

論文の内容の要旨

利用可能性の高い仮想計算機転送技術

A technique of virtual machine migration with high applicability

氏名 高橋 一志

既存の「遠隔仮想計算機転送システム」は「使う人と場所を選ぶ」システムであり、利用可能性が低いという問題がある。これは、仮想計算機転送システムの技術的な問題に起因する。本研究では、既存の3つの遠隔計算機転送システムに対して、それぞれ、3つの着想 (a) ほとんどすべてのデバイスに組み込み済みの Web ブラウザの JavaScript+HTML 上で動作するインタラクティブ性能の高いシンクライアントプロトコルの提案 (b) Linux / KVM ドライバを、Linux エミュレーションレイヤを用いて「簡単」かつ「性能劣化無し」で Windows 上に移植することにより、Windows と Linux 間での VM ライブマイグレーション (c) VM の物理計算機間の往来という新たな振る舞いに着目した高速な差分転送、を元にして、既存の遠隔計算機転送システムの利用可能性を向上する。

「クラウド・コンピューティング」や「VDI: Virtual Desktop infrastructure」といった技術の普及に伴って「遠隔仮想計算機転送システム」がより重要な存在となりつつある。これらのシステムは、仮想マシンモニタ (VMM: Virtual Machine Monitor) によって生成される仮想マシン (VM: Virtual Machine) を、ネットワークを経由で、ユーザの持つ手元の端末でも利用可能にする技術である。クラウド・コンピューティングや VDI: Virtual Desktop infrastructure の場合、すべてのユーザは、中央集権的に集約された仮想マシンを、この、遠隔仮想計算機転送システムを通して、ネットワーク経由で操作することになる。そのため、このシステムは欠かすことが出来ない。そのため、遠隔仮想計算機転送システムに対して利用可能性の向上という貢献を行うのは重要な意義がある。

今日使用される遠隔仮想計算機転送システムは、画面転送によるシンクライアントシステム、VM そのものを他の物理マシンに転送する VM ライブマイグレーション、および、VM

のストレージも含めて VM を他の物理マシンに転送する VM ストレージマイグレーションの 3 つでありそれぞれに利用可能性を低下させている技術的な問題点が存在する。

まず、シンクライアントシステムの問題点は「導入が面倒」という点である。既存のシンクライアントシステムは、すべて、専用のアプライアンスや専用ソフトウェアのインストールと言う作業が必要になる。計算機に対して詳しくない人にとって、この作業は困難である。シンクライアントの研究は古くから行われているが、すべて専用のソフトウェアやアプライアンスを用いたものである。

また、既存の VM マイグレーション技術にも 2 つの問題点がある。1 つ目の原因は「コンシューマ向け OS とサーバ OS を結びつける VMM が無い」という点である。既存の VMM では、複数の異なる HostOS 上の VMM 間で、VM ライブマイグレーションを達成することができず、利用環境が同一 OS 上の VMM と限定されている。特定の OS 上でしか VM ライブマイグレーションが達成できないというのは利用可能性を低めている。また、2 つ目の原因としては、VM のストレージ転送に時間が掛かるという問題である。従来の VM ライブマイグレーションでは、この問題を解決するため、VM のディスクイメージを共有するために分散ファイルシステムや共有ファイルシステムを利用する必要があった。今までの VM ライブマイグレーションではこの制約が問題になることはなかったが、データセンタ間をまたいだ広域 WAN 環境におけるライブマイグレーションや、個人用の VM に VM ライブマイグレーションを利用するという研究も進んでおり、状況によっては、分散/共有ファイルシステムを利用するのが困難な状況下での VM ライブマイグレーションも珍しくなくなってきている。これを解決するために、VM ストレージマイグレーション技術という、VM のストレージを含めて、すべての VM ステータスを転送するという技術も開発されているが、VM のストレージは巨大であるため、転送には多くの時間がかかり、また、転送に使用する帯域幅も大きなものになる。当然、低速なネットワーク環境であれば、さらに多くの時間がかかり、これでは利用可能性が高いとは言えない。

