

資源共有型並列計算機“砂丘”の アクセス競合緩和法

1L-2

(そのアーキテクチャについて)

¹武田泰明・¹金崎益巳・²加納尚之・¹井上倫夫・¹小林康浩¹鳥取大学工学部 ・ ²米子工業高等専門学校

1. はじめに

我々の研究室では、共有メモリ型マルチプロセッサシステムのプロトタイプとして、プロセッサユニット(PU)を64台実装した並列計算機“砂丘”を製作した。

このタイプの計算機では、アクセス競合によるオーバーヘッドはシステムが大規模になるほど増え、PUの台数に比例した処理速度の向上が得られなくなる。アクセス競合を回避するために、共有メモリの階層化、インターリーブ方式、多重ポート化を行った。さらにメインメモリは、マルチリードワライツ方式を採用し、リードアクセスバスとライトアクセスバスを分離した。

本報告では、並列計算機“砂丘”のシステム構成について述べる。

2. システム構成

図. 1は、当研究室で製作した並列計算機“砂丘”のシステム構成図である。本システムは、64台のPUを実装した密結合型マルチプロセッサシステムである。その構成はバスコントロールユニット(BCU)と4台のPU(単一バスで接続)を1グループとし、それらの4グループをまとめて1ブロックとする。システム全体で4ブロックをマトリクススイッチ(MTX)を介してメインメモリ(MMU)、オメガネットワーク、システムエリアに接続した。

2.1 共有メモリ

共有メモリは、メインメモリとシステムメモリから成る。

(1) 2ポート方式

共有メモリにおいては、バスの切り替え等による無駄時間を少しでも省くために2ポート化した。

メインメモリには、アクセスバスをライトアクセスバスとリードアクセスバスに分けて接続する。このときのアクセスタイミングは、メインメモリの制御部により時分割制御する。

システムメモリには、共有バスを、32台ずつの2グループに分けて接続する。各ポートは、メインメモリと同様に時分割される。

この方法は、PU1台当たりのアクセス間隔は長くなるが、1つのポートが使用されている間に、次のポ

ートのリクエストを受け付けられるので、すぐに次のポートのアクセスを実行できる。これにより、共有バス上での遅延時間がラップされるので効率よく共有メモリを利用できる。

(2) マルチリードワライツ方式

通常のプログラムでは、PUのメインメモリへのアクセスは、リードアクセスの方がライトアクセスより多く行われる。このためメインメモリを4つのユニットに分割し、各PUブロック毎に置く。各メモリユニットのリードアクセスバスは、2インターリーブ方式でMTXの2出力と接続した。このため同じPUブロック内の2台のPUが偶数と奇数のアドレスをリードアクセスする場合、4ユニット全体で8台のPUが競合なしにMMUにリードアクセスできることになる。

またライトアクセスバスは、4つのPUブロックをオメガネットワークで接続する。オメガネットワークの一斉放送モードを用いると、任意のPUが4つのメインメモリユニットに同時に同一の内容のデータを書き込むことができる。このようにすると、64台のすべてのPUで並列処理を行う場合には、PU間で同一のデータを共有できることになる。

2.2 共有バス制御方式

(1) マトリクススイッチ(MTX)

MTXは4×4のバス切り替えスイッチであり、4つの入力のうちどれか1つを4つの出力のどれか1つに接続するものである。4入力はすべてBCUに接続する。4出力は、1出力がシステムエリアのアクセスに、他の3出力がメインメモリのアクセスに使用される。メインメモリのアクセスは、2出力が2インターリーブ構成のリードアクセスバスとして使用され、残りの1出力がライトアクセスバスとしてオメガネットワークに接続される。

(2) オメガネットワーク

ライトアクセスバスにオメガネットワークを使用したことにより、次の3種類の基本利用形態が実現できた。オメガネットワークは、これらの接続モードをソフトでコントロールでき、ジョブまたはタスクレベルで固定することも、ライトサイクルごとに自由に切り替えることも可能である。

How to Decrease Memory Access Contention in The Tightly-Cuppled MIMD Type Parallel Processor System "SAKYU" (Architecture)

¹Yasuaki takeda, ¹Masumi Kanesaki, ²Naoyuki Kanou, ¹Michio Inoue and ¹Yasuhiro Kobayashi

