

1. はじめに マルチプロセッサにおいて、同期と通信のオーバヘッドを削減するために、メモリにプロセッサを直結した SHDRAM¹⁾ や同期と通信を同時に扱う I-structure²⁾ 及び同期通信用メモリ TCSM³⁾ などが提案されている。一方、バストラフィックを削減するために多くのプロセッサはキャッシュをもち、コヒーレント保持のためにバス制御にバックオフ機能をもつ⁵⁾⁶⁾。

ここでは、バスバックオフ機能の基本動作と TCSM に対する読出しおよび書込みブロックへの適用法を述べ、それを MTA/TCSM⁴⁾ 及び MTA/TCSMII に実装し、MTA/TCSMII での TCSM のブロックと同期及び通信に対する評価結果を報告する。

2. バスバックオフ機能の基本動作 バスバックオフ機能はプロセッサの現在のバス転送を中断させ、再実行させる機能である。バスバックオフ機能に関するプロセッサの状態遷移図を図1示す。S₀ はバスに関する制御を解放しているプロセッサの状態、S₁ は有効なアドレスとステータスを出力している状態を意味し、BOFFのアサートはバックオフの実行を、ネゲートはその解除を意味する。

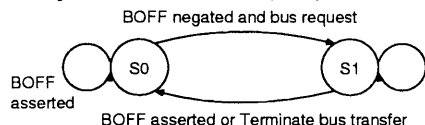


図1 バスバックオフ機能に関するプロセッサの状態遷移図

バスバックオフ機能の応用例としてライトバックキャッシュにおける無効化型の基本的なキャッシュ一致プロトコルとの関係を示す。図2はキャッシュラインの状態とプロセッサか

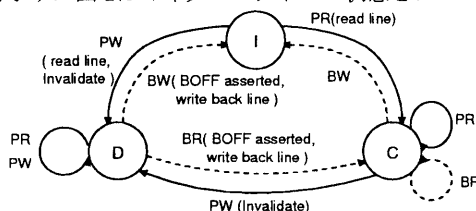


図2 基本プロトコルの状態遷移図

らの要求による状態遷移(実線)及びバスをスヌープした結果による状態遷移(破線)からなり、図2中の記号の意味を表1に示す。図2において、D状態からC状態への遷移途中で

表1 基本プロトコルでの記号の意味

記号	意味
I	無効
C	主記憶との一致
D	主記憶との不一致
PR	そのキャッシュラインを持つプロセッサによる読出し
PW	そのキャッシュラインを持つプロセッサによる書込み
BR	他のプロセッサによる読出しミス
BW	他のプロセッサによる書込みミスおよびヒット
()	状態遷移がバスに引き起こす動作

BOFFがアサートされている。D状態のラインに対して他のプロセッサによる読出しミスが起こった場合、そのラインを

主記憶に書戻す必要があるため、現在バスを占有しているプロセッサのバス転送をBOFFにより中断させる。D状態からI状態への状態遷移においては、現在バスを占有しているプロセッサによるキャッシュフィルが起こる前に、主記憶にD状態のラインを書戻す必要があるためBOFFがアサートされている。BOFFはD状態ラインの書戻し後にネゲートされる。

3. 同期通信用メモリのブロック TCSMは、マルチプロセッサオンチップでのマルチスレッド環境において、相互排除、1対多通信、条件同期、及びバリア同期を統一的に表現できる同期通信用メモリであり、概念図を図3に示す。タグはTCSMエントリの識別に用い、カウンタは通信回数を表し、その非ゼロ/ゼロによってデータの有効/無効を表す。

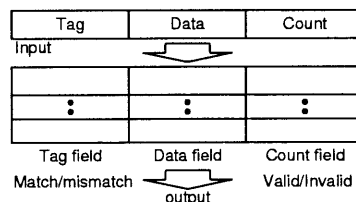


図3 TCSMの概念図

プロセッサがもつバスバックオフ機能によりTCSMを用いた生産者-消費者間の書込ブロックと読出しブロック⁴⁾を実現する。図4にバスバックオフ機能を適用したTCSMアクセスに関するプロセッサの状態遷移図を示す。図4はプロ

