

PCIバスを用いたマルチプロセッサ通信処理システムの設計

4M-11

井戸上 彰[†] 飯島 良行[‡]

加藤 聰彦[†] 鈴木 健二[†]

[†]KDD 研究所

[‡]電気通信大学

1. はじめに

通信ネットワークの高度化に伴い、複数の回線を収容し、高位層までのプロトコル処理を高いスループットで実現する、高機能な通信処理システムの必要性が高まっている。一方、高速な CPU や PCI バスを搭載し、コストパフォーマンスに優れたパソコンが普及しつつある。そこで、これらのパソコンに回線数に対応する複数の通信ボードを搭載し、通信処理システムを構築する方法が有効であると考えられる。しかしこの場合、パソコン本体の CPU ですべての通信処理を行わせる方式では、高速なネットワークや柔軟なシステム構築に十分対応できない恐れがある。そこで筆者らは、CPU を持つ複数の通信ボードをパソコンの PCI バスを介して接続し、各通信ボードが本体 CPU とは独立に通信処理を行うマルチプロセッサ通信処理システムの開発を行っている。本稿では、本通信処理システムのハードウェア/ソフトウェアの設計概要を述べる。

2. 設計方針

PCI バスを用いたマルチプロセッサ通信処理システムの設計方針を以下に示す。

- ① PCI バスをサポートするパソコンをベースとして利用し、それぞれが CPU を持つ複数の PCI バス通信ボードを搭載して、要求に応じた様々な通信処理を実現可能とする。
- ② 通信ボード間では、パソコン本体(ホスト)の CPU を介さずに、直接相互のメモリにアクセス可能なハードウェアとし、マルチプロセッサによる高速な通信処理を実現する。
- ③ 各通信ボードが独立に通信処理を実行可能とするために、マルチプロセッサ通信処理システム用カーネルを搭載する。カーネルは、通信プログラムの実行制御、メモリ管理、タスク間通信、ボード間の共有メモリや効率的なデータ転送などの機能を実現する。

④ ホスト上には、汎用のパソコン OS のもとで、システムの運転管理、通信ログなどの機能を提供するシステム管理プログラムを実装する。

3. マルチプロセッサ通信処理システムのハードウェア

マルチプロセッサ通信処理システムのハードウェア構成を図 1 に示す。

CPU を持つ複数の通信ボードをパソコン本体の PCI バスに搭載する。通信ボードは、CPU として Pentium (100MHz) を搭載し、256K バイトの 2 次キャッシュメモリ (S-RAM)、16M バイトのメインメモリ (D-RAM)、初期化用のブート ROM、周辺制御回路、PCI バスインタフェース、および対象ネットワークに応じた通信コントローラ (HDLC コントローラ等) を実装する。周辺制御回路は、メインメモリとキャッシュの制御、ハードウェアタイマ、割り込みなどの機能を提供する。

PCI バスインタフェースでは、バスマスタ機能

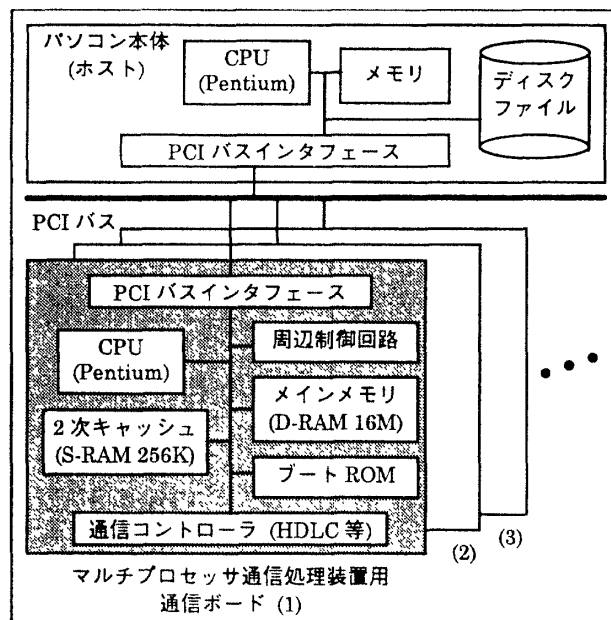


図 1 マルチプロセッサ通信処理システムのハードウェア構成

をサポートしており、これによりボードの CPU から他の通信ボード上のメモリやホスト上のメモリに対して直接アクセス可能となる。さらに、特定ボード上のメモリへの排他的アクセスを実現するバスロック機能や、ホストや他の通信ボードとの間の割り込み通知機能なども提供する。