このように、現存する遠隔仮想計算機転送システムでは利用可能性が低い。そこで、本研究では (1) JavaScript と HTML を用いたシンクライアント VoXY の実現 (2) Linux ドライバの Windows 移植による WinKVM の実装 (3) 差分転送による VM ストレージマイグレーションの高速化という 3 つの研究で、既存のシンクライアントシステム、VM ストレージマイグレーション、VM ライブマイグレーションのという 3 つの研究で、利用可能性を低下させている技術的な制約を取り除き、より多くの人が使えようとする遠隔仮想計算機転送システムを目指す。

(1) の研究により、Web ブラウザという、ほとんどすべてのデバイスの初期段階で導入

されているソフトウェア上で、新たなソフトウェアを導入することなく、使用出来る利用可能性の高いシンクライアントシステムを達成した。さらに、(2)の研究により、Windows と Linux 間といった異種 OS 間で動作する VMM 間で VM ライブマイグレーションを達成した。今までの VMM では、同一の HostOS 間で動作する同一の VMM 間でしか VM ライブマイグレーションを行うことができなかった。そのため、コンシューマ OS として広く普及する Windows や、サーバ用の OS として広く普及する Linux 間で VM ライブマイグレーションを達成することはできず、これは、利用可能性を低下させる技術的な要因であった。これを解決するため、本研究では、KVM と呼ばれる Linux ドライバとして実装されている Hybrid VMM を、Windows に移植することでこれを実現した。Windows と Linux 間でのライブマイグレーションを達成した。移植には Windows 上に Linux カーネルをエミュレーションするエミュレーションレイヤを構築することで KVM ドライバを「簡単」かつ「性能劣化」なく Windows 上で動作させることに成功した。最後の、(3)の研究では、一般に巨大な容量を持つ VM ストレージの転送時間と転送帯域を節約するために、VM の往来と呼ばれる今後 VM ライブマイグレーションに現れるであろう特質を利用した、新たな差分転送のメカニズムを提案した。これにより、従来の VM のストレージを含めた全転送時間を最大で 98%まで削減することに成功した。

この 3つの貢献により、現在のクラウドコンピューティングや VDI といった仮想計算機の提供システムを、技術に明るくない人であっても使用させることが可能になる。普段は、VoXY を使用して Web ブラウザ上で VM を操ることができる。VM はクラウド上にあり、ユーザは電源管理や、ハードウェアの故障によるデータロスといった心配に悩まされることなくコンピュータを使うことができる。また、より詳しい人であれば、シンクライアントソフトウェアで利用するには少々困難であるソフトウェア (3D ゲーム, CAD, CAM) を利用したい場合、WinKVM と VM ストレージマイグレーションの高速化を利用して、一時的にクラウド上から VM を、手元の物理計算機に、実行コンテキストを維持したまま高速に移動させることができる。使用が終わった後は、ふたたび、安全なクラウド上に VM を移動させる。このように、VoXY と WinKVM と VM ストレージマイグレーションの高速という研究を組み合わせることで、より便利なコンピューティング環境をユーザに対して提供することができる

以下、3つの研究についてより詳しく述べる。

(1) JavaScript と HTML を用いたシンクライアント VoXY では、Web ブラウザ上で動作するユーザに取って導入が容易なシンクライアントシステムを達成した。多数の OS インストールパッケージには、Web ブラウザが初期導入されているため、ほとんどすべての Web ブラウザ搭載端末で動作することが可能である。新たなアプライアンスやアプリケーション

