

密結合マルチプロセッサのバッファ合わせの評価

2Q-4

山本 登 (日本大学 工学部 情報工学科)

1. はじめに

処理装置群と主記憶装置群とをマトリクス・スイッチで接続した密結合マルチプロセッサシステムの実用化研究を進めている。図1はシステムの概要を示すものである。研究ツールとして、図1のシステムの動作をGPSで記述したシミュレータを開発¹⁾し、その上で並列プログラムを走行させている。スイッチ装置 (SWU) には、処理装置 (IPU) 群と主記憶 (MSU) 群のそれぞれに対応して接続制御部 (PIFCとMIFC) があり、P接続路とM接続路を介し接続される。

SWUには各IPUで共用する共有キャッシュ (Common Cache: CC) と、先取りを目的とした後置キャッシュ (Backend Cache: BC) とをそれぞれM接続路対応に設ける。IPU内のキャッシュ (固有キャッシュ: Private Cache: PC) はCCと共に論理アドレスで参照され、MSUのバッファ的性格のBCは実アドレスで参照される。

書き込み動作に対しては、PCはストアスルー方式、CCはストアイン方式を用いる。このため、書き込みはPCとCCに常に行われ、BCへはCCのブロック置換時にのみ行われる。

SWU内のメモリ参照制御部 (MAC) は、PIFCからの要求の受付、スイッチの開閉制御、メモリ参照動作の制御などを担当する。

2. バッファ合わせ機構

作動原理と実現上の問題について概要を述べる。

2.1 バッファ合わせ機構の作動原理

バッファ合わせ機構¹⁾の概要を以下に述べる。

(1) 複写表示語: CCの各ブロック単位にPC

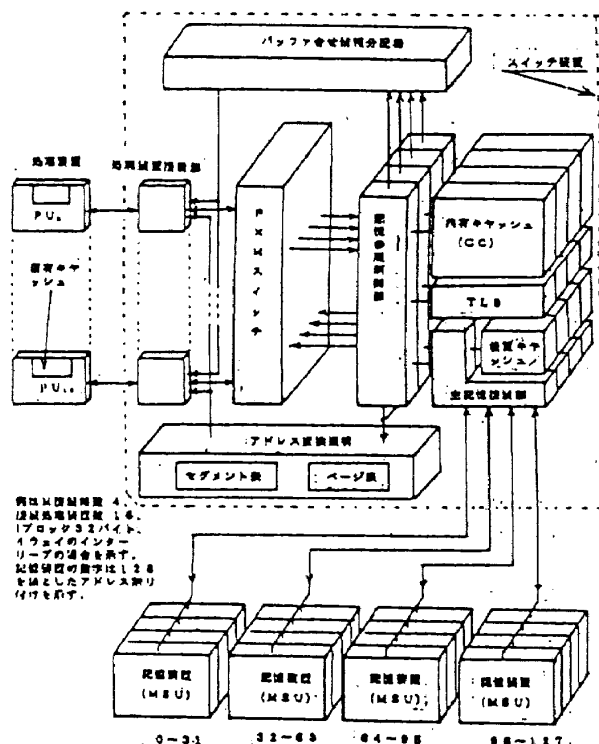


図1. マルチプロセッサシステムの構成

への複写の有無を示す情報をIPU単位に登録し複写表示語を構成する。複写表示記憶は複写表示語を記憶する記憶装置でCC単位に設ける。書き込み時は必ずCCが参照されるため、この複写表示語を利用したバッファ合わせが可能となる。

(2) バッファ合わせ検査: 書き込み時に当該ブロックの複写表示語を読み、複写をもつIPUにバッファ合わせ情報を転送する。

(3) 排他制御: 排他制御の対象となる全ての情報資源は一意的な番号で指定する。占有は Test & Set 指令、解除には Test & Reset 指令を用いる。この間、他IPUからの同一資源の占有は拒否され、当該IPUはハードウェア的に停止する。このための機構はSWU内にM接続路単位に設ける³⁾。

