

並列推論マシン PIM/i の概要

2L-7

大原輝彦 武田浩一 佐藤正樹
沖電気工業株式会社

1 概要

筆者らは、第5世代コンピュータ・プロジェクトの一環として、並列推論マシン PIM/i の研究開発を行っている。PIM/i は、第5世代プロトタイプ・ハードウェア・システムを構成するサブモジュール¹⁾の1つである。PIM/i は複数クラスタから構成されており、各クラスタは8台のプロセッシングエレメントから成る。本稿ではクラスタの概要について述べる。

2 アーキテクチャの特徴

PIM/i のクラスタは、共有メモリ型の並列マシンであり、KL1 向き RISC 型プロセッサ[1]とバススヌープ機能を持つ並列キャッシュ[2]を中心とするプロセッシングエレメントおよび共有メモリから構成されている。以下にそれらの特徴を示す。

2.1 プロセッサ

プロセッサは、RISC の概念を取り入れた1チップ CMOS プロセッサで、KL1 向きアーキテクチャを特徴としている。

(1) タグ付きアーキテクチャ

データ型判定の高速処理のためタグ付きアーキテクチャを採用した。

(2) バイブライン・アーキテクチャ

プロセッサは、命令フェッチ、レジスタ読み出しと命令実行、およびレジスタへの結果の書き込みの3段のステージからなるバイブライン構造をしており、実効的に1命令を1マシンサイクルで実行する。

(3) 排他制御

並列プロセッサ環境における排他制御機構を実現するためにメモリロック/アンロック制御命令を持つ。メモリのロックはワード単位に指定可能である。

(4) 複数フィールドの同時実行

命令は40ビット長であり、分岐、メモリアクセス、演算の3つのフィールドから成っており、これらのオペレーションを同時に実行できる。

(5) 汎用レジスタ

40ビットの汎用レジスタを32本もつ。バイブラインのインタロックを防ぎ、かつ3フィールドの命令の同時実行を可能とするため5ポートのレジスタを新たに開発し採用した。

(6) 割込み、例外処理

プロセッサのハードウェアコンテキストの退避/回復の最適化がはかれる機構をもつ。

2.2 並列キャッシュ

プロセッサに命令とデータを効率よく供給し、かつ共有バスのトラフィックを低くおさえるためにバススヌープ機能を備えた並列キャッシュを採用した。以下にその特徴を示す。

(1) 6状態キャッシュ

キャッシュエントリの制御は以下に示す6つ状態を用いて行なわれる。

・ I (Invalid) :

エントリが無効であることを示す。

・ EOC (Exclusive Owner Clean) :

唯一のコピーであり、書き戻しの必要がないエントリである。

・ EOD (Exclusive Owner Dirty) :

唯一のコピーであるが、書き戻しの必要があるエントリである。

・ SUC (Shared Unowner Clean) :

共有されている可能性のあるエントリである。

・ SOC (Shared Owner Clean) :

共有されている可能性のあるエントリであり、キャッシュ間通信の転送元となる。

・ SOD (Shared Owner Dirty) :

共有されている可能性のあるエントリであり、キャッシュ間通信の転送元となる。また書き戻しの必要もある。

(2) ライトバック型キャッシュ

キャッシュエントリのスワップアウトが必要になるまで共有メモリに書き戻しを行わない方式である。

(3) ブロードキャストプロトコル

キャッシュ間の一貫性制御のためにブロードキャストタイプのプロトコルを採用した。この方式では、共有状態のキャッシュエントリに書き込み(ロック操作を含む)を行うとバストラフィックが生ずる。高いヒット率を実現できる反面バスがボトルネックになる可能性がある。このため不必要なバストラフィックが生じないように以下に示すような最適化を行っている。

・ E (Exclusive) 状態を導入し共有状態と区別した。

・ ロック操作オペレーションの競合を検出する機構をつけた。

The Overview of Parallel Inference Machine
PIM/i

Teruhiko OOHARA, Koichi TAKEDA, Masaki SATO
Oki Electric Industry Co., Ltd.