ンを別途インストールする必要はない。

VoXY の開発を始めた 2007 年当初、Web ブラウザ上ですべてのコンピューティングを行おうという潮流が存在し、クラウドコンピューティングというキーワードも誕生した。Web ブラウザ上で動作する、いわゆる Web アプリケーションには、ユーザは新たにソフトウェアをインストールする必要がないという利点が存在し、利用可能性の向上に大きく寄与した。この状況に倣い、Web ブラウザ上でシンクライアントシステムを実現すれば、シンクライアントシステムの利用可能性も同じく向上できると考えた。

しかし、JavaScript と HTML のみでインタラクティブ性能の高いシステムを実現するにあたっては、既存のシンクライアントシステムが採用しているプロトコルをそのまま流用することが不可能であった。なぜなら、貧弱な画像描画システムしか持たない JavaScript と HTML では、従来の、変更点のみを転送するシンクライアントシステムのプロトコルでは高いインタラクティブ性を実現することが不可能であった。もちろん、Web ブラウザに拡張性を持たせる専用のプラグインを導入することで、Web ブラウザに高機能な画像描画システムをもたせることは可能であるし、事実そのような研究も存在している。しかし、専用のプラグインは規格化されているわけではないので、限定的な環境でしか動作せず、ユーザにとって導入が容易であるとは言えない。そこで、われわれは、純粋に規格化された JavaScript と HTML を利用して、いかなる環境でもインタラクティブ性の高いシンクライアントシステムを実現する手法を提案する。具体的には、VoXY では、計算機のフレームバッファをある程度の大きさを持つタイルに区切り、タイルを単位とした転送を行うことで、この問題を解決する。従来のシンクライアントプロトコルに比べると、使用帯域は大きくなるが、HTML と JavaScript 環境が動作する Web ブラウザであれば、いかなる環境でも動作するシンクライアントシステムであるため、ユーザにとって導入が容易であり、使いやすいものである。そのため、「使う人を選ばず」「使う環境が限定されていない」遠隔仮想計算機システムを VoXY で実現することに成功する。

また (2) Linux ドライバの Windows 移植による WinKVM の実装では、Linux カーネルに付属する VMM (KVM: Kernel-based Virtual Machine) の Windows 移植手法について論じる。利用可能性の高い仮想計算機転送技術を実現するためには、様々なオペレーティングシステム上で動作する Hybrid VMM の実現が重要である。コンシューマ向けのオペレーティングシステムである Windows と、サーバ/サービス構築用として利用される Linux との間で VM ライブマイグレーションを実現することで、ユーザは、自由に自分の作業環境である VM (仮想マシン) を実行コンテキストを維持したまま移動させることができる。これにより、ユーザは、シンクライアントシステム VoXY 経由で扱うことが難しい CAD、CAM や 3D ゲームといったソフトウェアを使いたいときは、自分の環境を直ちに手元の計算機に

転送することができる。これは、仮想マシンの利用可能性に寄与する。

しかし、VM ライブマイグレーションのプロトコルは、統一規格が存在せず、VMM 毎に独自プロトコルを用いている。そのため、現状、多くの VMM があるが、それらすべての VM ライブマイグレーションの VM ライブマイグレーションプロトコルを統一するのは現実的ではない。そのため、本研究では Linux カーネルに標準搭載されている VMM である KVM を Windows 上に移植することで、Windows と Linux 間で VM ライブマイグレーションを実現する。

KVM を Windows に移植するためには Linux のデバイスドライバを Windows に移植する必要がある。KVM は特権命令である CPU の仮想化支援命令を使用しているため、KVM のコアの部分は Linux のデバイスドライバとして実装する必要があるのである。したがって、WinKVM を実現するためには、Linux デバイスドライバを Windows ドライバとして移植する必要がある。

デバイスドライバというプログラムは OS に強く依存するため、ユーザ空間で動作するプログラムに比べると移植が難しい。そのため、プログラマがドライバのソースコードを解釈し、他の OS に移植するように書き換えるのは人的コストが高い。とくに問題となる (x) 実行可能な Privilege な命令の種類、(y) ユーザプログラムとのインタフェース (z) カーネルとのインタフェースという 3 つの OS 間の差異の 3 つである。

