

5X-7

ブロードキャストメモリマルチプロセッサシステムの性能評価

呉江荔* 黒岩実* 馬場光弘** 齊藤忠夫*

* 東京大学工学部 ** (株)日立製作所

1. まえがき

共通バスを利用してプロセッサ間の通信を行うマルチプロセッサシステムは今まで多くに研究され、開発されてきたが、バスの容量が限られているため、高々十何個しかプロセッサを搭載できない。このような密結合マルチプロセッサシステム(TCMS)の欠点を改善することを目的として、ブロードキャストメモリマルチプロセッサシステム(BMMS)の研究を進めている。(1)(2)

筆者等は既製のワークステーションを組み合わせたマルチプロセッサに外部的にブロードキャストメモリを付加して、ワークステーションのOSを生かしながらBMMSを実現する方式を検討している。この場合、ローカルメモリとブロードキャストメモリのアクセスの間に大幅な差が生じるので、このような場合のBMMSの性能評価を検討したので御報告する。

2. BMMSの構造

BMMSは図1に示した様に各プロセッサにローカルメモリと共通メモリのコピーを設け、共通メモリをリードしたいとき、共通バスを使わず、このメモリコピーをアクセスすれば済む。その代わりに、これらのメモリコピーの一貫性を保つために、共通メモリにライト時、すべてのプロセッサのコピーに放送する必要がある。

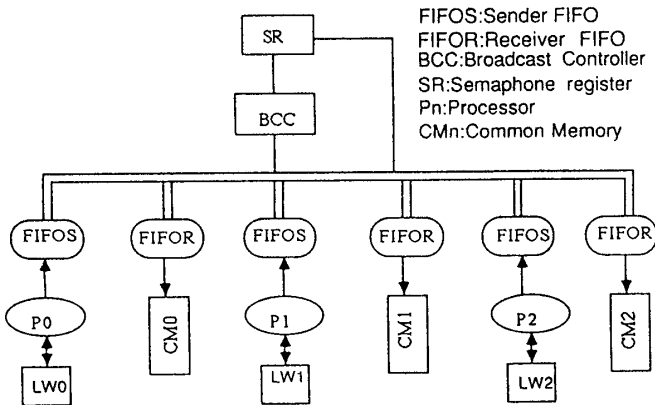


Fig.1. Broadcast Memory Multiprocessor system model

3. BMMSとTCMSの性能比較

BMMSとTCMSの性能比較のために次のような簡単な条件を設定する。

- リード/ライト比 (r) = 8;
- ローカル/グローバルアクセス比 (l) = 4;

- ローカルメモリアクセス間隔 (la) = 300ns;
- 共通メモリアクセス間隔 (ca) = 1μs;
- プロセッサ数 = n;
- 一つの命令が二回のアクセスを含む。

1) プロセッサの増加による性能評価

待合わせを無視して、以上の条件に基づいて、1秒間の実行する命令数がアクセス間隔によって決められる。すなわち、スループット(MIPS数)が以下の式で求められる。

$$\begin{aligned} \text{アクセス間隔 (t)} &= (l*la+ca)/(l+1)=440(\text{ns}); \\ \text{スループット} &= (n*T)/(2*t)=(n*10^3)/(2*440)=1.1*n \end{aligned}$$

(MIPS)

プロセッサ数が増えるにつれて、資源の競合が生じるのはやむを得ない。飽和状態に入ると、スループットの増加速度が小さくなるのは当然なことだが、プロセッサの数はどのぐらいに増やすと、飽和状態になることをBMMSとTCMSについてそれぞれ調べた。

BMMSについて、放送時にしかプロセッサの間の待合わせが生じないので、飽和状態に達する最少プロセッサ数がブロードキャスト間隔によって決められている。

$$\begin{aligned} \text{ブロードキャスト間隔} &= la*r*l=300*8*4(\text{ns}); \\ \text{飽和状態に達する最少プロセッサ数} &= la*r*l/ca=10 \end{aligned}$$

(個); (1)

飽和状態に達してから共通メモリアクセス間隔(ca')と平均的なアクセス間隔(t')を次の式で求められる。

$$\text{共通メモリアクセス間隔 (ca')} = (r*ca+(n-10+1)*ca)/(r+1);$$

アクセス間隔 (t') = (l*la+ca')/(l+1);
n (> 10) 個プロセッサのシステムのスループットは以下の式で求められる。

$$\text{スループット} = (n*T)/(t'*2);$$

TCMSについて、グローバルメモリにアクセスする(リードとライトを含む)時、すべてのプロセッサが共通のメモリをアクセスするので、前者と比べると、競合が生じ易いために、以下の式の通り飽和状態に達する最少プロセッサ数は遙かに小さい。

$$\begin{aligned} \text{グローバルメモリアクセス間隔} &= la*l=300*4(\text{ns}); \\ \text{飽和状態に達する最少プロセッサ数} &= la*l/ca=1; \end{aligned}$$

