

組込応用向き小型プロセッサカードの予備的評価

戸田賢二[†] 石綿陽一^{††} 関山守[†] 高橋栄一[†]

[†]産業技術総合研究所 ^{††}(株)ムービングアイ

計算パワーを必要とする組込応用向きに12cm×8cmの小型のプロセッサカードを開発した。同カードは、インテルのモバイルモジュールを搭載可能で、PCI、イーサネット、IDE、USB、RS-232Cのインタフェースを備える。フラッシュROMを1MB搭載しておりオンボードで書き込み可能である。メモリはDIMMソケットに128MB実装を実装している。OSは、Linux/ART-Linuxを想定しており、ディスクからスタンドアロンで立ち上げる方法の他、フラッシュROMから立ち上げてNFSでファイルをマウントするディスクレスの構成も可能である。単体で用いるほか、分散形態での組込応用、並列計算機の要素プロセッサなどに利用する。本稿では、Linuxでの利用を中心に、ブートの手順や、予備的な性能について述べる。

Preliminary evaluation of a small-size processor card for embedded applications

Kenji TODAX[†], Yoichi ISHIWATA^{††}, Mamoru SEKIYAMA[†], and Eiichi TAKAHASHI[†]

[†]National Institute of Advanced Industrial Science and Technology ^{††}MovingEye Inc.

We have developed a small-sized 12cm by 8cm processor card for (parallel/distributed) high performance embedded applications. It uses the Intel Mobile Module and has PCI, Ethernet, IDE, USB, and RS-232C ports. It is equipped with 1MB flush ROM and 128MB DIMM. The card is designed to boot Linux/ART-Linux OS from the flush ROM, disk, or NFS file system. It can be used as a single processor or an element processor of parallel/distributed systems. This manuscript describes booting process and preliminary performance on Linux.

1. はじめに

社会へのIT技術の浸透に伴い組込機器の応用範囲も拡大しており、プロセッサも家電や携帯製品に組み込む1チップの非常に小型のもの、他、車載や作業用ロボット搭載用とか小型の計算サーバ用のすこし大きめでもよいが計算パワーが必要な応用向きのものへの需要も増している。例えば、ヒューマノイドロボットの制御に必要な動的バランスを考慮した計算はリアルタイムでかつ複数プロセッサを動員する必要のある大きな計算パワーを必要とする。産業応用において

も、制御対象のプロセスなどの物理的配置や、故障発生時の対処、高度制御に必要なリアルタイム処理能力、ネットワークにより外部システムと接続連携を図る機能は必要不可欠であり、並列分散のリアルタイム処理に対応しかつ計算パワーとサイズとのバランスに優れるプロセッサが必要とされてきている。

本稿では、サイズの制限がありかつ高性能が必要とされる並列や分散のリアルタイム応用で用いることを想定した小型プロセッサカードについて、ハードウェア、OS、その用途、等について紹介する。

2. プロセッサカードの仕様

プロセッサカードは、組み込み機器応用に単独使用するほか、並列や分散システムの要素プロセッサとして用いることを想定している。計算能力は確保しつつ、極力機能を限定しコンパクトにすることに留意している。プロセッサとして、インテル社の Mobile Module Connector 1 と 2 の二つの版を作成した。括弧内は、MMC-1 の場合。

- サイズ：12cm×8cm×1.5cm (2cm)
- プロセッサ：Intel MMC-2 (MMC-1) PentiumIII 500MHz (PentiumII 300MHz)搭載 82443DX Host Bridge system, System bus 100MHz (66MHz)
- メモリ：DIMM スロット×1 128MB SDRAM 実装
- flush ROM：512KB×2 オンボード書き込み可能、ジャンパでアドレスのスイッチ可能
- ポート：PCI, Ether 10/100base-T, RS232-C, USB, IDE 装備
- 電源：5V 及び 3.3V 20W
- OS：Linux 系 OS を想定 flush ROM/IDE disk/NFS サーバからブート可能（グラフィックチップはなし）
- その他：状態表示用のプログラマブル LED, ユーザ用ディップ SW, リセット SW 装備

図1に小型プロセッサカードの構成を示す。プロセッサと North Bridge はモバイルモジュールに搭載されており回路設計が容易化されている。内部の PCI をブリッジを経由せず直接外に出しているが、これは組込用途であるため PCI 接続の距離が長くなるとの想定の下に小型化を図ったためである。

