

並列推論マシン PIM/c

— 概要 —

後藤厚宏^{*1}、瀧 和男^{*1}、中川貴之^{*2}、杉江 衛^{*2}

(^{*1} 新世代コンピュータ技術開発機構、^{*2} 日立製作所 中央研究所)

1. はじめに

ICOT では第5世代コンピュータの中核となる並列推論マシンの研究開発を進めている [1]。本研究においては、核言語 (KL1)、オペレーティングシステム (PIMOS) とともに、論理型言語の枠組の基で並列推論マシンを開発することにより、並列ソフトウェアからハードウェアまでの一貫性を高めることを重視している。これにより、システム開発全体の見通しが良くなると共に、高いシステム性能が可能となる。

並列推論マシン PIM/c (Parallel Inference Machine / type c) は、第5世代コンピュータのプロトタイプハードウェアシステムを構成するモジュールの一つである。本マシンの研究開発においては、並列処理にとって重要な負荷分散方式の研究を中心として、今後の本格的な並列処理研究のベースとして利用できるように、高い実質性能を目指している。本報告では、並列推論マシン PIM/c の全体構成と特徴について述べる。

2. PIM/c の全体構成

PIM/c においては、図1に示すように、要素プロセッサ (PE) をクラスタによって階層的に接続し、全体で256台までのPEを結合する。

各クラスタは、共有バス / 共有メモリによって密に結合された8台のPEとクラスタ間にわたるKL1の並列処理を実行するCC(Cluster Controller)からなる。各PEとCCは、タグアーキテクチャのプロセッサであり、マイクロプログラム制御によって、KL1の高機能機械語命令 (KL1-B) を実行する。また、KL1向きに設計された一貫性キャッシュを持ち、クラスタ内では一つのアドレス空間を共有し、レスポンスが高速で通信コストが小さいプロセッサ間の並列処理が可能である。

クラスタ間は、各クラスタのCCをクロスバスインターフェースを介して接続している。CCは、他クラスタとの間の負荷分散操作、および、各PEの要求に従って、メッセージ通信によるクラスタ間に渡る分散ユニフィケーション操作を行う。クラスタ毎にアドレス空間が別であるため、メッセージ通信によるクラスタ間の通信コストはクラスタ内における共有メモリを介した通信に比べて通信コストが高くなるが、クラスタ毎のメモリ管理の独立性、クラスタ

数の拡張性が得られる。

PIM/c のように、数100台のPEを接続する大規模並列マシンを本格利用するためには、システムの立ち上げ、保守、調整を効率的に行うことが重要である。PIM/cでは、SVP(Service Processor)からの指示に従って各PEとCCのメンテナンスを行うSU(SVP Interface Unit)を設け、メンテナンス・デバッグ機構を階層構造とすることにした [6]。

図2に示すように、PEおよびCCのCPU部(EU)とキャッシュ部(BU)は共通である。これに加え、CCはネットワークを介した通信処理のためのNU(Network Interface Unit)を持つ。一方、PEは標準のSCSIインタフェースを持ち、複数のPEにユーザインタフェースのためのFEP(Front End Processor)やディスク等の入出力機器を接続することができる。

各PEまたはCCは、各要素(EU, BU, NU, およびSU)のLSI、マイクロプログラム記憶、キャッシュデータメモリ等の高速RAMからなり、一枚のプリントボードに実装されている。

3. PIM/c の特徴

(1) 水平マイクロ命令制御のタグアーキテクチャ

PIM/cのPEおよびCCは、マシンサイクル50nsで実行される水平マイクロ命令制御のEU [3]を持つ。レジスタ、内部バスを含め8ビットタグ+32ビットデータのタグアーキテクチャを採用している。マイクロ命令長は104ビット幅とすることにより、KL1の高速実行に必要な動的データ型判定とデータ処理の同時実行を狙った。

(2) KL1向き一貫性キャッシュと排他制御機構

クラスタ内での処理効率を向上させるために、各PEとCCにライトバック型の一貫性キャッシュ(BU)を設けた [4]。本キャッシュは、KL1の並列処理の特性を活かしてバストラヒックを節約するキャッシュコマンドを有する [2]。また、各PEのキャッシュのブロック状態を利用することにより、共有メモリアクセスの排他制御オーバーヘッドを小さくしている。一方、共有バスの高いトラヒックに備えて、キャッシュブロック単位の2重化インターリーブ構成によってバスを二重化し、転送能力を高めた。

Parallel Inference Machine PIM/c - Global Structure -

*Atsuhiko Goto, *Kazuo Taki, **Takayuki Nakagawa, **Mamoru Sugie

*ICOT **Hitachi, Ltd.

