

ネストした VM を用いた仮想化システムの高速メンテナンス

大庭 裕貴[†] 光来 健一^{††}

1. はじめに

近年、計算機の高性能化に伴い、一台の計算機の中で複数の仮想マシン (VM) を動作させ、VM 内でサービスを提供するようになってきている。このような仮想化を行う場合、仮想化システム自体のメンテナンスを行う必要がある。例えば、仮想化ソフトウェアであるハイパーバイザに不具合が見つかった場合には、それを修正するためにパッチを当て、ハイパーバイザの再起動を行う。その際には、ハイパーバイザ上で動作しているすべての VM を一旦停止させ、ハイパーバイザの再起動後に再び VM を起動し直すことになる。多くの VM の停止や起動には時間がかかるため、その間、サービスを提供できないダウンタイムが生じる。

メンテナンス時のダウンタイムを削減するために、VM を別のホストにマイグレーションしてからハイパーバイザを再起動する手法が用いられている。しかし、VM のマイグレーションはシステムやネットワークに大きな負荷をかけるため、マイグレーション中の性能低下が問題となる。VM をマイグレーションするには、VM のメモリの内容をネットワークを介して別のホストに転送する必要がある。すべての VM をマイグレーションするには数 GB~数十 GB のデータを転送しなければならず、動作中の VM に対しても大きな影響を与える。そのため、システム全体への影響を考慮して、マイグレーションの速度を抑える必要があり、メンテナンスに時間がかかる原因となっていた。

本論文では、ネストした VM を用いて仮想化システムのメンテナンスを高速化する VMBeam を提案する。

2. VMBeam

VMBeam では、VM をネストさせることにより、同一ホスト上の二つの VM それぞれで別々の仮想化システムを動作させる。そして、一方の仮想化システムをメンテナンスする際には、もう一方の仮想化システ

ム上に VM をマイグレーションする。この際に、VM のメモリの内容をネットワーク経由で転送する代わりに、同一ホスト上にあることを利用して VM 間でメモリをスワップする。この手法により、マイグレーションの高速化および負荷の軽減を図る。

2.1 ネストした VM

ネストした VM を用いたシステム構成を図 1 に示す。ホスト・ハイパーバイザ上に作成したホスト VM 内でゲスト・ハイパーバイザを動作させ、その上でゲスト VM を動作させる。本研究では、ホスト・ハイパーバイザに Xen を使い、ゲスト・ハイパーバイザに Xen-Blanket¹⁾ を用いた。本研究でメンテナンスの高速化の対象とするのはゲスト・ハイパーバイザであり、ホスト・ハイパーバイザは対象としない。ホスト・ハイパーバイザは最低限の機能だけを提供することを想定しており、メンテナンスの頻度は低いと考えられる。

2.2 VM 間メモリスワップ

VMBeam では、ホスト・ハイパーバイザがホスト VM 間でメモリを高速に入れ替える機能を提供する。この VM 間メモリスワップを実現するために、ホスト・ハイパーバイザは二つのホスト VM へのメモリ割り当てを変更する。Xen において、ホスト・ハイパーバイザはマシンメモリと呼ばれるホスト全体で管理している物理メモリを扱う。また、ホスト・ハイパーバイザはマシンメモリの一部をホスト VM に割り当て、疑似的な物理メモリとして扱わせる。この疑似物理メモリからマシンメモリへの対応は、P2M テーブルと呼ばれる対応表で管理されている。そこで、ホスト・ハイパーバイザが P2M テーブルを書き換え、図 2 の