図2は、プロセッサカードの写真(表裏)である(MMC-1用であるがMMC-2用も同様)。上の写真の左側のソケットはPCI用、下の写真の中央

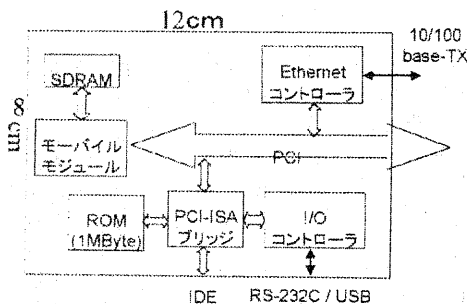


図1 小型プロセッサカードの構成

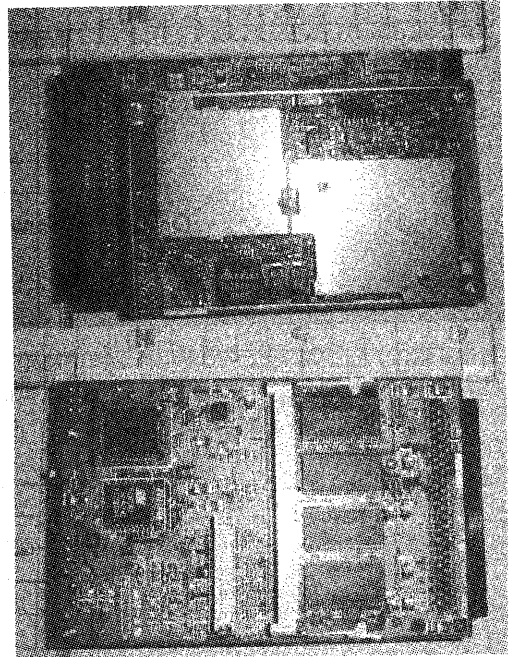


図2 小型プロセッサカード(写真)

部のソケットが IDE 用である。Ether や RS-232C を含めポートは小型化のため特殊コネクタを使用している。レイアウトの詳細や、チップセットの情報などは <http://www.etl.go.jp/~realtime> を参照されたい。

本プロセッサカードと同等のサイズで、PentiumII を搭載するカードプロセッサが PFU 株式会社から発売されている。これには、PCMCIA カードが装着可能であり Windows にも対応する BIOS が供給されている。ただし、Ether は標準ではなく、PentiumIII には対応していない。また、Pentium プロセッサを搭載したコンパクト PCI 用のプロセッサボードも各社から供給されているが本カードと比べかなりサイズが大きい。

3. ソフトウェア開発環境及び Linux のブート

小型プロセッサカードは、単独または、並列や分散形態で用いることを想定しているが、リアルタイム処理のサポートはいずれの場合でも必要である。このため各プロセッサカードには ART-Linux を搭載し、並列分散対応のリアルタイムタスクのスケジューラを搭載予定である。ART-Linux は、電総研で開発した Linux ベースの実時間 OS であり、プログラム開発環境と実時間制御機能の両方を提供できる。実時間化のための

OS の介入をポーリング方式として、Linux とのデバイスドライバのソースレベルの互換性を保持しており、アプリケーションレベルではバイナリ互換である。Linux の最新バージョンへの追従は容易であり、Linux の多くのソフト資産を活用することができる。また、実時間タスクのメモリ保護機能があるため、実行時の信頼性高く、プログラム開発も安全に行うことができる。実時間制御機能としては、実時間タスクの開始時間の保証、固定優先度による実時間スケジューリング、優先度継承の機能を有している。(ソースコードは <http://esd.etl.go.jp> で配布。)

また、リアルタイムの MPI への対応も検討している。以下では、Linux 系 OS のブートの方法を述べるが、むしろ他のリアルタイム OS を使用することも可能である。

Linux 系 OS のブートは、

- Flush ROM
- NFS
- Disk

のいずれからでも可能であり、ファイルシステムは、

- メモリ
- NFS
- Disk

を用いることが可能である。

Linux では、OS 起動後は BIOS は使用されないが、ブートのために、PCIブリッジやプロセッサのイニシャライズ、シリアルポート、IDE ポート、Ether コントローラなどの設定を行う必要がある。Pentium プロセッサの場合、PCIブリッジをはじめとするチップセットの設定が複雑であるが、FreeBIOS などの公開されているソースを参考に BIOS の開発 (Linux の起動に必要な初期設定のみ) を行った。現在、Flush ROM から Linux を起動し NFS を利用する方式で動作しており、各種パラメータのチューニングを行っているところである。小型プロセッサカードは、プロセッサとして、PentiumII の 300MHz と PentiumIII の 500MHz を搭載するバージョンがあり、これらについての ART-Linux 上での性能測定を進めている。

4. 小型プロセッサカードの応用

本小型プロセッサカードの用途は以下のものを想定している。

- 単独又はイーサ接続形態で
- リアルタイム並列計算機 CODA-RP の要素プロセッサとして

- ペリフェラルと PCI 接続して組込用途に
- PCI ブリッジを組み込んだバックプレーンで並列処理 (現在 3 台まで)
- Myrinet などを用いたデスクトップ並列計算機の要素プロセッサとして

以下主な使用方法について述べる。

4.1 単体又は小規模組込用

プロセッサカードの PCI ポートを利用して、コントローラやセンサなどを搭載するペリフェラルボードと接続する。図 3 はペリフェラル接続用 (2 枚まで) の PCI カードラックである。プロセッサカードが PCI マスタで、ペリフェラルがスレーブとなる。プロセッサカード同士を PCI 接続して小規模の並列システムを構成する場合は、(PCI ポートには内部バスを直接接続しているため) バックプレーンに PCI ブリッジを装着して接続する。PCI ポートは、32bits 33MHz であるため、132MB/s の転送速度となる。通信量が少ない場合は、10/100Mbps の Ether ポートや USB ポートの利用も手軽である。モバイルシステムに搭載し基地局とのリンクが必要な場合には、無線装置を USB 接続する方法も手軽である。

