優先制御ポリシデリゲート*

小畠 元† 小川 晃道‡ 杉浦 一徳‡
慶應義塾大学 環境情報学部† 慶應義塾大学 政策・メディア研究科‡

概要

インターネットにおいてボトルネックは無くならない．一般的に優先度の高い特定のトラフィックに対しトラフィックの優先制御を行う事で通信品質の低下を防ぐ．しかし，NATやIPv4-v6トランスレータを含むネットワークでは現状のトラフィック優先制御は有効に機能しない．本研究では，NATやIPv4-v6トランスレータといったトラフィック情報を隠蔽化する機構の存在するネットワークにおいて有効なトラフィック優先制御を行う機構を構築した．

1 背景

経路制御技術の発展や運用経験の蓄積によってインターネットの安定性は着実に向上し，社会基盤としての重要な位置を占めるようになった．ネットワークの規模や利用者数，利用者層は，今なお急速な拡大を続けている．その結果，利用者のニーズは多様化し，様々なサービスがインターネットに求められるようになった．

そのひとつに，通信品質への要求を挙げられる．多種多様なアプリケーションがインターネット上で展開される事により，単なる到達性だけではなく，通信の品質の向上が求められている．

一方，ネットワーク資源である物理回線は広帯域化しているが有限である．トラフィックの偏り，回線帯域格差などの理由から混雑した回線，ボトルネックの存在は解消できない．特にインターネットのバックボーンへの接続を行う回線の利用料金は内部インフラに比べ高額で，広帯域な回線を確保しにくいためボトルネックになりやすい．

ボトルネックではパケットロスが発生し，通信品質が低下する．パケットを転送するルータは標準でF IFOキューをそれぞれOutputとInputに持ち，それぞれにおいてtail dropで転送能力以上のトラフィックは破棄されてしまう．

* "Delegation of priority control policy"
Gen KOBATAKE† Akimichi OGAWA‡ Kazunori SUGIURA‡
E-mail:gen@sfc.wide.ad.jp
Faculty of Environmental Information, Keio University†
Graduate School of Media and Governance, Keio University‡
Keio University Shonan Fujisawa Campus. 5322, Endo, Fujisawa, Kanagawa 252, Japan

2 トラフィック優先制御技術

ボトルネックにおいて通信品質を保つためには二つの解法が考えられる．

- キューの長さを伸ばす

  キューの長さに限りがあることでtail dropが発生するので，キューの長さを伸ばすことでパケットロスを発生させない解法．ルータのメモリに限界があること，ボトルネックにおいてはキューの長さが延びる一方になってしまい，結果多大な遅延を発生させる可能性があるなど，実現は困難を伴う．

- トラフィックに優先度をつける

  多様化したトラフィックに優先度をつけ，優先度の高いトラフィックは低いトラフィックに比べtail dropの発生しにくい環境をつくる解法．

一般的にボトルネックに直接つながるルータにおいてトラフィックの優先制御をおこない，優先度の高いトラフィックに対し通信品質を保証する．図1にトラフィック優先制御の一例である優先度キューを示す．特定のトラフィックに対し別のキューを用意する事により，tail dropから発生するパケットロスによる通信品質の低下を防ぐ．

左のFIFOではパケットを順番にキューにいれて入りきらなかったパケットは全て破棄している．それに対し右の優先度キューはパケットの優先度を判断し，優先度に応じたキューに配分する．優先度の高いキューからパケットは送信される．結果特定のトラフィックにボトルネックを占有されることがなくなる．

