

1 L - 8

PIE64 のネットワーク・インタフェース・プロセッサの
シミュレーションによる性能評価

清水 剛, 小池 汎平, 田中英彦

{shimizu,koike,tanaka}@mtl.t.u-tokyo.ac.jp

東京大学 工学部*

1 はじめに

現在、我々の研究室では、並列推論マシン PIE64[1]の開発をすすめている。

PIE64 の 64 台の推論ユニット - IU (Inference Unit) - それぞれには、他の IU との間で相互結合網を介して、データ交換、プロセス間同期、一括ガベージ・コレクション支援といった並列処理機能を支援する専用プロセッサとしてのネットワーク・インタフェース・プロセッサ [2] [3] - NIP (Network Interface Processor) - が置かれる。

NIP は最大 1 クロック / 1 ワードのデータ転送、および、5 クロック / 1 単位処理のプロセス間同期処理を行なえるように設計されているが、その実効性能は、PIE64 の IU 内のメモリ・バスへのアクセス権の獲得率に左右される。

本稿では、メモリ・バスのアクセス権の獲得率が NIP のコマンド実行時間に与える影響を NIP のクロック・レベルのシミュレータによって評価した結果について報告する。

2 PIE64 の IU 内メモリ・バス

PIE64 の IU 内のプロセッサ (UNIRED [4], SPARC, 4 つの NIP - PAN-Master-NIP, PAN-Slave-NIP, DAN-Master-NIP, DAN-Slave-NIP) は、3 本のメモリ・バスを介して、4 つのメモリ・バンクを共有しているが、それぞれのプロセッサのメモリ・ポートがどのバスに接続されているかは固定であるので、その同じバスを共有している他のプロセッサのメモリ・アクセス頻度の実効効率が左右されることになる。

そこで、NIP のクロック・レベルのシミュレータにより、メモリ・バスのアクセス権の獲得率が NIP のコマンド実行時間に与える影響を測定した。

PIE64 のメモリ・バスは 1 クロック毎にアービトレーションが行なわれ、アクセスはアービトレーションとデータ読み書きの 2 フェーズでパイプライン的に行なわれ、アービトレーション・フェーズで、バスとメモリ・バンクの両方から許可が得られればメモリをアクセス可能となる。

なお、それぞれのシミュレーション結果は、乱数により変化させたパターンを 1000 回測定した後の平均値を示してある。

*Performance Evaluation through simulations of the Network Interface Processor of PIE64

Takeshi SHIMIZU, Hanpei KOIKE, Hidehiko TANAKA, the University of Tokyo

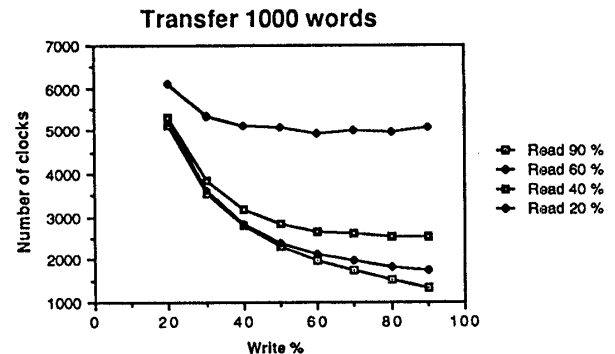


図 1: 1K ワードのデータ転送に要するクロック数の変化

3 データ転送処理の性能

PIE64 のネットワークは 回線交換型の多段網であり、NIP はネットワークの両側でパイプラインを形成する形を取りながら、最大 1 ワード / 1 クロック のデータ転送を可能にしている。メモリ・バスのアクセス権が得られなかった時の乱れは、両側 (読み出し側と書き込み側) のバッファリング用のレジスタを用いて解消する。このバッファリングは 読み出し側で 2 ワード、書き込み側で 3 ワードである。

