

WS クラスタを用いた CORErouter プロトタイプの評価

2Bb-7

丸山 充 三栄 武 小倉 毅 高橋 直久

NTT ソフトウェア研究所

1 はじめに

我々は、応答性、信頼性、サービス性、可用性、保守性、安全性の高いネットワーク制御機能およびネットワークサービス機能を備えた高並列ルータ (CORErouter) の研究を進めており、第一段階として WS クラスタを用いたプロトタイプを作成した^{1), 2)}。本稿では、プロトタイプの子備評価としてルーティングプロトコル処理がフォワード性能に与える影響をとりあげ、市販ルータとの比較により評価を行なう。

2 ルーティングプロトコルが与える影響

ルータの主要な機能は、ルーティングテーブルに基づき、パケットをフォワードする処理とルータ間で経路情報を交換し、ルーティングテーブルを作成するルーティングプロトコル処理に大別される。

現在、外部組織とのルーティングに使用されるプロトコルとしては、BGP (Border Gateway Protocol) が主に使われている³⁾。BGP は、経路情報を交換する BGP スピーカ同士でルート情報と該ルートが伝達されてきた組織番号のパス情報を交換する。また情報交換は TCP を使って行ない、コネクション開設時に、それぞれのルータが持つ全経路情報を交換し、その後はそれぞれのルータのテーブル情報に追加/削除がある毎に、更新メッセージを通知しあう。また定期的な keepalive により、コネクションの確認も行なう。

ある経路情報の追加/削除が頻繁に繰り返されることを経路のフラップと呼ぶ。経路のフラップがおきると受けとった BGP スピーカでは、そのたびに最適経路の再計算を行なうために、CPU の使用率が上がる。また削除の場合にはキャッシュエリアのクリアなどを行なうために、検索処理とテーブルアクセスが競合し、フォワード性能に影響を及ぼす。この現象は幹線のルータのように、複数箇所の BGP スピーカと経路情報の交換をしている部分では顕著である。実際、ある幹線ルータでは、1秒間に70回/秒の追加/削除が報告されており、現在の高性能な市販ルータを使っても、CPU ボトルネックを起している。

Evaluation of a Parallel IP Router Prototype CORErouter-I.
Mitsuru MARUYAMA (mitsuru@ntt-20.ntt.jp), Takeshi MIEI (take@slab.ntt.jp), Tsuyoshi OGURA (ogura@slab.ntt.jp) and Naohisa TAKAHASHI (naohisa@slab.ntt.jp)
NTT Software Laboratories
9-11 Midori-cho 3-Chome Musashino-sh, Tokyo 180 Japan.

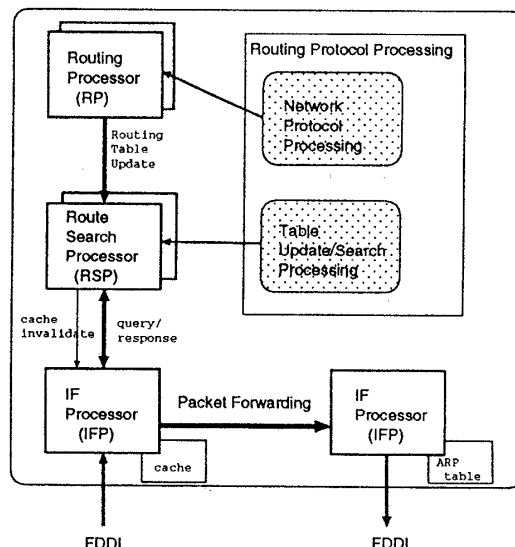


図1: ルーティングプロトコルの分散処理

3 ルーティングプロトコルの分散処理

ルーティングプロトコルの処理が、フォワード性能に影響を与えることは、QoSの観点からも問題であるため、我々の提案している CORErouter では、図1に示すように、ルーティングプロトコルの処理をネットワークプロトコル処理機能とテーブル管理/検索機能に分離し、前者をルーティングプロセッサ (RP)、後者をルートサーチプロセッサ (RSP) に機能分散して実行している。また、対応する BGP スピーカの箇所が増加する場合には、RP と RSP の組を複数台用意し、ネットワークプロトコル処理機能を複数の RP に分散し CPU 負荷を下げ、それとともに起きるルーティングテーブルの頻繁なアクセス競合に対しては、ルーティング情報を更新中の RSP と検索中の RSP を物理的に分け、次々に切替えることで、インタフェースプロセッサ (IFP) からの検索要求が待たされない構成を採る。その結果、ルーティングテーブル検索、キャッシュ格納、パケットフォワード転送という一連のパイプライン動作の性能があがり、フラップに対しても所要の性能を得るようになっていく。