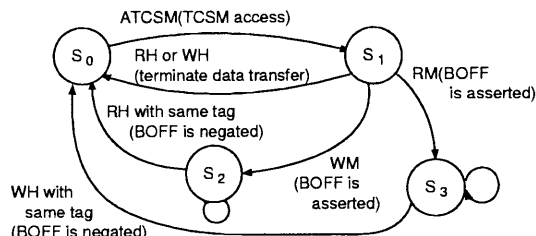


図4 TCSMアクセスに関するプロセッサの状態遷移図

セッサの状態からなっており、記号の意味を表4に示す。S₀のプロセッサがTCSMへのアクセスを開始するとS₁に遷移し、書込みミスが発生した場合S₂に、読出しミスが発生した場合S₃に遷移する。このときBOFFのアサートにより、プロセッサはバスを駆動できない状態になる。S₂状態のプロセッサに対してブロック時と同じタグによるTCSMの読出しが発生し、かつカウンタのデクリメントによって読出し回数がゼロになった場合、BOFFがネゲートされ、プロセッサはS₀に遷移し再びTCSMへの書込みを実行できる。S₃状態のプロセッサに対して、同一タグによるTCSMへの書込

表2 図4における記号の意味

記号	意味
S ₀	ノンブロック状態
S ₁	TCSMアクセス状態
S ₂	書込みブロック状態
S ₃	読出しブロック状態
RM	TCSMに対する読出しミス
WM	TCSMに対する書込みミス
RH	他のプロセッサによるTCSM読出しかつカウンタのゼロ
WH	他のプロセッサによるTCSM書込み
ATCSM	プロセッサによるTCSMアクセス

みもバックオフのネゲートによりプロセッサは S_0 に遷移し、再び TCSM の読出しが可能となる。

4. マルチプロセッサへの適用 MTA/TCSMII は MTA/TCSM において L1 キャッシュの無効化信号を L1 キャッシュに書き込んだプロセッサには送らないよう変更し、かつ TCSM をメモリマップド I/O にした改良機である (図 5)。MTA/TCSMII は 486DX2 を 8 台搭載し、動作周波数は

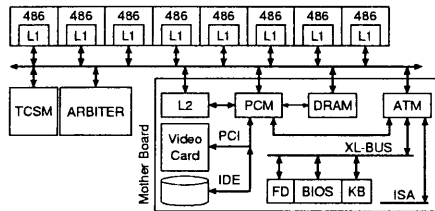


図 5 MTA/TCSMII のブロック図

16MHZ である。ホストバス上に L1, L2 キャッシュ、主メモリ、チップセット (PCM, ATM), TCSM, 集中型バスアービタ (ARBITER) を接続している。L1 キャッシュはライトスルー、L2 キャッシュはライトバックであり、TCSM はキャッシングを行わない。TCSM および ARBITER は Xilinx 社製の FPGA (XC4013E) を用いて実装し、TCSM のエンタリは 16, Tag は 8 ビット、データは 32 ビット、カウンタは 3 ビットである。ARBITER のバス調停に関する優先度は回転式であり、バス使用权の譲渡には HOLD-HLDA プロトコルを用いている。TCSM の書き込み/読み出しブロックに 486DX2 が用意するバスバックオフ機能 (BOFF#) を使用しており、そのコントローラは ARBITER の中に実装した。それらのハードウェア量は、TCSM が約 7000 ゲート、ARBITER が約 2500 ゲート規模である。

5. 実験 複数スレッドによる 1 対多通信と相互排除を TCSM 及びメモリで実装し MTA/TCSMII 上で評価を行った。実験環境は L1, L2 とともにオンであり、プログラムコードとともに L1 キャッシュ、メモリにおける共有データは L2 キャッシュに置いて測定した。MTA/TCSMII の基本データを表 3 に示す。実験で使ったプログラムを図 6 に示す。括弧内の表 3 MTA/TCSMII のリード/ライト性能

	クロック数
TCSM 書込み	4
TCSM 読出し	4
TCSM 書込み/読出しミス	3
L2 キャッシュ書込み	4~5
L2 キャッシュ読出し	4~10
DRAM 書込み	4~7
DRAM 読出し	5~16