2.2 バッファ合わせ機構実現上の問題

実現する場合の主要な問題を以下に示す。

(1) SWU内のバッファ合わせ情報の転送：

マトリクス・スイッチはバッファ合わせ情報の転送に適しない。IPUからのメモリ参照時のスイッチ動作は1対1接続だが、バッファ合わせ情報の転送は1対n接続が必要となるからである。しかも転送要求元が通常の主記憶参照はP I F Cだが、バッファ合わせ情報の場合はM A C部となるため、同期の点でスイッチ部の制御を複雑にする。

(2) P接続路の排他使用管理： P接続路は主記憶への読み書きに伴う情報転送のほか、排他制御指令やバッファ合わせ情報の転送に使われる。バッファ合わせ情報の転送と前2者とは互いに非同期に発生する。このため同期化への工夫を必要とする。

(3) IPU内の同期： バッファ合わせ情報を受信したIPU中では、主記憶への書き込み動作やブロック転送、資源の排他制御指令などの処理が進行中の可能性がある。これらの処理の終了を待ってバッファ合わせを始めるのでは、遅延による誤動作の恐れがある。このため、バッファ合わせ操作を上述した処理サイクルの間隙で行えるようにした。

3. これ迄に評価した事項

これ迄に評価した事項の概要を述べる。

(1) バッファ合わせ頻発試験： バッファ合わせが頻発する環境を意識的に作り、その状況に耐えられることを確認した²⁾。

(2) 排他制御による共用領域の更新： 共用情報の更新は更新権を得て行う。そこで、Test & Set 命令で更新権を得て更新し、Test & Reset 命令で更新権を放棄する操作が正しくなされること、権利を得ないまま複数のプロセッサで更新すると正しい結果が得られないことなどを確認した²⁾。

4. メモリの資源の型

バッファ合わせ情報の転送遅延の影響を調べるため、情報資源を参照・更新操作の許可の要否により分類してみる。

形式1：随時参照・随時更新資源 参照・更新共に自由に実行できる資源である。メッセージを他プロセスに渡す場合のように、その領域の更新権が

特定のプロセッサに独占的に与えられている場合である。メッセージの到着を通知されたプロセッサも許可を得ることなく情報を参照してよい。

形式2：随時参照・許可式更新資源 参照は随時可能だが更新は許可を得て行う資源であり、更新は順次的となる。参照時に同一資源を更新中の場合があるため矛盾の生じない用法が要求される。

形式3：許可式参照・随時更新資源 更新は随時可能だが、参照には許可が必要な場合である。このような資源の必要性はあっても稀であろう。

形式4：許可式参照・許可式更新資源 参照・更新とも権利が必要な極めて重要な資源である。

形式5：許可式更新資源 更新権を得て更新すべき資源だが、形式2や4と異なり参照利用はない。本システムでは、待行列・IPU管理表・プロセス管理表など頻繁に更新される情報が該当する。

5. おわりに

密結合マルチプロセッサ用バッファ合わせ方式の概要と実現上の問題点を述べ、機能・性能面の評価例を示した。

(1) バッファ合わせ機構はシミュレータ上で走行する各種並列プログラムで正しく動作している。

(2) バッファ合わせを頻発させても、プログラムもシステムも正しく動作することを確認している。

(3) Test and Set による資源占有制御を利用した共用情報の利用法は正しく動作する。

(4) 遅延の少ないバッファ合わせ情報転送方式の実現に意を用いたが、共用情報の更新は順次的で、しかも更新権取得後行うため、 μ 秒程度の転送遅延で誤動作することはないと考える。

【文献】 (1) 山本：マトリクス・スイッチ結合式マルチプロセッサシステムのバッファ合わせの一方式、情学論、Vol.26、o.3、pp.429-437 (1985)。 (2) 山本：密結合マルチプロセッサシステムの cache consistency 維持機構の評価、平成9年度電気関係学会東北支部連大講演予稿集、2H16 (1997)。 (3) 山本：密結合マルチプロセッサ用排他制御方式の一例、平成10年度電気関係学会東北支部連大講演予稿集(1998)。 (4) 山本：密結合マルチプロセッサシステムシミュレータの開発、情学論、Vol.30、No.7、pp.878-894 (1989)。