

仮想マシンモニタにおけるリアルタイム通信を保証する仮想 NIC の研究

鈴木 健一† 宮田 宏†
佐藤 未来子† 並木 美太郎†

1. はじめに

近年、仮想マシン (VM) にリアルタイムアプリケーションを適用する研究¹⁾²⁾の進展にともない、音声通話や制御通信など通信に対してもリアルタイム性を確保可能な仮想マシン環境が求められている。しかし、VM 間でリアルタイム通信を実現するためには、二つの大きな課題がある。一つ目は、ゲスト OS が送信したトラフィックの帯域の合計が物理 NIC(Network Interface) を超えるときに仮想マシンモニタ (VMM) 層において必要となるパケットスケジューリングによる遅延である。二つ目は、VM 間 (End-to-End) における通信のデッドライン保証である。

本稿では、VMM 層に付加するリアルタイム通信制御機構によりこの課題を解決し、VM 間のリアルタイム通信を実現する通信基盤 RTvNIC システムの設計を述べ、試作したシステムの評価結果について考察する。RTvNIC システムでは、ゲスト OS が VMM の提供するリアルタイム通信向け仮想 NIC(RTvNIC) を利用することで、ゲスト OS の改変無しに VM 間のリアルタイム通信を行うことができる。VM 管理者は、VM 起動時のパラメータとして各 RTvNIC にパケットの転送遅延時間のデッドライン時間と最低帯域を設定することができる。RTvNIC を持つ VM 間の通信は、VMM 層の通信制御によりこの設定値を守った通信が保証される。なお、RTvNIC システムでは、DiffServ(Differentiated Services)³⁾ による QoS 制御が行われているネットワークを対象としている。

2. RTvNIC システムの設計

図 1 に全体構成を示す。DiffServ で制御されたネットワークを介して二台の物理マシンが接続され、RTvNIC を持つ VM 間でリアルタイム通信を行う。VMM 層

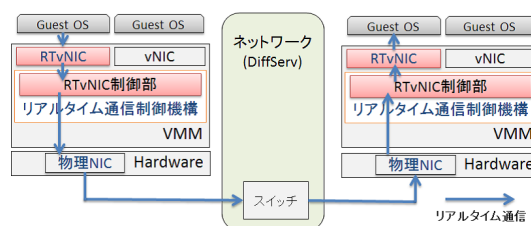


図 1 RTvNIC システムの全体構成

のリアルタイム通信制御機構では、RTvNIC およびリアルタイム性を持たない通常の仮想 NIC(vNIC) の両方を制御対象とし、ゲスト OS が送受信する全てのパケットを統一的に制御する。

リアルタイム通信制御機構における RTvNIC 制御部では、デッドライン時間を守るパケットスケジューリングとデッドラインミス情報を基にしたフィードバック制御の両方により、VMM 層におけるパケットのリアルタイム性を確保する。これらの設計については、2.1, 2.2 節で述べる。また、デッドライン時間に応じて DiffServ の優先度情報をパケットにマーキングすることで、ネットワークにおけるパケットのリアルタイム性を確保する。この優先度の決定は 2.2 節で述べるフィードバック制御と連携して行う。

RTvNIC 制御部では、デッドラインミス検出のため、パケットの送信時刻を各送信パケットに付加する。また、ネットワーク上の優先制御のため、パケットの IP ヘッダーに対して DiffServ の DSCP(Differentiated Services Code Point) 値を設定する。なお、付加した情報はパケットを受信した VMM で取り外されるため、ゲスト OS に対して隠蔽される。

2.1 リアルタイムパケットスケジューリング

RTvNIC 制御部では、RTvNIC ごとに設定されたデッドライン時間と帯域を守るために、各 VM の送信パケットに対して、EDF アルゴリズムを用いたパケットの送信順序の制御 (リアルタイムパケットスケジューリング) を行う。リアルタイムパケットスケジューリングでは、RTvNIC に設定された帯域を超えない範囲で、VM の送信したパケットが VMM 層の送信キュー

* Study on virtual network device on hypervisor to guarantee real-time communication

† 東京農工大学 (Tokyo University of Agriculture and Technology)