4. マルチプロセッサ通信処理システムのソフトウェア

マルチプロセッサ通信処理システムのソフトウェアは、通信ボード上のカーネル、カーネルのもとで動作する通信プログラム、およびホスト上のシステム管理プログラムから構成される(図2)。

(1) カーネル

カーネルは、個々の通信ボード内において、通信プログラム(タスク)の管理、すべてのタスクからアクセス可能な共有バッファを含むメモリ管理、キューを介して共有バッファへのポインタを転送するタスク間通信、タスク間通信によるメッセージ転送を契機としたスケジューリングなどの機能を提供する^[1]。

また、ボード間にまたがる機能として、ボード間共有メモリ機能と、ボード間データ転送機能を提供する。ボード間共有メモリ機能は、個々の通信ボード上に存在する D-RAM の特定領域に対して、複数の異なる通信ボード上のカーネルやタスクからの書き込み/読み出しを可能とし、一貫性を保証したデータを保持する。ただし、PCI バスを介したメモリアクセス速度は、ボード内のローカルな D-RAM へのアクセスと比較して遅いため、各ボードの D-RAM にキャッシュを保持して共有メモリの読み出しの高速化を図っている。本領域は、カーネルが他のボード上のカーネルとの間で共有する管理用データ領域や、通信プログラムが使用するルーチングテーブルなどに利用される。

ボード間データ転送機能は、受信側のバッファの確保、データ転送、イベント通知などの手順により、ボード内部と同様なキューによるタスク間通信インタフェースを提供する。本機能を実現するための具体的な通信方式に関しては、バス上のやり取りやメモリ管理などのオーバーヘッドを考慮して、最適な方式を採用することとした^[2]。

(2) 通信プログラム

要求に応じた様々な通信処理を実現するため、通信プログラムは、カーネルによって動的にホストからダウンロードされ、それぞれ独立したタス

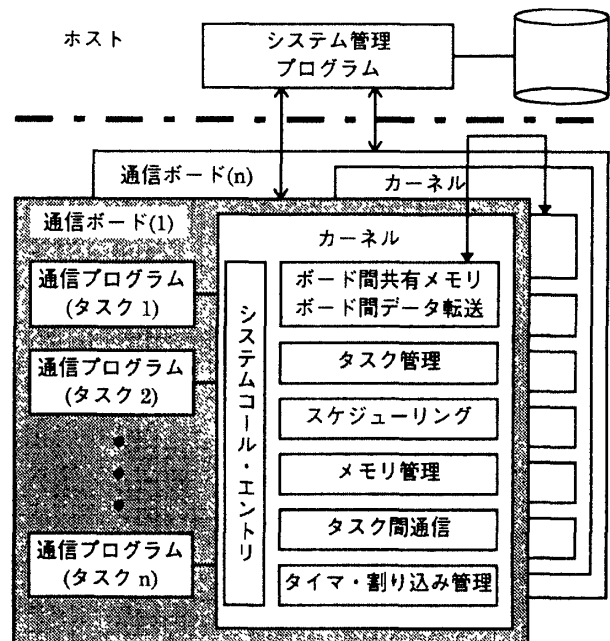


図2 マルチプロセッサ通信処理システムのソフトウェア構成

クとして動作する。通信プログラムは、カーネルが提供するボード間共有メモリやボード間データ転送などの機能を利用し、他のボード上の通信プログラムと連携しながら、各層のプロトコル処理やゲートウェイ処理などを実現する。

(3) システム管理プログラム

ホスト上のシステム管理プログラムは、通信ボードの初期化、カーネルおよび通信プログラムのダウンロード、通信ボードからのデバッグメッセージの表示、通信ログなどをサポートするためのファイル入出力機能を提供する。

5. おわりに

本稿では、PCI バスを介して接続された複数の通信ボードによって構成されるマルチプロセッサ通信処理システムの設計について述べた。これまでに、本システム用の通信ボード・ハードウェアの開発を完了しており、現在、ボードに搭載するカーネルの実装とともに、ボード間データ転送方式の評価を進めている。最後に、日頃御指導頂く KDD 研究所村上所長に感謝する。

参考文献

- [1] 井戸上, 加藤, 鈴木, 小野, “OSI 7 層ボードのためのオペレーティング・システム,” 情報処理学会論文誌, Vol. 35, No.5, May 1994.
- [2] 飯島, 井戸上, 加藤, 鈴木, “マルチプロセッサ通信処理システムにおけるボード間通信方式,” 本大会予稿