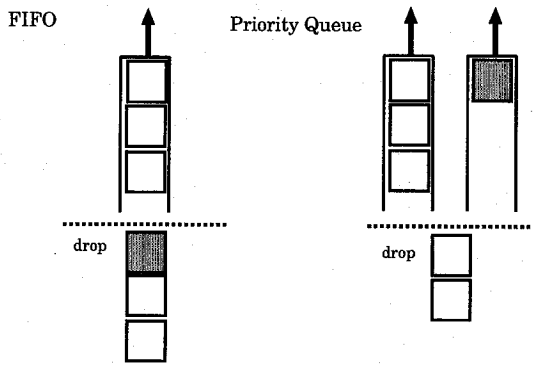

図 1: キューイング

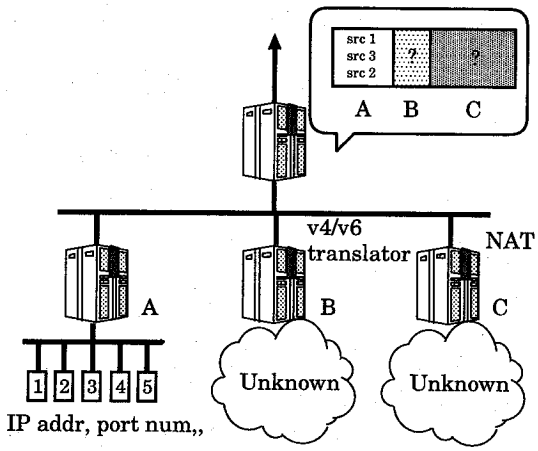

図 2: ネットワーク構成例

ボトルネックにおいてすべてのトラフィックを tail drop を発生させずに転送するのは不可能である．よって以上に示される様にトラフィックを優先度別に区別化し，優先度の高い順番に転送を行う．こういったトラフィック優先制御機構は大きく二つのプロセスに分けられる．

- トラフィックの識別

  優先度をトラフィックの種類から識別する．トラフィックの種類は送信先 IP アドレス，送信元 IP アドレス，送信先ポート番号や送信元ポート番号により識別する．

- 優先度別転送

  優先度に対応したパケット転送を行う．

3 問題点

トラフィック優先制御プロセスのひとつであるトラフィックの識別は下位ルータがトラフィックに対して変更を行う等といった事を行わない際には実行できる．しかし IPv4 アドレスが欠乏している場合，セキュリティーなどの理由から Firewall として NAT を導入する事があり，おなじ理由や IPv6 推進のために IPv4-v6 トランスレータを導入している事がある．これらトランスレータはトラフィックの情報を隠蔽化してしまうため，トラフィックの識別をする際障害となる．

図 2 にこれらトランスレータを導入したネットワークの例を示す．

アップリンク，それに繋がる上位ノード，下位ノードが三つあるネットワークを示している．それぞれの下位ノードはエンドノードが直接繋がっている．そのなかにはトランスレータを利用するもの (NAT, v4/v6) とトランスレータを利用しない下位ノードがある．

上位ノードはトランスレータを利用しないノード A からのトラフィックは分類を容易にできるため優先制御できる．しかしトランスレータを利用するノード B，C から発生するトラフィックは分類できない。これはトランスレータが IP アドレス、ポート番号などを変換し上位ルータに対し隠蔽化してしまうからである．

図 2 の様な構成は企業などネットワークにおける末端の組織にみられる構成である．問題点を簡略化し，解決策を明確にするために本章以降では図 2 のネットワーク構成を想定して説明する．

4 解決手法と設計

特定の帯域を下位ノードに割り当て、その帯域に対するポリシの決定権限を下位ノードに委譲する。この手法により下位ノードで隠蔽化されたトラフィックへの優先制御が可能となる．

4.1 ポリシの決定権限を委譲

図3に委譲の過程を示す.

1. まず上位ノードは一階層下のノードに対して帯域の割り当てをアップリンクの帯域を考慮した上で決定する.
2. 決定した帯域の割り当てを下位ノード ABC に伝える.
3. またノード C の様に更に下層がある場合は下層にポリシの決定権限を委譲する.
4. それを受けた下位ノードは割り当てられた帯域内でトラフィック優先制御を行う.

この際, 下位ノードからは上位ノードとボトルネックまでのリンクに仮想的に物理線を与えられた様に見える.

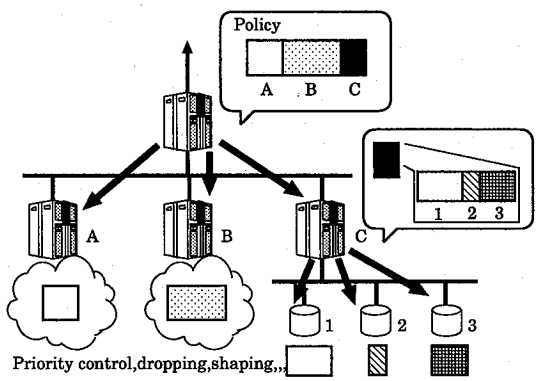

図 3: ポリシの委譲

ポリシーの決定権限を委譲することで各ノードは一段階上のリンクに対する制御を一階層下の情報を監視する事だけでトラフィック優先制御が成り立つ. 下層がどのトラフィックを優先するかを上層では考慮する必要がなくなる.