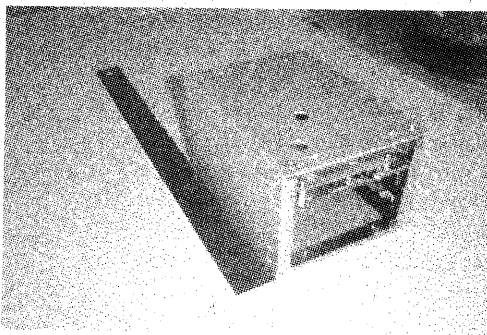


図 3 ペリフェラル接続用 PCI カードラック

4.2 スケーラビリティをもった小型並列分散システム

応用によって、多数のプロセッサを用いた並列や分散のシステムが必要とされる場合は、Ether のスイッチングハブの利用や PCI インタフェースで Myrinet を利用する等の方法がある。また、次に述べるリアルタイム相互接続網を利用した専用ネットワークによる接続により、ハイエンド処理の応用に対応することができる。この専用機

低コストリアルタイム通信技術の開発

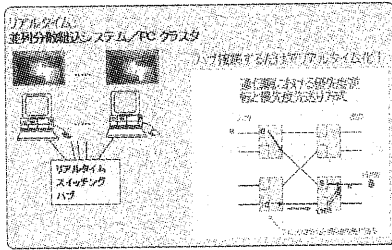


図 4 リアルタイムスイッチングハブ

での研究成果の活用として、時間制約の厳しいリアルタイム応用に対しては、スイッチングハブに優先度先送り方式を導入し、既存の Ether プロトコルとコンパチビリティを保持しつつ、簡便に低コストでリアルタイムシステムを構築できるようにすることをねらって研究を進めている。

4.3 専用システム: リアルタイム並列計算機 CODA-RP

CODA-RPは、独自の優先度先送り方式と呼ばれる通信網における優先度逆転(注1)を抑制する機構を実装した実時間並列計算機プロトタイプである。同方式は、通信網で行き先が詰まっていた場合、自分の優先度情報を次段に先に送り、優先度を受け取ったスイッチは、送られてきた優先度と自分のバッファ内の最大優先度を比較し、送られてきた方の優先度が高ければ、その優先度を

行き先がぶつかった際の調停に用いるほか、更に次段のスイッチに送る。これを順次繰り返していけば、段数が多くても一番優先度の高いパケットが、いつまでもブロックされ続けることはなく、最悪遅延が保証される(図5-1及び図5-2は、2入力2出力のルータを用いた場合の例である。図中“0d100”は、出力ポートの番号0に行くべき優先度100の packets を示す。この最悪ケースでも優先度先送りが多段に渡って行われ、優先度逆転を解消している)がまた、ネットワークにおけるホットスポット(注2)を解消する方向で制御が働くため(込み合ったところは後ろから高い優先度のパケットが来れば押し出されて解消する)、リアルタイム制御性のみならず全体のスループットの向上も見込めるとう評価結果を実機で実証している(注3)。この実現では、通信プロトコルは独自のものであったが、パケットをスイッチで一旦蓄えて送信するストアアンドフォワード方式の通信には、優先度先送り方式が適用可能であるため、ATMやイーサのスイッチングハブに用いることができる。

- 注1) 優先度逆転: 低い優先度のパケットが、高い優先度のパケットをブロックしてしまう現象で、優先度による追い越しができないシステムや、優先度による追い越しができるバッファなどを用いても、それが一杯になってしまった場合発生する。
- 注2) ホットスポット: ネットワークのトラフィックが輻輳して、流れが阻害されている場所。道路交通にたとえると、渋滞ポイントのこと。
- 注3) 優先度先送り方式を実装した4入力4出力の通信用LSI(ルータ)を開発し、並列システムで評価を行った。クロックレベルのシミュレ

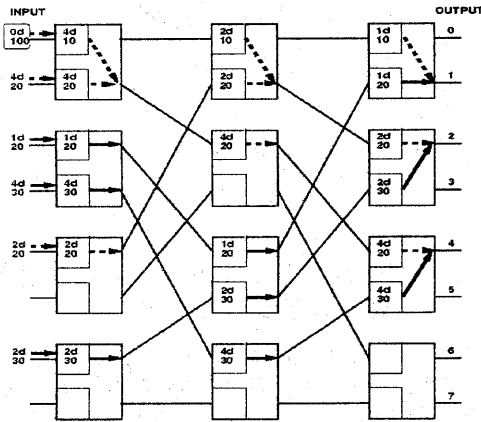


図 5-1 最悪ケース

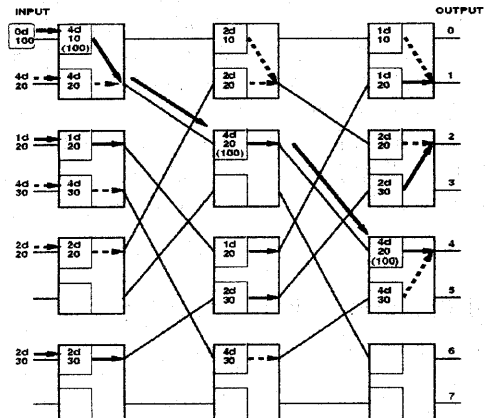
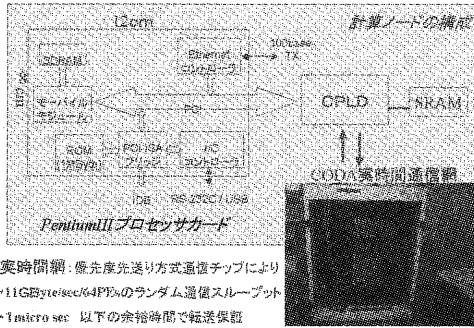


