高並列推論エンジン PIE64 研究経過報告 − ハードウェア −

7M−3

日高　康雄　小池 汎平　高橋 栄一　島田 健太郎　清水 剛　田中　英彦
{hidaka,koike,eiichi,shimada,shimizu,tanaka}@mtl.t.u-tokyo.ac.jp
東京大学 工学部 電気工学科

1　はじめに

我々は，大規模知識処理のための高並列推論エンジン − PIE64[1] − の研究開発を，従来より実機ベースで進めてきたが，このほどハードウェアの開発をほぼ終了し，情報処理学会 第46回全国大会 (本大会) と前後して，全システムが稼働状態に入る運びとなった．本稿では，このPIE64の開発方針と特色をまとめ，ハードウェア開発状況の経過，現状，予定について報告する．

2　PIE64の開発方針

我々の研究目標は，知識処理をはじめとする複雑な記号処理と，数値計算等の性質の良い計算処理とを，区別することなく高速に処理可能な，汎用の高並列計算機システムを実現することである．そこで我々はまず，「**高い並列性の抽出**」と「**記号処理を含む汎用の並列処理**」の二つを，基本方針とした．

我々はこの基本方針に基づき，以下の開発方針を採った．

- **並列処理オーバヘッドの低減**
- **ローカリティ依存度の低減**
- **記号処理オーバヘッドの低減**

並列処理オーバヘッドとは，**通信や同期，負荷分散やスケジューリング**など，並列処理に伴う本質的なオーバヘッドのことである．一般には，並列性の向上と並列処理オーバヘッドの増大はトレードオフの関係にある．システム全体が全力稼働状態になる適度な並列性が抽出できる場合は，さらに並列性を抽出しても無駄であり，粒度を大きくして並列性を適度なところまで抑え，オーバヘッドを低減させることが重要である．しかし，そのような並列性を容易に抽出できるのは，大変性質の良いプログラムだけであり，一般的なプログラムでは，並列性の抽出が困難であったり，例え抽出できたとしても，無理な抽出のために並列処理オーバヘッドが大きくなり過ぎることが多い．

我々は汎用の並列処理を目的としており，むしろ，**適度な並列性の抽出が困難な場合に主眼をおいた**．並列処理オーバヘッドを十分に低減させることで，「並列性が不足する場合は，多少無理をしてでも高い並列性の抽出を優先させる」という**積極的な負荷分散戦略**が可能となる．

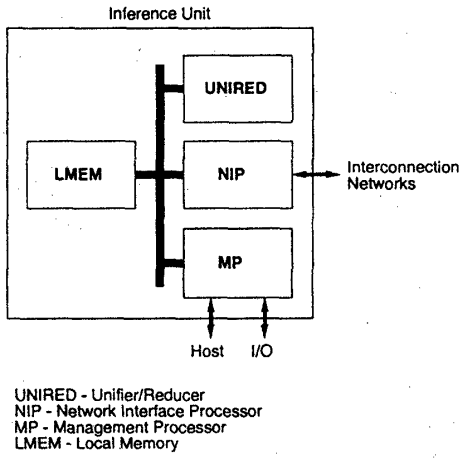

図 1: 推論ユニットの概略アーキテクチャ

一方，並列性の向上は，ローカリティの減少ともトレードオフの関係にある．並列性の抽出が容易なプログラムでは，並列性を抽出しても高いローカリティを保てるが，そうでないプログラムで並列性を無理に抽出した場合は，**ローカリティは期待できない**．また，記号処理においては，ヒープ上を多数のポインタが飛び交うため，数値計算などに比べてローカリティが期待できない．我々は，もしあれば，**当然ローカリティを有効利用**するが，ローカリティの少ない場合にも満足のいく性能が得られるように，**性能のローカリティ依存度**をできるだけ減らす，という方針を採った．

記号処理オーバヘッドの低減には，二つの側面がある．すなわち我々は，**記号処理の実行効率**向上をはかる一方，記号処理以外の処理を行なう際には，その記号処理サポートが**反ってオーバヘッドとならないように**注意する，という方針を採った．

3　PIE64の特色

PIE64は，64台の推論ユニットを2系統の相互結合網で接続した構成を持ち，各推論ユニットは，図1に示すように，推論プロセッサ，通信プロセッサ，管理プロセッサの三種類のプロセッサとローカルメモリなどから構成される．

次にPIE64の特色を，上記の方針に照らして述べる．

相互結合網[2]　負荷分散戦略のローカリティ依存度低減のために，推論ユニット間が等距離となる**多段結合網**を採用し，通信オーバヘッド低減のために，回線交換，ノンバッファリング，32ビット並列双方向データ伝送，4×4クロ

Progress report of highly-parallel inference engine PIE64 − Hardware −
Yasuo HIDAKA, Hanpei KOIKE, Eiichi TAKAHASHI, Kentaro SHIMADA, Takeshi SHIMIZU, and Hidehiko TANAKA
Department of Electrical Engineering, The University of Tokyo

スパーによる段数削減，クロック複相化による接続時間短縮，といった方式で低レイテンシ化をはかった．一方，レイテンシよりスループットを重視する通信用途も考慮し，2系統の相互結合網を用意した．また，負荷分散の際に負荷の転送先を自動選択する動的負荷分散機能を備えることにより，負荷分散オーバヘッドの低減をはかった．