そこで、この研究では、WinKVM を実現するために、Windows カーネル上に Linux デバイスドライバを動作させるためのエミュレーションレイヤを手法を提案する。エミュレーションレイヤは上述した 3 つの差異を吸収し「簡単」かつ「性能劣化無し」で Windows 上で Linux ドライバである KVM ドライバを動作させることができた。しかし、この手法には制限もあり、Linux ドライバとして実装された KVM には 10 行程度の修正を加える必要がある。しかし、Linux / KVM ドライバを 10 行程度の修正でかつ性能劣化なく移植可能な本手法は、「使う人を選ばず」「使う環境が限定されていない」という限定的な遠隔仮想計算機転送技術である VM ライブマイグレーションに貢献する技術である。

最後に (3) 差分転送による VM ストレージマイグレーションの高速化では、VM が特定の物理マシンの間を行ったり来たりする VM の往来と呼ばれる挙動に着目し、VM を含めた VM ストレージの転送を高速化の手法を提案した。従来、VM ライブマイグレーションでは、分散ファイルシステムや共有ファイルシステムといったシステムを併用し、転送先と転送元で、予め VM ストレージを共有しておくのが一般的であった。しかし、複数のデータセンタ間で VM を転送する、広域ライブマイグレーションや、個人用の VM ライブマイグレ

ーションでは、分散ファイルシステムや共有ファイルシステムを利用することが困難となる。そこで、この問題を解決すべく、VM のストレージを含めて他の物理マシンに VM を転送する、VM ストレージマイグレーションと呼ばれる技術が提案されているが、VM のストレージは数十 GB になることも珍しくない。これほど巨大なデータを転送するには、VM が完全に転送し終わるまでに、多くの時間と帯域を要求する。そこで、本研究では、データセンタ間での広域ライブマイグレーションや、個人用の VM ライブマイグレーションでは、VM が特定の物理マシンの間を往来する可能性が高いという点に着目し、この性質を利用した、VM の差分転送メカニズムを開発する。

具体的には、VMM 内に DBT (Dirty Block Tracking) と、世代番号と呼ばれる概念を追加し、いつどの物理マシンで、ディスクの土の汚れたかを追跡しておき、VM ライブマイグレーション時には、差分のみを正確に取り出し、差分転送を行うメカニズムを開発する。差分転送というアイデアは、古今東西様々な場面で使用されてきた古典的な手法ではあるが、今回は、特定の物理マシン間を VM が往来するという現象に着目した差分転送を提案したことが貢献となる。ベンチマークの結果、90%以上の転送時間を削減することができ、既存の VM ストレージマイグレーションと比べて高速な VM 転送を達成した。

以上、(1) VoXY では JavaScript+HTML で利用できる新たなシンククライアントプロトコルを開発することで、「誰でも容易に使える」Web ブラウザ上で動作するシンククライアントシステムを実現し、利用可能性を向上する。さらに、(2) WinKVM では、Linux/KVM ドライバを簡単かつ性能劣化なく Windows 上で動作する事ができるエミュレーションレイヤを開発し、使う環境を選ばない VM ライブマイグレーションを達成し、利用可能性を更に向上する。最後に(3) 差分転送による VM ストレージマイグレーションの高速化では、分散/共有ファイルシステムを使用できない環境下における VM ライブマイグレーションでは、VM が特定の物理計算機上を往来するという挙動を利用した高速な差分転送メカニズムを提案する。

この3つの研究によって、本博士論文で解決すべき問題である**現在の限定的である VM とユーザと結びつける遠隔計算機転送システムを、より多くの利用シーンでも利用できるような、遠隔仮想計算機転送システムを提案する**を達成した。「使う人を選ぶ」かつ「使う環境を選ぶ」遠隔仮想計算機の利用可能性を向上させたことが本博士論文の貢献である。