4.2 必要な機構

以上の一連の流れを成立させるため3つの機構が必要となる.

1. 割り当てた帯域とそれに対する権限を下位ノードに委譲する機構

   上位ノードと下位ノードで取り引きされるメッセージは帯域の上限を下位ノードに伝える, 下位ノードから受理したという返答の2種類が考えられる. 帯域の上限を下位ノードに伝えそれ以降の処理を委任する.

2. 上位ノードからメッセージを受け取る機構

   常に上位ノードからのメッセージを受け取れるよう待つ機構が必要. そのためにはプロトコルを取り決める必要がある.

3. 受け取ったメッセージをトラフィック優先制御機構に反映する機構

   帯域制限メッセージをうけ上限の値を変える. キューイング技術, 例えば ALTQ/HFSC の設定ファイルを書き換えることで反映する.

4.3 詳細

4.3.1 例外の考察

ボトルネックを利用しない内部同士のトラフィックに対して制限を誤って加える事があってはならない. この様なトラフィックは例外処理しなければならない.

4.3.2 帯域割当の決定

帯域の割り当てを静的に行うには以下の二つの点において問題が残る.

- 静的な設定ならばポリシのデリゲートを必要としない

  ネットワークの管理者が静的に NAT などのトランスレータに対してのみ静的にトラフィック優先制御の設定を行えばよい.

- ボトルネックを有効に利用できない.

  優先的なトラフィックを選ぶことで限りある回線資源を利用する事が目的なのに対し, ボトルネックを静的に割り振る場合, 使用できる回線帯域が余剰に存在する瞬間にも割り振られた帯域の制限により余剰を利用することができない.

以上を考慮した上でポリシのデリゲートをイベントドリブン, 又は定期的に行い, 静的ではなく動的に帯域の割り当てを行っていく必要がある.

4.3.3 付加価値

動的な変更は以前述べた仮想回線を付加価値のある物に変える．ボトルネックとなる回線を複数の組織で共用している場合を想定する．既存の機構ではボトルネックの最大帯域幅をそれぞれの組織の使える上限帯域としてまとめて提供するか，または料金体系などからボトルネックの帯域を静的に割り振るということが考えられる．こういった場合，前者後者共に各組織において保証のできない利用できる帯域上限を提示する事しかできない．しかし，動的な仮想回線を与えることで下限を設定でき，品質の保証された帯域として提示することが可能となる．

5 実装

5.1 実装環境

本機構を以下のような環境で実装をおこなった

- OS: FreeBSD4.4-Release
- 言語: C 言語
- 使用ライブラリ: libpcap,libaltq
- トラフィック優先制御機構: altq-hfsc

5.2 回線使用状況の抽出機構

libpcap を用い layer2 のトラフィック量を監視した．下位ルータ毎のトラフィックを抽出し，目安量以上のトラフィックが発生した場合に割り当て決定機構を呼び出す．

5.3 帯域資源の割り当て判断機構

まずそれぞれ下位ルータ毎に余剰帯域を計算する．目安量以上のトラフィックを発生させているルータ使用帯域目安にその余剰帯域を加算し，下位ルータにある転送機構へ適応させる機構に渡す．

5.4 転送機構への適応を行う機構

各ルータにおいてトラフィック優先制御機構として altqd を作動させた．ポリシのデリゲートに対応させるため libaltq を用い，altqd を止めること無く動的に数値変更可能にした．

6 実験

以上の機構を利用し，2001 年 9 月の WIDE 合宿において実験を行った．WIDE 合宿とは 300 人程度の参加者が集まり毎年 2 回行われるイベントである．