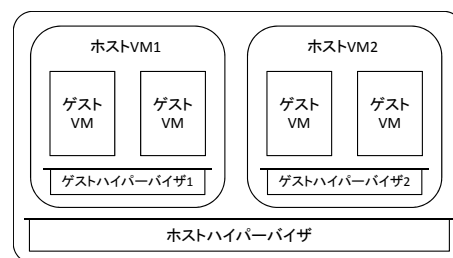


図 1 ネストした VM のシステム構成

[†] 九州工業大学

^{††} 九州工業大学, CREST

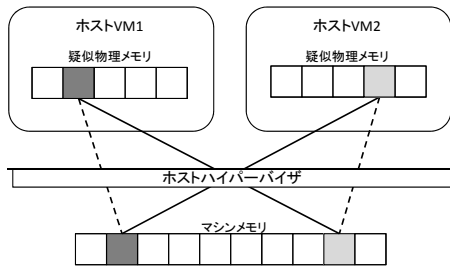


図 2 ホスト VM 間でのメモリスワップ

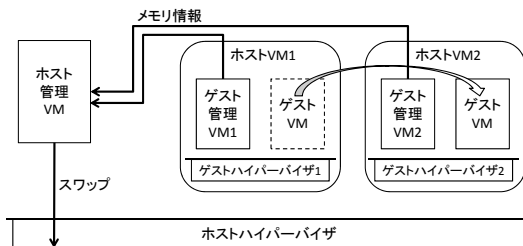


図 3 メモリスワップを用いたマイグレーション

ようなメモリ割り当てに変更することで、二つのホスト VM 間でメモリを入れ替える。

2.3 メモリスワップを用いたマイグレーション

VMBeam では、ホスト VM 間のメモリスワップを用いてゲスト VM のメモリを入れ替えることによって、高速なマイグレーションを実現する。このマイグレーションの流れを図 3 に示す。

ゲスト VM のマイグレーションはゲスト管理 VM と呼ばれる VM 間で行われる。従来のマイグレーションでは、マイグレーション元のゲスト管理 VM がゲスト VM のメモリの内容をマイグレーション先に送っていた。VMBeam ではその代わりに、ゲスト VM に割り当てられている疑似物理メモリの情報だけをホスト管理 VM に送る。同時に、マイグレーション先のゲスト管理 VM は新たに作成したゲスト VM に割り当てた疑似物理メモリの情報をホスト管理 VM に送る。これらのメモリ情報を受け取ったホスト管理 VM は、ホスト・ハイパーバイザを呼び出すことによってメモリのスワップを行う。

3. 実験

ゲスト VM に割り当てるメモリサイズを変えながら、メモリスワップを用いてホスト VM 間でメモリ転送を行うのにかかる時間および CPU 負荷を計測した。比較のために、ホスト VM 間の仮想ネットワークを用いた場合についても計測した。実験には Intel Xeon E5-2665 2.40GHz (8 コア) の CPU、32GB のメモ

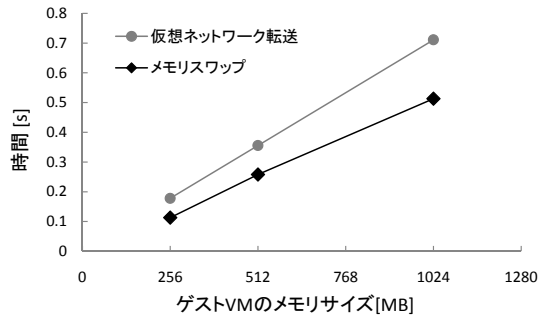


図 4 メモリスワップとネットワーク転送の比較

リを搭載したマシンを使用した。ホスト・ハイパーバイザとして Xen 4.1.3, ゲスト・ハイパーバイザとして Xen-Blanket 4.1.1 を動作させ、ホスト管理 VM では Linux 3.2.0, ゲスト管理 VM では Xen-Blanket 用に拡張された Linux 3.1.2 を用いた。ホスト管理 VM には 2 CPU と 8GB のメモリ、ゲスト管理 VM には 4 CPU と 4GB のメモリを割り当てた。

メモリスワップ時間と仮想ネットワーク転送時間を 10 回ずつ計測した平均値を図 4 に示す。この結果より、メモリスワップにかかる時間はゲスト VM のメモリサイズに比例することが分かる。割り当てるメモリサイズが 1024MB の場合でも、ゲスト VM 間のメモリスワップにかかる時間は 0.5 秒となった。仮想ネットワークを用いた転送と比較しても、メモリスワップの方が 28% 高速に行えることが分かった。

一方、CPU 負荷はメモリスワップの場合は 98%、仮想ネットワーク転送の場合は 346% となり、メモリスワップの方が CPU 負荷を大幅に抑えられることが分かった。

4. まとめ

本論文では、ネストした VM を用いて仮想化システムのメンテナンスを高速化する VMBeam を提案した。VMBeam では、同一ホスト上で二つの仮想化システムを動作させ、仮想化システムのメンテナンス時には VM をもう一方の仮想化システムにマイグレーションする。このマイグレーションは VM 間メモリスワップを用いて高速化する。今後の課題は、VM 間メモリスワップをマイグレーションに組み込む実装を完成させることである。

参考文献

- 1) Williams et al, The Xen-Blanket: Virtualize Once, Run Everywhere. Proc. EuroSys, 2012.