1 K ワードのデータ転送に要するクロック数を、両側のアクセス権獲得率を変えて測定した結果を 図 1 にしめす (論理的な最小実行クロック数はパイプラインの立ち上がり時間を含めて 1006 クロック である)。

データ転送処理の実行クロック数はネットワークの閉塞、接続先の NIP の状態等にも左右されるが、ここでは、それは考慮せず、単に回線接続後の状態のみを考える。

また、このデータを 1 クロックあたりの実効転送ワード数 (ネットワーク接続後の使用効率) で表したものを 図 2 により示す。

読み出し側と書き込み側では、バッファリングのワード数、状態遷移の仕方が異なるのであるが、それぞれの側のアクセス権獲得率が効率に与える影響にはほとんど差異はない。

4 プロセス間同期処理の性能

プロセス間同期処理では、サスペンション・レコード管理のためのリスト処理時にメモリ・アクセスが必要となる。このようなリスト処理を繰り返し行なうアクティブ処理のサイ

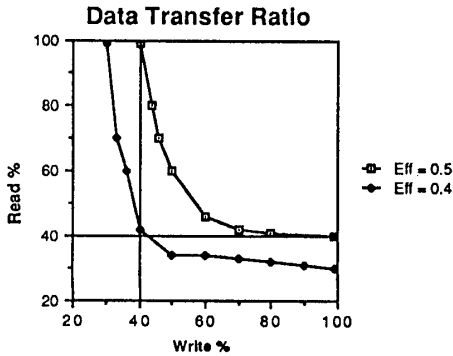


図 2: アクセス権獲得率とネットワークの使用効率

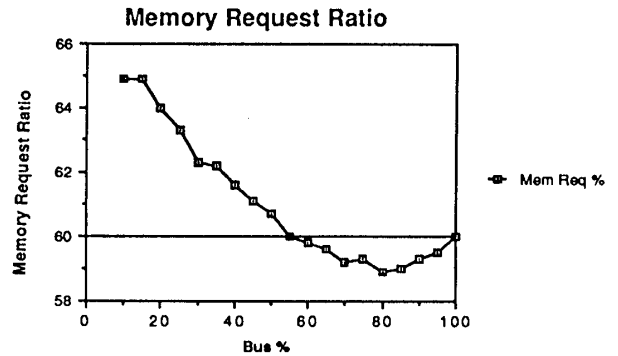


図 4: アクティブ処理時のメモリ・バス要求率

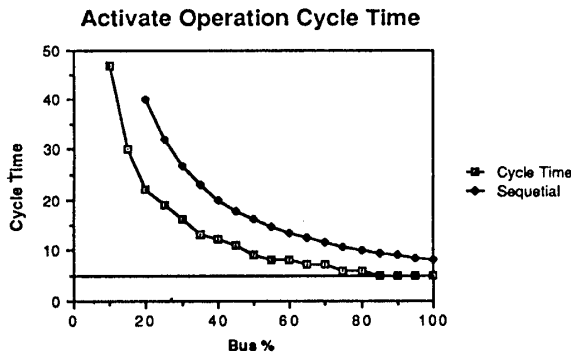


図 3: アクティブ処理のサイクル・タイムの変化

クル・タイム (1つのリスト・セルを処理するのに要するクロック数) がメモリ・バスのアクセス権獲得率によって変化する様子を図3に示した。

NIP は、メモリ・アクセスの順序を可能な範囲で動的に変え、PIE64 で採用したパイプライン式のメモリ・アクセス方式に適合させて、リスト処理の高速化を図っている。これにより、アクセス順序を変えない場合に比較し、平均して 40% 程度、実行サイクル数が減少している。

また、プロセス間同期処理で行なわれるリスト操作とデータ転送が同時に処理される場合、プロセス間同期処理のメモリ・アクセスが優先されるが、プロセス間同期処理のうちのアクティブ処理では、メモリ・アクセス要求の確率はメモリのアクセス権獲得率に対して図4で示すように 60% から 70% 程度におさまっており、優先度固定でも大きな問題は生じていないことがわかる。

5 各 NIP のメモリ・アクセス・パターン

PIE64 のハードウェア上は等価である 2 系統のネットワークのうち、PAN - Process Allocation Network - はゴール負荷分散と、プロセス間同期処理に、DAN - Data Access / Allocation Network - はデータ負荷分散とリモート・データ・リードに基本的に使うものと仮定をすると、IU 上に 4 個あるそれぞれの NIP がメモリ・バスを使用するのは以下のような場合である。

- PAN Master: ゴール負荷分散のベクタ書き込み
- PAN Slave: ゴール負荷分散のベクタ書き込み、プロセス間同期リスト処理
- DAN Master: リモート・データ・リード (長さ ≥ 2)、データ負荷分散書き込み (長さ ≥ 2)
- DAN Slave: リモート・データ・リード (長さ ≥ 1)、データ負荷分散書き込み (長さ ≥ 2)

この仮定と PIE64 での Fleng 実行時の各処理の頻度に関する大まかな仮定とから、各 NIP のメモリへのアクセス頻度は次のような順位になっていると思われる。

$$PANSlave \gg DANSlave \\ > DANMaster \gg PANMaster$$

これらの情報と UNIRED の 3 本あるメモリ・ポートのアクセス頻度情報などを考慮して IU のバス接続を決定する。

6 まとめ

NIP のクロック・レベルのシミュレータにより、メモリのアクセス権の獲得率が NIP のコマンド実行時間に与える影響について報告した。

今後は、応用プログラムを用いたシミュレーションによる評価あるいは、実際に PIE64 を用いた性能測定等を順次行っていく予定である。

なお、本研究は文部省の特別推進研究 No. 62065002 の一環である。

参考文献

- [1] 小池, 田中, “並列推論マシン PIE64 の概要”, 情報処理学会第 37 回全国大会 5N-4, Sep. 1988.
- [2] 清水, 小池, 島田, 田中, “並列推論マシン PIE64 のネットワーク・インタフェース・プロセッサ”, 並列処理シンポジウム '89 A2-2, 情報処理学会, Feb. 1989.
- [3] 清水, 小池, 田中, “並列推論マシン PIE64 の推論ユニット間通信”, 情報処理学会計算機アーキテクチャ研究会 79-5, Nov. 1989.
- [4] 島田, 下山, 清水, 小池, 田中, “推論プロセッサ UNIRED II のアーキテクチャ”, 情報処理学会計算機アーキテクチャ研究会 77-2, July 1989.