

4F-2

磁気ディスク制御システム向け
マルチプロセッサOSの性能評価

林 秀房 石川 知雄

武蔵工業大学

1 はじめに

近年、マイクロプロセッサは半導体技術により、高速化、集積度の向上が進んでいる。一方、システムやアプリケーションは、機能を充実させていくにつれ負荷が増大するので、より処理能力の高いプロセッサを要求している。このような要求に答えるため、複数の処理装置(CPU)を有機的に結合することによって、システムを構成するマルチプロセッサシステムの利用が有効である。しかし、一般に(密結合)マルチプロセッサシステムはプロセッサの多重度が高くなるにつれ、プロセッサ1台あたりの処理性能が低下する。そこで、各プロセッサに仕事を効率良く分散させること、すなわちタスクのスケジューリングが重要なものとなってくる。

本研究では、マルチプロセッサシステムとして磁気ディスク制御システムをシミュレーションモデルとして用い、密結合型マルチプロセッサシステム用OSをスケジューリングアルゴリズムの観点から性能評価を行なうことを目的とする。

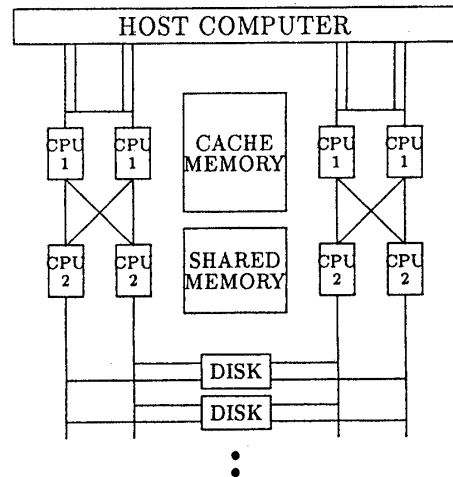


図1: システム構成図

2 磁気ディスク制御システム

密結合型マルチプロセッサシステムのモデルとして、磁気ディスク制御システム^{[1][2][3]}を用いる。このシステムは、計算機のファイルアクセスの高速化の要求に応えるため、その制御システム内にキャッシュメモリを設けることにより、データアクセスを高速化するものである。

このシステムはリードキャッシュ機能、高速ライト機能、スルー処理機能、先読み機能といった特徴を持つ。

2.1 システム構成

チャンネルプロセッサ(CPU1)とデバイスプロセッサ(CPU2)はデータ転送路ごとに独立に配置する。チャンネル、CPU1、CPU2、ディスク装置間はすべて結線されており、任意のCPU1、CPU2を用いてデータ転送を行なうことができる。キャッシュヒット処理はCPU1のみで実行され、PRE-LOAD処理、WRITE AFTER処理はCPU2のみで実行される。構成情報は共有メモリ上に置き一元管理する。このため、タスクを他のCPUへ移動するときのオーバーヘッドを軽くすることができる。

このシステムの構成図を図1に示す。

HOSTからの入出力要求に同期して実行するオンライン処理(HOST-CACHE間)と、HOSTからの入出力要求とは非同期に実行するオフライン処理(CACHE-DISK間)の並列実行を効率良く行なうことによって、システム全体の性能を上げている。

3 シミュレーションモデル

3.1 シミュレーションモデルの構成

2.1のシステム構成に基づいてシミュレーションモデルを作成する。ポアソン分布に従ってジョブを発生させ、任意のチャンネルに送る。チャンネルにジョブが到着すると任意のCPU1に割り込みがかかり、割り込みハンドラがタスクを生成し、ジョブを渡す。

起動されたタスクはジョブの内容に従って、各処理ごとのタスクを生成して処理を行なう。

以下、ジョブとタスクについて説明する。

3.2 ジョブ

シミュレータ上では、READ処理、WRITE処理、WRITE AFTER処理、PRE-LOAD処理といったジョブを扱う。READ処理とWRITE処理を生成し、チャンネルに送る。WRITE AFTER処理とPRE-LOAD処理はCPU2のみで処理されるジョブであり、全体の負荷をみて負荷の少ない時にシステムが生成するものである。

3.3 タスク

3.2で述べたジョブを実行するために、次のようなタスクを起動する。

チャンネルにジョブが到着すると、CPU1に割り込みがかかり割り込みハンドラを起動、この割り込みハンドラが適当なCPU1にRECEIVEタスクを生成する。RECEIVEタスクは受け取ったジョブがREAD処理かWRITE処理かを判断する。

Performance Evaluation of Multi-Processor OS
for a Disk Control System
Hidefusa HAYASHI, Tomo ISHIKAWA
Musashi Institute of Technology

READ 処理の時は、READ タスクを起動する。このタスクは、ジョブの属性を見て HIT タスクを起動してキャッシュヒット処理を、または THROUGH タスクを起動し CPU2 側にも THROUGH タスクを起動してスルー処理に入る。

