

バッファチェック命令により変動するレイテンシに対応する 1 Q-3 プロセッサの性能検討

三竹 大輔* 清水 尚彦**

*東海大学大学院工学研究科 **東海大学工学部

1 はじめに

近年、数値計算分野において分散共有メモリマルチプロセッサなどの技術が普及している。これにより、ソフトウェアから予測する事が難しいメモリレイテンシの変動要因が浮き彫りになり、メモリレイテンシを隠蔽する技術が重要となっている。

ハードウェアによる対応方法として、マルチスレッドプロセッサによるスレッド切替えが考案されてきたが、一般的な単一プログラムの性能向上を期待する向きには至っていない。

マルチスレッドを用いずに単一のプログラムで対応する方法として著者らが検討を進めている SCALT[1][2][3]がある。SCALTは、ソフトウェアにキャッシュメモリ相当の資源としてバッファを開放し、バッファへのデータ到着の有無をチェックする専用命令により、レイテンシの変動に対応しようとするものである。

本報告では、SMP構成とした場合のイベントドリブンシミュレータによる性能評価について説明する。

2 SCALT

SCALTは、メモリレイテンシの隠蔽のためにプロセッサ実装の容易性とシステムリソースにスケラブルな性能を目的としたアーキテクチャである。その構成は、図1となっている。

図において、データバッファ(SCALT バッファ)としてキャッシュと同一のメモリ素子をキャッシュと並行に配置する。バッファは、ソフトウェアコンテキストであり、キャッシュにあるメモリとのコンシステンシは要求されない。また、特別な仮想アドレスにマッピングし、通常のロード/ストア命令によってアク

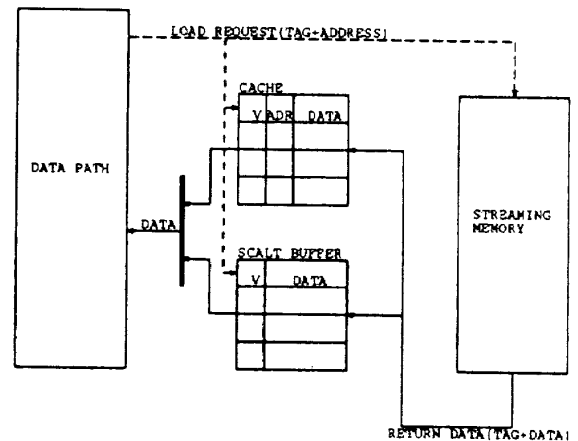


図1: An example load path configuration with SCALT

セスする。この時、バッファ領域はエントリ番号をタグとして付けることにより管理する。プロセッサは、主記憶からの戻りデータを区別するためにタグをメモリ要求に付加して発行し、主記憶からの戻りデータは、タグの値によってバッファの対応エントリに入る。バッファへの転送要求(ロード)またはバッファからの転送要求(ストア)は、主記憶のアドレス、バッファエントリ番号、転送量を指定して発行する。バッファへの転送要求が発行されると、バッファの対応するエントリの有効ビットが解除され、転送が終了すると有効ビットがセットされる。有効ビットが解除されているエントリへのロード命令は、プロセッサをストールさせる。バッファへのデータの到着の有無をチェックするため、有効ビットの値を返す命令として SCALT バッファチェック命令を導入した。バッファを使用する前にバッファの有効ビットを調べる事でプログラムレベルでレイテンシの変動に対応できる。

Performance evaluation of the processor corresponding to the latency deviation with buffer check instruction
Daisuke Mitake Graduate School of Eng., Toukai Univ.
Naohiko Shimizu Faculty of Eng., Toukai Univ.

3 SCALT バッファチェック命令

SCALT バッファチェック命令により、どのように変動するレイテンシに対応するかを図 2 にフローチャートとして示す。

レイテンシの変動によりデータが未到着であるバッファエントリが生じる。バッファチェック命令によりデータが未到着と判別されるエントリに対しては、演算処理を後回しに、全てのエントリに対してチェックが終了した後に演算処理を行う。プロセッサが、未到着リストに載ったエントリに対して処理を行うころにはデータは到着していると考えられ、データの待ち合わせの時間を隠蔽する事になる。

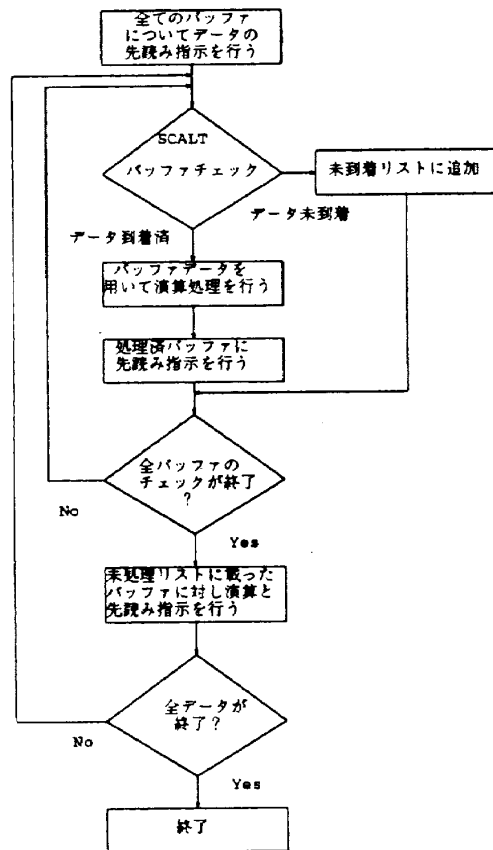


図 2: バッファチェック命令を用いたフローチャート

4 性能検証

ハードウェアの性能検証におけるシミュレーションは、C 言語により記述されたイベント・ドリブン・シミュレータを用いる。評価の対象としては、リバモアループを用い、アドレス生成パターンはハンドコ

ンパイルにより決定する。シミュレーションに必要なハードウェアパラメータは次のように仮定する。

CPU サイクルタイム	5nS
メモリバンク数	16
LSI 間データ転送時間	10nS
SC リクエストピッチ	10nS
メモリバンク幅	16B
バッファエントリ長	32B
メモリサイクルタイム	100nS
メモリアクセス時間	60nS
バッファエントリ数	M

プロセッサ台数は、パラメータとして 1 台から 16 台までを用いる。主記憶装置とプロセッサはクロスバ接続とし、主記憶装置にはメモリバンク毎に 4 つのリクエストキューを用意する。バッファのエントリ数は、パラメータとして複数構成のシミュレーションに用い、リバモアループのシステムスループットを求める。シミュレーションではプロセッサの動作をイベントドリブンに記述し、リバモアループを繰り返し実行させる。また、SMP の各プロセッサは確率分布に従った時間遅れを伴ってループを実行する。SMP のメモリのバンクコンフリクトは頻繁に起こり、SCALT フェッチのレイテンシは動的に変更される。

5 まとめ

SMP 構成時のシミュレーション内容に関して述べた。今後は、性能シミュレーションの実施とプロセッサの設計を進めていく。

参考文献

- [1] 清水, "スケーラブル・レイテンシ・トレラント・アーキテクチャ" IPSJ Sig Notes, Vol.97, No.21, 1997
- [2] 清水, "SCALT / SMP の性能評価" IPSJ Sig Notes, Vol.97, No.76, 1997
- [3] 三竹, 清水 "予測できないメモリレイテンシにソフトウェアで対応するためのハードウェア機構" 1998 年 3 月信学会総合大会