4 ルーティングプロトコルによる影響

ここでは BGP をとりあげ、経路情報のフラップがもたらす影響を定量的に測定する。本測定のために、特

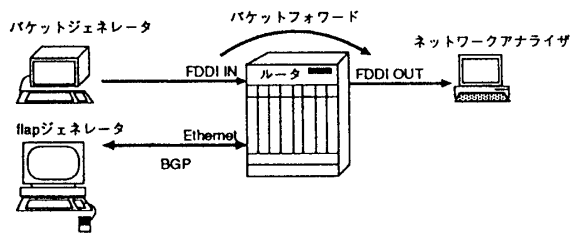


図 2: フラップ試験環境

定の経路情報の追加/削除を意図的に起こすようなフラップジェネレータを gated (GateDaemon)⁴⁾ の改造により作成した。

オリジナルの gated は独自の radix tree 構造のテーブルを持っており、このテーブルを変更すると接続された他の BGP スピーカに対して、更新フレームを送出する。そこで、一定時間おきにテーブルに経路を追加/削除する特別な flap タスクを追加することで、任意の間隔/経路数の更新フレームを発生できるようにしている。

このフラップジェネレータを図 2 のように試験ルータのイーサネットに接続し、2 つ持つ FDDI インタフェースの入力側から 80Mbps 程度の一定レートの IP パケットを入力し、出力側の FDDI インタフェースからフォワードされたパケットをネットワークアナライザにより観測した。

この結果、図 3 に示すように市販ルータでは、80Mbps 程度では問題なくフォワードできるが、フラップジェネレータによる 70 回/秒程度のフラップ情報を入力すると、すぐにフォワード性能がほぼ 1/2 まで減少する様子が見られる。

一方、我々の CORErouter においては、同様に 80Mbps 程度のフォワードをおこない、フラップを入力しても全く性能に影響がないことがわかる。

5 おわりに

ルータのルーティングプロトコル処理がフォワード性能に与える事を述べ、一例として BGP による経路のフラップの問題点を示した。またこの問題を解決するために、我々の並列ルータにおけるルーティングプロトコル処理の分散処理方式の概略を示し、その有効性を検証するために、実際に経路のフラップをエミュレートできる環境を構築し、フォワード性能への影響を現在の市販ルータとの比較により明らかにした。

今後、BGP 以外の様々なルーティングプロトコルの実装を行ない、ルーティングプロトコル処理の分散処理方式が同様に適用できることを明らかにすると共に、限界性能などの詳しい評価を進める予定である。

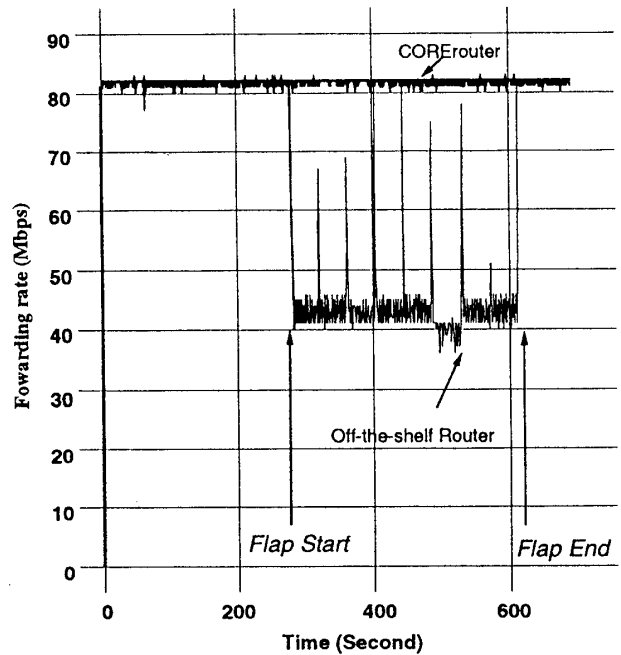


図 3: 評価結果

謝辞

本研究の機会を与えていただいた、NTT ソフトウェア研究所 広域コンピューティング研究部の後藤滋樹部長、ならびに日頃御討論いただく超並列コンピュータグループの皆様へ感謝いたします。

参考文献

- 1) 高橋, 丸山, 三栄, 小倉, “柔軟でスケラブルな高性能並列ルータ CORErouter の基本構想”, 情処第 52 回大会, 2B-5, 1996.
- 2) 三栄, 小倉, 丸山, 高橋, “WS クラスタを用いた CORE router プロトタイプの構成”, 情処第 52 回大会, 2B-6, 1996.
- 3) Y. Rekhter, T. Li: “A Border Gateway Protocol 4 (BGP-4)”, IETF RFC-1771 03/21/1995.
- 4) GateDaemon Project and the GateDaemon Consortium: “GateDaemon (Gated)”, <ftp://ftp.gated.merit.edu/research.and.development/gated/gated-r3.5alpha.11.tar.gz>