数字はプログラム単体の実行にかかるクロック数で、バス転送分も含んでいる。Thread1 は生産者の実行するプログラム、Thread2, ..., p-1 は消費者の実行するプログラムである。各プログラムをスレッドとしそれぞれをプロセッサに割り当てて実行した結果を図 7, 図 8 に示す。図中の実線は、TCSM 版のメモリ版に対する速度向上比、破線および一点破線はプログラムを実行した際のバスアクセスの総クロック数で、前者はメモリ版、後者は TCSM 版である。

双方の結果において TCSM 版はメモリ版に対して良い結果を得た。1 対多通信では、メモリ版はデータ生成のフラグチェックを行うが、TCSM 版はカウンタにより同期が自動的にとられるためフラグチェックによるバス競合を削除できたからである。相互排除では、メモリ版は TEST&SET 命令のスピロロックによりバストラフィックが増大するが、TCSM 版は計算対象データをロック変数にして相互排除自体のオーバーヘッドを削減でき、変数のロックに失敗したスレッドはバスバックオフ機能によりブロックされ、バス競合が低減した

multicast		
	thread1	thread2,...,p-1
TCSM	mov [tag_count],eax (4.5)	mov eax,[tag] (4.5)
Memroy	mov [data],eax mov [flag],1 (9)	A: cmp [flag],0 je A mov eax,[data] (12)

p : number of processors

mutual exclusion	
TCSM	Memroy
mov eax,[tag] ; LOCK	Spin: lock
imul ebx	bts [lock_var],0 ; LOCK
mov [tag_count1],eax ; UnLOCK	jc Spin
(16)	mov eax,[data]
	imul ebx
	mov [data],eax
	mov [lock_var],0 ; UnLOCK (31.5)

ebx register stores thread ID

図 6 実験プログラム

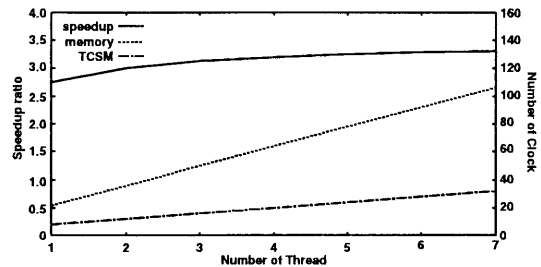


図 7 1 対多通信の結果

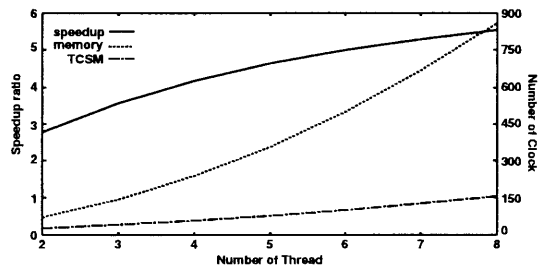


図 8 相互排除の結果

ためである。

6. むすび バスバックオフ機能を同期通信用メモリ TCSM のブロッキングに適用し、MTA/TCSM と MTA/TCSMII 上に実装した。MTA/TCSMII 上で複数スレッドによる 1 対多通信と相互排除の実験を行った結果、TCSM は同期通信のオーバーヘッドとバストラフィックの削減に有効であることがわかった。今後の課題は、MTA/TCSMII におけるマルチプログラミング環境下での評価を行うことである。

参考文献

- 1) K.Hirano: "A New Multiport Memory for High Performance Parallel Processor System with Shared Memory," Proc.of SASIMI '97, pp168-175, 1997.
- 2) Arvind and R.S.Nikhil: "I-Structure: Data Structures for Parallel Computing," Trans. on Prog. Lang. and Sys. ACM, pp.598-639, oct. 1989.
- 3) 岩根他: "CAM による同期通信機構を持ったマルチプロセッサの開発," 九州工業大学研究報告 (工学), pp.43-50, 1999.
- 4) 岩根他: "マルチプロセッサオンチップにおける CAM を用いた同期通信用メモリ," 電子情報通信学会論文誌, J83-D-I, pp.317-328, 2000.
- 5) Intel Corp: "Pentium プロセッサデータブック", 1993.
- 6) Intel Corp: "インテル 486 マイクロプロセッサデータブック", 1992.