

Xen によるドメイン間にまたがるプロセススケジューラ

田所 秀和[†] 光来健一[†] 千葉 滋[†]

1. はじめに

現在の Xen 仮想マシンモニタでは、物理マシン全体でプロセスを意識したスケジューリングができない。なぜなら、Xen 仮想マシンモニタは Xen における仮想マシンであるドメイン単位でしかスケジューリングできず、各ドメインの中のゲストオペレーティングシステムはその中のプロセスのスケジューリングしかできないからである。

例えば、ドメイン U1 内に重要度の高いデータベースプロセス DB と重要度の低いバックアッププロセス bkup があり、さらにドメイン U2 に重要度が中のウェブサーバプロセス web があるとすると、ここでシステム全体で DB>Web> バックアップという優先度をつけたい。ドメインに U1>U2 という優先度をつけ、かつ U1 内で DB>bkup という優先度をつける場合はうまくいかない。なぜなら、DB プロセスが止まると、bkup が U1 の CPU 時間をすべて使い、bkup と web の優先度が逆転してしまうからである。また、ドメインに U1<U2 という優先度をつけてもうまくいかない。なぜなら、プロセスの優先度は DB<Web になってしまうからである。

そこで本研究では、ドメイン間にまたがるプロセススケジューラを提案する。Xen 仮想マシンモニタがすべてのプロセスをスケジューリングすることにより、ゲストオペレーティング内のプロセスの優先度を物理マシン全体の中で決めることができる。物理マシン全体の中で優先度を決めることにより、ドメインの優先度に依存してその中の優先したくないプロセスまで優先してしまう問題を解決した。

2. 提案と実装

本研究で提案するシステムは、ゲストオペレーティングシステムのランキューを Xen 仮想マシンモニタか

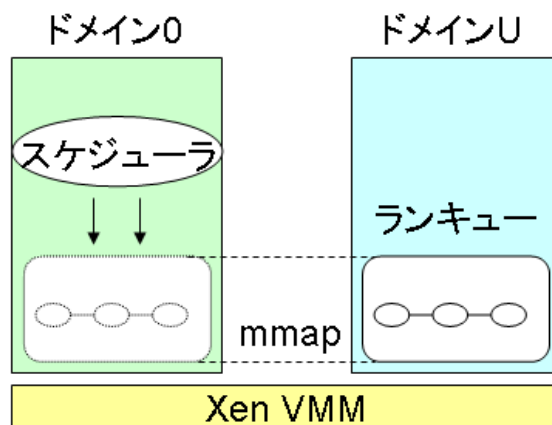


図 1 ドメイン U のランキューをドメイン 0 のプロセスのアドレスにマップする様子

ら操作することで実装している。ゲストオペレーティングシステムのメモリには、Xen 仮想マシンモニタの機能を用いて、コントロール用の仮想マシンであるドメイン 0 のプロセスからアクセスする (図 1)。

あるプロセスの優先度を低くしたい場合、そのプロセスをランキューから外すことで実行を止める。また、必要に応じて外したプロセスをランキューに戻すことで、そのプロセスの実行を再開させる。

ゲストオペレーティングシステムのカーネルイメージを操作するために必要な型情報は、カーネルイメージのデバッグ情報とソースコードから取得する。

ランキューを操作する際、ドメインのカーネルの一貫性を保つように操作する必要があり、カーネルがスケジューラのロックを取っていないことをカーネルのメモリを参照し確認することで実現している。また、外部からカーネルのランキューを操作するので、カーネルのソースコードを変更する必要がない。

3. 実験

実験では、ドメイン間にまたがるプロセススケジューラとして、他にプロセスが動いていない時のみ特定の

[†] 東京工業大学 情報理工学研究所 数理・計算科学専攻
Department of Mathematical and Computing Sciences,
Tokyo Institute of Technology

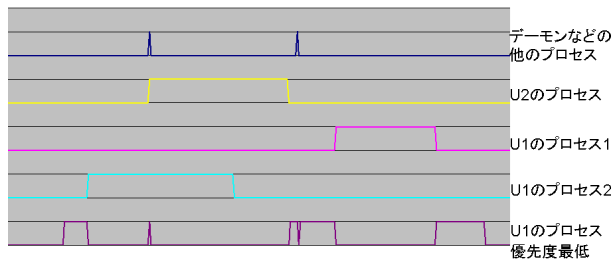


図 2 スケジューリングの様子

プロセスを動かすというポリシーを実装し、実際にスケジューリングが行なえていることを確認した。図 2 は、実験の結果である。グラフは、プロセスが各ドメインのランキューにのっているかどうかを表わしている。線が上になっている時はランキューにのっていることを表し、線が下になっている時はランキューにのっていないことを表す。ドメイン U1 でプロセス 1、プロセス 2、優先度最低のプロセスを動かす、ドメイン U2 でプロセスの一つ動かす。優先度最低のプロセスが他のプロセスが止った時だけ動いている事が確認できる。

4. ま と め

本研究では、複数のドメインにまたがるプロセススケジューリングを提案した。その実装方法として、ドメイン 0 からドメイン U のメモリを操作する手法を用いた。メモリ操作によって、ドメイン U のスケジューラのランキューを操作してドメイン U のスケジューリングを変更する。また、この手法を用いて簡単なスケジューリングポリシーを実装し、実際にスケジューリングが行なえることを確認した。スケジューラはドメイン 0 のプロセスとして実装したので、Xen やドメイン U のゲストオペレーティングシステムのカーネルソースの変更をしていない。

5. 今後の課題

現在の実装では、ゲストオペレーティングシステムのランキューを操作してスケジューリングしているため、ランキューでは掴まえないプロセスを制御するのは難しい。よって、このようなランキューでは掴まえないプロセスを制御するのが今後の課題である。また、これらのプロセスを制御する技術を用いて、より実用的なスケジューリングポリシーの実装をする予定である。