¹⁾ 交信管理サブモジュールと位置付けられている。

- ・プロセッサ側とスヌープ側との競合がなくなるようにキャッシュメモリに2ポートメモリを用いた。

(4) ダイレクトマッピング

アソシアティビティの増大はキャッシュ制御回路の複雑化を招くという問題がある。これにはキャッシュの大容量化とバスサイクルの高速化により対処することにし、ダイレクトマッピング方式を採用した。

2. 3 バスシステム

システムの性能を高めるため共有バスに用途別の2種類のバスを用意した。1つはキャッシュと共有メモリが使用する高スループットバスであり、大きなデータ転送能力を有する。もう1つは、1対多のオンデマンドな非同期メッセージ通信をサポートするバスである。プロセッサ間通信は共有メモリを使用するのが原則であるが、非同期メッセージ通信が必要となる場合がある。そのような場合、このバスを使用することにより、高スループットバスのトラフィックや、並列キャッシュに不要な共有領域ができるのを軽減することができる。

2. 4 メモリシステム

PIM/iではKL1を効率よくサポートするため以下に示すようなメモリシステムを構築した。

(1) 共有メモリ

プロセッシングエレメント間で共有するメモリである。さらに、メモリ管理専用のハードウェアを用意しセグメント・ページング方式により、1TB（テラバイト）の記憶空間を提供している。なお、メモリのロックビットはデータのタグフィールドに割り当てた。

(2) アドレス変換

システム・ワイドな単一仮想アドレス空間を導入し物理的なメモリ管理と論理的なメモリ管理を分離できるように考慮した。これによりプロセススイッチ時にTLBやキャッシュのバージを不要とした。

(3) 局所メモリ（命令、データ）

共有バスのトラフィックの低減のため命令、データ局所メモリを用意した。高速に実行したいプログラムや、局所性の高いデータを格納する。

(4) 命令、データキャッシュを分離

命令、データのアクセス競合が最小となるようにキャッシュを分離した。

3 システム構成

PIM/iのハードウェア・システム構成を図1に示す。PIM/iはG-BUSと呼ぶ共有バスに8台のプロセッシングエレメントが接続されている構成をとる。各プロセッシングエレメントは、プロセッサ、データおよび命令キャッシュメモリ、局所メモリから構成され、それらをA3大の基板に実装した。

4 おわりに

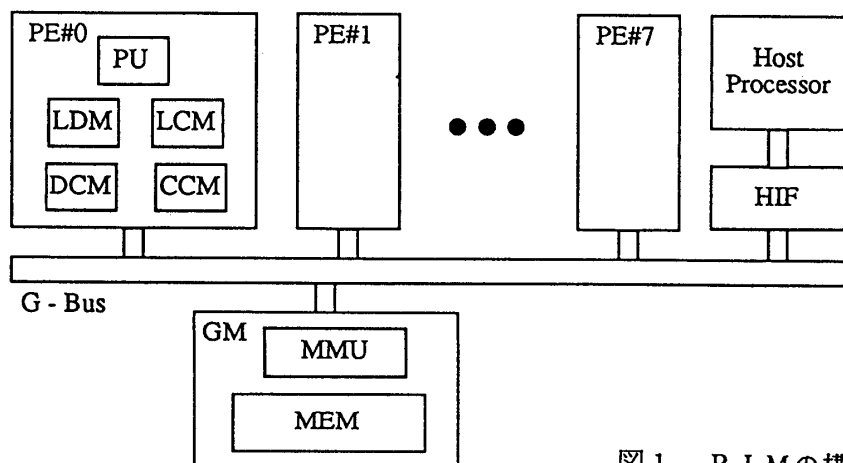
並列推論マシンPIM/iのクラスタの概要について述べた。現在、筆者らは、システムの開発を進めるとともに、ソフトウェア・シミュレータによるシステムの性能評価を行っている。PIM/iおよびシミュレータには性能評価のための機能が多数用意されており、これらを有効に活用し、アーキテクチャを評価し今後の研究に役立てたいと考えている。評価プログラムによる性能評価は別の機会に発表する予定である。

謝辞

日頃助言をいただくICOT第4研究室、および沖電気のPIM担当諸氏に感謝する。

参考文献

- [1] 武田他 並列推論マシンPIM/iの要素プロセッサのアーキテクチャ 本大会予稿集
- [2] James Archibald and Jean-Loup Baer 'Cache Coherence Protocols: Evaluation Using a Multiprocessor Simulation Model', ACM Trans. on Computer Systems, Vol 4, No 4, Nov. 1986



PE : Processing Element
 PU : Processing Unit
 LDM : Local Data Memory
 DCM : Data Cache Memory
 LCM : Local Code Memory
 CCM : Code Cache Memory
 GM : Global Memory
 MMU : Memory Management Unit
 MEM : MEMory
 HIF : Host InterFace

図1 PIM/iの構成図