また CPU2 側では、スルー処理のほかにも、全体の負荷(ジョブの総数)に応じて WRITE AFTER タスクや PRE-LOAD タスクを走らせる。

4 シミュレーション結果

4.1 入力パラメータ

以下のパラメータを使って、READ 処理のジョブに対するシミュレーションを行なった。平均 30[u.t.] のポアソン分布にしたがって発生させたジョブをチャンネルに与える。チャンネル、DISK は 20 個、CPU2 は 4 個とする。リードヒット率を 10, 50, 80% と変化させた場合と、CPU1 を 4, 8, 12 台と変えた場合について、CPU 使用率、ジョブの応答時間、システムのスループットについて調べた。

4.2 スケジューリングアルゴリズム

4.2.1 ノンプリエンプトスケジューリング

最初に、すべてのタスクの優先度を同じにして、プリエンプトが起こらないようにしてシミュレーションを行なった。

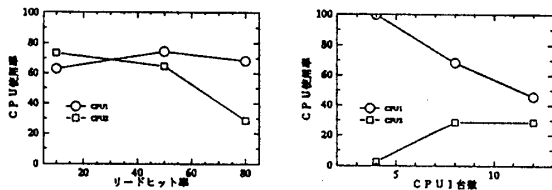


図 2: CPU 使用率

CPU1 の台数が多くなるにつれて CPU1 の使用率は低くなっている。また、CPU1 が 4 台の時は、CPU1 の使用率がほぼ 100% になっている。これは、CPU1 が 4 台では負荷が大きすぎて処理しきれなくなっている状態を表している。このとき、CPU1 が常に Busy なためスルー処理がほとんどできず CPU2 の使用率は非常に小さい値となっている。

4.2.2 優先度スケジューリング

次に各タスクに優先度をつけてスケジューリングした結果を示す。各タスクの優先度は次の通りである。

表 1: タスクの優先度

| | | | |
|---------|----|---------|--------|
| RECEIVE | 10 | HIT | 14 |
| READ | 12 | THROUGH | 14 → 5 |

スルー処理の後半部の優先度を他のタスクより上げているのは、スルー処理は DISK アクセスを伴うので、DISK の SEEK が完了した時にすぐに THROUGH タスクが起動できないとまた DISK 待ちに入ってしまうオーバーヘッドが大きくなってしまいますので、これを防ぐためである。

この結果、応答時間が多少悪くなったものの、スループットが上がっていることがわかった。

4.2.3 タスクマイグレーションが可能なスケジューリング

応答時間を良くするために、タスクマイグレーションを可能にした。4.2.2の方法では、ある CPU1 に RECEIVE タスクが起動すると HIT または THROUGH タスクが終了するまでその CPU1 上で実行されてしまう。そこで、タスクが終了し次のタスクを生成した時、スケジューリングテーブルを見て、一番早く実行されそうな CPU1 に割り当てるようスケジュールした。その結果、わずかだが応答時間は良くなった。

4.2.4 待ち時間を考慮したダイナミックスケジューリング

タスクマイグレーションを行なっても CPU 上の待ち行列にタスクが溜まってくると、優先度の低いタスクはなかなか実行されなくなりレスポンスが悪くなる。そこで待ち時間を考慮したスケジューリングを行なった。タスクが待ち状態になってからの時間を計測し、一定時間以上の待ちとなったタスクは待ち時間に応じて優先度が上げられる。このようにしてタスクが増えてもレスポンスが悪くなるのを防ぐ。

この結果を次に示す。応答時間、スループットとも 4.2.3 とほとんど同じ結果となった。しかし、CPU1 4 台の時の応答時間、スループットがかなり良くなっている。よって、このスケジューリングは、負荷が大きい時に効果があることがわかった。(表 2)

表 2: シミュレーション結果(応答時間とスループット)

| | A | B |
|-------|------------------------------|-----------------------------|
| ヒット処理 | 20099.5 [u.t.] 1544.1 [個] | 8050.4 [u.t.] 2207.4 [個] |
| スルー処理 | 21159.2 [u.t.] 378.9 [個] | 2180.4 [u.t.] 26.4 [個] |

A: 待ち時間を考慮しない
B: 待ち時間を考慮する

5 まとめ

応答時間を見てみると、ノンプリエンプトスケジューリングが一番良くなっている。しかし、スループットを比べると優先度スケジューリングの方が良い。特に、負荷が大きい時は待ち時間考慮の方法が効果的である。よって、このような環境では待ち時間を考慮したダイナミックスケジューリングアルゴリズムが適している。

参考文献

- [1] 佐藤孝夫, 北嶋弘行, 神林公咲: キャッシュ付き磁気ディスク制御装置のマルチプロセッサ制御方式, 情報処理学会第 41 回全国大会 7N-2, 1990
- [2] 横畑静生, 阿部行雄, 中村勝憲: 磁気ディスク制御装置用マルチプロセッサ対応実時間 OS, 情報処理学会第 41 回全国大会 7N-3, 1990
- [3] 猪股宏文, 佐藤孝夫, 北嶋弘行, 四谷守彦: キャッシュ付き磁気ディスク制御装置におけるマルチプロセッサ制御方式の性能評価, 情報処理学会第 41 回全国大会 7N-4, 1990