

マルチプロセッサ性能評価システムの開発 (2)

～ソフトウェアトレーサ～

沖 健治

酒井 浩

(東芝ソフトウェアエンジニアリング㈱)

(東芝 総合研究所)

1. はじめに

われわれは通産省第5世代コンピュータプロジェクトの一端として知識ベースマシンMu-X [1]を開発している。Mu-Xは8台のプロセッシングエレメント(以下PE)と共有メモリとして通常の共有メモリとマルチポート・ページメモリを持つ(図1)。PEには汎用のプロセッサMC68020を用いている。また知識ベース処理の処理系はアセンブラで記述している。

今回、Mu-Xの性能評価を専用のハードウェアによる方法とソフトウェアによる方法の2つの方法で行なうこととし、前者のために性能評価ボードを開発し[2]、後者のためにソフトウェアトレーサを開発した。本稿ではこのソフトウェアトレーサについて述べる。

ソフトウェアトレーサは主にMu-Xの知識ベース処理の処理系・並列演算ソフトウェアの評価・改良を目的とした解析ツールで、マルチプロセッサシステムに対応するよう設計した点に特徴がある。以下、このソフトウェアトレーサの設計方針、その機能および構成、実現上の問題点、収集データの妥当性に関する検討を述べる。

2. ソフトウェアトレーサの設計方針

(1) マルチプロセッサシステム対応

Mu-Xはマルチプロセッサシステムであるから、システム全体の評価のために各PEで、同時にデータ収集が行えるようにする。

(2) 操作性

利用者がシステムコンソールを通し、トレーサを集中的に制御して操作性良く使えるようにする。

(3) 評価・改良のツール

ソフトウェアの評価・改良のため下記の項目の測定ができるようにする。

① ビットフィールド操作命令や高機能のアドレッシングモードの使用頻度

これにより並列演算ソフトウェアをアセンブリ言語で記述した妥当性を評価する。(これは性能評価ボードでは測定できない。)

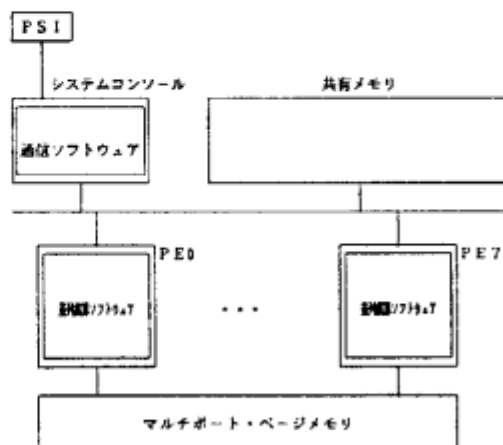


図1. Mu-Xシステム構成図

② データ領域ごとのアクセス回数

これは性能評価ボードでも測定できるが、ここでは例えば、個々のフラグ単位で測定するなどよりきめ細かい設定を可能にする。

③ モジュールごとの実行命令数

並列演算ソフトウェアは多くのモジュールからなる比較的大きなプログラムなので(約5万ステップ)改良する際の指針とする。また知識ベース処理のどこが重い処理になるか評価する。

(4) 並列システムのデバッグ手段の提供

プログラムの実行過程に関するデータを保存して並列システムのデバッグに使う。またこの機能はトレースプログラム自身のデバッグに役立つ。

(5) プログラムの開発の容易化

各PEで動作させるトレースプログラムを同一にする。

3. ソフトウェアトレーサの機能と構成

上記の設計方針を満たすため、図2に示すように8台のPEに対し自分のPE番号を認識して共有メモリの割り当てられたエリアへデータを書き込む同一のトレースプログラムを置く。トレースプログラムには次のデータを収集する機能を持たせる。

① 命令種類ごとの実行回数

② データ領域ごとのアクセス回数

③ モジュールごとの実行命令数

④ プログラムの分岐ログ

さらにそれらを制御するモニタプログラムをシステムコ

ンソール上に置く。モニタプログラムが共有メモリ上にあるフラグを立てると各P Eはトレースモードに入り、それぞれのP Eにおける並列演算ソフトウェアの実行過程に関するデータを同時に収集する。収集したデータは共有メモリのP Eごとに割当てられたエリアに書き込む。書き込まれたデータはシステムコンソールでモニタプログラムによりシステムの動作を妨げることなく表示、もしくはプリンタに出力できる。共有メモリには各P Eにおける並列演算ソフトウェアのその時点での動作状況（アドレス/データレジスタ、プログラムカウンタ、コンデションコードレジスタの内容）も格納しモニタできるようにしている。