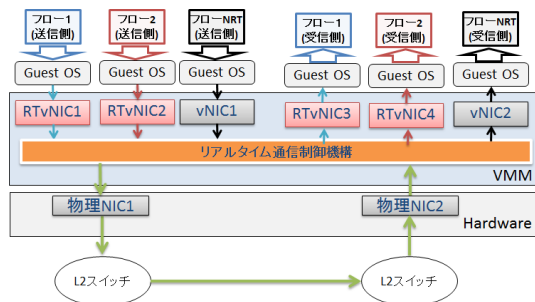


図 2 評価実験の構成

に到着した時刻に各パケットのデッドライン時間を加えたデッドライン時刻に近い順に物理 NIC から送信を行う。できるだけ多くのパケットのデッドライン時間を守るようパケットスケジューリングを行い、VMM 層におけるパケットのリアルタイム性を確保する。

2.2 リアルタイムパケットに対するデッドライン制御方式

RTvNIC 制御部では、End-to-End での通信のリアルタイム性が確保されていることを確認、保証するために、VMM 層において各パケットの転送遅延時間の計測を行う。この転送遅延時間がデッドライン時間以内であれば、リアルタイム性の確保が確認できる。しかし、ネットワーク上の遅延などの影響によりデッドライン時間を超えた場合デッドラインミスとして検出される。デッドラインミスを検出した場合、デッドラインミスしたパケットの送信元 VMM はデッドライン超過時間のフィードバック情報を受け取り、さらなるデッドラインミスの発生を回避する対処を行う。送信元 VMM では、受け取ったフィードバック情報をもとにデッドラインミスしたフローに対するリアルタイムパケットスケジューリングのデッドライン時間を短く設定することで、全体の転送遅延時間を削減しデッドラインミスを回避する。また、当該フローの DiffServ における優先度を上げることで、ネットワーク内の転送遅延時間を削減しデッドラインミスを回避する。

3. 評価実験と考察

基礎評価として、RTvNIC システムを適用した仮想マシン環境における VM 間通信のパケットの転送時間の計測を行う。これを RTvNIC システムを用いない場合の結果と比較し、リアルタイム通信に重要な要素である最悪遅延時間の保証を確認する。同一の物理マシン上に RTvNIC を持つ VM と通常の NIC(vNIC) を VM を図 2 のように配置し、この RTvNIC 間の UDP 通信のパケット転送時間を計測する。RTvNIC1, 2,

表 1 パケットの転送遅延時間の計測結果

遅延時間	RTvNIC 適用		RTvNIC 適用なし	
	フロー 1	フロー 2	フロー 1	フロー 2
最大 (ms)	2.1	8.5	49	38
最小 (ms)	0.9	1.7	4.0	5.5
平均 (ms)	1.5	3.2	9.8	10.2

vNIC1 が送信するパケットは物理 NIC1 から送信され、理想状態のネットワークとした L2 スイッチを経由し物理 NIC2 を介して RTvNIC3, 4, vNIC2 で受信する構成となっている。VMM 層では、各 VM の送受信パケットに対してリアルタイムパケットスケジューリングを行い、各パケットが設定されたデッドライン時間以内に転送されるよう優先制御する。この構成で次の条件で実験を行った。計測用の二つのフローとして、RTvNIC1 から RTvNIC3 へ転送されるフロー 1 は 5ms, RTvNIC2 から RTvNIC4 へ転送されるフロー 2 は 10ms のデッドライン時間を設定する。さらに、vNIC 間では物理 NIC の帯域である 1Gbps のノンリアルタイムフローを常に転送し、物理 NIC 部分が輻輳する状態とする。評価結果を表 1 に示す。提案手法である RTvNIC システムを適用した場合はフロー 1, 2 ともに最大遅延時間がデッドライン時間以内となり、適用しない場合はどちらもデッドライン時間を越えている。これは、提案手法ではフロー 1, 2 のデッドライン時間に応じたパケットスケジューリングが行われ、ノンリアルタイムフローより優先的に送信された結果、最大遅延時間を抑えられたと言える。これより、VMM 層のリアルタイム通信制御機構におけるリアルタイムパケットスケジューリングが、VM 間通信のリアルタイム性保証に有効であることを確認した。

4. おわりに

本稿では、VMM 層のリアルタイム制御通信機構によりリアルタイム通信を実現する通信基盤 RTvNIC システムを提案、評価した。現在、OpenFlow と VMM の協調動作によるネットワークまで含めたリアルタイム性の保証手法を検討しており今後研究を進めていく。

参考文献

- 1) Luwei Cheng, et al. Defeating Network Jitter for Virtual Machine, UCC IEEE, p.65-72, 2011
- 2) 太田貴也, 他, 組込みマルチコア向け仮想化環境における性能低下抑止手法, 情報処理学会研究報告 2012-EMB-27(12) p.1-8, 2012
- 3) S.Blake: An Architecture for Differentiated Services, IETF RFC Standard 2475, 1998