

仮想マルチコア状態の Linux におけるスピロックの影響

The influence of spinlocks on Linux on virtualized CPUs

安川 要平† 杵淵 雄樹† 中島 達夫†

早稲田大学理工学部 コンピュータ・ネットワーク工学科†

1. 導入

近年、携帯電話や情報家電などの組み込み機器の高性能化に伴い、組み込み機器の限られた資源を有効に活用することが求められてきている。この限られた資源を有効利用する1つの手法として、仮想化CPUを用いたマルチOS環境に注目が集まっている[1]。既存の組み込み機器向けのマルチOS環境として、早稲田大学のSPUMONEがある。SPUMONEとは、シングルコアプロセッサおよびマルチコアプロセッサ上でマルチOS環境を実現するCPU仮想化層のことである。このCPU仮想化層では、1つの物理CPUから複数の仮想化CPUを生成し、ゲストOSに生成した仮想化CPUを提供することが出来る。これにより、例えば、RTOS(Real-Time Operating System)とLinuxを1つのCPU上で並列に動かすことができるようになる。結果として、RTOSがCPU資源を使っていないときに、Linuxがその余ったCPU資源を活用出来るようになる[2]。

しかしながら、SPUMONEを用いてRTOSとSMP(Symmetric Multi Processor) LinuxとでCPUを共有させたところ、RTOSが高負荷な処理をする場合において、SMP Linuxの処理性能が著しく低下した[2]。具体的には、以下のCPU非共有時(図1)と、CPU共有時(図2)とでSMP Linuxの処理性能を比較したところ、RTOSのCPU負荷率が70%を超える状況下では、CPU非共有時よりも、CPU共有時の方が処理性能が低くなった。これは、組み込み機器のマルチOS化の目的であった、CPU資源の有効利用に反している。

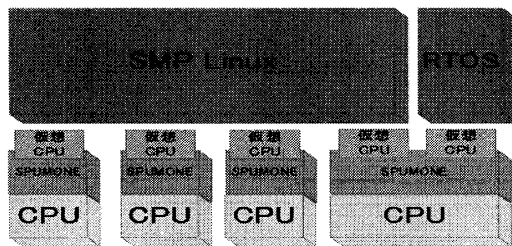


図1. SMP Linux 3 CPUs + RTOS 1 CPU

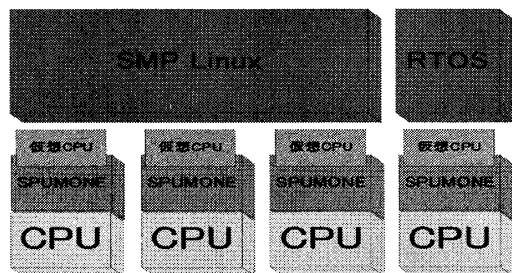


図2. SMP Linux 3 CPUs + 1 Shared CPU

以上の結果を踏まえて、我々はこの処理性能低下の原因について、以下の3つの仮説を立て、評価実験を行った。

仮説1: プリエンプト時のタイマ割込みによる誤差

仮説2: ベンチマークソフト依存の処理性能低下

仮説3: スピロックによる処理性能低下

本稿では、これらの仮説を検証し、処理性能低下の原因について考察する。

2. 評価環境

処理性能低下の原因を調査するため、以前の評価環境[2]と同様に、マルチコアCPUが搭載されているターゲットボードで評価する必要がある。具体的な評価環境を、以下の表1に示す。

表1. 評価環境

ターゲットボード	RP1
CPUアーキテクチャ	SH4A
CPU	600MHz × 4
メモリ	128 MB
ゲストOS	Linux 2.6.16 TOPPERS/JSP 1.3
ベンチマークソフト	Hackbench

3. 各仮説の検証結果

3.1 プリエンプト時のタイマ割込みによる誤差

本仮説では、SMP Linux側のタイマ割込みが行われるCPUが、RTOS側にプリエンプトされたためタイマ割込みが正常に行われず、結果として実験結果に誤差が生じた可能性を検証する。

SPUMONEでは、固定優先度スケジュールに基づいてOS間スイッチが行われるため、RTOSが動作するときは、SMP Linuxの一部のタスクは強制的にプリエンプションされてしまう。このため、OSを共有しているCPU上でSMP Linuxのタイマ割込みが行われた場合、時間が正常に計測できなくなると考えた。この現象を検証するため、OS間で共有しているCPU上でタイマ割込みを発生させて処理性能を評価したところ、RTOSの負荷に応じてベンチマークソフトの実行時間が短くなることが分かった(図3)。したがって、以前の評価実験では、RTOSによるプリ

The influence of spinlocks on Linux on virtualized CPUs

† Yohei Yasukawa, Dept. of Computer Science,
Faculty of Science and Engineering

エンプト時のタイマ割込みによって正しく評価出来ていなかったことが分かる。

この現象を防ぐために、タイマ割込みを OS 間で非共有の CPU 上で明示的に発生させるように変更した。その状況下で再び処理性能を計測したところ、RTOS の CPU 負荷率が 60%以上のとき、CPU 非共有時よりも性能が低下した(図 4)。これは、以前の評価実験時において共有 CPU 上で行われた一部のタイマ割込みが、全て非共有 CPU 上で行われるようになったため、正常に動かなかった分の時間が加算されたからだと考えられる。