¹Tottori University ²Yonago National College of Technology

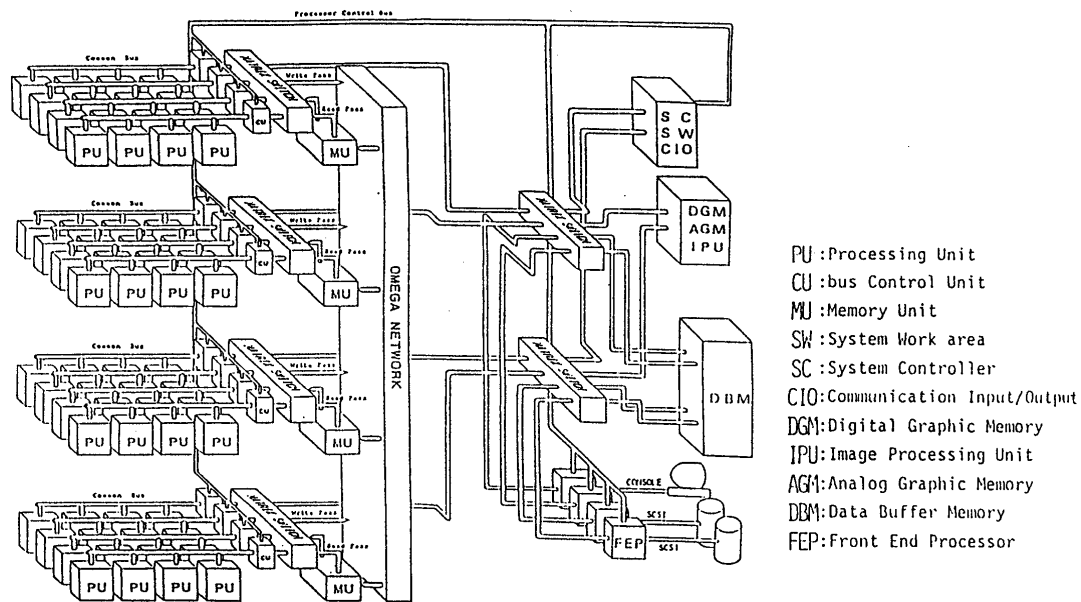


図. 1 システム構成図

a) 一斉放送モード

4つのメインメモリユニットの内容を、同一に保持することができる。したがって、見かけ上のメモリ容量は4分の1になるが、全PUで一斉に同一データに対し処理する場合リードアクセスのアクセス競合を緩和することができる。しかし、全てのPUが1つのメモリユニットにライトアクセスすると同等の動作になるためライトアクセスの競合に注意を要する。

b) バイブラインモード

あるPUブロックで処理した結果を、任意のとなりのメモリユニットにライトアクセスする。そしてライトされたデータに対し新たな処理を行い、結果を更にとり元のメモリユニットにライトアクセスする。この動作を繰り返すことによりタスクレベルでのバイブライン処理ができる。

c) 個別接続モード

それぞれのPUブロックでリードアクセスすると同じメモリユニットに、ライトアクセスする。この形態は、各ブロックが4つのメモリユニットをそれぞれ独立して使用することになり、PUブロック別に分散処理が可能となる。

(3) 非同同期2進木バスアービタ

本システムでは、2入力基本アービタを3つ、2進木状に組み合わせて4入力バスアービタを構成したものを、MTXとBCUに採用している。

具体的には、

・アクセス要求を出しているプロセッサを、速やかに選択する。

・全てのプロセッサが、平等に共有バスを利用できるようにする。

以上のことを目的とし、システムの各部で共有バスのアクセス権を2進木バスアービタによって調停している。

3. おわりに

以上、共有メモリ型並列計算機“砂丘”のアクセス競合緩和方法について述べた。今後の課題は、本システムを使用して各種の具体的な問題に対する評価をすることである。

参考文献

- [1] 加納 他：“マルチマイクロプロセッサシステムの大容量共有メモリの一構成法”
情処研報 Vol.87, No. 45
CA-66-2, pp.1-8, July (1987)
- [2] 井上 他：“マルチマイクロプロセッサシステム“砂丘”の共有メモリアーキテクチャについて”
情処研報 Vol.89, No. 99
ARC-79-2, pp.9-16, Nov (1989)
- [3] 荒川 他：“資源共有型並列計算機“砂丘”
情処研報 Vol.90, No. 90
ARC-85-1, pp.1-6, Nov (1990)
- [4] 瀬崎 他：“資源共有型並列計算機“砂丘”の制御方式の検討
情処研報 Vol.91, No.100
ARC-91-10, pp.77-84, Nov (1991)