図 5-2 優先度先送り方式の動作(最悪ケース)

CODA-RPの全体構成(32台版)



実時間網: 優先度先送り方式通信チップにより
 ・11GByte/sec(64PF)のランダム通信スループット
 ・1micro sec 以下の発給時間で転送保証

図 6 リアルタイム並列計算機 CODA-RP

ータにより数百台規模まで正確に性能予測可能。

優先度先送り方式は、4入力4出力のルータチップとしてLSI化しており、計算ノードとしてFPGAをもちいたCODA-Rの通信網として実装されている。性能評価の結果、クロックレベルのシミュレータとも非常に良く整合する評価結果を得ている。ポートあたり190MB/s(入出力合わせると×2となる)、64台システムで1GB/sのスループット(ランダム通信,ランダム優先度),どの通信ノードへの転送も最高優先度の場合1ミリ秒以下で到達する。また、台数が増加したときのスケラビリティも非常に高いことが実証された。

CODA-Rの計算ノードを拡張し、小型プロセッサカードを付加したものが、CODA-RPであり、図6の構成となっている。FPGAがルータインタフェース、プロセッサインタフェース、SRAMをもつ(図7参照)。FPGAの果たす機能を以下にまとめる。

- パケット送信: プロセッサからSRAMの特定のアドレスへの書き込みで実行

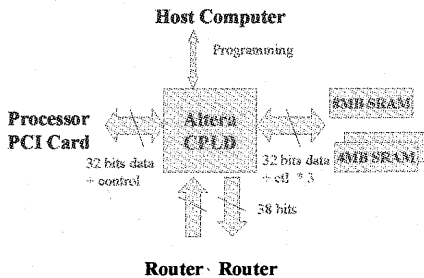


図 7 CODA-RP のFPGA 回りの構成

Interconnection Network of CODA-RP

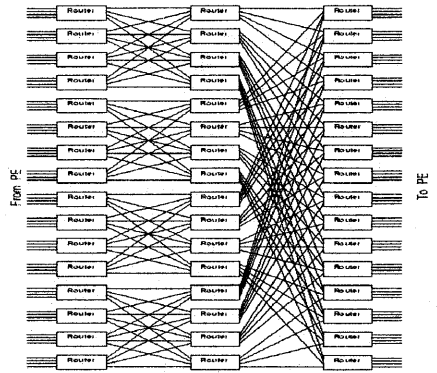


図 8 CODA-RP の相互接続網 (リーバースペースライン)

- パケット受信: プロセッサへの割り込み/フラグチェックによるパケット到着 (SRAMへのバッファリング) の通知
- 入札: 実時間タスクの実行可否を他のプロセッサに問い合わせる
- 応札: 実行可能な場合は実行コストを入札プロセッサに返す
- スケジューリング: タスクの価値に従って、実行すべきタスクをプロセッサに教える(入札/応札も含む)

CODA-RPの相互接続網は、リーバースペースライン網(図8)であり、4入力4出力のルータを用いているため64台までは3段階網となる。256台までは4段階網、1024台までは5段階網となり、台数規模の増加に対するスケラビリティが高い。

5. おわりに

小型で高性能なプロセッサカードへの需要はある程度あると思われるので、市販化を検討したい。次のバージョンを作成する場合は、消費電力の削減も重要であるので、省電力向きのプロセッサの使用やFPGAをカード上に搭載し、応用の自由度を増すことも考えたい。

謝辞

プロセッサカードのBIOSの開発でお世話になった東京大学の長嶋 功一氏, カスタマシステムの中島俊夫氏, 明電舎の宍道洋氏に感謝いたします。

参考文献

- 1) 戸田, 石綿, 関山, 高橋, 山口, “市販プロセスッサとFPGAから構成される計算機ノードをもつ実時間並列システムアーキテクチャ”, 電子情報通信学会, 信学技報 FTS99-35 pp.31-36, 平成11年8月.
- 2) Yamaguchi, Toda, Takahashi, Nishida, “CODA-R: A Reconfigurable Testbed for Real-Time Parallel Computation”, Procs of Int’l Workshop on Real-Time Systems and Applications, pp.252-259, IEEE 他 1997.10.
- 3) 戸田, 西田, 高橋, 山口, “A Priority Forwarding Scheme for Real-Time Multistage Interconnection Networks and Its Evaluation”, 電子情報通信学会和文論文誌D-I 「実時間処理システムとその応用」特集号, Vol. J78-D-I, No. 8, pp.724-734 1995.
- 4) 戸田, 西田, 高橋, Michell, 山口, “優先度先送り方式による実時間相互結合網用ルータチップの実現と性能”, 情報処理学会論文誌 Vol. 36 No. 7, pp.1619-1629, 1995.7.
- 5) ART-Linux : <http://esd.etl.go.jp>
- 6) CODA & Compact Processor Card :
<http://www.etl.go.jp/~realtime>

