

並列計算機用ハードウェアモニタの実現

5 L-9

Implementation of Hardware Monitor for Parallel Computers

應海 阿刀田央一 富澤眞樹 田村仁

東京農工大学大学院工学研究科

1. 計算機性能測定の実現

計算機システム性能測定では、主に特殊なソフトを被測定システム上で動作させ、システムの状態データを収集するソフトウェアモニタと特別なハードウェア装置で被測定システムのハードウェア状態信号を直接取り出すハードウェアモニタという技術がある。この二種類のモニタはそれぞれ長短がある。

ソフトウェアモニタは実現と変更がしやすい反面、モニタ自身が被測定システムのリソースを使用するため、被測定システムの動作に影響を及ぼす欠点がある。特に集計データの量が多くなると、この影響も大きくなる。一方、ハードウェアモニタでは、被測定システムと分離したハードウェア装置で被測定システム内部のハードウェア信号を収集するため、被測定システムへの影響はほとんどない。

2. 被測定システム

本研究で実現したハードウェアモニタはVME (Versa Module Europe) バスを使う分散共有メモリを持つ密結合型並列計算機を測定対象とした。共有メモリは、各PE (Processor Element) から通常のメモリアクセスで利用できるメモリである。この共有メモリの存在のため、密結合型並列計算機の各PEが共有メモリにある同一変数を利用することができる。しかし、複数のPEが同時に同じ分散共有メモリに参照する時、バス競合が起こる。これは並列計算機の処

理能力を低下させる重要な原因である。このバス競合の発生原因を究明し、並列計算機のハードウェア改良とソフトウェアの最適化に重要な参考データを提供することが、本研究の目的である。

3. システム測定の具体的手法

ハードウェアモニタで計算機システムの評価をする時、通常次の設計方法が採用される。

(1) ハードウェアイベントの検出

被測定システムの測定したいハードウェア状態が変化した時、ハードウェアイベントが発生したと定義する。本研究では並列計算機がVMEバスで特定なメモリ範囲内のバスサイクルを起こすとハードウェアイベント発生とする。このハードウェアイベントの発生を検出し、記録する。この時記録されるデータはすでにハードウェアイベント情報である。そうではない場合は、後でソフトウェアでイベントを割り出す必要がある。

(2) 記録データの保存

記録されるイベントデータの保存は通常二つの方法がある。

- (a) イベントデータをラッチすると同時に集計する計算機に転送する。この方法は記録メモリが少ないが、ハードウェアモニタの設計が複雑になる。
- (b) イベントデータを一時ハードウェアモニタのメモリに記録させ、後に集計する計算機に転送する。この方法は大容量な記録メモリが必要になる。しかし、大容量なメモリが安くなった現在は、実用的な方法である。本研究に作成するハードウェアモニタはこの方法を採用した。

Implementation of Hardware Monitor for Parallel Computers

Hai Ying, Oichi Atoda, Masaki Tomisawa
Hitoshi Tamura

Tokyo University of Agriculture and Technology
2-24-16 Nakamachi, Koganeishi, Tokyo 184, Japan

4. ハードウェア構成と機能

4.1 構成

本研究は VME バスを結合網とする並列計算機用のハードウェアモニタを作成するため、ハードウェアイベント信号とモニタ制御信号はすべて VME バスの信号を利用する。本ハードウェアモニタは次の各部により構成される。Fig.1 は本モニタの構成図である。

- (1)VME バスに接続するコネクタ
- (2)VME バス信号のラッチバッファ
- (3)ハードウェアイベント発生時刻タイマ
(精度 62.5ns, 24bit)。
- (4)イベントデータ記録メモリアドレス生成カウンタ (20bit)
- (5)イベントデータ記録メモリ
(SRAM 512K×64bit)
- (6)制御回路
- (7)集計計算機へのデータ転送回路
- (8)評価結果を表示するソフト

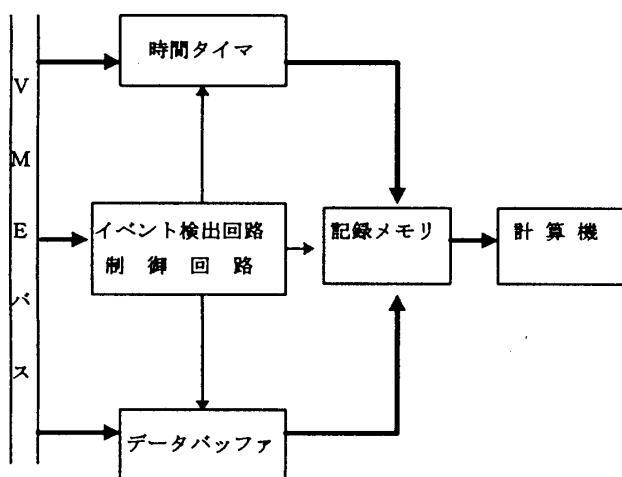


Fig.1 ハードウェアモニタ構成図

4.2 測定機能と方法

本モニタは次のハードウェアイベント情報を記録する。

- (1)共有メモリアドレス (24bit)
- (2)アクセス発生時刻と時間 (24bit)

(3)アクセス属性 (11bit)

本モニタでは3つの測定ルートがある。

- (1)ハードウェアイベントの発生頻度、または発生時刻とイベント経過時間を測定する。そしてイベント経過時間と標準メモリアクセス時間の比較によって、バス競合が発生したかを判断する。
- (2)ハードウェアイベントのステップ実行に対応する。この場合は、イベントが発生したごと、その情報を集計する計算機に送られ、表示される。この方法によって、イベント発生したときの並列計算機ソフトウェアの進行情報を得ることができる。
- (3)特殊なイベントの発生を検出する。検出したイベントの情報を集計計算機のソフトに記入し、(2)の方法を利用して実現できる。

本研究のハードウェアモニタは通常の測定の時、毎 512K 個のイベントデータを記録達した時、割り込みが発生させ、記録データを集計計算機に転送した後測定を再開する。

6. むすび

本研究は計算機システム測定の方法について述べた。そして、VME バスを結合網とする分散共有メモリ型並列計算機システム評価用のハードウェアモニタを試作した。この装置は性能測定だけではなく、測定方法によって、並列計算機のソフトウェアデバッグ道具としても使える。ハードウェアモニタが被測定システムへの影響が少ないが、製作周期が長いと汎用性が乏しい欠点がある。そして、被測定システムを正確に評価するため、記録するデータの選択の重要性が痛感した。これらはこれからの課題である。

7. 参考文献

- (1)箱崎 勝也, 小野 隆喜: ハードウェア・モニタによるシステム測定, 情報処理, Vol.13, No.11, pp782-788(1972)