4. 実現上の問題点

本ソフトウェアトレサでは命令をトレースするために通常その命令を別の場所へコピーして実行する方式とした。また実効アドレスの計算にはL E A命令を構成して用いる。これらは処理の速さを考慮してのことであるが、自分で自分のプログラムを書き換えて実行するためそのままではキャッシュに捕捉されていた書き換える前の命令が実行されることがある。そのため上で述べた方式でソフトウェアトレサが命令キャッシュの影響無しで正常に動作するように、内蔵キャッシュの構成と動作を検討し、命令を書き換える部分から256バイト後の番地でNOP命令を実行する事により確実に書き換える前の命令を追い出すようにした。

5. 収集データの妥当性に関する検討

一般にトレサ動作時の見かけのCPUの処理速度はトレース無しの時と比べ数10倍遅くなるのに対し、OSの実行や入出力装置の動作速度は変わらない。そのため、ソフトウェア・トレサを組み込んだシステムでは、収集したデータが実システムの動作を正しく反映しているかどうか検討する必要がある。

Mu-Xの場合、下記の理由により収集したデータは、実システムの動作を比較的良く反映していると考えられる。第1の理由は、制御ソフトウェアは各P E上でシングルタスクでビジーウェイトを行なうプログラムであり、マルチタスクの場合と異なりタスクスケジューリングの影響を受けないことである。第2の理由は、制御ソフトウェアが使用するOS機能はファイル操作（ディスク装置のアクセス）のみであり、OSの実行速度の影響が分離しやすいことである。第3の理由は、Mu-Xの入出力装置にはディスク装置とマルチポート・ページメモリがあるが、このうち後者の動作は実システムについても十分高速であり、それがさらに数10倍は高速になったとしてもほとんど影響が無視できるためである。

従って、トレース付きのシステムと実システムの動作の

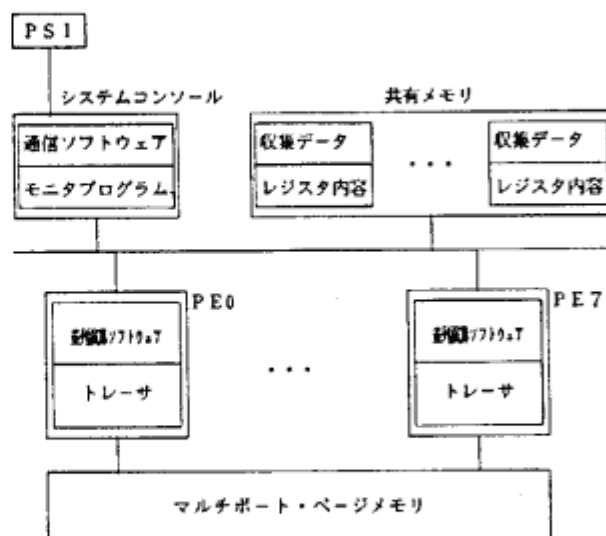


図2. トレーサによる評価システム構成図

違いは、ファイル操作（ディスク装置のアクセス）が見掛上数10倍高速になる点と、ホストマシンからのキューリ到着待ちのビジーウェイト回数が数10倍少なくなる点に限られると思われる。いずれにせよ今後、トレースつきシステムで収集したデータと性能評価ボードによる実システムの収集データを突き合わせるにより、実システムに対する詳細な評価を行ないたいと考えている。

6. おわりに

Mu-Xの性能評価用解析ツールとしてソフトウェアトレサを開発した。また実際にこのツールを使用した結果性能評価を行なうのに有効であることを確認した。今後性能評価ボードとこのソフトウェアトレサを用いて詳細にMu-Xの評価を行なっていく予定である。

参考文献

- [1] 酒井ほか, “知識ベースマシンMu-X (2) 並列処理のための基本機能”, 第36回情報処理全国大会講演論文集
- [2] 浅野ほか, “マルチプロセッサ性能評価システムの開発 (1) 知識ベースマシンMu-Xへの適用”, 第38回情報処理全国大会講演論文集
- [3] “MC68020ユーザズ・マニュアル”,

CQ出版社