組込応用向き小型プロセッサカードの予備的評価

戸田賢二[†] 石綿陽一^{††} 関山守[†] 高橋栄一[†][†]産業技術総合研究所 ^{††}(株)ムービングアイ

計算パワーを必要とする組込応用向きに 12cm×8cm の小型のプロセッサカードを開発した。同カードは、インテルのモバイルモジュールを搭載可能で、PCI、イーサネット、IDE、USB、RS-232C のインタフェースを備える。フラッシュ ROM を 1MB 搭載しておりオンボードで書き込み可能である。メモリは DIMM ソケットに 128MB 実装を実装している。OS は、Linux/ART-Linux を想定しており、ディスクからスタンドアロンで立ち上げる方法の他、フラッシュ ROM から立ち上げて NFS でファイルをマウントするディスクレスの構成も可能である。単体で用いるほか、分散形態での組込応用、並列計算機の要素プロセッサ、などに利用する。本稿では、Linux での利用を中心に、ブートの手順や、予備的な性能について述べる。

Preliminary evaluation of a small-size processor card for embedded applications

Kenji TODAX[†], Yoichi ISHIWATA^{††}, Mamoru SEKIYAMA[†], and Eiichi TAKAHASHI[†][†]National Institute of Advanced Industrial Science and Technology ^{††}MovingEye Inc.

We have developed a small-sized 12cm by 8cm processor card for (parallel/distributed) high performance embedded applications. It uses the Intel Mobile Module and has PCI, Ethernet, IDE, USB, and RS-232C ports. It is equipped with 1MB flush ROM and 128MB DIMM. The card is designed to boot Linux/ART-Linux OS from the flush ROM, disk, or NFS file system. It can be used as a single processor or an element processor of parallel/distributed systems. This manuscript describes booting process and preliminary performance on Linux.

1. はじめに

社会への IT 技術の浸透に伴い組込機器の応用範囲も拡大しており、プロセッサも家電や携帯製品に組み込む 1 チップの非常に小型のもの、他、車載や作業用ロボット搭載用とか小型の計算サーバ用のすこし大きめでもよいが計算パワーが必要な応用向きのものへの需要も増している。例えば、ヒューマノイドロボットの制御に必要な動的バランスを考慮した計算はリアルタイムでかつ複数プロセッサを動員する必要のある大きな計算パワーを必要とする。産業応用において

も、制御対象のプロセスなどの物理的配置や、故障発生時の対処、高度制御に必要なリアルタイム処理能力、ネットワークにより外部システムと接続連携を図る機能は必要不可欠であり、並列分散のリアルタイム処理に対応しかつ計算パワーとサイズとのバランスに優れるプロセッサが必要とされてきている。

本稿では、サイズの制限がありかつ高性能が必要とされる並列や分散のリアルタイム応用で用いることを想定した小型プロセッサカードについて、ハードウェア、OS、その用途、等について紹介する。

2. プロセッサカードの仕様

プロセッサカードは、組み込み機器応用に単独使用するほか、並列や分散システムの要素プロセッサとして用いることを想定している。計算能力は確保しつつ、極力機能を限定しコンパクトにすることに留意している。プロセッサとして、インテル社の Mobile Module Connector 1 と 2 の二つの版を作成した。括弧内は、MMC-1 の場合。

- サイズ：12cm×8cm×1.5cm (2cm)
- プロセッサ：Intel MMC-2 (MMC-1) PentiumIII 500MHz (PentiumII 300MHz)搭載 82443DX Host Bridge system, System bus 100MHz (66MHz)
- メモリ：DIMM スロット×1 128MB SDRAM 実装
- flush ROM：512KB×2 オンボード書き込み可能、ジャンパでアドレスのスイッチ可能
- ポート：PCI, Ether 10/100base-T, RS232-C, USB, IDE 装備
- 電源：5V 及び 3.3V 20W
- OS：Linux 系 OS を想定 flush ROM/IDE disk/NFS サーバからブート可能 (グラフィックチップはなし)
- その他：状態表示用のプログラマブル LED, ユーザ用ディップ SW, リセット SW 装備

図1に小型プロセッサカードの構成を示す。プロセッサと North Bridge はモバイルモジュールに搭載されており回路設計が容易化されている。内部の PCI をブリッジを経由せず直接外に出しているが、これは組込用途であるため PCI 接続の距離が長くなるとの想定の下に小型化を図ったためである。