FEP: Front-End Processor
CE: Controlling Engine
PE: Processing Element

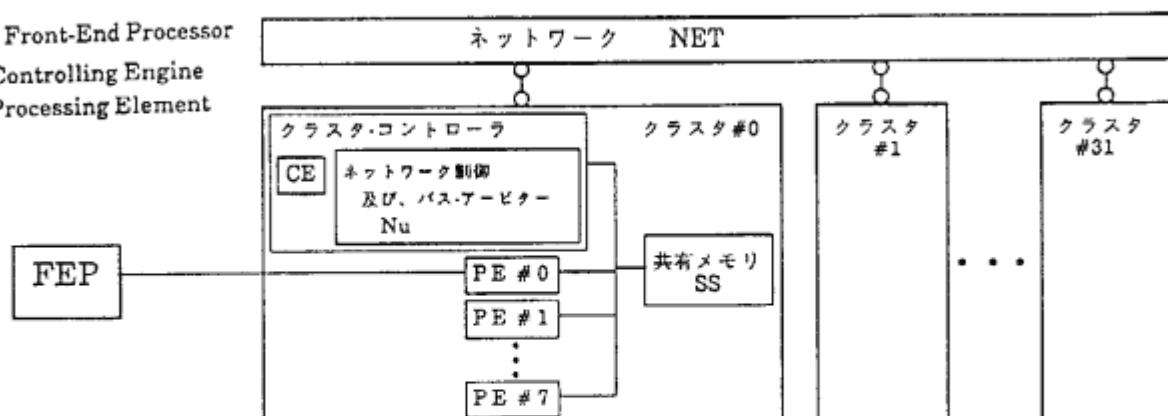


図 1: 並列推論マシン PIM/c の全体構成

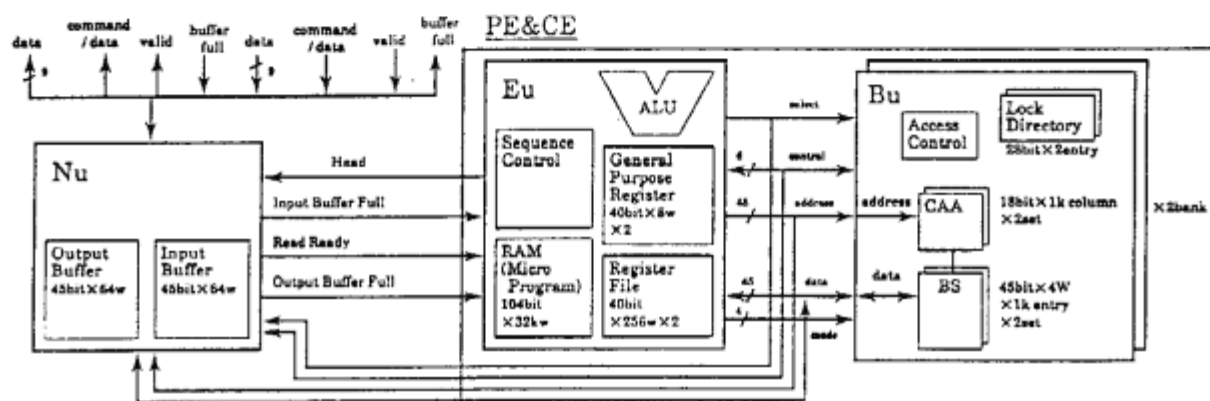


図 2: PIM/c の要素プロセッサおよびクラスタコントローラ

(3) クラスタ内のシグナル通信の支援

クラスタ内の KL1 並列実行では、負荷分散要求や、ガベージコレクションの起動要求等のために、クラスタ内の PE (CC) 間のシグナル通信が必要となる。PIM/c では、共有メモリの排他制御を用いずに、一種のソフトウェア割り込みとしてシグナル通信を行う機構を設けている。シグナルは 16 種類まで使い分け可能であり、PE 間の 1 対 1 通信とクラスタ内での放送が指定できる。

(4) クラスタ間負荷分散支援

クラスタ間に渡る非同期メッセージ通信によって KL1 の並列処理を進める上では、クラスタ間の効率的な負荷分散が重要である。このために、CC 内の NU (Network Interface Unit) と RU (Router) の双方に、負荷情報を高速に伝搬する機構を設けている。

4. おわりに

並列推論マシン PIM/c の全体構成と主な特徴について述べた。今後、試作機を用いて試験評価、改良、拡張を進め、目標とする大規模システムの開発を進める予定である。

謝辞: 日頃、ご指導ご助言を頂く ICOT 内田第 4 研究室室長、ならびに ICOT と日立の PIM 研究開発メンバーに感謝致します。

参考文献

- [1] A. Goto et al. Overview of the Parallel Inference Machine Architecture (PIM). In *FGCS 1988*, pages 208 - 229, Nov. 1988.
- [2] A. Goto et al. Design and Performance of a Coherent Cache for Parallel Logic Programming Architectures. In *16th ISCA*, pages 25 - 33, May 1989.
- [3] 田中 他. "並列推論マシン PIM/c - CPU について -". 本大会発表予定.
- [4] 垂井 他. "並列推論マシン PIM/c - キャッシュメモリについて -". 本大会発表予定.
- [5] 井門 他. "並列推論マシン PIM/c - 負荷分散支援機構 -". 本大会発表予定.
- [6] 前田 他. "並列推論マシン PIM/c - メンテナンス・デバッグ機構について -". 本大会発表予定.