6.1 WIDE 合宿での実験

実験ネットワークでは，一本の対外線の帯域を共有した．そこで，共有の割合を動的に変える仕組みを構築した．割当は流量の監視，情報の集約，優先制御ポリシの適応の 3 段階で成り立つ．今回優先制御ポリシの決定には各セグメントに存在する人数とトラフィックの流量を利用した．図 3 にその構成を示す．

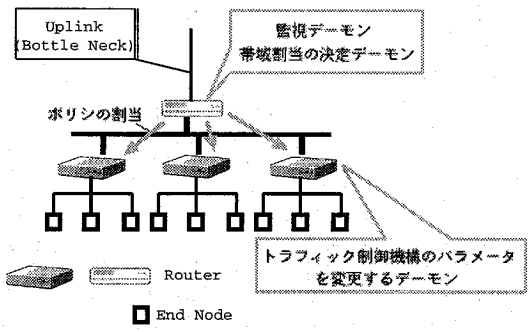

図 4: 合宿ネットワーク

6.2 実験内容

WIDE 合宿での実験を元に動的な帯域割り当て判断機構，および割り当てを実行する機構の設計および実装を行った．

実験は図 3 に示されたネットワークにおいて行った．以下の点については別途合宿のネットワークに適応させる仕組みとして設計，実装した．

- ポリシによる帯域割当の決定デーモン

  イベントの関係上，セグメント間の人の移動が多かった．公平性を求めるため，各セグメントに存在するノードの数による帯域割当を一つのポリシとした．ノード数を 5 秒毎に報告させ，その都度帯域制限の計算を行った．

- ルータ間通信を行うデーモン

  監視デーモンを動作させているルータと各セグメントにあるルータで，必要なメッセージの交換を行う機構を実装した．メッセージは人数の通知と帯域割当の二種類とした．

7 評価

本機構を用いる事によって第3章でのべた問題点であるトランスレータを介したトラフィック優先制御を行うことが可能になった．静的割り当てによる帯域の無駄を省く機構ができた．

WIDE合宿における実験のデータ解析は現在進行中であるが，ノード数による帯域割り当てを採用したことにより特定のノードがボトルネックを占有することが不可能になった．よってDoSアタックなどによる被害を最小限に押えることができた．

7.1 トランスレータを介したトラフィック優先制御

トランスレータの下に位置するエンドノードのトラフィックに対しボトルネックを通過する上で優先制御できているか観測する．実際にトランスレータを導入した場合，トラフィックの情報が隠蔽化されてしまうのでトランスレータと仮定して実験し，定量評価とする．

7.2 ボトルネックの回線利用率

WIDE合宿の実験では5秒毎にノード数の調査に合わせて帯域割り当てを計算することで動的な制御を行った．このインターバルを最適化させ，回線利用率が最大になる状態を目指す．

8 結論

本研究ではトラフィック優先制御のポリシを上位ルータから下位ルータに委譲する手法を提案した．本手法によりNATやIPv4-v6トランスレータを含むネットワークにおいてトラフィックの優先制御を行うことが可能となった．

9 今後の展望

第4章で述べた機構の内以下の点については実現していない．

- 内部ルーティングのトラフィックに対する例外を設置する事

  宛先IPアドレス別にトラフィック優先制御機構を設定することで回避は可能である．

- 双方向トラフィックの優先制御

  ボトルネックを利用するトラフィックの半分にあたる上位ルータからボトルネックに流れるトラフィック優先制御は実現したが，反対方向のトラフィックの制御を行っていない．

参考文献

[1] P.Ferguson,G.Huston," インターネット QoS"pp.74-94, オーム社,May 2000

[2] S.Blake,D.Black,M.Carlson,E.Davies,Z.Wang,W.Weiss, "An Architecture for Differentiated Services", RFC2475,December 1998

[3] S.Floyd,V.Jacobson, "Link-sharing and Resource Management Models for packet Networks",IEEE/ACM Transactions on Networking, August 1995

[4] Y.Tamura,Y.Tobe,H.Tokuda, "NBQ:Neighbor-state Based Queuing for Adaptive Bandwidth Sharing", IEEE International Conference on Network Protocols(ICNP'99),Nov 1999

[5] G.Kobatake,A.Ogawa,K.Sugiura, "Delegation of Priority Control Policy", 情報処理学会第63回全国大会論文集 (3)pp.245-246(2001)