図2は、プロセッサカードの写真(表裏)である(MMC-1用であるがMMC-2用も同様)。上の写真の左側のソケットはPCI用、下の写真の中央

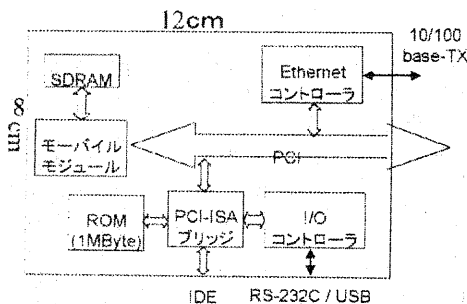


図1 小型プロセッサカードの構成

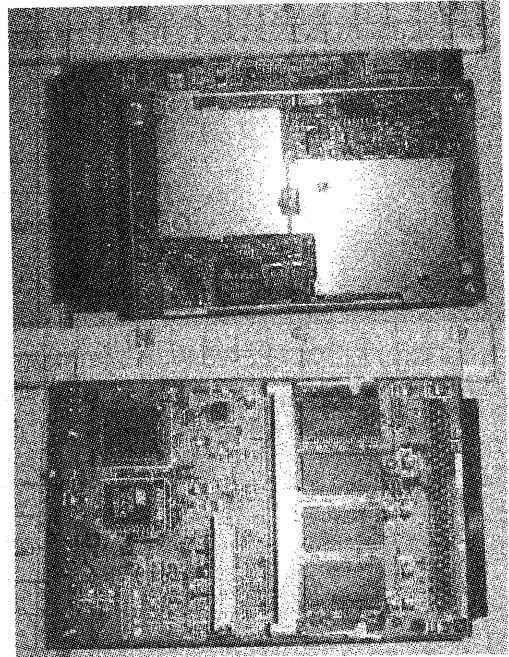


図2 小型プロセッサカード(写真)

部のソケットが IDE 用である。Ether や RS-232C を含めポートは小型化のため特殊コネクタを使用している。レイアウトの詳細や、チップセットの情報などは <http://www.etl.go.jp/~realtime> を参照されたい。

本プロセッサカードと同等のサイズで、PentiumII を搭載するカードプロセッサが PFU 株式会社から発売されている。これには、PCMCIA カードが装着可能であり Windows にも対応する BIOS が供給されている。ただし、Ether は標準ではなく、PentiumIII には対応していない。また、Pentium プロセッサを搭載したコンパクト PCI 用のプロセッサボードも各社から供給されているが本カードと比べかなりサイズが大きい。

3. ソフトウェア開発環境及び Linux のブート

小型プロセッサカードは、単独または、並列や分散形態で用いることを想定しているが、リアルタイム処理のサポートはいずれの場合でも必要である。このため各プロセッサカードには ART-Linux を搭載し、並列分散対応のリアルタイムタスクのスケジューラを搭載予定である。ART-Linux は、電総研で開発した Linux ベースの実時間 OS であり、プログラム開発環境と実時間制御機能の両方を提供できる。実時間化のための

OS の介入をポーリング方式として、Linux とのデバイスドライバのソースレベルの互換性を保持しており、アプリケーションレベルではバイナリ互換である。Linux の最新バージョンへの追従は容易であり、Linux の多くのソフト資産を活用することができる。また、実時間タスクのメモリ保護機能があるため、実行時の信頼性高く、プログラム開発も安全に行うことができる。実時間制御機能としては、実時間タスクの開始時間の保証、固定優先度による実時間スケジューリング、優先度継承の機能を有している。(ソースコードは <http://esd.etl.go.jp> で配布。)

また、リアルタイムの MPI への対応も検討している。以下では、Linux 系 OS のブートの方法を述べるが、むしろ他のリアルタイム OS を使用することも可能である。

Linux 系 OS のブートは、

- Flush ROM
- NFS
- Disk

のいずれからでも可能であり、ファイルシステムは、

- メモリ
- NFS
- Disk

を用いることが可能である。

Linux では、OS 起動後は BIOS は使用されないが、ブートのために、PCIブリッジやプロセッサのイニシャライズ、シリアルポート、IDE ポート、Ether コントローラなどの設定を行う必要がある。Pentium プロセッサの場合、PCIブリッジをはじめとするチップセットの設定が複雑であるが、FreeBIOS などの公開されているソースを参考に BIOS の開発 (Linux の起動に必要な初期設定のみ) を行った。現在、Flush ROM から Linux を起動し NFS を利用する方式で動作しており、各種パラメータのチューニングを行っているところである。小型プロセッサカードは、プロセッサとして、PentiumII の 300MHz と PentiumIII の 500MHz を搭載するバージョンがあり、これらについての ART-Linux 上での性能測定を進めている。

4. 小型プロセッサカードの応用

本小型プロセッサカードの用途は以下のものを想定している。

- 単独又はイーサ接続形態で
- リアルタイム並列計算機 CODA-RP の要素プロセッサとして

- ペリフェラルと PCI 接続して組込用途に
- PCI ブリッジを組み込んだバックプレーンで並列処理 (現在 3 台まで)
- Myrinet などを用いたデスクトップ並列計算機の要素プロセッサとして

以下主な使用方法について述べる。

4.1 単体又は小規模組込用

プロセッサカードの PCI ポートを利用して、コントローラやセンサなどを搭載するペリフェラルボードと接続する。図 3 はペリフェラル接続用 (2 枚まで) の PCI カードラックである。プロセッサカードが PCI マスタで、ペリフェラルがスレーブとなる。プロセッサカード同士を PCI 接続して小規模の並列システムを構成する場合は、(PCI ポートには内部バスを直接接続しているため) バックプレーンに PCI ブリッジを装着して接続する。PCI ポートは、32bits 33MHz であるため、132MB/s の転送速度となる。通信量が少ない場合は、10/100Mbps の Ether ポートや USB ポートの利用も手軽である。モバイルシステムに搭載し基地局とのリンクが必要な場合には、無線装置を USB 接続する方法も手軽である。