飽和状態に達してから平均的なアクセス間隔(t')とn (> 1) 個プロセッサのシステムのスループットは次の式で求められる。

アクセス間隔 (t') = $(1+la+ca+n)/(l+1)$;
スループット = $(n*T)/(t'+2)$;
以下の図2が以上の式の結果を示す。

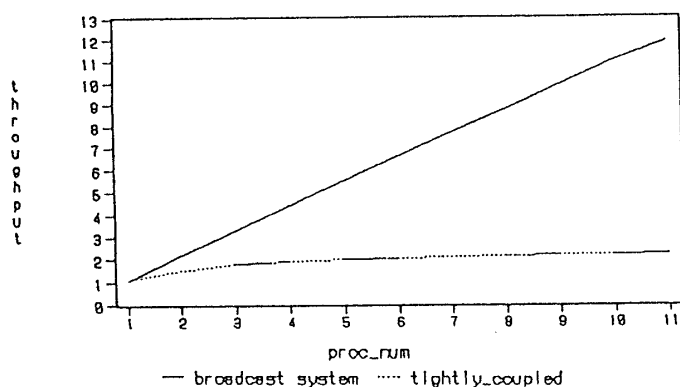


Fig. 2. Comparison of throughput of BMMS and TCMS(1)

2) リード/ライト比による性能評価

以上の仮定に基づいて、飽和状態ではなければ、リード/ライト比がスループットに影響を与えないが、飽和状態になってから、リード/ライト比が小さければ、小さいほど、スループットの増加速度が遅くなる。プロセッサ数が5と15の場合を例として、述べる。上の(1)式によると、プロセッサ数が5の場合では、リード/ライト比が4以下だったら、システムが飽和状態になるが、4以上になれば、システムが飽和状態にならない。プロセッサ数が15の場合では、リード/ライト比が13以上になれば、システムが飽和状態にならない。

TCMSの場合なら、リード/ライト比がスループットに影響を与えないから、プロセッサ数が5と15の場合とも、スループットを示す線は水平になる。両方の場合ともシステムが飽和状態になっているので、プロセッサ当りのスループットは1.1(MIPS)より小さい。

以上の考えに基づいて、下の図3にリード/ライト比によるBMMSとTCMSのスループットの増加傾向を示す。

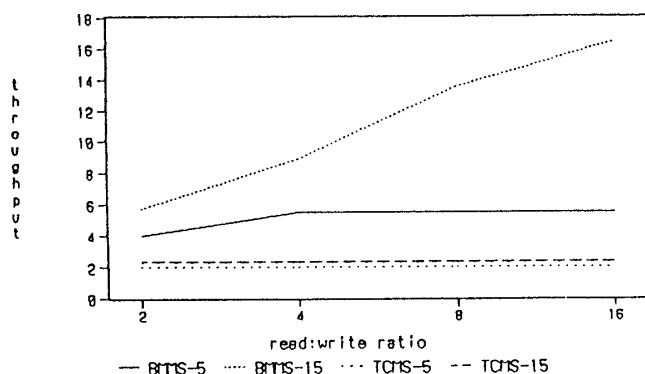


Fig. 3. Comparison of throughput of BMMS and TCMS(2)

3) 具体的な応用例を用いて、両システムを評価

GAUSS-SEIDELによる連立一次方程式を解く例を用いて、両システムの実行時間を評価した。

a) メモリの割当て

求める変数 $x(i)$ はグローバルメモリに格納され、プログラムと $A(i, j)$ (方程式の係数) はローカルメモリに格納される。プロセッサ当りに一つの方程式の計算を分配する。一つの方程式を繰り返し計算するので、プログラムとその方程式の係数を方程式の担当プロセッサのローカルメモリに割り当てる。 $x(i)$ は毎回毎回更新されるので、グローバルメモリに置く。

b) アクセス回数

プロセッサ当り一回繰り返しに当たって、グローバルリード (GR; $x(i)$ をリード) $n+1$ 回、ローカルリード (LR; プログラムと $A(i, j)$ をリード) $3n$ 回、グローバルライト (GW; $x(i)$ をライト) 1回を実行する。

c) 実行時間

前提条件:

1 クロック = $1000/16.67 = 60$ (ns);

かけ算サイクル時間 = 70 (クロック) * 60 (ns) = $4.2 \mu s$;

加算サイクル時間 = 4 (クロック) * 60 (ns) = 240 (ns);

繰り返し回数 = M ; 方程式数 = 変数個数 = N ;

繰り返し毎に実行時間 = $(2*240 + 1000 + 240 + 4200 + 240) * N = 6.16N$ (μs); すなわち、グローバルライトの間隔も $6.16N \mu s$ になる。

トータル実行時間 (競合なし) = $6.16N * M$ (μs)

d) 結論

BMMSの場合では、グローバルライトする時しか競合が生じないので、以上の前提条件を満たしていれば、飽和状態になる最少プロセッサ数 ($p1$) が以下の式で求められる。

$$p1 = 6160 * N / 1000 = 6N \text{ (個)} \quad (>N)$$

上の式から分かるように、もしプロセッサ数が方程式数より多ければ、システムは飽和状態にはならない。

TCMSの場合では、グローバルアクセスであれば、プロセッサの間に待合が生じる可能性が出てくる。同様の前提条件に基づいて、TCMSの場合の飽和状態になる最少プロセッサ数 ($p2$) を求めれば次のようになる。

$$p2 = 6160 / 1000 = 6 \text{ (個)}$$

4. むすび

以上には、理論式と具体例でBMMSとTCMSのスループットを比較したが、あるアプリケーションにたいして、前者のほうは後者より遙かに性能が優れていることがわかる。

参考文献

- (1) 小畑正貴、金田悠紀夫、前川貞男: ブロードキャストキャストメモリ結合形マルチプロセッサシステムの試作、情報処理学会論文誌 Vol. 24 No. 3.
- (2) 阿江忠、高橋浩一、松本建治: 共用メモリ結合によるマルチマイクロプロセッサの並列動作について、電子通信学会論文誌、82/3 Vol. J65-D No. 3.