推論ユニット (IU)[3] 並列処理オーバヘッドの本質的な削減をはかるため，本来の計算を行なうUNIREDに，通信・同期処理を行なうNIP，負荷分散やスケジューリングなどの並列処理管理を行なうMPを加えた**三種類のプロセッサによる協調処理モデル**を採用した．また，プロセッサ間の高速コマンドバスによって，コマンド発行オーバヘッドの低減をはかり，3ウェイ4バンク構成のローカルメモリによって，メモリアクセス競合の低減をはかった．なお，ローカリティ依存度を減らすという観点から，IU間に渡るリモートメモリ参照のキャッシュは設けなかったが，ローカルメモリは常にキャッシュヒット時間でアクセス可能となっている．

推論プロセッサ (UNIRED)[4] これは本来の計算を行なうプロセッサで，通信オーバヘッド低減，ローカリティ依存度低減のために，メモリアクセス命令においてローカル/リモートをハードウェアで判定する**分散共有メモリ**，リモートメモリ参照時のパイプライン充填率向上をはかる**サイクル毎の多重コンテクスト処理**といった方式を採った．そして，記号処理の実行を支援する**タグアーキテクチャ**をとる一方，記号処理以外の時のオーバヘッドを抑えるために，マイクロプログラムは使わず，どの命令もパイプラインピッチで実行可能なRISC型命令セットを採用した．また，リモートメモリ参照は，間接参照テーブルを使わない**直接リモート参照**で行ない，テーブル参照・維持オーバヘッドの削減，ローカリティ依存度低減をはかった．間接テーブルを使わないと，記号処理において必須のガーベジコレクション (GC) をローカルに行なうのが困難となるが，その代わりに我々は，効率の良い**グローバル**GCアルゴリズムを開発した[5]．さらに，参照カウンタを用いた実行時即時回収は，本来GCのいらない処理の際にカウンタ維持のオーバヘッドを生じるため採用せず，代わりに，静的解析に基づく**コンパイル時回収**により，ヒープ消費量の低減，GC回数の低減をはかった．

通信プロセッサ (NIP)[6] これは通信・同期オーバヘッド低減のために設けたプロセッサで，**通信処理・単一代入変数による同期処理**をハードウェアで実行し，リモートIU上では，MPやUNIREDによらない**自律的なリモート動作**を行なう．また，グローバルGCにおけるリモートポインタ処理も支援する．

管理プロセッサ (MP)[7] これは負荷分散やスケジューリングなどの**並列処理管理**のオーバヘッド低減のために設けたプロセッサである．現時点ではまだ，様々な管理アルゴリズムの実験が必要であると我々は考え，高速RISCプロセッサであるSPARCとそのファームウェアである並列処理管理カーネルで，MPを構成した．MPはシステム全体で適度な並列性が得られているかどうかを判断し，不足している場合は積極的な負荷分散，足りている場合は保守的な負荷分散と，**負荷分散戦略の動的な切替**を行なう．また，積極的な負荷分散においても，むやみな負荷分散ではなく，コンパイラの出す**静的解析情報**に基づいた戦略をとる[8]．

4 開発の経過，現状，予定

以下に，これまでの開発の経過，現状，予定を記す．

'87年　7月 概要設計開始
　　9月 相互結合網 LSI 詳細設計開始
'88年　7月 相互結合網 LSI 完成
'89年　1月 相互結合網 詳細設計開始
　　4月 NIP LSI 詳細設計開始
　　10月 UNIRED LSI 詳細設計開始
　　12月 IU PCB 詳細設計開始
'90年　4月 相互結合網 完成
'91年　2月 IU PCB(試作版) 完成
'92年　4月 NIP LSI 完成
　　7月 UNIRED LSI 完成
　　9月 IU PCB 量産開始
'93年　1月 25日現在，32台の IU が稼働中
　　2月 IU PCB 量産終了 (予定)
　　3月 全システム稼働 (予定)
　4月以後 チューンアップ，評価等

5 おわりに

以上，PIE64の開発方針と，その方針に基づいて採用されたPIE64の特色，そして，これまでの開発の経過，現状，予定について述べた．

なお，相互結合網のLSI, UNIRED, NIPを製造して頂いた富士通株式会社に感謝する．また，本研究の一部は，文部省科学研究費補助金 特別推進研究 No.62065002 の補助を受けて行なわれたもので，現在は，試験研究 No.03555071 の補助を受けて行なわれている．

参 考 文 献

1) 小池，田中: 並列推論エンジン PIE64，bit 増刊 並列コンピュータアーキテクチャ，Vol.21, No.4, 1989.
2) 高橋，小池，田中: 並列推論マシン PIE64 の相互結合網の作成及び評価，情処論，Vol.32, No.7, 1991.
3) 日高，小池，田中: 並列推論エンジン PIE64 の推論ユニットのアーキテクチャ，信学技報 CPSY90-44, Vol.90, No.144, 1990.
4) 島田，小池，田中: 推論プロセッサ UNIREDII: プロセッサ・アーキテクチャの評価，情処論，Vol.33, No.3, 1992.
5) 小池，田中: 分散メモリ並列計算機上でのジェネレーションスキャベンジング GC，JSPP'90 論文集，1990.
6) 清水: Design Specification of Network Interface Processor，東京大学情報工学専攻テクニカルレポート，TRIE-92-1, 1992.
7) 日高，小池，田中: PIE64 の並列処理管理カーネルのアーキテクチャ，情処論，Vol.33, No.3, 1992.
8) 日高，小池，田中: 実行プロファイルに基づくコミティッドチョイス型言語の静的負荷分割手法，情処論，Vol.32, No7, 1991.
9) Nilsson,Tanaka: Massively Parallel Implementation of Flat GHC on the Connection Machine, Proc. of FGCS88, 1988.