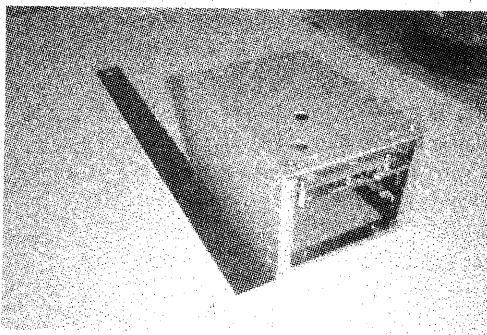


図 3 ペリフェラル接続用 PCI カードラック

4.2 スケーラビリティをもった小型並列分散システム

応用によって、多数のプロセッサを用いた並列や分散のシステムが必要とされる場合は、Ether のスイッチングハブの利用や PCI インタフェースで Myrinet を利用する等の方法がある。また、次に述べるリアルタイム相互接続網を利用した専用ネットワークによる接続により、ハイエンド処理の応用に対応することができる。この専用機

低コストリアルタイム通信技術の開発

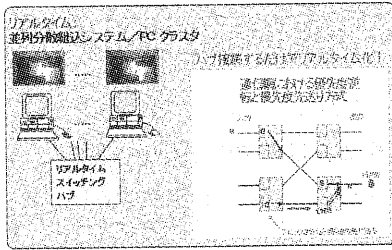


図 4 リアルタイムスイッチングハブ

での研究成果の活用として、時間制約の厳しいリアルタイム応用に対しては、スイッチングハブに優先度先送り方式を導入し、既存の Ether プロトコルとコンパチビリティを保持しつつ、簡便に低コストでリアルタイムシステムを構築できるようにすることをねらって研究を進めている。

4.3 専用システム: リアルタイム並列計算機 CODA-RP

CODA-RPは、独自の優先度先送り方式と呼ばれる通信網における優先度逆転(注1)を抑制する機構を実装した実時間並列計算機プロトタイプである。同方式は、通信網で行き先が詰まっていた場合、自分の優先度情報を次段に先に送り、優先度を受け取ったスイッチは、送られてきた優先度と自分のバッファ内の最大優先度を比較し、送られてきた方の優先度が高ければ、その優先度を

行き先がぶつかった際の調停に用いるほか、更に次段のスイッチに送る。これを順次繰り返していけば、段数が多くても一番優先度の高いパケットが、いつまでもブロックされ続けることはなく、最悪遅延が保証される(図5-1及び図5-2は、2入力2出力のルータを用いた場合の例である。図中“0d100”は、出力ポートの番号0に行くべき優先度100の packets を示す。この最悪ケースでも優先度先送りが多段に渡って行われ、優先度逆転を解消している)がまた、ネットワークにおけるホットスポット(注2)を解消する方向で制御が働くため(込み合ったところは後ろから高い優先度のパケットが来れば押し出されて解消する)、リアルタイム制御性のみならず全体のスループットの向上も見込めるとい評価結果を実機で実証している(注3)。この実現では、通信プロトコルは独自のものであったが、パケットをスイッチで一旦蓄えて送信するストアアンドフォワード方式の通信には、優先度先送り方式が適用可能であるため、ATMやイーサのスイッチングハブに用いることができる。

- 注1) 優先度逆転: 低い優先度のパケットが、高い優先度のパケットをブロックしてしまう現象で、優先度による追い越しができないシステムや、優先度による追い越しができるバッファなどを用いても、それが一杯になってしまった場合発生する。
- 注2) ホットスポット: ネットワークのトラフィックが輻輳して、流れが阻害されている場所。道路交通にたとえると、渋滞ポイントのこと。
- 注3) 優先度先送り方式を実装した4入力4出力の通信用LSI(ルータ)を開発し、並列システムで評価を行った。クロックレベルのシミュレ

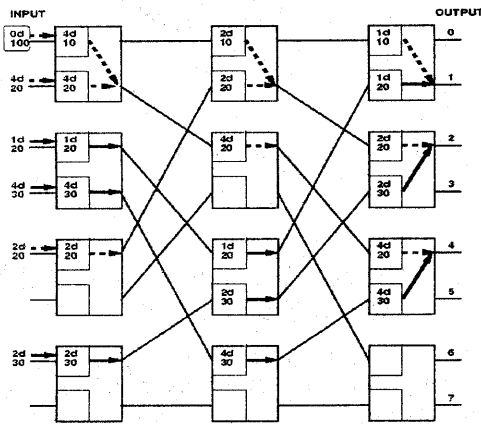


図 5-1 最悪ケース

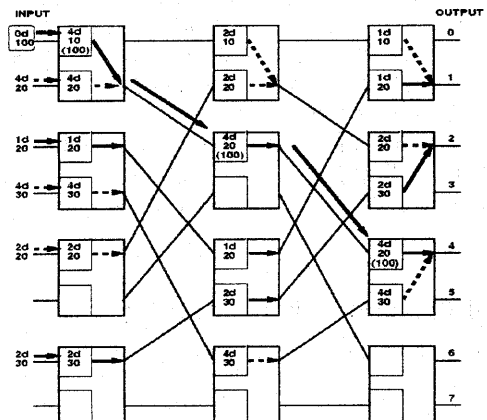
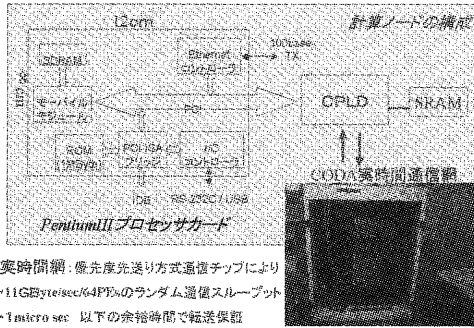


