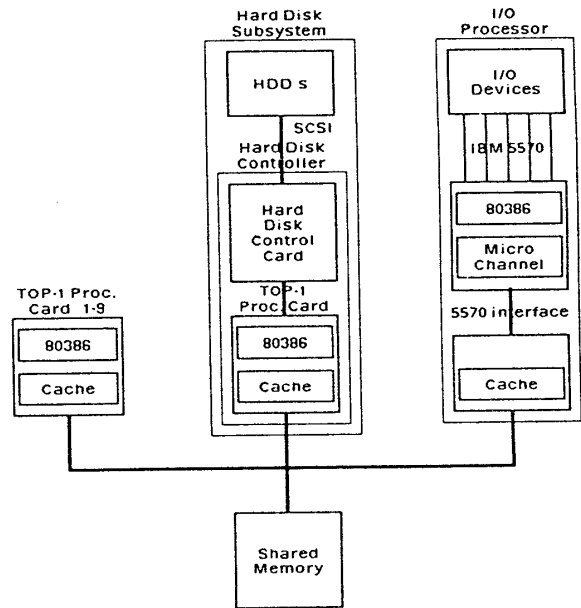


図1. 入出力構成

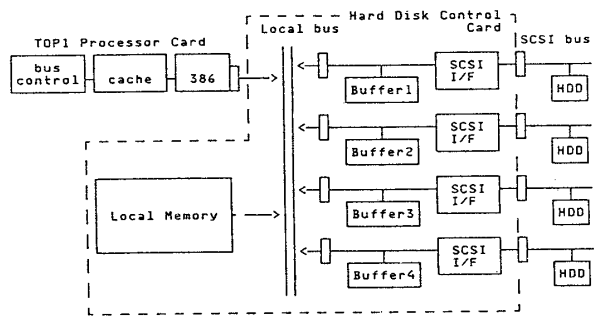


図2. ハードディスク・サブシステム

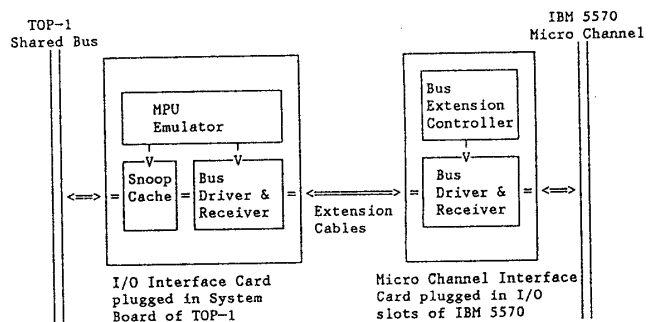


図3. インターフェース・カード