しかし、図 4. では、依然として処理性能が低下する現象が見られる。したがって、処理性能が低下した原因は、仮説 2 および仮説 3 のような、プリエンプト時のタイマ割込みによる誤差以外の原因であることが分かる。

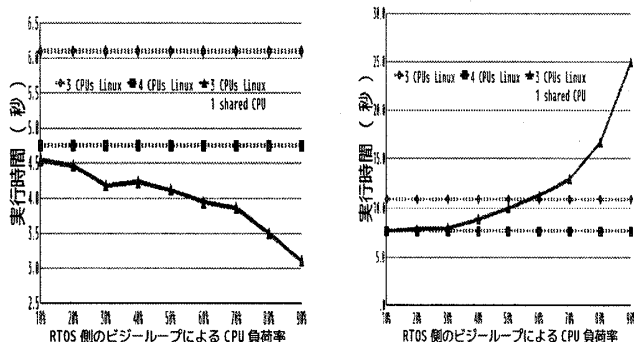


図 3. 共有 CPU 上の計測 図 4. 非共有 CPU 上の計測

3.2 ベンチマークソフト依存の処理性能低下

以前の処理性能の測定では、Hackbench を使って性能の評価を行ったが、処理性能が低下した原因として、Hackbench の内部仕様に強く依存していた可能性も考えられる。

Hackbench の性能評価では、sender と receiver のプロセスを大量に発行し、対となる各プロセス間で IPC(Inter Process Communication)を行う。そして、全ての対となるプロセス間で IPC を終えるまでの実行時間を計測し、出力する。このため、OS 間で共有している CPU 上に割り当てられたプロセスが、IPC のボトルネックとなってしまう、全体の処理性能が低下したとも考えられる。このことから、IPC 以外の方法で処理性能を計測するベンチマークソフトを使い、処理性能が同様に低下するかどうかを確かめる必要がある。

以上の理由から、Dhrystone という整数演算を用いたベンチマークソフトを用いて、再び処理性能の計測を行った。Dhrystone では、Dhrystone 数を指標として処理性能を評価する。この Dhrystone 数が高ければ高いほど、処理性能が高いことを示している。

この Dhrystone を使って処理性能を評価した結果、Hackbench と同様に、RTOS 側が高負荷な処理を行うにつれて、SMP Linux 側の処理性能が低下する現象が見られた(図 5)。したがって、処理性能が低下した原因は、ベンチマーク依存の現象ではないことが分かった。

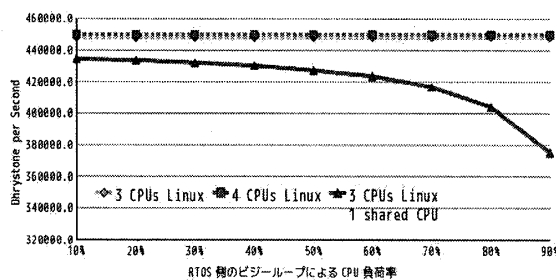


図 5. Dhrystone を用いた処理性能の測定結果

3.3 スピンロックによる処理性能低下

最後に、スピンロックによる処理性能低下が原因である可能性を検証する。3.1 節で述べたように、SPUMONE は固定優先度スケジュールで動いているため、優先度の高い RTOS は、SMP Linux に対してプリエンプションを行うことができる。このため、OS 間の共有 CPU 上に発行されたプロセスがスピンロックを取得すると、RTOS によってプリエンプトされたとき、スピンロックの解放までの時間が長引いてしまう。結果として、そのスピンロックの解放を待つ他 CPU のプロセスにも影響を及ぼしてしまい、全体の処理性能が低下してしまうと考えられる。

なお、Xen を使って仮想化 CPU を生成し、処理性能を評価した関連研究では、スピンロックを取得したプロセスがプリエンプトされることにより、処理性能が著しく低下することが分かっている[3]。ただし、評価環境として使用しているターゲットボードや Hypervisor が異なるため、本稿における処理性能低下の主な原因が、スピンロックによる影響かどうかは未だ断定できない。

4. 将来課題

本稿での検証結果から、スピンロックを取得したプロセスがプリエンプトされたことによって、処理性能が著しく低下した可能性が高い。今後の将来課題として、スピンロックによる待ち時間や、スピンロックが発生した回数を計測するプログラムを実装し、スピンロックのより正確な影響を測る必要がある。

5. まとめ

本稿では、SPUMONE を用いた組み込み機器のマルチ OS 環境において、OS 間で CPU を共有した時に発生する処理性能低下の原因について調査を行った。結果として、スピンロックによる影響により、処理性能が低下した可能性が高いことが分かった。

6. 参考文献

- [1]. An Operating System Architecture for Future Information Appliances, Tatsuo Nakajima et al, Waseda University, 2008
- [2]. 組み込み向けマルチコア CPU 仮想化技術の設計と実装 神田 渉, 早稲田大学大学院基幹理工学研究科, 2008
- [3]. How to Deal With Lock Holder Preemption, Thomas F. and Sebastian B., Advanced Micro Devices Operating System Research Center, 2008