図 5-2 優先度先送り方式の動作(最悪ケース)

CODA-RPの全体構成(32台版)



実時間網: 優先度先送り方式通信チップにより
 ・11GByte/sec(64PF)のランダム通信スループット
 ・1micro sec 以下の発給時間で転送保証

図 6 リアルタイム並列計算機 CODA-RP

ータにより数百台規模まで正確に性能予測可能。

優先度先送り方式は、4入力4出力のルータチップとしてLSI化しており、計算ノードとしてFPGAをもちいたCODA-Rの通信網として実装されている。性能評価の結果、クロックレベルのシミュレータとも非常に良く整合する評価結果を得ている。ポートあたり190MB/s(入出力合わせると×2となる)、64台システムで1GB/sのスループット(ランダム通信,ランダム優先度),どの通信ノードへの転送も最高優先度の場合1ミリ秒以下で到達する。また、台数が増加したときのスケラビリティも非常に高いことが実証された。

CODA-Rの計算ノードを拡張し、小型プロセッサカードを付加したものが、CODA-RPであり、図6の構成となっている。FPGAがルータインタフェース、プロセッサインタフェース、SRAMをもつ(図7参照)。FPGAの果たす機能を以下にまとめる。

- パケット送信: プロセッサからSRAMの特定のアドレスへの書き込みで実行

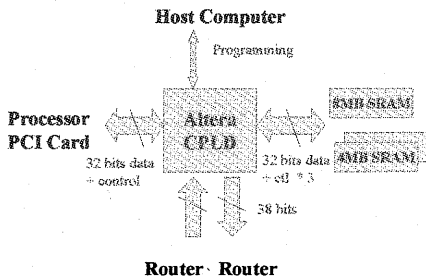


図 7 CODA-RP のFPGA 回りの構成

Interconnection Network of CODA-RP

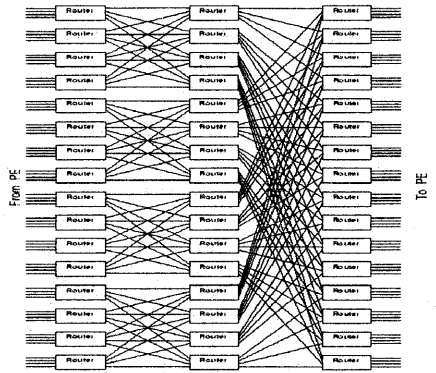


図 8 CODA-RP の相互接続網 (リーバースペースライン)

- パケット受信: プロセッサへの割り込み/フラグチェックによるパケット到着 (SRAMへのバッファリング) の通知
- 入札: 実時間タスクの実行可否を他のプロセッサに問い合わせる
- 応札: 実行可能な場合は実行コストを入札プロセッサに返す
- スケジューリング: タスクの価値に従って、実行すべきタスクをプロセッサに教える(入札/応札も含む)

CODA-RPの相互接続網は、リーバースペースライン網(図8)であり、4入力4出力のルータを用いているため64台までは3段網となる。256台までは4段網、1024台までは5段網となり、台数規模の増加に対するスケラビリティが高い。

5. おわりに

小型で高性能なプロセッサカードへの需要はある程度あると思われるので、市販化を検討したい。次のバージョンを作成する場合は、消費電力の削減も重要であるので、省電力向きのプロセッサの使用やFPGAをカード上に搭載し、応用の自由度を増すことも考えたい。

謝辞

プロセッサカードのBIOSの開発でお世話になった東京大学の長嶋 功一氏、カスタマシステムの中島俊夫氏、明電舎の宍道洋氏に感謝いたします。

参考文献

- 1) 戸田, 石綿, 関山, 高橋, 山口, “市販プロセスサと FPGA から構成される計算機ノードをもつ実時間並列システムアーキテクチャ”, 電子情報通信学会, 信学技報 FTS99-35 pp.31-36, 平成 11 年 8 月.
- 2) Yamaguchi, Toda, Takahashi, Nishida, “CODA-R: A Reconfigurable Testbed for Real-Time Parallel Computation”, Procs of Int’l Workshop on Real-Time Systems and Applications, pp.252-259, IEEE 他 1997.10.
- 3) 戸田, 西田, 高橋, 山口, “A Priority Forwarding Scheme for Real-Time Multistage Interconnection Networks and Its Evaluation”, 電子情報通信学会和文論文誌 D-I 「実時間処理システムとその応用」特集号, Vol. J78-D-I, No. 8, pp.724-734 1995.
- 4) 戸田, 西田, 高橋, Michell, 山口, “優先度先送り方式による実時間相互結合網用ルータチップの実現と性能”, 情報処理学会論文誌 Vol. 36 No. 7, pp.1619-1629, 1995.7.
- 5) ART-Linux : <http://esd.etl.go.jp>
- 6) CODA & Compact Processor Card :
<http://www.etl.go.jp/~realtime>