

研究課題別研究評価

1. 研究課題名： 自律最適化を支援する資源割り当て方式の研究

2. 研究者名： 松本 尚

3. 研究のねらい

単体の計算機の能力を越えた性能を安価に実現するために、複数の計算機を一つの大きなシステムとして使用する計算機クラスタシステムが実用化されつつある。計算機クラスタシステムは、オフィスや研究所にある計算機を安価な標準品のネットワークによって接続して構成される。このクラスタシステムの上で効率良くかつ公平に仕事を行うことを可能にする計算資源の割り当て方式を構築することが本研究の目標である。

4. 研究結果及び自己評価

研究結果

- 1) 各アプリケーションプログラムが自己責任で負荷分散スキームを決断する自由市場原理に基づくスケジューリング方式 (FMM方式: Free Market Mechanism方式) を提案した (*1)。 FMM方式は中央集権的なシステム提供のスケジューラを必要とせず、各マシンレベルで公平性を保つことができる。各マシンレベルの資源競合においてプロセッサ以外の資源競合をプロセッサスケジューリングに反映させることにより、通信やメモリの競合も考慮に入れた公平な資源割り当てが実現できる。
- 2) 上記FMM方式をオペレーティングシステムに実装して有効性を実証するために、研究者が以前より研究開発していたSSS-COREオペレーティングシステムをSun Microsystems社のUltra60に移植し、その上にマイグレーション機能を新規に実装し (*3)、FMM方式を実現した。
- 3) イーサネットの競合回避アルゴリズムであるCSMA/CDからヒントを得たユーザレベルのスケジューリング戦略を考案して (*1)、ライブラリとしてSSS-COREオペレーティングシステムに実装した。
- 4) レイトレーシングとポリゴンレンダリングの2種類の並列画像生成アプリケーションを作成し、上記ユーザ戦略を使って、SSS-CORE上でFMM方式における資源割り当ておよび負荷分散がうまく行くことを実証した。
- 5) FMM方式はプロセッサのみではなくI/O装置を過剰に使うタスクに対してペナルティを課することで、マシン内での公平性を維持しているため、FMM方式はユーザレベルでI/O装置が直接制御できて、I/O装置を高頻度かつ大量に使用する高性能システムの制御に向いている (*1)。そこで、ユーザレベルで直接制御できてマルチタスクに対応可能なユーザレベルI/Oアクセス方式を考案して (*2)、その方式に基づいたマルチメディア対応高速通信カードを設計開発した。

自己評価

- 1) 各アプリケーションに適した資源割り当てを実現するために、市場メカニズムに見られる自由競争原理を計算機の資源管理に導入した。研究者の研究以外にもマーケットドリブンの名乗っている資源管理方式もあるようであるが、それらは入札等の手続きを使うもので、筆者の考案したシステム全般にわたる調停者がまったくいない資源管理システムとはまったく異なるものである。システム全般の状況が低コストで把握できる情報開示機構と各マシンにおけるローカルかつ公平なスケジューリングの組み合わせで実用性のある分散並列システム用の資源管理システムを構築できたことは大きな成果であると考えている。
- 2) マシン内のローカルなスケジューリングに関しても、「使用要求が飽和している資源に関して使用要求を多く出しているタスクにペナルティを課す」という非常に単純な指導原理でプロセッサのみではなくI/O装置まで含めたスケジューリングの公平性を実現したことの意義は大きいと考えている。
- 3) I/O装置を高頻度で使用するシステムにも適用するため、低コストでアクセス可能でありマルチタスク対応可能なマルチメディア対応高速通信カードを自ら作製したのは、ちょっと手を広げすぎであったと反省している。過去に自分で行ったプロジェクトの成果を流用しているとは言え、MPEG2処理LSIの部品入手等に非常に手間取って、FMM方式の実証テストを研究期間内に十分に行う余裕がなくなってしまった。ただし、このカード開発の元となったユーザレベルI/Oアクセス方式の発明は次の研究課題「高性能組込マイクロプロセッサ」への大きなステップになったため、カード作製を計画したことは結果的には非常にプラスではあったと現時点では考えている（計画期間的には無謀だった）。
- 4) クラスタ用の自作汎用オペレーティングシステムを64bit化して、最新鋭機に移植して、マイグレーション機能を追加して、FMM方式を実装した。その一方で、マルチメディア対応高速通信カードのハードならびにソフトの設計と開発を行った。研究期間内にFMM方式の実証試験ならびに評価が十分に行えていない点と、通信カードの全機能がまだ使えていない点（ファームとドライバのソフトが未完成）は、当初計画とは異なるが、FMM方式に関してより深い理解を得て、実用レベルのシステムに実装できたので、当初予定の一番大きな目標は達成できたものと考えている。ユーザレベルI/Oアクセス方式が次の研究に結びついたのみではなく、さきがけで作成したマイグレーション機能からSSS-COREのノンストップコンピューティング化という新たな研究開発課題への広がりが生まれ、こちらも現在プロジェクトが進行中である。さきがけに参加できたことは、他の研究者たちからの刺激も含め、私の研究者人生にとって非常に価値があったと考えている。

5. 領域総括の見解

計算機アーキテクチャにおけるプロセス等の計算資源の効率的利用に関する新しいアルゴリズムを開発し、効率よくかつ公平に仕事を行わせることを可能する計算資源割り当て方式をモデル化し、設計し、実装したことは高く評価できる。アーキテクチャ分野の若手研究者として今後の研究発展が期待できる。

6. 主な論文等

- * 1 ・松本 尚, 平木 敬: 自由市場原理に基づくスケジューリング方式. 電子情報通信学会研究会報告, Vol.99, No.251, pp.63--70 (August 1999).
- * 2 ・松本 尚: アクセス方法及びアクセス処理プログラムを記録した記録媒体. 科学技術振興事業団, 特願平11-255272 (September 1999).
 - ・松本 尚: プロセッサ. 科学技術振興事業団, 特願平11-354203 (December 1999).
 - ・Sasaki, S., Matsumoto, T., Hiraki, K.: On the Schedulability Conditions on Partial Time Slots. Proceedings of 6th Int. Conf. on Real-Time Computing Systems and Applications (RTCSA'99), pp.166--173 (December 1999).
 - ・松本 尚: NICを活用したネットワークRAID方式の提案. 情報処理学会研究会報告 Vol.2000, No.74, pp.79--84 (August 2000).
 - ・大平 怜, 松本 尚, 平木 敬: 汎用クラスタ上の資源情報を用いた HTTP サーバにおける負荷分散性能の評価. 情報処理学会研究会報告 Vol.2000, No.75, pp.31--38 (August 2000).
 - ・松本 尚: 高性能組み込み用プロセッサアーキテクチャの検討. 電子情報通信学会技術研究報告, CPSY, Vol.100, No.248, pp.17--24 (August 2000).
 - ・田中 清史, 松本 尚: 実時間処理 RISC コア Casablanca の評価. 電子情報通信学会技術研究報告, CPSY, Vol.100, No.248, pp.25--32 (August 2000).
- * 3 ・Matsumoto, T.: A Study on Memory-Based Communications and Synchronization in Distributed-Memory Systems. Dissertation Thesis, Graduate School of Science, University of Tokyo (February 2001).
 - ・丹羽 純平, 松本 尚, 平木 敬: ソフトウェアDSM機構を支援する最適化コンパイラ. 情報処理学会論文誌 Vol.42, No.4, pp.879--897 (April 2001).
 - ・松本 尚: メモリベース通信ファシリティの評価. 電子情報通信学会技術研究報告, CPSY, Vol.101, No.329, pp.31--40 (October 2001).